Biostatisztika
Biostatisztika
Tartalom

I. Első rész ................................................................. 1
  1. Bevezetés ................................................................... 5
  2. Statisztikai alapfogalmak ............................................. 8
      1. Véletlen jelenségek leírása, valószínűségi változók ............... 8
      2. Mérési skálák ........................................................... 9
      3. Gyakoriság, relatív gyakoriság, valószínűség ..................... 9
      4. Eseményalgebra, a valószínűség grafikai értelmezése .......... 11
      5. Feltételes valószínűség ............................................. 13
      6. Teljes valószínűség tétele ........................................ 14
      7. Bayes tétel ................................................................ 14
      8. Diagnosztikus eljárások a biostatisztika szemszögéből .......... 15
  3. Leíró statisztika: a populáció és a minta jellemzése ............. 20
      1. Adatok csoportosítása .............................................. 20
          1.1. A közép jellemzésére használt statisztikák .................... 20
          2. A közép jellemzésére használt statisztikák összehasonlítása 21
          3. A szóródás jellemzésére használt statisztikák ............... 21
      4. Az eloszlások leírására használható módszerek ............... 23
      5. Az eloszlás alakjának tömör jellemzése ........................ 29
      6. Standardizálás ....................................................... 30
      7. Eloszlások grafikus megjelenítése .............................. 30
      8. Bevezetés a klinikai vizsgálatokba .............................. 37
      9. A klinikai vizsgálatok fázisai .................................... 40
  4. Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások (Szőlőssi János dr.) ....... 42
      1. Diszkrét eloszlások ................................................. 42
          1.1. Binomiális eloszlás .............................................. 42
          1.2. Poisson-eloszlás ............................................... 44
      2. Folytonos eloszlások .............................................. 46
          2.1. Egyenletes eloszlás ............................................ 46
          2.2. Exponenciális eloszlás ....................................... 47
          2.3. Normális eloszlás ............................................. 49
          2.4. Lognormális eloszlás ....................................... 58
          2.5. ?2 eloszlás ..................................................... 60
          2.6. A t-eloszlás ..................................................... 61
          2.7. Az F-eloszlás .................................................. 63
  5. Becslések, referencia értékek, megbízhatósági tartományok ......... 66
      1. A statisztikai következtetésről ................................. 66
      2. A statisztikai becslés ............................................. 67
          2.1. A fejezet tanulmányozásához szükséges előismeretek ........ 67
          2.2. A pontbecslés .................................................. 67
          2.3. Az intervallumbecslés ....................................... 67
            2.3.1. Intervallumbecslés -ra és alapján ....................... 68
            2.3.2. Intervallumbecslés -re az és alapján ................... 69
            2.3.3. Az ismeretlen helyett használjuk az ismert sx-et! Hogy változik a megbízhatóság? ......................... 70
            2.3.4. Az ±t*sx/n intervallumbecslés a -re .................. 73
            2.3.5. Mintafeladatok a -re vonatkozó ±t*sx/n intervallumbecslésre .......... 75
      3. Összefoglalás .................................................... 75
          3.1. A gondolatmenet összefoglalása ........................................ 75
          3.2. A várható érték megbízhatósági intervalluma: az intervallumbecslés felidézése 76
  4. Ismeretlenörző és gyakorló kérdések és feladatok .................. 77
  6. A hipotézisvizsgálat ............................................. 79
      1. A fejezet tanulmányozásához szükséges előismeretek ........ 79
      2. A hipotézisvizsgálatról ......................................... 79
      3. A hipotézisvizsgálat gondolatmenete. A „hétköznapi” hipotézisvizsgálat, mint modell. 80
4. A hipotézisvizsgát a biometriában ................................................................. 81
  4.1. Az egymintás t próba ........................................................................ 81
  4.2. Három fontos megjegyzés .................................................................. 82
      4.2.1. A nullhipotézisről ...................................................................... 82
      4.2.2. A döntési szintről ...................................................................... 83
      4.2.3. A döntésről ............................................................................. 83
  5. A hipotézisvizsgát további két alkalmazása .............................................. 83
    5.1. Páros t próba ..................................................................................... 83
    5.2. Az előjel-proba ............................................................................... 84
  6. Összefoglalás ............................................................................................. 85
      6.1. A hipotézisvizsgálat gondolatmenete ............................................. 85
      6.1.1. A hipotézisvizsgálat működése: az egymintás t próba ............ 85
  7. Ismeret-ellenőrző és gyakorló kérdések és feladatok .............................. 86
  8. Csoportok összehasonlítása folytonos eloszlásból származó adatok esetén 98
    8.1. A fejezet tanulmányozásához szükséges előismeretek ................. 98
    8.2. Bevezetés .......................................................................................... 98
      8.2.1. Összertartozó és független minták .......................................... 98
      8.2.2. Paraméteres és nemparaméteres próbák .............................. 98
    9. Két összetartozó minta összehasonlítása ............................................. 99
      9.1. Példa ............................................................................................... 99
      9.2. Paraméteres módszer – páros t-próba ....................................... 99
          9.2.1. A példafeladat megoldása páros t-próbával ...................... 100
      9.3. Nemparaméteres módszerek, előjelpróba, Wilcoxon-féle előjelrangpróba 101
          9.3.1. A példafeladat kiértékelése előjelrangpróbával ............. 102
    4. Két független csoport összehasonlítása ................................................. 103
      4.1. Példa ............................................................................................... 103
      4.2. Paraméteres módszer – kétmintás t-próba .................................. 104
          4.2.1. Atlagok összehasonlítása, kétmintás t-próba ................. 104
          4.2.2. A variánciák összehasonlítása ............................................ 105
      4.3. A példafeladat kiértékelése kétmintás t-próbával ....................... 105
      4.4. Nemparaméteres módszer – Mann-Whitney próba .................. 106
          4.4.1. A példafeladat kiértékelése Mann-Whitney U-próbával ...... 107
    5. Két csoport összehasonlítására vonatkozó feladatok és kérdések ......... 108
    9. Kettőnél több csoport összehasonlítása folytonos adatok esetén ......... 111
      9.1. Paraméteres eljárások, varianciaanalízis ..................................... 111
          9.1.1. Egyszempontos varianciaanalízis .................................... 112
          9.1.1.1. Példa .............................................................................. 112
          9.1.2. Az egyszempontos varianciaanalízis elve ....................... 112
          9.1.3. Többszörös összehasonlítások ......................................... 115
          9.1.4. Az egyszempontos varianciaanalízis feltételei ................. 116
          9.1.5. A varianciaanalízis elvégzéséhez, a számításokhoz szükséges képletek. 116
                9.1.5.1. A példafeladat megoldása ...................................... 118
      2. Nem normális eloszlású sokaságok összehasonlítása, nemparaméteres próbák ...... 120
          2.1. Több összertartozó minta nemparaméteres összehasonlítása: Friedman próba 120
             2.1.1. Példa ............................................................................. 120
          2.2. Több független minta nemparaméteres összehasonlítása: Kruskal-Wallis próba 121
3. Több csoport összehasonlítására vonatkozó feladatok és kérdések 123
  3.1. Feladat, számítógépes „játék” az Excel programmal 123
  3.2. Egy varianciaanalízis eredménye a következő táblázatban látható 124
  3.3. Három félé tanítási módszer tesztelésére a hallgatókat véletlenszerűen 3 csoportba sorolták. A maximum 100 pontos teszteredmények a következők voltak: 125
4. Összefoglalás 125
5. Irodalomjegyzék 9. fejezethez 125

10. KORRELÁCIÓN ÉS REGRESSZIÓ ANALIZIS 127
  1. Korreláció számítás 127
  1.1. Bevezetés 127
  1.2. Lineáris korreláció (parametriktus korreláció) 127
    1.2.1. Korrelációs együttható szignifikancia 134
  1.3. Népparaméteres korreláció 137
    1.3.1. Spearman–féle rangkorreláció 137
    1.3.2. Kendall–féle rangkorreláció 138
  2. Regresszióanalízis 139
    2.1. Bevezetés 139
    2.2. Kétváltozós lineáris regresszió 139
      2.2.1. Együttható konfidencia intervalluma 141
      2.2.2. A regresszió ANOVA táblázata 142
  11. Döntések előfordulási gyakoriságokról. Kontingencia táblázatok 146
  1. Nominális változók és a kontingencia táblázat 146
  2. Függetlenségvizsgálat khi-négyzet próbával 2-szer 2-es táblázat esetén 148
  3. Függetlenségvizsgálat khi-négyzet próbával nagyobb méretű gyakorisági táblázatok esetén 153
  4. Az összefüggés erősségének jellemzésére használt mérőszámok 156
  12. Egy adatsor illeszkedése eloszláshoz 158
  1. A statisztikai illeszkedésvizsgálat célja és fajtái 158
  2. Egyenletes eloszlásra történő illeszkedésvizsgálat 158
    2.1. 1. példa. 158
    2.2. 2. példa. 159
  3. Tiszta illeszkedésvizsgálat 159
II. Második rész 167
  13. Két faktor szerinti keresztosztályozás 169
    1. Példa 169
  14. Két faktor szerinti keresztosztályozás 180
    1. Rögzített és véletlen faktor 182
    2. Egy véletlen faktor szerinti varianciaanalízis 182
    3. A példa megoldásának folytatása (5. táblázat) 184
    4. Keresztosztályozás két véletlen faktor szerint 184
  15. Keresztosztályozás egy rögzített és egy véletlen faktor szerint: véletlen blokk 188
    1. Példa 188
    2. A példa folytatása 190
    3. Hierarchikus osztályozás 191
    4. Példa 191
      4.1. A varianciakomponensek becsleése 193
    5. Megjegyzés a kísérleti tervek és a modellek alkalmazásához 193
  16. Többváltozós elemzések: korreláció és regresszió 195
    1. Bevezetés 195
    2. Többszörös korreláció 195
    3. Többváltozós regressziószámítás 196
      3.1. Regressziós eljárások 197
      3.2. Logisztikus regresszió 197
      3.3. Probit regresszió 199
      3.4. Poisson regresszió 199
    4. Nem-lineáris regresszió 199
III. Harmadik rész 201
  17. Szoftverhasznlati útmutató 203
    1. A statisztikai programokról 203
18. Az interneten elérhető szabad szoftverek, shareware szoftverek, kereskedelmi programcsomagok ................................................................. 204
   1. Ingyenes programok .................................................................................................................. 204
      1.1. Rice Virtual Lab in Statistics: Analysis Lab .......................................................... 204
      1.2. JavaStat .............................................................................................................................. 204
      1.3. StatCrunch (WebStat) ...................................................................................................... 204
      1.4. VassarStats (Richard Lowry) ............................................................................................ 204
      1.5. The R Project for Statistical Computing ........................................................................ 204
   2. A statisztikai próbák ereje ........................................................................................................... 205
19. Milyen feladathoz milyen szoftvert használjunk? ................................................................. 206
   1. Adat-tárak az interneten .................................................................................................................. 206
20. Az egyszerű statisztikai számításokhoz optimális szoftverek köre ........................................ 207
   1. Általános célú programok .............................................................................................................. 207
   2. Mintaszám meghatározás ............................................................................................................. 207
   3. Adat amputáció ............................................................................................................................ 207
21. Feladatmegoldás táblázatkezelővel ...................................................................................... 208
   1. Az első fajta hiba megadása ........................................................................................................ 208
22. Biostatisztika, szimulációk az interneten (angol nyelven) .................................................. 209
   1. Oktató programok az interneten ............................................................................................. 209
   2. Introductory statistics: concepts, models, and applications (D.W. Stockburger) ... 210
   23. Biostatisztikai szimulációk statisztikai szoftvercsomagokkal ............................................. 210
      1. Oktató programok az interneten ............................................................................................. 210
      2. Introductory statistics: concepts, models, and applications (D.W. Stockburger) .... 210

Irodalomjegyzék ........................................................................................................................... 211
A táblázatok listája

2.1. Emberek magasságának vizsgálatára vett 15 elemű minta eloszlása ................................. 18
2.2. A testmagasság eloszlása nemenként .................................................................................. 18
2.3. A 2.2. táblázatban feltüntetett magasságeloszlás feltételes valószínűségek felhasználásával ... 19
2.4. Diagnosztikus eljárások lehetséges eredményei a betegség fennállásától függően. .......... 19
2.5. A .......................................................................................................................................... 19
3.1. Látóterenkénti sejtszám-adatok csoportosítása ....................................................................... 41
3.2. A közép és a szóródás jellemzésére használt statisztikák szemléltetése ............................ 41
4.1. ................................................................................................................................................ 43
4.2. Élesztő sejtek eloszlása a hemocitométerben. (Student adatai) ................................................. 46
5.1. $= 170 \text{ cm} = 8$ paraméterű normális eloszlás esetén adott intervallum esés valószínűssége, és adott mintaelemszámnál az intervallumon belül és kívül eső elemek várható aránya. ........................................ 68
5.2. $= 170 \text{ cm} = 8$ paraméterű normális eloszlás esetén 16 elemű minták átlagainak adott intervallumba esésének valószínűsége, és adott csoportoszámokkal az intervallumon belül és kívül eső elemek várható aránya. ........................................ 68
5.3. ................................................................................................................................................ 70
5.4. ................................................................................................................................................ 71
5.5. ................................................................................................................................................ 74
6.1. A diétával kombinált tréning hatása ....................................................................................... 83
8.1. Egy diéta-kísérlet adatai .......................................................................................................... 99
8.2. Az előjeles rangpróba elvégzése a 8.1 táblázat adatain. ......................................................... 102
8.3. Diéta kísérlet adatai ................................................................................................................ 103
8.4. A 8.3. táblázat adatai a testsúlyváltozás (diéta előtt és után különbsége) szerint rendezve, rangszámokkal. ................................................................. 107
8.5. A 8.3. táblázat adatai az eredeti sorrendben, rangszámokkal. .............................................. 107
8.6. ................................................................................................................................................ 109
8.7. ................................................................................................................................................ 109
9.1. ................................................................................................................................................ 112
9.2. A varianciaanalízis táblázata ............................................................................................... 117
9.3. A varianciaanalízis táblázata a mintapélda adataira ................................................................ 118
9.4. A Dunnett próba eredménye a mintapélda adataira .............................................................. 119
9.5. Többszörös összehasonlítások LSD módszerrel a mintapélda adataira ................................. 119
9.6. Többszörös összehasonlítások Bonferroni módszerrel a mintapélda adataira ................... 119
9.7. Mintaadatok a Friedman próbahoz ....................................................................................... 120
9.8. A próbaadatok ragsorolása ................................................................................................... 121
9.9. Rangsúszmok, rangsúszmósszegek és átlagok a mintapélda adataira ................................. 122
9.10. ............................................................................................................................................... 123
9.11. Többszörös összehasonlítás a mintapélda adataira. ............................................................. 123
9.12. Egy varianciaanalízis eredménye ....................................................................................... 124
10.1. .............................................................................................................................................. 133
10.2. Egy tantárgyra fordított felkészülési idő és az elért eredmény .............................................. 135
10.3. Regressziós ANOVA tábla .................................................................................................. 142
10.4. Adatok az ANOVA táblazathoz ......................................................................................... 144
10.5. ANOVA tábla ........................................................................................................................ 144
11.1. "Hétköznapi" táblázat ........................................................................................................ 146
11.2. Gyakorisági táblázat .......................................................................................................... 146
11.3. Normális valószínűségi változó ............................................................................................ 147
11.4. Kétdimenziós gyakorisági táblázat ..................................................................................... 147
11.5. Kontingencia táblázat a megfigyelt (observed) gyakorisági értékekkel ............................... 148
11.6. Segéd-kontingencia táblázat a várható (expected) gyakoriság értékekkel .......................... 149
11.7. Kontingencia táblázat a megfigyelt gyakorisági értékekkel ............................................... 152
11.8. Segéd-kontingencia táblázat a várható gyakorisági értékekkel ........................................... 153
11.9. $r$-szer $s$ mezős kontingencia táblázat megfigyelt gyakoriságokkal a vastagított részben, valamint fejssorral, fejesszöppel és peremgyakoriságokkal ............................................................... 153
11.10. $r$-szer $s$ mezős segéd-kontingencia táblázat a várat gyakoriságokkal ................................. 154
11.11. Kontingencia táblázat a megfigyelt gyakorisági értékekkal ............................................... 155
11.12. Segéd-kontingencia táblázat a várható gyakorisági értékekkel ........................................... 155

Created by XMLmind XSL-FO Converter.
11.13. Kontingencia táblázat a megfigyelt gyakorisági értékekkel .................................................. 156
12.1. .................................................................................................................................................. 158
12.2. .................................................................................................................................................. 159
12.3. .................................................................................................................................................. 159
12.4. .................................................................................................................................................. 160
12.5. .................................................................................................................................................. 164
13.1. Kísérleti eredmények .................................................................................................................. 169
13.2. Az ANOVA-tábla általánosan ................................................................................................. 174
13.3. .................................................................................................................................................. 175
13.4. 4. táblázat .................................................................................................................................. 178
14.1. A reciprok értékek táblázata, kiegészítve a sor- és oszlop-átlagokkal .................................. 180
14.2. Az ANOVA-táblázat .................................................................................................................. 180
14.3. Kémiai elemzés .......................................................................................................................... 183
14.4. Az ANOVA-táblázat .................................................................................................................. 183
14.5. .................................................................................................................................................. 184
14.6. .................................................................................................................................................. 184
14.7. .................................................................................................................................................. 185
14.8. A példa adatainak kiértékelése ............................................................................................... 186
15.1. .................................................................................................................................................. 188
15.2. .................................................................................................................................................. 189
15.3. .................................................................................................................................................. 190
15.4. .................................................................................................................................................. 191
15.5. ANOVA táblázat ....................................................................................................................... 192
15.6. A varianciaanalízis eredményei .............................................................................................. 193
16.1. .................................................................................................................................................. 195
I. rész - Első rész
Tartalom

1. Bevezetés .......................................................................................................................... 5
2. Statisztikai alapfogalmak .................................................................................................. 8
   1. Véletlen jelenségek leírása, valószínűségi változók ......................................................... 8
   2. Mérési skálák .................................................................................................................... 9
   3. Gyakoriság, relatív gyakoriság, valószínűség ................................................................... 9
   4. Eseményalgebra, a valószínűség grafikai értelmezése ...................................................... 11
   5. Feltételez valószínűség .................................................................................................... 13
   6. Teljes valószínűség tétele ................................................................................................ 14
   7. Bayes tétele ..................................................................................................................... 14
   8. Diagnosztikus eljárások a biostatisztika szemszögéből .................................................... 15
3. Leíró statisztika: a populáció és a minta jellemzése .......................................................... 20
   1. Adatok csoportosítása ....................................................................................................... 20
       1.1. A közép jellemzésére használt statisztikák ................................................................. 20
   2. A közép jellemzésére használt statisztikák összehasonlítása ............................................. 21
   3. A szóródás jellemzésére használt statisztikák ................................................................. 21
   4. Az eloszlások leírására használható módszerek ............................................................... 23
   5. Az eloszlás alakjának tömör jellemzése .......................................................................... 29
   6. Standardizálás .................................................................................................................. 30
   7. Eloszlások grafikus megjelenítése .................................................................................... 30
   8. Bevezetés a klinikai vizsgálatokba .................................................................................. 37
   9. A klinikai vizsgálatok fázisai .......................................................................................... 40
4. Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások (Szöllősi János dr.) .......................................................... 42
   1. Diszkrét eloszlások ........................................................................................................... 42
       1.1. Binomális eloszlás ...................................................................................................... 42
       1.2. Poisson-eloszlás ...................................................................................................... 44
   2. Folytonos eloszlások ........................................................................................................ 46
       2.1. Egyenletes eloszlás .................................................................................................... 46
       2.2. Exponenciális eloszlás ............................................................................................... 47
       2.3. Normális eloszlás ....................................................................................................... 49
       2.4. Lognormális eloszlás ............................................................................................... 58
       2.5. ?2 eloszlás ................................................................................................................ 60
       2.6. A t-eloszlás ............................................................................................................... 61
       2.7. Az F-eloszlás ............................................................................................................. 63
5. Becslések, referencia értékek, megbízhatósági tartományok .................................................. 66
   1. A statisztikai következtetésről ........................................................................................ 66
   2. A statisztikai becslés ........................................................................................................ 67
      2.1. A fejezet tanulmányozásához szükséges előismeretek .................................................... 67
      2.2. A pontbecslés ............................................................................................................. 67
      2.3. Az intervallumbecslés ............................................................................................... 67
         2.3.1. Intervallumbecslés -ra és alapján ............................................................... 68
         2.3.2. Intervallumbecslés -re az és alapján ............................................................ 69
         2.3.3. Az ismeretlen helyett használjuk az ismert sx-et! Hogy változik a megbízhatóság? ................................................................. 70
         2.3.4. Az ±t*sx/n intervallumbecslés a-re ................................................................. 73
         2.3.5. Mintafeladatok a-re vonatkozó ±t*sx/n intervallumbecslésre ................................................. 75
   3. Összefoglalás .................................................................................................................... 75
      3.1. A gondolatmenet összefoglalása ................................................................................ 75
      3.2. A várható érték megbízhatósági intervalluma: az intervallumbecslés felidézése ... 76
   4. Ismeretlenörző és gyakorló kérdések és feladatok ................................................................ 77
6. A hipotézisvizsgálat ........................................................................................................... 79
   1. A fejezet tanulmányozásához szükséges előismeretek .................................................... 79
   2. A hipotézisvizsgálatról .................................................................................................... 79
   3. A hipotézisvizsgálat gondolatmenete. A „hétköznapi” hipotézisvizsgálat, mint modell. ... 80
   4. A hipotézisvizsgálat a biometriában ................................................................................ 81
      4.1. Az egymintás t próba .................................................................................................. 81
4.2. Három fontos megjegyzés ................................................................. 82
  4.2.1. A nullhipotézisról ............................................................... 82
  4.2.2. A döntési szintről ............................................................... 83
  4.2.3. A döntéskérdésről ............................................................... 83
5. A hipotézisvizsgáló összehasonlítása folytonos adatok esetén .................. 83
  5.1. Páros t-próba .......................................................... 83
  5.2. Az előjel-próba ............................................................. 84
6. Összefoglalás .......................................................................................... 85
  6.1. A hipotézisvizsgáló gondolatmenete ............................................ 85
  6.1.1. A hipotézisvizsgáló működése: az egymintás t-próba .................. 85
7. Ismeret-ellenőrző és gyakorló kérdések és feladatok ................................ 86
8. Összefoglalás .......................................................................................... 90
9. Előjel- és a másodfajú hiba .................................................................... 92
  9.1. Hogyan csökkenthetjük ezt a kockázatot? ..................................... 92
  9.2. Mi az optimális? (döntési szint)? ................................................. 93
     9.2.1. Egy hétköznapi példa: a bírósági ítéllet ................................. 93
     9.2.2. Egy új gyógyszer bevezetése és mellékhatása. Az? megválasztása .... 93
5. Összefoglalás .......................................................................................... 94
6. Ismeret-ellenőrző és gyakorló kérdések és feladatok ................................ 94
8. Csoportok összehasonlítása folytonos eloszlásból származó adatok esetén ........................................................................................................ 98
  8.1. A fejezet tanulmányozásához szükséges előismeretek ...................... 98
  8.2. Bevezetés ...................................................................................... 98
     8.2.1. Összetartozó és független minták ........................................ 98
     8.2.2. Paraméteres és nemparaméteres próbák ................................ 98
3. Két összetartozó minta összehasonlítása ................................................ 99
     3.1. Példa ..................................................................................... 99
     3.2. Paraméteres módszer – páros t-próba ................................... 99
        3.2.1. A példafeladat megoldása páros t-próbával ....................... 100
     3.3. Nemparaméteres módszerek, előjelpróba, Wilcoxon-féle előjeles rangpróba ... 101
        3.3.1. A példafeladat kiértékelése előjeles rangpróbával .............. 102
4. Két független csoport összehasonlítása .................................................. 103
     4.1. Példa ..................................................................................... 103
     4.2. Paraméteres módszer – kétmintás t-próba ................................ 104
        4.2.1. Atlagok összehasonlítása, kétmintás t-próba ....................... 104
        4.2.2. A varianciák összehasonlítása ....................................... 105
     4.3. A példafeladat kiértékelése kétmintás t-próbával ...................... 105
     4.4. Nemparaméteres módszer – Mann-Whitney próba .................... 106
        4.4.1. A példafeladat kiértékelése Mann-Whitney U-próbával .......... 107
5. Két csoport összehasonlítására vonatkozó feladatok és kérdések ............ 108
9. Kettőnél több csoport összehasonlítása folytonos adatok esetén ............... 111
  9.1. Paraméteres eljárások, varianciaanalízis ....................................... 111
     9.1.1. Egyszempontos varianciaanalízis .................................. 111
        9.1.1.1. Példa ......................................................................... 112
     9.2. Az egyszempontos varianciaanalízis elve .................................. 112
     9.3. Többszörös összehasonlítások ......................................................... 115
     9.4. Az egyszempontos varianciaanalízis feltételei ................................ 116
     9.5. A varianciaanalízis elvégzéséhez, a számításokhoz szükséges képletet. 116
        9.5.1. A példafeladat megoldása .............................................. 118
  2. Nem normális eloszlású sokaságok összehasonlítása, nemparaméteres próbák ................................................................. 120
     2.1. Több összetartozó minta nemparaméteres összehasonlítása: Friedman próba ... 120
        2.1.1. Példa ............................................................................. 120
     2.2. Több független minta nemparaméteres összehasonlítása: Kruskal-Wallis próba . 121
3. Több csoport összehasonlítására vonatkozó feladatok és kérdések ......... 123
     3.1. Feladat, számítógépes „játék” az Excel programmal. ...................... 123
     3.2. Egy varianciaanalízis eredménye a következő táblázatban látható .... 124
3.3. Három féle tanítási módszer tesztelésére a hallgatókat véletlenszerűen 3 csoportba sorolták. A maximum 100 pontos teszteredmények a következők voltak: 125

4. Összefoglalás 125

5. Irodalomjegyzék 9. fejezethez 125

10. KORRELÁCIÓ ÉS REGRESSZIÓ ANALIZIS 127

1. Korreláció számítás 127

1.1. Bevezetés 127

1.2. Lineáris korreláció (parametríkos korreláció) 127

1.2.1. Korrelációs együthato szignifikanciája 134

1.3. Nemparaméteres korreláció 137

1.3.1. Spearman–féle rangkorreláció 137

1.3.2. Kendall–féle rangkorreláció 138

2. Regresszióanalyzis 139

2.1. Bevezetés 139

2.2. Kétváltozós lineáris regresszió 139

2.2.1. Együtthatók konfidenciaintervalluma 141

2.2.2. A regresszió ANOVA táblázata 142

11. Döntések előfordulási gyakoriságokról. Kontingencia táblázatok 146

1. Nominális változók és a kontingencia táblázat 146

2. Függetlenségvizsgálat khi-négyzet próbával 2-szer 2-es táblázat esetén 148

3. Függetlenségvizsgálat khi-négyzet próbával nagyobb méretű gyakorisági táblázatok esetén 153

4. Az összefüggés erősségének jellemzésére használt mérőszámok 156

12. Egy adatsor illeszkedése eloszláshoz 158

1. A statisztikai illeszkedésvizsgálat célja és fajtáí 158

2. Egyenletes eloszlásra történő illeszkedésvizsgálat 158

2.1. 1. példa. 158

2.2. 2. példa. 159

3. Tisza illeszkedésvizsgálat 159

1. fejezet - Bevezetés

A biostatisztika az élővilággal kapcsolatos jelenségeket matematikai módszerekkel elemző tudomány. Módszerei lefedik az élő természettudományok legkülönbözőbb elméleti és gyakorlati ágait, a kísérlet megtervezésétől a kísérleti eredmények értékeléséig.

A körírtörténeti leírásokban, a laboratóriumi naplókban, a kísérleti jegyzőkönyvekben igen sok számszerű adat szerepel. Ismeretes pl., hogy hazánkban igen magas a koraszülések (2500 g alatti magzat) száma, de a szív- és érrendszer eredetű halálokok is magas arányt (54%) képviselnek a hazai halálzati statisztikákban, más országbeli adatokkal összevetve. Véletlent-e ez vajon, vagy esetleg az alkoholfogyasztásnak, a dohányzásnak, étkezési szokásoknak (többek között) is köze van ehez statisztikai adatok ilyenek alakulásához? Ilyen és hasonló kérdések megválaszolásához szolgáltat módszereket a biostatisztika. Fontos, hogy az adatok helyes értékelését, a helyes következtetések levonásának módszereit és az eredmények megbízhatóságának vizsgálatát az orvosok, fogorvosok, gyógyszerészek megismerjék és munkájukban alkalmazni tudják.

Itt és a továbbiakban a kísérlet (mérés) kifejezést általánosabb értelemben használva: kísérletnek (mérésnek) olyan – akár mesterségesen előidézett, akár spontán lejátszódó – jelenség (állapot, illetve állapotot jellemző valamilyen fizikai mennyiség) megfigyelését (megmérést) nevezzük, amelynek lefolyása (kialakulása) a véletlentől is függ. A mérések célja mindig valaminek a megismerése, valamilyen kérdésnek a megválaszolása, és akár egy pénzértéknél vagy egy dobókockának a feldobása is.

A kísérlet (mérés) eredményét általában számértékkel fejezzük ki, de vannak olyan kísérletek, amelyek során valamilyen esemény bekövetkezett, vagy elmaradását, esetleg valamilyen tulajdonság megléte vagy hiányát észlelünk, ez utóbbi esetben az esemény vagy tulajdonság előfordulásának a gyakoriságát számszerűen (pl. a korábban érettségizettek száma egy hallgatói csoportban).

Válasszuk ki példaképpen a pulzusszámmérést, amely könnyen elvégezhető. A pulzusszám folytonos változó és csak az egyszerűség kedvéért, illetve megszokásból használunk diszkrét (egész) értékeket. Mértékegysége az 1/perc. Ezek tisztázása után a továbbiakban elegendő csupán a mérőszámokkal dolgoznunk. Sokféle kérdésünk lehet, amire ettől a méréstől várjuk a választ.

Például:

1. **MEKKORA** X.Y. páciens pulzusszáma? Normális érték ez? (**MEKKORA** a normális pulzusszám?)

2. **VÁLTOZIK** a pulzusszám mondjuk valamilyen szívritmus szabályozó gyógyszer hatására?

3. **VAN-E KÜLÖNBSÉG** a nők és a férfiak pulzusszáma között?, stb.

A feltett kérdések egyikére sem tudunk korrekt választ adni egyetlen mérés alapján, ezért többször mérünk. Az első kérdés esetében a méréseket ugyanazon a személyen ismételjük meg, azonban a normális érték megállapításához már több különböző személyen kell mérnünk pulzusszámot. A további kérdésekre pedig csak még bonyolultabb módon tudunk választ adni.

A „statisztika” szó több jelentéssel is bír. E könyvben használt értelmét leginkább a latin eredetre visszanyúlva a „status” szóból ismerhetjük meg. A status eredeti jelentése: állapot, helyzet, a dolgok állásának módja. Ennek megismeréséhez, illetve leírásához biztosítanak lehetőséget az adatok. Általánosságban a környező világ minőségi és mennyiségi jellemzőit nevezhetjük adatoknak. Mindennapi életünkben leggyakrabban előforduló adatok például a személyi adataink: nevünk, születési helyünk, születésünk időpontja, ..., vagy akármilyen boltozatának az árucikkek neve, az árucikkek ára, ..., de egészségi állapotunkkal kapcsolatban is mondhatunk példát: arcunk sápadtsága, vérnyomásunk, hőmérsékletünk, de akármilyen laboratóriumi diagnosztikai vizsgálat eredménye is adat.

Normális esetben adatokat csak valamilyen cél érdekében gyűjtünk. Például megkérdezzük valakinek a telefonszámát, hogy később fel tudjuk hívni.

5

Created by XMLmind XSL-FO Converter.
Az összegyűjtött, de rendezetlen adatok önmagukban, sok esetben teljesen használhatatlanok. Gondoljunk például arra, hogy egy telefonkönyv adatait a központba való beérkezésük sorrendjében adnád közre, mire mennénk vele?

Sokszor kell az adatokat értékelni, például fontosságuk szerint. Ezt teszi az orvos is, amikor felállít egy diagnózist, vagy amikor a beteg állapotáról nyilatkozik. Amikor úgy ítéli meg, hogy említésére érdemes változás nem történt a beteg állapotában, akkor ezt orvostársainak a „statusa változatlan” (status idem) kifejezéssel jelzi.

Az adatokat tehát össze kell gyűjteni, legtöbbször fel kell dolgozni őket, és szükség esetén következtetéseket kell levónni belőlük, illetve időnként döntéseket kell hozni azok alapján. Mindez egy adott helyzet értékelésére, statisztikának helyes megállapítását jelenti. A statisztika az a tudomány, amely ezen eljárások szabályait, korlátait vizsgálja, korrekt módszereket fogalmaz meg.

A „statisztika” szó még ezek után is egy nagyon elvont, a közönséges halandók számára meglehetősen érthetetlen fogalmat takar. Mindezt annak ellenére állíthatjuk, hogy aki Magyarországon az utódban néhány évtizedben járt iskolába, az életében már néhányszor biztosan használt statisztikai módszereket. Elég csak azt az esetet említeni, amikor egy kisiskolász gyorsan a nagyobb tárgyakból szerzett érdemjegyeinek átlagát abból a célból, hogy megtudja, milyen osztályzatot számított a félévi vagy év végi bizonyítványában. Ilyenkor a tanuló a statisztikai becslés egy tipikus esetét alkalmazta, nagy valószínűséggel anélkül, hogy tudott volna róla.

A „statisztika” szó előtt esetszámok szerelével a „bio-” előtag. A biostatisztika, amelynek színvonalaként gyakran a biometria felfedezését is használják, azt a többlet informatíciót tartalmazza, hogy ezt a tudományt elsősorbaként az élettudományosok segédtudományaként alkalmazzák. Tehát előbb vagy utóbb mindazok találkoznak vele, akik az élettudományokkal valamilyen kapcsolatba kerülnek, így például az egyetemi hallgatók, doktorandusok, kutatók, gyakorló orvosok.

Lássunk néhány példát az alkalmazási területekről:

A szakirodalom értő tanulmányozása, értékelése. Talán ez a legáltalánosabban, legtöbbek által használt és emiatt a legfontosabb terület. Hogy a doktorandusok illetve a kutatók esetében erre miért van szükség, azt nem kell külön hangsúlyozni, hiszen elsősorban így lehet hozzájutni a legújabb tudományos eredményekhez. A szakirodalom tanulmányozásának elhangyagolása az ő esetükben a kutatómunka folytatásának akadályát jelentheti. Ugyanis, ha valaki nagyon leszakad a tudományos élvonaltól, az saját eredményeit sem tudja publikálni és ez végős soron a kutatómunka anyagi fedezetének megszűnéséhez vezethet.

„A jó pap holtig tanul” szókált mondanak, a jó orvos pedig holtig képezi magát, mondhatják, hiszen egyetlen végzett orvos sem gondolhatja azt, hogy az egyetemi évek alatt nem kis fáradtsággal megszerzett tudása örökérvényű. Ezért a szakirodalom rendszeres, értő olvasása a klinikai orvosok esetében sem nélkülözhető. Aki az élettudományok szakirodalomát olvassa, annak azért kell érteine a biostatisztikához, hogy őröl kétségézetet fenntartva el tudja dönteni egy publikációról, hogy az abban közölt eredmények mennyire megbízhatók, és ezért értékelésre szokták megpróbálni azokat, amelyek nincsenek megfelelően megalapozva, azért elég sok olyan is átjut a szűrőn, ahol például formális statisztikai analízis alkalmazott, mert olyanok is könnyen átvonhatók a biostatistika tekintetében.

Az orvosi gyakorlatban nagyon fontos szerepe van a rendszeres továbbképzés mellett a rendszeres fejlesztésnek is. Egy gyakorló orvosnak illik tudnia, hogy a legújabb diagnosztikai eljárások közül melyek a legjobbakként az adott célokra és az adott körülményekhez nézve. Ezért az orvosnak rendszeresen értékelnie kell az újabb és újabb gyógyszereteket, eszközöket, berendezéseket azzal a reklám árazattal szemben is, ami napjainknál körülveszi a gyártó cégeket. Ugyanis, ha ezt nem tudják megfogalmazni, nem lehetnek bizonyos statisztikai eredmények kétmintás tervének, nem van bizonyítékként, hogy a bizonyítékként használt kifejezés szerint érvényes, ha ez nem a legújabb diagnosztikai technika, ha nem a legalkalmasabb módszer, ha nem a legelsők közé tartozik.

Az összegyűjtött, de rendezetlen adatok önmagukban, sok esetben teljesen használhatatlanok. Gondoljunk például arra, hogy egy telefonkönyv adatait a központba való beérkezésük sorrendjében adnák közre, mire a páciens jobb, korszerűbb ellátásának segítése.

Az orvosi gyakorlatban nagyon fontos szerepe van a rendszerez továbbképzés mellett a rendszeres fejlesztésnek is. Egy gyakorló orvosnak illik tudnia, hogy a legújabb diagnosztikai eljárások közül melyek a legjobbakként az adott célra és melyek az optimális kezelési módszerek. Azt is ismernie kell, hogy ezekhez milyen új anyagok és eszközök tartoznak, illetve, alkalom adtán, neki kell választania a lehetőségeket közül. Ezért az orvosnak rendszeresen értékelnie kell az újabb és újabb gyógyszereteket, eszközöket, berendezéseket azzal a reklám árazattal szemben is, ami napjainknál körülveszi. A gyártó cégek ugyanis mind azt szeretnék bizonyítani, hogy az ő készülékeik az ő anyagaik jobbak, mint a másiké, azért hogy tőlük válsárolják meg a termékeket. Az orvosnak ilyenkor azért szükségesek a biostatisztikai
ismeretek, hogy gyorsan és jól tájékozódjon az összehasonlító diagramok, grafikonok, táblázatok rengetegében
annak érdekében, hogy a szebbnél szebb kiállítású kiadványokat, illetve a termékeket népszerűsítő előadásokat
kritikusan, tartalmi szempontból tudja értékelni. Pl. „A reklám készítője nem mondott igazat.”

_Napi orvosi munka._ Az legtöbb orvos számára a mindennapi, „rutin” munkához is szükséges a statisztikai
szemlélet. Az orvosok ugyanis tevékenységük során csaknem állandóan értékelnek, adatot dolgoznak fel.
Például a páciens laboratóriumi diagnosztikai eredményeit összehasonlítják az ún. normál érlekekkel. Ezt az
összehasonlítást lehet mechanikusan is végezni, tehát mindaz, ami a tartományon kívül van kóros, ami pedig
belül az egészséges. Ha azonban ennél árnyaltabb döntéshozatlara van szükség, akkor ismerni kell azokat a
módszereket is, amelyek alapján a normális értékek tartományát meghatározták. A módszerek ismeretében a
döntéshozatalkor előforduló hibákról is ismeretünk lesz, és mindez alapvetően megváltoztathatja a kezelés
menetét is.

Könyvünk megírásához azzal a szándékkal fogtunk hozzá, hogy segítségére legyünk az élettudományok
területéhez tartozó diákoknak, kutató vagy gyógyító tevékenységet folytató szakembereknek. Könyvünk
olyanoknak szól, akik érzik az igényt a statisztikai alapismeretek elsajátítására, azonban eddig nem találtak
olyan szakkönyvet, amelyet forgatva önmaguktól további magyarázatok, külső segítség nélkül eljutottak volna
arra a tudásszintre amely az alapvető statisztikai kiértékelések elvégzésére elegendő. Igyeksztünk az anyagot
példákon keresztül megvilágítani, a matematikai formulákat értelmezve beláthatóvá, megjegyzhetővé tenni.
Reméljük, erőfeszítéseink sikerre vezetnek.
2. fejezet - Statisztikai alapfogalmak

Nagy Péter dr.

1. Véletlen jelenségek leírása, valószínűségi változók

A valószínűségszámítás azon matematikai elméletek összessége, melyek olyan jelenségek leírására alkalmazhatók, melyek kimenetelét nem lehet teljes pontossággal előre megjósolni. Ez legtöbbször akkor fordul elő, amikor a vizsgált jelenséget olyan tényezők határozzák meg, melyeket nem ismerünk vagy nem tudjuk, hogy milyen hatást fejtjenek ki. Más esetben ismerjük ugyan a befolyásoló tényezőket, de nem tudunk determinisztikus, függvényeszerű összefüggést találni ezek és a végeredmény között. Így pl. egy ember aktuális vércukorszintjét befolyásoló tényezők közül sokat ismerünk, sokat viszont nem, és egy konkrét esetben még azokat a tényezőket sem tudjuk pontosan számba venni, amelyek hatását többé-kevésbé ismerjük. A statisztika ilyen jellegű véletlen jelenségek leírásával foglalkozik, melyek kimenetelét a fenti okok miatt nem lehet 100% biztonsággal megjósolni. Ehhez a statisztika a valószínűségszámítás tételeit használja fel, melyek közül delkori szempontból talán legfontosabb a nagy számok törvénye. A statisztikában a következtetéseket megfigyelések vagy kísérletek elvégzésére alapozzuk, melyeket általánosan elfogadott megegyezés szerint kísérleteknek nevezzünk attól függetlenül, hogy a jelenség kiváltásában részt vettünk-e vagy csak megfigyeltük azt. A kísérlet lehetséges kimenetelének eseményekének nevezzük. A kísérletek kimenetelét egyes esetekben véletlen tényezők befolyásolják; de ha a kísérletet kellően nagy számban végezzük el, akkor a sok kísérlet átlagában a véletlen szerepe csökken. Így tudjuk pl. azt kijelenteni, hogy egy egészséges ember vércukorszintje legnagyobb valószínűséggel 5 mM, de egyedi megfigyelések alkalmával ez lehet pl. 4 vagy 6 mM is.

A valószínűségszámításban a véletlen jelenségek leírásához számadatok kellene. Egyes esetekben ezen számok jelentése triviális (pl. kockadobás során a dobás értéke vagy egy véletlenszerűen kiválasztott ember vércukorszintje), más esetekben a kísérlet kimenetele nem számszerű. Az utóbbi esetben az egyes kimenetekhez numericus értéket rendelünk (pl. 1 - férfi, 2 - nő). Ezek a számok véletlenszerű ingadozást mutatnak. Az ilyen véletlentől függő mennyiségeket valószínűségi változónak nevezzük. Egzaktabb megfogalmazás szerint ha egy véletlen eseményhez számszerű értéket rendelünk, egy véletlentől függő változót, valószínűségi változót kapunk. Így pl. kísérlet egy alany kiválasztása egy klinikai vizsgálatban, az esemény az, hogy éppen az adott embert választottuk ki. Ha ehhez hozzárendeljük a nemet (1 - férfi, 2 - nő) vagy vércukorszintjét, akkor a nem vagy a vércukorszint véletlenszerű ingadozását jellemző valószínűségi változót kapunk. A valószínűségi változó lehetséges értékei alapján megkülönböztettünk diszkret és folytonos valószínűségi változókat. Diszkret valószínűségi változókról beszélünk abban az esetben, ha a valószínűségi változónak véges számú vagy megszámlálhatóan végetelen lehetséges értéke van. Ilyen pl. a fentebb említett példák közül a születendő gyermek neme, illetve a valószínűségszámításban gyakran példaként emlegetett születési kisérlet. A folytonos valószínűségi változókra, hiszen az elméletileg lehetséges értékek a 0 és a 1 között vesznek fel például a vércukorszintet vagy a magasságot. A folytonos valószínűségi változó lehetséges értékei egy folytonos skálát alkotnak (tehát megszámlálhatatlanul végtelen a lehetséges kimenetek száma). Ha például egy embernek a folytonos valószínűségi változója a vércukorszintje, akkor az értékei változhatnak belülről a 0 és a 10 mille mellé, csak a változás értéke változhat belül a kimenetek számában.

Mint korábban említettük, a statisztikai következtetéseket nagy számban elvégzett kísérlet alapján tesszük. A statisztikai megfigyelés tárgyát képező egyedek összességét populációk nevezzük. Ha az emberek átlagos várható élettartamát vizsgáljuk, a megfigyelésünk tárgya, tehát a populáció, az emberiség egészéje. Ha a megfigyelés célja csak a magyar nők vizsgálata, a populációt az összes magyar nő alkotja. Mivel a populáció a legtöbb statisztikai vizsgálatban túl nagy elemszámú ahhoz, hogy minden egyes egyedet mindon lenne megfigyelni, ezért csak egy részével foglalkozunk, melyet mintának hívunk. Gyakran halljuk, hogy a közvéleménykutatások során pl. 10000 ember kérdeztek meg azért, hogy kiderítsek, milyik párt fogja a választásokat nyerni. Ebben az esetben a populáció a választásra jogosultak összessége, a minta az a 10000 ember, akinek a véleményét megkérdeztek. A minta alapján a teljes populációra vonatkozóan kívánunk következtetéseket lezárni, így a mintát alkalmas módon kell megválasztani. A mintának kellően nagyszámtól elemből kell állnia, hogy a statisztikai következtetés kellően megbízható legyen. Másrészt representatívnek kell lennie, ami azt jelenti, hogy a populáció minden egyede ugyanolyan valószínűséggel kerülhet a mintába. Ez a véletlen mintavételezés fontos következménye. Tehát a minta nem attól következik, hogy nagy elemszámú, hanem attól, hogy véletlen mintavétel eredménye.
2. Mérési skálák

A statisztikai megfigyelések során kapott eredményeket megfelelően módon prezentálni kell. A megjelenítés módszerét meghatározza egyrészt a megfigyelt jelenség természeté (diszkrét vagy folytonos valószínűségi változó), illetve a vizsgálat célja. Ennek megfelelően az alábbi három fontosabb skálatípust különböztetjük meg:

1. **Nominális skála**: a statisztikai vizsgálat eredményeit osztályokra, kategóriákra osztjuk. A nominális skálára példa a születendő gyermek neme (fiú/lány) vagy a tüdődaganatok szövettani beosztása (kissejtes rák, nagysejtes rák, mirigyhám eredetű rák, laphámrák). Ezekben az esetekben az egyes kategóriák között nincsen mennyiségi összefüggés, nem lehet azt mondani, hogy az egyik kategóriába tartozó elem nagyobb, több, stb., mint a másikba tartozó.

2. **Ordinális (sorrenden alapuló skála)**: ez a nominális skála rokonának tekinthető, de ebben az esetben az egyes kategóriák kvantitatív alapon sorba rendezhetők, meg tudjuk mondani, melyik a „jobb” vagy „több”. Sok betegséget, pl. a daganatokat vagy a szívelégtelenséget előrehaladottságuk szerint stádiumokra szokták osztani. A daganatok stádiumbeosztását általában egy 0 és IV közötti skálán végzik, és az előrehaladottság mértéke a stádiumnál nő.

3. **Numerikus skála**: ebben az esetben a prezentált számoknak numerikus, kvantitatív jelentése van. Pl. ha két ember magassága 150 illetve 75 cm, akkor egyikijük kétszer olyan magas, mint a másik. Ordinális skála esetében ugyanez nem mondható el (egy IV. stádiumú daganat nem kétszer olyan súlyos vagy előrehaladott, mint egy II. stádiumú). A numerikus skála lehet folytonos (pl. vércukor szint) vagy diszkrét (pl. csontrőlésnek száma), tehát gyakran tükröző azt, hogy folytonos vagy diszkrét valószínűségi változó megjelenítésére használják. Azonban a vizsgálat vagy az adatmegjelenítés céljainak megfelelően folytonos valószínűségi változót is meg lehet diszkrét numerikus skálán jeleníteni, ha a folytonos skálat intervallumokra osztjuk. Az 2.1. táblázat - Emberek magasságának vizsgálatára vett 15 elemű minta eloszlása 2.1. táblázatban a vizsgált emberek magasságát hat intervallumra osztjuk. Az intervallumokra osztás megkönnyítí az adatok összefoglalását és megjelenítését, ugyanakkor információveszteséssel jár, hiszen egy folytonos változót diszkrét transzformálunk.

3. Gyakoriság, relatív gyakoriság, valószínűség

A biostatisztikában a valószínűségnek két eltérő megközelítése, értelmezése van.

1. **Relatív gyakoriság szerinti megközelítés**: Egy statisztikai vizsgálat során egy kísérletet mindig többször végeznek el. Az egyes események bekövetkezési számát az esemény gyakoriságának nevezzük. A gyakoriság önmagában kevés kísérlet sorozatától független információt hordoz, mert füg a kísérletek számától. Ha a gyakoriságot elsőszorzjuk a kísérletek számával, a relatív gyakoriságot kapjuk (2.1. táblázat - Emberek magasságának vizsgálatára vett 15 elemű minta eloszlása 2.1. táblázat). Ha egy kockát hatvanszor dobunk fel, és az egyes számot tizenkét szter kaptuk, akkor 12 az esemény gyakorisága, 12/60 pedig a relatív gyakorisága. Ha a kísérlet sorozatot újból és újból végrehajtjuk, azt figyelhetjük meg, hogy az egyes dobásának relatív gyakorisága egy meghatározott érték, ebben az esetben 1/6, körülfog ingadozni. Ezt az értéket az esemény valószínűségének nevezzük. A relatív gyakoriság annál jobban közelíti ezt az elméleti értéket, minél nagyobb számú kísérletet, mérést végzünk el (}
2.1. ábra). Ez a nagy számok törvényének egyik megnyilvánulása. Ezt a következőképpen fejezhetjük ki:

$$RGy(E) = \frac{m}{N} \rightarrow P(E), \quad \text{ha} \ N \rightarrow \infty$$

$RGy(E)$ a vizsgált $E$ esemény relatív gyakorisága, $N$ a kísérletek száma, $m$ az $E$ esemény bekövetkezésének száma (gyakorisága). A $P(E)$ érték, amelyet a relatív gyakorisággal közeltünk az esemény valószínűsége.

Mind a valószínűségre, mind a relatív gyakoriságra teljesül, hogy nem lehet nagyobb, mint egy, és kisebb, mint nulla. A köznyelvben a matematikai valószínűség és relatív gyakoriságot (%)-ban kifejezve használjuk.

A gyakoriság, relatív gyakoriság és valószínűség fenti értelmezése nem alkalmazható folytonos valószínűségi változóknak, mert ezek esetében annak végletlenül kicsiny a valószínűsége, hogy a változó pontosan egy adott értékét vegyen fel, pl. annak, hogy egy ember magassága pontosan $178,2846437...$ cm. A folytonos valószínűségi változók esetében a valószínűség értelmezésére a sűrűségfüggvény tárgyalása során térünk majd vissza. Most is meg kell azonban jegyezni, hogy a folytonos valószínűségi változó lehetséges értékeit intervallumokra oszthatjuk (1. táblázat), és ebben az esetben a gyakoriság, relatív gyakoriság és valószínűség fentebb ismertetett definíciói már alkalmazhatók.

2. A valószínűség klasszikus értelmezése: Ha egy kísérletnek meghatározott számú, azonosan valószínű kimenetéle lehetséges, egy esemény valószínűségének kiszámításához nem kell a kísérletet többször elvégezni, csak számba kell venni a lehetséges kimeneteleket. Ha egy kísérletnek $n$ különböző és egyenlően valószínű kimenetele van, és ebből $q$ darab rendelkezik egy adott $F$ tulajdonsággal, az $F$ valószínűségét a

$$RGy(E) = \frac{m}{N} \rightarrow P(E), \quad \text{ha} \ N \rightarrow \infty$$

képlettel lehet kiszámítani. A valószínűség ezen megközelítését azért nevezik klasszikusnak, mert a valószínűségelmélet történetére, Pascal és Fermat dolgozták ki a XVII. században szerencsejátékok (kockadobás, kártyajátékok) tudományos leírására. Pl. a kockadobás esetében minden dobásnak ugyanannyi a valószínűsége, így ha G-vel jelöljük azt az eseményt, hogy 3-nál többet dobunk, ennek valószínűségét úgy számíthatjuk ki, hogy a feltételezett teljesitő események számát (**3) osztjuk az összes kimenetel számával (**6), tehát $P(G) = 0,5$. A relatív gyakoriságon alapuló megközelítéssel ellentétben a klasszikus értelmezés
az esemény valószínűségét közelítés nélkül, teljes pontossággal adja meg, azonban alkalmazhatóságához igazolni kell azt, hogy a kísérlet lehetséges kimenetelei egyenlően valószínűek.

4. Eseményalgebra, a valószínűség grafikai értelmezése

A valószínűségszámítás és a statisztika alkalmazása során gyakran van szükség arra, hogy egy kísérlet lehetséges kimenetelei között műveleteket végezzünk, ezért röviden áttekintjük ezeket és a rájuk vonatkozó legfontosabb definíciókat. *Elemi eseménynek* hívjuk egy kísérlet lehetséges kimeneteit, *összetett eseménynek* pedig azt, amely több elemi eseményre bontható. Az elemi események halmazát eseménytérnek nevezzük (2.2. ábra). Pl. a kockadobás során a dobás lehetséges értékei az elemi események, összetett esemény pl. az, hogy a dobás eredménye páros szám.

Azt az eseményt, melynek bekövetkezési valószínűsége 1, *biztos eseménynek*, aminek valószínűsége 0, *lehetetlen eseménynek* nevezzük. Az események olyan halmazát, amely egymást kölcsönösen kizáró eseményekből áll, és az események összessége a biztos eseményt adja, *teljes eseményrendszernek* hívjuk. Mivel az elemi események egymást kölcsönösen kizárják (a kockadobás értéke nem lehet egyszerre 1 és 2) és egyikük biztosan bekövetkezik (a kocka biztosan valamelyik lapjára esik), összességük a biztos eseményt adja. Tehát az elemi események halmaza is teljes eseményrendszer.

Az összetett események, illetve a hozzájuk rendelhető valószínűségek tárgyalását nagymértékben megkönnyíti, ha az eseményeket grafikusan jeleníti jük meg. Ebben az esetben az eseményhez egy idomot rendelünk oly módon, hogy az esemény valószínűsége egyenesen arányos legyen az idom területével. Így könnyen értelmezhetjük a következő műveleteket (2.2. ábra).
2.3. ábra):

Azt az eseményt, amely akkor következik be, ha az A és a B események közül legalább az egyik bekövetkezik, az A és B események összegének vagy uniójának nevezzük, és A+B-vel vagy AB-vel jelöljük.

Azt az eseményt, amely akkor következik be, ha mind az A, mind a B esemény bekövetkezik, az A és B események szorzatának vagy metszetének nevezzük, és AB-vel vagy AB-vel jelöljük.

Ahogyan a valós számok között mind az összeadás, mind pedig a szorzás kommutatív művelet, az események összegére és szorzatára is igaz, hogy A+B = B+A illetve AB = BA.

Az A és B események A-B-vel jelölt különbségén azt az eseményt értjük, hogy az A esemény bekövetkezik, de a B nem.

Azt az eseményt, hogy valamely A esemény nem következik be, A komplementer vagy ellentett eseményének nevezzük, és A felülvonással jelöljük.

Az események grafikus megjelenítése alapján könnyen értelmezhetjük az A+B esemény valószínűségét:

\[
P(A+B) = P(A) + P(B) - P(AB)
\]

A P(AB) esemény valószínűségét azért kell levonni, mert az mind a P(A), mind a P(B) valószínűségekben szerepel, így a kivonás nélkül kétszer lenne számolva.
5. Feltételes valószínűség

Gyakran van arra szükség, hogy egy esemény bekövetkezési valószínűségét a teljes eseményrendszer egy részére vonatkoztatva számoljuk ki. Másképpen megfogalmazva ilyenkor arra vagyunk kiváncsiak, hogy egy esemény milyen valószínűséggel következik be a teljes eseményben, ha a másik esemény bekövetkezik. Így pl. megvizsgálhatjuk azt, hogy az A esemény milyen valószínűséggel következik be a B esemény esetén, ha a B esemény bekövetkezik, tehát ha csak azokat a kísérleteket vesszük figyelembe, amikor B bekövetkezik (2.4. ábra).

2.4. ábra. Ezt a valószínűséget az A esemény B-re vonatkoztatott feltételes valószínűségének nevezzük, jele: \( P(AB) \). Kiszámításához induljunk ki abból, hogy ha csak azokat a kísérleteket vesszük számításba, amikor B bekövetkezik, az összes lehetséges események számát leszűkítettük a B eseményre. Az A és B esemény együttes bekövetkezésének valószínűsége \( P(AB) \), az összes eset valószínűsége pedig \( P(B) \). Így a \( P(AB) \) valószínűségét a kedvező esetek valószínűségének az összes eset valószínűségével való osztásával kaphatjuk meg (vö. (02_02_keplet)):

\[
P(A|B) = \frac{P(AB)}{P(B)}
\]

A feltételes valószínűséget felhasználhatjuk a független események definiálására. Egy A esemény akkor független a B eseménytől, ha a B bekövetkezése nem befolyásolja az A bekövetkezési valószínűségét. Ebben az esetben

\[
P(A|B) = P(A)
\]

A feltételes valószínűség definicióját és a fenti egyenlőséget felhasználva

\[
P(AB) = P(A|B)P(B) = P(A)P(B)
\]
A fenti törvény a valószínűségek szorzási tétele, amely szerint független események együttes bekövetkezésének
valószínűségét úgy kapjuk meg, ha az egyes események valószínűségeit összeszorozzuk.

A 2.2. táblázat - A testmagasság eloszlása nemenként 2.2. táblázat adatból felhasználva az M₁ esemény (egy ember magassága 160 és 169 cm között van) valószínűsége 0,3. Annak, hogy valaki nő és 160-169 cm között van a magassága, 0,22 a valószínűsége. Ez megfelel az M₁N szorzateseménynek. Ha arra a kérdésre keressük a választ, hogy mi a valószínűsége annak, hogy egy vizsgálati alany magassága 160-169 cm között van, ha csak a nőket vizsgáljuk, az M₂ eseményre vonatkoztatott feltételes valószínűségét keressük. Kiszámítani ezt úgy tudjuk, hogy a P(M₁N) valószínűséget osztjuk a P(N) valószínűséggel:

\[ P(M_2 | N) = \frac{P(M_2 \cap N)}{P(N)} = \frac{0,22}{0,5} = 0,44 \]

A táblázat alapján válaszolhatunk arra a kérdésre, hogy a testmagasság és az emberek neme független-e egymástól. Mivel

\[ P(M_2 \cap N) = 0,22 \neq P(M_2) P(N) = 0,3 \cdot 0,5 = 0,15 \]

ezért a két esemény nem független egymástól.

6. Teljes valószínűség tétele

Gyakran fordul elő, hogy egy esemény valószínűsége helyett annak feltételes valószínűségeit ismerjük. A 2.3. táblázat - A 2.2. táblázatban feltüntetett magasságeloszlás feltételes valószínűségeket felhasználásával 2.3. táblázatban a 2.2. táblázat - A testmagasság eloszlása nemenként 2.2. táblázat adatokat tüntetett fel, csak ebben az esetben nem a szorzat események (pl. M₁N) valószínűségét, hanem a feltételes valószínűségeket, pl. a P(M₁|N) valószínűségét. Az N és F események teljes eseményrendszert alkotnak, hiszen egymást kölcsönösen kizárják és összegük a biztos eseményt adja. Ilyenkor gyakran kell válaszolni olyan kérdésre, hogy pl. mi annak a valószínűsége, hogy egy véletlenszerűen kiválasztott ember magassága 160 és 169 cm közé esik. Ezt az M₁ esemény peremvalószínűségének nevezzük. A kérdés megválaszolásához a teljes valószínűség tételét kell felhasználni. A 2.3. táblázat - A 2.2. táblázatban feltüntetett magasságeloszlás feltételes valószínűségeket felhasználásával 2.3. táblázat alapján láthatjuk, hogy az M₂ esemény két elemi eseményből tevődik össze, az M₁N és M₁F eseményekből. Ezek egymást kölcsönösen kizárják, mivel P(NF) = 0. Így valószínűségeik összeadódnak:

\[ P(M_2) = P(M_2 \cap N) + P(M_2 \cap F) \]

A feltételes valószínűség definícióját felhasználva

\[ P(M_2 \cap N) = P(M_2 | N) P(N), \quad P(M_2 \cap F) = P(M_2 | F) P(F) \]

\[ P(M_2) = P(M_2 | N) P(N) + P(M_2 | F) P(F) = 0,44 \cdot 0,5 + 0,16 \cdot 0,5 = 0,3 \]

Általánosan megfogalmazva, ha a B₁, B₂, ..., Bₙ események teljes eseményrendszert alkotnak, és ismerjük egy A esemény Bᵢ eseményekre vonatkoztatott feltételes valószínűségeit, az A esemény valószínűségét (amit ebben az esetben peremvalószínűségnek nevezzünk) a következő képlettel lehet kiszámítani:

\[ P(A) = \sum_{i=1}^{n} P(A|B_i) P(B_i) \]

7. Bayes tétel

Szintén a 2.3. táblázat - A 2.2. táblázatban feltüntetett magasságeloszlás feltételes valószínűségek felhasználásával 2.3. táblázatban levő adatsorba vonatkozóan, válaszoljuk meg azt a kérdést, hogy ha véletlenszerűen kiválasztunk egy 160-169 cm magasságú embert, mi annak a valószínűsége, hogy az illető férfi. Tulajdonképpen a P(FM₁) feltételes valószínűséget keressük. A feltételes valószínűség definícióját felhasználva:
\[ P(F|M_2) = \frac{P(FM_2)}{P(M_2)} \]

Az előzőekben a \( P(M_2) \) valószínűségét már meghatároztuk a teljes valószínűség tételének felhasználásával (2.11. egyenlet). A kifejezés számlálóját a következő összefüggésből nyerjük:

\[ P(M_2|F) = \frac{P(M_2F)}{P(F)} \]

Mivel \( P(M_2F) = P(FM_2) \), ezért

\[ P(FM_2) = P(M_2|F)P(F) \]

A fenti egyenletet a 13. egyenlet számlálójába, a 11. egyenletet a 13. egyenlet nevezőjébe helyettesítve:

\[ P(F|M_2) = \frac{P(FM_2)}{P(M_2)} = \frac{P(M_2|F)P(F)}{P(M_2|N)P(N) + P(M_2|F)P(F)} = \frac{0.16 \cdot 0.5}{0.3} = 0,266 \]

Ez Bayes tételének alkalmazása. Általánosan megfogalmazva, ha a \( B_1, B_2, \ldots, B_n \) események teljes eseményrendszert alkotnak, a \( P(B|A) \) valószínűséget az alábbi képlet segítségével lehet kiszámolni:

\[ P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)} \]

8. Diagnosztikus eljárások a biostatisztika szemszögéből

Az orvos a betegségek diagnózisának felállítása közben fizikális, műszeres és laboratóriumi vizsgálatokat végez. Minden egyes diagnosztikus lépés szüköti a lehetséges diagnózisok számát, megerősíti vagy elveti a korábbi feltételezéseket. A vizsgálati eljárások gyakran számszerű eredményt adnak (pl. egy hormon koncentrációja a vérben), de a legtöbb esetben ezt a numerikus adatot nominális skálán jeleníti meg, tehát az eredmény vagy megerősíti egy betegség fennállását (pozitív) vagy elveti azt (negatív). Egy eredményt akkor tekintenek pozitívnak, ha a vizsgálat eredménye jelentősen eltér az egészséges emberekben mért erőtől. Az egészséges emberekben tapasztalható eloszlás ismeretében a pozitív és negatív eredmények közötti küszöbértéket úgy választjuk meg, hogy lehetőleg minél kevesebb egészséges egyén kerüljön a pozitív tartományba. Már ebből is látható, hogy egy diagnosztikus eljárás csak akkor működik tökéletesen, ha az egészséges és beteg emberekhez mért erők teljesen elkülönülnek egymástól, azaz nincs átfedés. Ez a legritkábban teljesül, hiszen az egészséges egyének egy meghatározott arányának értékei a küszöbérték felett lesznek ( 
Ha ehhez hozzávesszük azt, hogy a diagnosztikus eljárások minden esetben mérési hibával terheltek, tovább romlik a módszerek megbízhatósága. Optimális esetben egy teszt pozitív eredményt ad a vizsgált betegség fennállása esetén (valós pozitív) és negatív eredményt akkor, ha a betegség nem áll fenn (valós negatív), tehát sosem ad álnegatív (negatív teszeredmény a betegség fennállása esetén) vagy álpozitív (pozitív eredmény a betegség hiányában) eredményt. Azt, hogy egy adott diagnosztikus eljárás mennyire optimális, két paraméter megadásával jellemzük:

specificitás: annak a valószínűsége, hogy a diagnosztikus teszt értéke negatív lesz egy olyan pácienszen, akiben nem áll fenn a vizsgált betegség. A specificitás tehát azt jellemzi, hogy a teszt milyen megbízhatóan azonosítja azokat, akikben nem kóros a vizsgált paraméter. A 2.4. táblázat - Diagnosztikus eljárások lehetséges eredményei a betegség fennállásától függően. 2.4. táblázat jelöléseit felhasználva:

\[
\text{specificitás} = \frac{VN}{VN + AP}
\]

Másként megfogalmazva a specificitás a negatív teszt eredmény (T-) feltételes valószínűsége arra vonatkoztatva, hogy a betegség nem áll fenn (B-): P(T-B-).

szenzitivitás: annak a valószínűsége, hogy a diagnosztikus teszt értéke pozitív lesz egy olyan pácienszen, akiben fennáll a betegség. A szenzitivitás azt jellemzi, hogy a teszt milyen megbízhatóan detektálja a betegség fennállását. A 2.4. táblázat - Diagnosztikus eljárások lehetséges eredményei a betegség fennállásától függően. 2.4. táblázat jelöléseit szerint:

\[
\text{szenzitivitás} = \frac{V}{V + AN}
\]

Biostatisztikai nömenklatúra szerint a szenzitivitás a pozitív eredmény (T+) feltételes valószínűsége arra vonatkoztatva, hogy a betegség fennáll (B+): P(T+B+).

A 2.4. táblázat - Diagnosztikus eljárások lehetséges eredményei a betegség fennállásától függően. 2.4. táblázatban használt változók segítségével további két fontos paraméter definiálható:

pozitív prediktív érték: egy pozitív teszeredmény megbízhatósága, tehát annak a valószínűsége, hogy egy pozitív eredmény esetében a betegség valóban fennáll:

\[
\text{pozitív prediktív érték} = \frac{VP}{VP + AP}
\]

negatív prediktív érték: egy negatív teszeredmény megbízhatósága, tehát annak a valószínűsége, hogy egy
negatív eredmény esetében a betegség nem áll fenn:

\[
\text{negatív prediktív érték} = \frac{VN}{VN + \dot{AN}}
\]

2.5. ábrán látható, hogy egy teszt szenzitivitását és specificitását a beteg és egészséges populációk eltérései is befolyásolják: sokkal könnyebb nagy specificitású és nagy szenzitivitású tesztet alkotni abban az esetben, ha a beteg és egészséges populáció jelentősen különbözik egymástól, mert ilyenkor a két eloszlás átfedő része kisebb. Szintén az

2.5. ábrán, illetve az abból készített 2.5. táblázat - A, 2.5. táblázat adatai alapján látható, hogy a szenzitivitás és a specificitás a küszöbérték megválasztásával változik: ha nő a specificitás, csökken a szenzitivitás. Ezért a küszöbérték megválasztása alapvető jelentőségű a diagnosztikában. Ezt nemcsak a teszt paraméterei és a beteg, valamint egészséges populációk eloszlásai befolyásolják, hanem a teszt célja is. Ha pl. egy diagnózis kizárása a
cél, egy nagy szenzitivitású teszt a hasznos, mert ezt a betegség fennállása esetén nagy valószínűséggel pozitív eredményt ad. Diagnosztikus eljárások összehasonlítására, illetve az optimális küszöbérték megválasztására használt eljárás a ROC (receiver operating characteristic) görbe készítése, azaz a szenzitivitás ábrázolása az (1-specificitás) függvényében (2.5. ábra), Minél közelebb van a görbe a bal felső sarokhoz (100% specificitás, 100% szenzitivitás), annál megbízhatóbb az adott eljárás.

2.1. táblázat - Emberek magasságának vizsgálatára vett 15 elemű minta eloszlása

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gyakoriság</td>
<td>2/15</td>
<td>4/15</td>
</tr>
<tr>
<td>Relatív gyakoriság</td>
<td>2/15</td>
<td>4/15</td>
</tr>
</tbody>
</table>

2.2. táblázat - A testmagasság eloszlása nemenként

<table>
<thead>
<tr>
<th>Magasság (cm)</th>
<th>Valószínűségek</th>
<th>Férfi (F)</th>
<th>Nő (N)</th>
<th>Peremvalószínűség</th>
<th>Peremvalószínűség</th>
</tr>
</thead>
<tbody>
<tr>
<td>150-159 (M₁)</td>
<td>0,02</td>
<td>0,15</td>
<td>0,17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>160-169 (M₂)</td>
<td>0,08</td>
<td>0,22</td>
<td>0,3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>170-179 (M₃)</td>
<td>0,2</td>
<td>0,1</td>
<td>0,3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>180-189 (M₄)</td>
<td>0,2</td>
<td>0,03</td>
<td>0,23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0,5</td>
<td>0,5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A magasságokat intervallumokra osztottuk, és az egyes intervallumokat az M1-M4 eseményekhez rendeltük. Az emberek neme az F (férfi) és N (nő) események valamelyikébe esik.

2.3. táblázat - A 2.2. táblázatban feltüntetett magasságeloszlás feltételes valószínűségek felhasználásával

<table>
<thead>
<tr>
<th>Magasság (cm)</th>
<th>Feltételes valószínűségek</th>
<th>Peremvalószínőség</th>
</tr>
</thead>
<tbody>
<tr>
<td>150-159 (M₁)</td>
<td>0,04</td>
<td>0,3</td>
</tr>
<tr>
<td>160-169 (M₂)</td>
<td>0,16</td>
<td>0,44</td>
</tr>
<tr>
<td>170-179 (M₃)</td>
<td>0,4</td>
<td>0,2</td>
</tr>
<tr>
<td>180-189 (M₄)</td>
<td>0,4</td>
<td>0,06</td>
</tr>
<tr>
<td>Peremvalószínőség</td>
<td>0,5</td>
<td>0,5</td>
</tr>
<tr>
<td>P(MxF)</td>
<td>P(MxN)</td>
<td></td>
</tr>
</tbody>
</table>

2.4. táblázat - Diagnosztikus eljárások lehetséges eredményei a betegség fennállásától függően.

<table>
<thead>
<tr>
<th>Teszt eredmény</th>
<th>Betegség</th>
<th>Fennáll (B⁺)</th>
<th>Nem áll fenn (B⁻)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pozitív (T⁺)</td>
<td>VP</td>
<td>ÁP</td>
<td></td>
</tr>
<tr>
<td>Negatív (T⁻)</td>
<td>ÁN</td>
<td>VN</td>
<td></td>
</tr>
<tr>
<td>VP = valós pozitív, ÁP = álpozitív, VN = valós negatív, ÁN = álnegatív</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2.5. táblázat - A

<table>
<thead>
<tr>
<th>Küszöbérték</th>
<th>130</th>
<th>150</th>
<th>170</th>
</tr>
</thead>
<tbody>
<tr>
<td>Szenzitivitás</td>
<td>99,9%</td>
<td>98%</td>
<td>84%</td>
</tr>
<tr>
<td>Specifikitás</td>
<td>70%</td>
<td>94%</td>
<td>99%</td>
</tr>
</tbody>
</table>
3. fejezet - Leíró statisztika: a populáció és a minta jellemzése

Nagy Péter dr.

A statisztikai vizsgálatok egyik első célja a vizsgált populáció vagy a populációból vett minta tömör jellemzése. Ezek egy része egy szám, mások táblázatba foglalható vagy grafikusan megjeleníthető jellemzések. Statisztikának nevezzük az adatson értelmezett függvényeket, melyek az adatok jellemzését szolgálják.

1. Adatok csoportosítása

Minél nagyobb egy vizsgálendő statisztikai adatsor, annál nehezebb az adatok áttekintése. Mind az adatsor áttekintését, mind egyes statisztikák kiszámítását jelentősen megkönnyíti az adatok csoportosítása. Mint a 3.1. táblázat - Látóterenkénti sejtszám-adatok csoportosítása 3.1. táblázatból látható, ez azt jelenti, hogy az adatok gyakoriságát, illetve relatív gyakoriságát táblázatba foglaljuk. Bizonyos esetekben a valószínűségi változó lehetséges értékeinek száma túl nagy a vizsgált adat sor elemszámához képest, így a legtöbb lehetséges érték gyakorisága 0 vagy 1 lenne. Ilyen esetben célszerű az adatok intervallumokba osztása, hogy az egy intervallumba eső átlagos elemszám nőjön (2.1. táblázat).

1.1. A közép jellemzésére használt statisztikák


\[
\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{2+3+5+1+2+4+5+2+4}{9} = 3,11
\]

ahol \(x_i\) jelöli a minta elemeket, \(n\) a minta elemszámát, a számtani közép. Az adatok csoportosításának felhasználásával a számítás egyszerűbben is elvégezhető:

\[
\bar{x} = \frac{\sum_{j=1}^{J} n_j \bar{x}_j}{N} = \frac{1\cdot1+3\cdot2+1\cdot3+2\cdot4+2\cdot5}{9} = \frac{1+6+3+6+8+10}{9} = 3,11
\]

ahol \(j\) a lehetséges értékek száma, \(n_j\) az adott érték gyakorisága, \(p_j\) pedig az adott érték relatív gyakorisága. Az egyszerűsítés alapja az, hogy \(n_j/n = p_j\).

Medián: az az érték, amelynél a minta fele kisebb vagy egyenlő. Megtalálásához a nyers adatsorból el kell készíteni a rendezett mintát, azaz a mintaelmeletet növekvő sorrendbe kell rendezni. A 3.1. táblázat - Látóterenkénti sejtszám-adatok csoportosítása 3.1. táblázatból tárgyalt adatokra a rendezett minta: 1, 2, 2, 2, 3, 4, 4, 5, 5 A medián az az érték, amely úgy osztja két részre a rendezett mintát, hogy a mediánnál nagyobb-egyenlő elemek száma egyenlő a mediánnál kisebb-egyenlő elemek számával, tehát a fenti mintában az aláhúzott 3-as szám. A medián meghatározása a fenti módon csak akkor lehetséges, ha az elemek száma páratlan, hiszen ekkor létezik középső elem. Páros elemszámú esetben a két középső elem átlagát nevezzük mediánnak. Pl. ha a fenti adatsort (3.1. táblázat, 3.2. táblázat A minta) kiegészítjük még egy elemmel \((x_{w} = 50, 3.2. táblázat B minta)\), akkor a rendezett minta a következő: 1, 2, 2, 2, 3, 4, 5, 5, 50, így a medián \((3+4)/2 = 3.5\).

Módusz: a legnagyobb gyakorisággal előforduló elemet nevezzük módusznak, tehát mindkét fenti adatsorban a módusz 2.

A közép jellemzésére ismertetett három statisztikát (átlag, medián, módusz) ugyanúgy kell kiszámolni akár egy populációval, akár egy populációból vett mintával van dolgunk. A minta alapján számolt átlag a populáció átlagának a közelítése. A populáció-átlagot várható értékek is nevezzük és általában -vel jelöljük. Hasonlóan az átlaghoz, a minta alapján számolt medián és módusz a populáció megfelelő paramétereinek közelítése, ún. torzítatlan becslése (bővebben l. Becsléselmélettel foglalkozó fejezet).
2. A közép jellemzésére használt statisztikák összehasonlítása

A számtani közép hátrányos tulajdonsága az, hogy egy-egy extrem érték jelentősen befolyásolja, maga felé húzza. A 3.2. táblázat - A közép és a szóródás jellemzésére használt statisztikák szemléltetése 3.2. táblázat B mintája egyetlen kiugróan nagy elemmel (xₘₐ = 50) térr el az A mintától, de ez a minta átlagát 3,11-ről 7,8-re módosítja. Figyelembe véve, hogy mind a diagnostikus eljárások, mind a biológiai kísérletek során kapott adatsorok gyakran tartalmaznak olyan mérési hibákat, melyek jelentősen eltérnek az átlagtól, ez azt eredményezi, hogy az átlag nagyon érzékeny a hibás mérési eredményekre. A medián nem rendelkezik ezzel a hátrányos tulajdonsággal, hiszen a fenti két minta mediánja (3, illetve 3,5) sokkal kevésbé térr el egymástól, mint az átlagok. Az átlag ezen hátrányos tulajdonságának kiküszöbölésére bevezették a megnyesett (rendeleti) átlag (angol: trimmed mean) fogalmát: meghatározásához az elemek meghatározott százalékát, illetve kis elemszám esetén meghatározott számú elemet, kihagynak a rendezett minta felső és alsó részéből, és a maradék alapján határozzák meg az átlagot. A már tárgyalt két adatsorra (3.2. táblázat - A közép és a szóródás jellemzésére használt statisztikák szemléltetése 3.2. táblázat A és B minta) a legnagyobb, illetve legkisebb elem elhagyásával a következő megnyesett átlagok adódnak:

\[
\frac{2 + 2 + 2 + 3 + 4 + 4 + 5}{7} = 3,14
\]

\[
\frac{2 + 2 + 2 + 3 + 4 + 4 + 5 + 5}{8} = 3,375
\]

Ebből látható, hogy a megnyesett átlag sem érzékeny az extrémény nagy vagy kicsi értékekre.

3. A szóródás jellemzésére használt statisztikák

Ha a 3.2. táblázat - A közép és a szóródás jellemzésére használt statisztikák szemléltetése 3.2. táblázat C mintájára kiszámítjuk a közép jellemzésére leggyakrabban használt statisztikákat, az a benyomásunk támadhat, hogy a C minta ugyanolyan, mint az A. Az adatokra pillantva azonban önkéntelenül is észrevesszük, hogy a C mintában az elemekek jobban centráltak az átlag körül. Az elemek átlag körüli szétterülésének, szóródásának jellemzésére a következő statisztikákat használjuk:

*Terjedelem:* a legnagyobb és a legkisebb értékek közötti különbség. Önmagában elég kevés értékkel bír, ugyanis az átlaghoz hasonlóan a terjedelem is nagyon érzékeny az extrém kicsi vagy nagy értékekre.

*Variancia és standard deviáció (SD):* a szóródás jellemzésére leggyakrabban használt statisztikák. A varianciát a következő képlet szerint kell kiszámítani egy minta esetében:

\[
\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}
\]

ahol

- a minta átlaga, \(x\), az i-dik mintaelem, \(n\) a minta elemszáma. A minta fenti képlettel számolt varianciája a populáció varianciájának közelítése. A variancia kiszámításának móda eltér, ha a teljes populácioról rendelkezésre álló adatok alapján kívánjuk meghatározni a populáció varianciáját (
Leíró statisztika: a populáció és a minta jellemzése

3.1. ábra). Ebben az esetben a következő képletet kell használni:

\[ V_{\text{populáció}} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n^2} \]

ahol \( x \) a populáció elemeit, a populáció átlagát, \( n \) a populáció elemszámát jelöli. Mivel a variancia mértékegysége a minta- vagy populációelemek mértékegységének négyzete, ezért legtöbbször a variancia négyzetgyökét, a standard deviációt vagy SD-t használjuk. A 3.2. táblázat - A közép és a szóródás jellemzésére használt statisztikák szemléltetése

3.2. táblázat A mintájára alkalmazva a képletet

\[ SD = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(1-3,311)^2 + (2-3,311)^2 + ... + (5-3,311)^2}{8}} = 1,45 \]

A minta SD-jének magyar szakirodalomban elterjedt elnevezése a korrigált empirikus szórás. A 3.2. táblázat - A közép és a szóródás jellemzésére használt statisztikák szemléltetése 3.2. táblázat A és C adatsorainak SD-jét összehasonlítva láthatjuk, hogy az SD jól jellemzi azt a tendenciát, hogy a C minta esetében a mintaelemek jobban centráltak a minta átlaga körül. A (25. egyenlet)-(26. egyenlet) nevezőiben szereplő \( n \) illetve \( (n-1) \) közötti különbösg nagy \( n \)-nek esetén elhanyagolható.

Variációs koefficiens (angol elnevezés (coefficient of variation) rövidítése alapján CV): az SD értéke az átlag százalékában kifejezve:

\[ CV = \frac{SD}{\bar{x}} \cdot 100 \]

A variációs koefficiens standardizálja az SD értékét, így különböző mérési skálán végrehajtott mérések SD értékei összehasonlíthatóvá váltnak.

Percentilis, kvartilis: Egy eloszlás \( x% \)-os percentilisénak nevezzük azt a számot, amelynél kisebb vagy egyenlő az elemek \( x% \)-a. Néhány percentilisnek kitüntetett szerepe van. Ezek közül az 50%-os percentilissel már találkoztunk, ez megegyezik a mediánnal. A 25%-os percentilist első kvartilisnek (Q1), a mediánt második
kvartilisnek (Q_1), a 75%-os percentilist harmadik kvartilisnek (Q_3) is nevezzük. Az első és harmadik kvartilis közötti különbséget interkvartilis terjedelem (IQT) hivják:

\[ IQT = Q_3 - Q_1 \]

Az eloszlás interkvartilis része az elemek centrális 50%-át tartalmazza. Szintén fontos az 5% és 95%-os, valamint a 2,5% és 97,5%-os percentilisek közötti rész, melyek az eloszlás centrális 90%-át, illetve 95%-át tartalmazzák. A 2,5%-os, illetve 5%-os percentilsnél kisebb, valamint a 97,5%-os, illetve 95%-os percentilsnél nagyobb értékek ritka események számítanak, hiszen az elemeknek csak 2,5, illetve 5%-a esik ezen tartományokba. Az 5 és 95%-os percentiliseket használják csesemők növekedésének megítéléésére. Azon gyermekek, akiiknek súlya az 5%-os percentilsnél kisebb vagy a 95%-os percentilsnél nagyobb, valószínűleg valamilyen növekedési rendellenességben szennednek. Laboratóriumi mérések esetében a referencia tartományt, tehát azon értékeket, amiket normálisnak tekintenek, általában a 2,5% és 97,5%-os percentilis közötti tartománnyé definiálják.

Középérték közepes hibája (az angol „standard error of the mean” illetve „standard error” elnevezések alapján SEM vagy SE): Populációból történő mintavétel során az egyik legfontosabb törekvés az, hogy a minta átlaga megbízható becslést adjon az egész populáció átlagára. Matematikailag ezt a középérték közepes hibájával jellemezhetik, melyet az alábbi módon kell kiszámítani:

\[ SEM = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n(n-1)}} = \frac{SD}{\sqrt{n}} \]

ahol \( x \) a minta i-dik elemét, a minta átlagát, \( n \) a minta elemszámát és az SD a standard deviációja jelenti. Formalag tehát a SEM és az SD kiszámítása hasonló, azonban nagyon fontos kiszámítások során az egyik legfontosabb törekvés az, hogy a minta átlaga megbízható becslést adjon az egész populáció átlagára. Matematikailag ezt a középérték közepes hibájával jellemezhetik, melyet az alábbi módon kell kiszámítani:

\[ F(k) = P(x < k) \]

Az SD viszont nem csökken, hanem egyre jobban közelíti a populáció szórását.

4. Az eloszlások leírására használható módszerek

Az előzőekben tárgyalt függvények a populáció vagy a minta egy-egy jellegzetességet írják le. Ezek pontos jellemzése akkor válik lehetővé, ha megadjuk az összes lehetséges érték valószínűségét (vagy relatív gyakoriságát; 2.1. és 2.5. táblázat). Ezt a valószínőségi változó vagy a minta eloszlásának nevezzük. Ez a megközelítés alkalmazható diszkrét valószínőségi változó vagy egy tetszőleges minta eloszlásának leírására, bár néha célzott a lehetséges értékeket egyenrangúaknak vagy intervallumokba osztani (lásd Az adatok csoportosítása fejezet). Ezt a következőképpen láthatjuk be: egy folytonos valószínőségi változó esetén az eloszlás pontos értéke semmutható, hiszen az egész intervallum számítható eseménnyé. Ezt a következőképpen láthatjuk be: egy folytonos valószínőségi változó esetén az eloszlás pontos értéke nem mutatható, hiszen az egész intervallum számítható eseménnyé. Ezt a következőképpen láthatjuk be: egy folytonos valószínőségi változó esetén az eloszlás pontos értéke nem mutatható, hiszen az egész intervallum számítható eseménnyé.
Leíró statisztika: a populáció és a minta jellemzése
3.2.a. ábra). A sűrűségfüggvényre teljesül az, hogy - és + közötti grafikon alatti területe egyel egyenlő, hiszen ez annak a valószínűségét adja meg, hogy a valószínűségi változó értéke - és + között van, ami a biztos esemény, így valószínűsége egy.

Az eloszlás jellemzésére használt másik függvény az eloszlásfüggvény. Az $F(k)$ eloszlásfüggvényre teljesül az, hogy $k$ helyen felvett értéke megadja azt a valószínűséget, hogy az $x$ valószínűségi változó $k$-nál kisebb:

Teljesen hasonló megközelítés alkalmazható akkor is, ha az eloszlásfüggvényt nem egy valószínűségi változó, hanem egy minta jellemzésére használunk, csak akkor a valószínűségek helyett relatív gyakoriságokkal kell számolni.

Annak a valószínűségét, hogy egy valószínűségi változó egy adott intervallumba esik, az eloszlásfüggvény segítségével is ki lehet számítani. A
Leíró statisztika: a populáció és a minta jellemzése
Leíró statisztika: a populáció és a minta jellemzése

3.2.a. ábrán jelölt valószínűség kiszámításához vegyük észre, hogy ha $F(x)$-szel jelöljük az eloszlásfüggvényt, akkor

\[
F(60) = P(x < 60)\\
F(80) = P(x < 80)\\
F(80) - F(60) = P(x < 80) - P(x < 60) = P(60 \leq x < 80)
\]

Folytonos valószínűségi változó esetében $P(60 \leq x \leq 80) = P(60 \leq x < 80)$, mert $P(x = 60) = 0$.
Leíró statisztika: a populáció és a minta jellemzése

A

Sűrűségfüggvény

B

Eloszlásfüggvény

0,000

0,005

0,010

0,015

0,020

0,025

0,030

0,0

20

4

1,0

0,8

0,6

0,4

0,2

0,0

0,0

20

4

0,9786

0,7577
3.2.b. ábra).

Az eloszlásfüggvény („cumulative distribution function”) és sűrűségfüggvény („probability distribution” vagy „probability density function”) elnevezéseket angol megfelelőikkel összevetve zavaró és félreértésekre ad okot, hogy a sűrűségfüggvény egyik angol megfelelőjének („probability distribution”) magyar fordítása az eloszlás („distribution”) szót tartalmazza. Tovább bonyolítja a helyzetet, hogy az angol nyelvű szakirodalomban az eloszlásfüggvény („cumulative distribution function”) értelmezése eltér a fentebb ismertetettől, ugyanis ott következő definíciót használják:

\[ F(k) = P(x \leq k) \]

Az eloszlásfüggvény elkészítésének jobb megértéséhez a

![Diagram A](image1)

![Diagram B](image2)

3.4. ábrán a kockadobás és a 3.2. táblázat - A közép és a szóródás jellemzésére használt statisztikák szemléltetése 3.1. táblázatban feltüntetett minta eloszlásfüggvényét adjuk meg. Így példát látunk diszkrét eloszlásfüggvényére is.

5. Az eloszlás alakjának tömör jellemzése

Az eloszlás alakjának tömör numerikus jellemzésére két statisztika áll rendelkezésünkre.

Ferdeség (angol: skewness): az eloszlás aszimetriájának mértékét jellemzi. Kiszámításának alapja az, hogy képezzük a mintaelemek átlagtól való, SD-vel normalizált eltérésének harmadik hatványát, és ezt összegezzük minden mintaelemre:
Leíró statisztika: a populáció és a minta jellemzése

\[
\frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left( \frac{x_i - \bar{x}}{SD} \right)^3
\]

A pozitív ferdeség azt jelenti, hogy az eloszlásnak a magas értékek felé elnyúló része van, míg a negatív ferdeség esetében az átlagnál alacsonyabb értékek felé nyúlik el az eloszlás. Ha a ferdeség 0, az eloszlás szimmetrikus (3.3. ábrán). Az aszimmetria felmérésének egy igen egyszerű módja az, ha összehasonlítjuk a minta átlagát és mediánját. Ha az átlag nagyobb, mint a medián, pozitív ferdeségről beszélünk, ha kisebb, akkor a ferdeség negatív, ha a két statisztika értéke egyenlő, az eloszlás szimmetrikus.

Kurtózis: az eloszlás csúcsosságának mértékét jellemző statisztika:

\[
\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left( \frac{x_i - x}{SD} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}
\]

A kurtózis a vizsgált minta csúcsosságát egy referencia eloszláshoz, a normális eloszláshoz képest jellemzi. (A normális eloszlásról a következő fejezetben lesz szó.) Pozitív kurtózis esetében a vizsgált eloszlás csúcsosabb, mint a referencia eloszlás, tehát a módusz jobban kiemelkedik az eloszlás többi részéből. Negatív kurtózis esetében az eloszlás laposabban.

6. Standardizálás

Egy mintaelem értékének megadása önmagában nem mond sokat arról, hogy az adott elem hol helyezkedik el az eloszlásban. Ha pl. azt halljuk, hogy valakinek a szisztolés vérnyomása 160 Hgmm, tudjuk, hogy ez az érték relatív magas, hiszen a 160-at összehasonlítjuk a „normális” értékkel, illetve figyelembe vesszük azt is, hogy a normális értéktől való 40 Hgmm eltérés az sok. Matematikailag ezt a standardizálás vagy a „z” érték kiszámításával érjük el, ami megadja, hogy a mintaelem eltérése az átlagtól az SD hányszorosa:

\[
z = \frac{x_i - \bar{x}}{SD}
\]

ahol \( x \), a vizsgált mintaelem és a minta- vagy populáció átlag. A standardizálás segítségével különböző mérési eredmények átlagtól való eltérése összehasonlithatóvá válik.

7. Eloszlások grafikus megjelenítése

Az eloszlások reprezentálásának legszemléletesebb módja a grafikus megjelenítés. A következőkben a legelterjedtebb ábrázolási típusokat tárgyaljuk a teljesség igénye nélkül:
átlag vagy az átlag és a szórás (SD) ábrázolása: az egyik tengelyen a különböző mintákat tüntetjük fel, és erre merőlegszen az átlaggal arányos magasságú oszlopokat rajzolunk. Ha a szórást is meg kívánjuk jeleníteni, az oszlop tetejére a minta SD-jével vagy a középérték közepes hibájával arányos magasságú hibajeleket rajzolunk (3.5.a. ábra). Átlag vagy átlag-szórás ábrát nemcsak oszlopdiagram formájában, hanem vonalábra formájában is készíthetünk (}
Leíró statisztika: a populáció és a minta jellemzése

3.5.b. ábra.

doboz ábra: az átlag és a szórás megjelenítésére használt ábrázolási mód. A vízszintes tengelyen a különböző mintákat tüntetjük fel. Erre merőlegesen egy doboz kell rajzolni, aminek első, illetve felső határa az első, illetve harmadik kvartilisnek megfelelően helyezkedik el. A dobozba egy vízszintes vonalat kell húzni a második kvartilisnek (medián) megfelelően. A doboz ábrák szerkesztésének elvei nem egységesek. Egyes leírások szerint a doboz tetejétől egy vonalat kell húzni a legnagyobb, az aljától pedig egy másikat a legkisebb elemig. Mások szerint a lefelé és felfelé húzott vonalaknak a 10%-os, illetve 90%-os percentilisig kell érniük (
Leíró statisztika: a populáció és a minta jellemzése

3.5.c. ábra). arányok megjelenítése: ha a vizsgálati eredmények egy populáción vagy mintán belül a különböző osztályok arányát jelentik, ezeket egymás tetejére rajzolt oszlopok segítségével vagy körábrával jeleníthetjük meg. A körábrában az egyes körökkélyek nagysága az adott osztály százalékos arányával arányos (}
3.5.d.-e. ábra.
az eloszlás ábrázolása: folytonos valószínűségi változó esetében a sűrűségfüggvényt ábrázoljuk (
Leíró statisztika: a populáció és a minta jellemzése

A

B

Sűrűségfüggvény

Eloszlástömegfüggvény

0,000

0,005

0,010

0,015

0,020

0,025

0,030

0

20

4

1,0

0,8

0,6

0,4

0,2

0,0

0

20

4

0,7577

0,9786
3.2.a. ábra). Diszkrét valószínűségi változó esetében vagy ha a folytonos valószínűségi változó numerikus skáláját intervallumokra osztjuk, akkor a relatív gyakoriságot vagy a valószínűséget ábrázoljuk hisztogram formájában. A hisztogram olyan oszlopidagram, ahol az oszlopok magassága a relatív gyakorisággal vagy a valószínűséggel arányos.

3.5.f. ábra). pont diagram: olyan esetekben használjuk, amikor egy vizsgálati alany esetében két különböző paraméter mérését végezzük el, és a paraméterek összefüggését, korrelációját kívánjuk tanulmányozni. A két tengelyen a két mennyiség kerül feltüntetésre, és a koordináta rendszerben minden ábrázolt pont egy vizsgált egyedet jelent.
3.5.g. ábra).

8. Bevezetés a klinikai vizsgálatokba

Annak eldöntése, hogy egy új terápiás eljárás hasznosabb-e, mint a már meglevők, nem egyszerű feladat, mert a legtöbb esetben az új eljárás csak a betegek egy részében lesz hatásosabb, mint egy régebbi. Hasonlóan nehéz annak felmérése, hogy egy betegség kialakulásában feltételezett szerepet játszó rizikófaktor valójában milyen szerepet tölt be. Ilyen és ehhez hasonló kérdések megválaszolására klinikai vizsgálatokat szoktak végezni, melyek megtervezése és kiértékelése a biostatistika egyik legfontosabb feladata. Az alábbiakban nagyon röviden áttekintjük a legfontosabb vizsgálattípusokat. Mivel ezek magyar elnevezése nem közismert, megadjuk a megfelelő angol neveket is.
• **Megfigyeléses vizsgálatok** (observational studies): ezekben a vizsgálatokban a részt vevő betegek és egészséges emberek adatait rögzítik és analizálják. Legtöbbször a betegségek kialakulásáért felelős tényezők felderítésére használják (pl. a dohányzás valóban növeli-e a gégerák kockázatát). Négy alapvető típusát különböztetjük meg:

• Esetsorozat vizsgálat (case series study): ebben az esetben csupán néhány érdekes és tanulságos eset ismertetéséről van szó. Mivel statisztikai elemzést nem foglal magába, ezért sokan nem is sorolják a klinikai vizsgálatok közé, és csak arra alkalmazható, hogy a további vizsgálatok céljára releváns kérdéseket vessen fel.

• Eset-kontroll vizsgálat (case-control study): egy ilyen vizsgálat során egy betegségben szenvedő és egy kontroll csoportot hasonlítanak össze, és azt vizsgálják, hogy milyen tényezők vagy rizikófaktorok vannak jelen a betegcsoportban és hiányoznak a kontroll csoportban, a lehetséges tényezőket vagy rizikófaktorokat pl. kórlapokból keresve vissza. Ezen vizsgálatok tehát a „Mi történt?” kérdésre kerisik a választ, ezért retrospektív vizsgálatoknak is hívják. A vizsgálatok előnye, hogy rövid idő alatt olesön végrehajthatók. Hátrányuk, hogy a kontroll csoport csak az adott betegség fennállása tekintetében különbözzőn a beteg csoporttól. Mivel a legtöbb esetben a kontroll csoport olyan betegek képezik, akik nem a vizsgált betegség miatt kerültek kórházba, a fenti feltétel teljesülése gyakran megkérdőjelezhető. A legtöbb esetben azért nem valóban egészséges egyéneket választanak kontrollnak, mert egészséges emberekről nem állnak rendelkezésre körtörténeti adatok.

• **Keresztmetszeti vizsgálat** (cross-sectional study): a megfigyeléses vizsgálatok harmadik típusa során egy vizsgált csoport adatait elemzik egy adott időben, tehát a „Mi a helyzet jelenleg?” kérdésre kerisik a választ. Azért, hogy a betegségek vagy valamilyen paraméter időbeli változását követni lehessen, egyszerre több különböző csoporton végezik el ugyanazt a vizsgálatot (pl. olyan csoportokban, akikben 5, 10 illetve 20 éve áll fenn egy betegség). Ennek az a hátránya, hogy nem ugyanazt a csoportot követik végig, ami szisztematikus hibához vezethet (3.6. ábra).
Leíró statisztika: a populáció és a minta jellemzése

- **Kohorsz tanulmány** (cohort study): A vizsgálat irány fordított, mint az eset-kontroll vizsgálatoknál, tehát két csoportban (1. kontroll csoport; 2. olyan csoport, amely egy rizikófaktornak kitett) vizsgálják, hogy az idő előrehaladtával milyen betegségek alakulnak ki. Tehát ezen típusú vizsgálat a „Mi fog történni?” kérdésre keresi a választ, ezért prospektív tanulmánynak is nevezik. A megfigyeléses vizsgálatok közül a legmegbízhatóbb, hiszen nem érintik azon hátrányok, melyeket a keresztmetszeti vagy eset-kontroll vizsgálatoknál említettünk. Hátránya, hogy a vizsgált betegség kialakulásának sebességétől függően a tanulmány akár évekig is elhúzódhat. A leghíresebb ilyen vizsgálat a Framingham tanulmány volt, amelyben 1948-ban a Massachusetts állambeli Framingham város 6000 lakosát kezdték el figyelni évtizedeken keresztül a keringési betegségek és az érzelmeszesedés rizikófaktorait után kutatva.

- **Klinikai kísérletek** (clinical trials): terápiás eljárások hatékonyságának tesztelésére használt vizsgálattípusok. Hiszen a vizsgált terápiás eljárások hatékonysága, sőt egyes esetekben biztonságos alkalmazhatósága sem ismert a vizsgálat megkezdésekor, tulajdonképpen embereken végrehajtott kísérleteknek felelné meg, amelyekre szigorú etikai szabályok vonatkoznak. Két alapvető típusát különböztetjük meg:
  - **Kontrollal nem rendelkező vizsgálatok**: ebben az esetben egy vizsgált betegcsoporton alkalmazzák a tesztelni kívánt eljárást, de nincs olyan kontroll csoport, amihez ezt hasonlíthatjuk. A kontroll hiánya miatt ezen tanulmányok értéke a legalacsonyabb. Hívják őket egy ágú (one-arm) vizsgálatoknak is, hogy csak kezelt csoport (ág) van. Általában egy új eljárás kezdeti tesztelésére használják.
  - **Kontrollal rendelkező vizsgálatok**: Három típusát különböztetjük meg:
    - **parallel kontroll**: két betegcsoportot hasonlítanak össze. Az egyik csoportot a vizsgált eljárással kezelik, a kontroll csoportot vagy nem kezelik, vagy egy régebbi módszer szerint kezelik (amihez az új eljárást hasonlítani kívánják) vagy placebóval kezelik. Placebónak nevezik azt az anyagot, ami csak abban különbözik a vizsgált gyógyszerétől, hogy nem tartalmaz hatóanyagot (ugyanolyan színű, ugyanolyan
adagolásban adják, stb.). A szubjektív tényezők kiszűrése miatt kell használni. Ha a betegek véletlenszerűen kerülnek a kezelt vagy a kontroll csoportba, randomizált vizsgálatról beszélünk. A nem randomizált vizsgálatok értéke kisebb, mert a kezelő orvos szubjektíven dönheti el, hogy ki melyik csoportba kerül, ami szisztematikus hibához vezethet. Ha a betegek nem tudják, hogy a kontroll vagy kezelt csoportba tartoznak, vak (blind) vizsgálatról beszélünk. Ha sem a beteg, sem az orvos nem ismeri, hogy a beteg melyik csoportba tartozik, kettős vak (double-blind) vizsgálatról beszélünk. Ebben az esetben a klinikai kísérletet szervező személyzet tudja a betegeket azonosítani. A kettős vak vizsgálatok kiküszöbölők azt, hogy a beteg vagy az orvos azt tapasztalja a vizsgálat során, amit vár (tehát pl. azt, hogy az új eljárás jobban működik). Ha a beteg és az orvos is tudja, hogy a páciens milyen kezelést kap, nyílt jelölésű (open-label) vizsgálatról beszélünk. A bioistatistika jelenlegi állása szerint a legmegbízhatóbb klinikai vizsgálatok a randomizált, kettős vak, placebo kontrolloso tanulmányok.

- **belső kontroll:** az ilyen vizsgálat lehet önkontrollos, amikor a kísérlet kezdődése előtt meghallgatják a betegeket (pl. mennyire magas a vérnyomásuk), majd végrehajtják rajtuk a beavatkozást (pl. sóssegény dietája 2 éven keresztül), végül ellenőrizzük ugyanazon betegcsoporton a vizsgálat perspektíváját változását. A belső kontroll és a parallel kontrolloso vizsgálatok leveréke a keresztezett (crossover) vizsgálat. Ebben az esetben a vizsgálat első részében az egyik betegcsoportot kezelik, a másik a kontroll csoport, majd mind a két csoporton rövid, ún. kimosási (washout) időre felfüggesztik a kezelést, majd a korábbi kontroll csoport most a kezelt csoport lesz és fordítva. Ezen vizsgálat előnye, hogy kiküszöbölők a két betegcsoportba tartozó emberek egyéni variabilitása miatt hibát.

- **történalmi kontroll:** a vizsgálat során kezelt betegek eredményeit korábbi vizsgálatok eredményeihez hasonlíthatják. Hátránya, hogy nehéz ellenőrizni azt, hogy a korábbi és mostani vizsgálatok körülményei mennyire voltak azonosak.

- **Meta-analízis:** Az orvosi statisztikai vizsgálatok széleskörű elterjedésével egyre gyakoribbá vált, hogy egy gyógyszer vagy rizikófaktor hatását több vizsgálat is elemzett, és ezek nemrítkán eltérő következtetésekre jutottak. A meta-analízis a különböző vizsgálatok eredményeinek statisztikailag egzakt összesítése, melynek részletes módszereire nincs lehetőségünk kitérni. A meta-analízis célja lehet i) statisztikai megbízhatóság növelése azáltal, hogy több tanulmány összetevőivel az effektív mintamért növekszik; ii) az ellentmondó konklúzióra jutó vizsgálatok ellentmondásainak feloldása. A meta-analízis értékes statisztikai módszer, de hátrányaival tisztában kell lenniünk, melyek közül a legfontosabbak: i) a különböző vizsgálatok eltérő tervezése miatt ezeket összevetni, eredményeiket összehasonlítani nehéz és sok hibaforrása van; ii) az orvosi irodalomban sokkal gyakoribb a pozitív eredmények leírása, mint a negatívoké, ezért a meta-analízis eredménye könnyen lehet torzított a pozitív eredmények felé.

9. **A klinikai vizsgálatok fázisai**

Egy gyógyszer kifejlesztése a sejt- és molekuláris biológiai laboratóriumiakban kezdődik, és több év kell ahhoz, hogy klinikai kipróbálásra, majd mindennapi használatba kerüljön. A klinikai kipróbálás veszélyes folyamat, hiszen embereken végzett kísérletekről van szó, ezért meghatározott időrendben lehet csak elvégezni. Ennek megfelelően a következő fázisokról beszélünk:

1. **I. fázis:** Általában kevés (50-100) önként vállalkozó egészséges emberen vizsgálják a tolerált dózist, a gyógyszer metabolizmusának sebességét, kiürülését a szervezetből, azaz farmakokinetikáját. Ha potenciálisan súlyos mellékhatásokkal rendelkező gyógyszer tesztselnek (pl. daganatok ellen használt citosztatikumokat), akkor nem egészséges, hanem a kezelendő betegségben szenvedő önként vállalkozókon hajtják végre a megfigyeléseket.

2. **II. fázis:** Néhány száz önként vállalkozó betegen hajtják végre. Célja a gyógyszer hatásosságának kezdeti felmérése, a dózis-hatás összefüggések felmérése.

3. **III. fázis:** Néhány ezer önként vállalkozó betegen, több helyen végrehajtott (ún. multicentrikus) vizsgálatok, melyeknek célja a gyógyszer hatásosságának nagyobb beteganyagony történő demonstrálása, mellékhatások megfigyelése, és a gyógyszer hatásosságának speciális betegcsoportokban történő vizsgálata. A multicentrikus kísérő az egy-egy kórház speciális adottságairól eredő variabilitást.

4. **IV. fázis:** A gyógyszer engedélyezése után történő megfigyelések tartoznak ide. A gyógyszer szélsőkörű használata a megfigyelhető kezelt betegek számát megsokszorozza, így ritka mellékhatások is megfigyelhetők. E vizsgálatok hozzájárulnak továbbá a gyógyszer dozírozásának pontosabb beállításához.
3.1. táblázat - Látóterenkénti sejtszám-adatok csoportosítása

<table>
<thead>
<tr>
<th>Sejtek száma</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gyakoriság</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Relatív gyakoriság</td>
<td>1/9</td>
<td>3/9</td>
<td>1/9</td>
<td>2/9</td>
<td>2/9</td>
</tr>
</tbody>
</table>

Egy mikroszkópos kísérlet során azt vizsgálták, hogy látóterenként hány sejt fejezte ki az epidermális növekedési faktor receptort a sejtmembránban. A táblázatban feltüntettük a feldolgozatlan nyers adatsort, és a csoportosított adatokat a gyakoriság és a relatív gyakoriság feltüntetésével.

3.2. táblázat - A közép és a szóródás jellemzésére használt statisztikák szemléltetése

<table>
<thead>
<tr>
<th>Rendezett minta</th>
<th>Átlag</th>
<th>Medián</th>
<th>Megnyesett átlag</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3,11</td>
<td>3</td>
<td>3,14</td>
<td>1,45</td>
</tr>
<tr>
<td>B</td>
<td>7,8</td>
<td>3,5</td>
<td>3,38</td>
<td>14,89</td>
</tr>
<tr>
<td>C</td>
<td>3,11</td>
<td>3</td>
<td>3,14</td>
<td>0,6</td>
</tr>
</tbody>
</table>

A feltüntetett rendezett mintákon az átlag, medián, megnyesett átlag és SD statisztikákat tüntettük fel. A megnyesett átlagot a legnagyobb és legkisebb elemek elhagyásával határoztuk meg.
4. fejezet - Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások (Szöllősi János dr.)

Valószínűség számítás során számos jelenségről megállapítható, hogy valószínűség szempontjából hasonlóan viselkednek, pl. ugyanazt az eloszlást követik, vagyis a valószínűségi változó x, értékeihez a p, valószínűségek azonos eljárással, (függvényel, képlettel) számítható ki, függetlenül az x, jelentésétől. Ezekben az esetekben a paraméterértékek különbözőek lehetnek, de ugyanazt azt sűrűség vagy eloszlásfüggvényt alkalmazhatjuk. Az alábbiakban azt írjuk le, hogy bizonyos feltételeknek eleget tevő valószínűségi változóknak milyen jellegzetes eloszlásfüggvényei vannak. Külön fogjuk tárgyalni a diszkrét, illetve a folytonos nevezetes valószínűség-eloszlásokat.

1. Diszkrét eloszlások

1.1. Binomiális eloszlás

Tekintsünk egy független kísérletekből álló kísérletsorozatot, amelyben minden kísérletnek két lehetséges kimenetele lehet, A illetve ennek ellentett eseménye A felülvonás. Ilyen típusú kísérlet pl. az érme feldobása, a születendő gyernek neme, súlyos fertőzód megbetegedésnél a betegség kimenetele. Legyen az A esemény bekövetkezésének valószínűsége minden kísérletben állandó \( P(A) = p \) és az ellentett esemény bekövetkezésének valószínűsége pedig \( P(\bar{A}) = 1 - p = q \). Végezzük el \( n \)-szer egymástól függetlenül a kísérletet és legyen az a diszkrét valószínűségi változó, amely azon esetek számát számítjuk, amelyeknél az A esemény következett be. Annak valószínűségét, hogy az A esemény éppen \( k \) alkalommal következett be, az alábbi összefüggéssel számolhatjuk ki:

\[
P(k = k) = \binom{n}{k} p^k q^{n-k} = \frac{n!}{(n-k)!k!} p^k q^{n-k}
\]

ahol \( k = 0, 1, 2, \ldots, n \) lehet. Az ilyen diszkrét valószínűségi változót \( n \)-ed rendű és \( p \) paraméterű binomiális eloszlású változónak nevezzük. (A binomiális szó jelentése „két névből álló” vagy „két kategóriából álló”.) A binomiális valószínűségi változó várható értéke:

\[
\mu = np
\]

varianciája pedig:

\[
\sigma^2 = npq
\]

Példa: Vizsgáljuk meg a fiú és leány gyermekek születésének valószínűségi eloszlását egy családban, adott gyermekszámánál. A statisztikai adatok alapján a fiú születésének valószínűsége \( p = 0.515 \), és ebből adódóan, a leány születésének valószínűsége \( q = 1 - p \) összefüggés alapján, \( q = 0.485 \). (Ennek biológiai magyarázata részben az Y kromoszómát hordozó spermiumok könnyebbségéből, nagyobb mozgékonyságából adódik.) Tekintsünk olyan családot, amelyekben 6 gyermek van, azaz \( n = 6 \), és a binomiális eloszlás alapján számoljuk ki annak valószínűségét, hogy a családban 0, 1, 2, 3, 4, 5, 6 fiú van. A számított valószínűségeket az 4.1. táblázat - 4.1. táblázatban tüntettük fel, a binomiális eloszlás hisztogramja pedig az

42

Created by XMLmind XSL-FO Converter.
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szöllősi János dr.)

4.1. ábrán látható.

4.1. táblázat -

<table>
<thead>
<tr>
<th>Fiúk száma</th>
<th>Lányok száma</th>
<th>Valószínűség</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>?</td>
<td>64</td>
<td>~1,00</td>
</tr>
</tbody>
</table>

A fiú és leány gyermekek születésének valószínűségi eloszlását egy családban, adott gyermekszámnál.
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szőllősi János dr.)

A lehetséges valószínűsége összege 1-et ad, ami a biztos esemény bekövetkezési valószínűsége, ugyanis a felsorolt események közül az egyik biztosan bekövetkezik.

1.2. Poisson-eloszlás

Nézzük meg, mi lesz a binomiális eloszlás határesete olyan körülmények között, ha a kísérletek számát (n) minden határon túl növeljük úgy, hogy közben az A esemény bekövetkezési valószínűsége (p) minden határon túl csökken, és az np szorzat egy g 0 határértékhez tart. Ebben az esetben a diszkrét valószínűségi változót paraméterű Poisson-eloszlású valószínűségi változónak nevezzük, ilyenkor a valószínűségi változó a 0, 1, 2, 3, .... n értékeket

\[ P_k = \frac{\lambda^k}{k!} e^{-\lambda} \]

valószínűséggel veszi fel. (Az eloszlás a nevét a francia matematikusról Siméon Poissonról kapta, aki 1837-ben módosította a binomiális képletet izolált, véletlenszerű előfordulások leírására, amennyiben korábbról ismertük az átlagos előfordulások számát.) A Poisson-eloszlás paramétere szigorú kapcsolatban áll a valószínűség-eloszlás várható értékével és variációjával.

\[ M(\xi) = \mu = \lambda \]

\[ D^2(\xi) = \sigma^2 = \lambda \]
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szöllősi János dr.)

A Poisson-eloszlás jelentőségét az adja, hogy sok esetben fordul elő természeti jelenségeknél, biológiai problémák nál, így ezek előfordulási valószínűsége a Poisson-eloszlás segítségével tárgyalható. A valóságban persze az n értéke nem mindig végként, de ilyenkor is jó közelítést ad a Possion-eloszlás, ha az n elegendően nagy. A különböző Poisson-eloszlásoknál az a közös vonás, hogy az események véletlenszerűen oszlanak el, időben, síkban, vagy térben. Ha az egyes események egymástól függetlenül azonos valószínűséggel eshetnek bármelyik, egyébként azonos nagyságú tér- sík-, vagy időrészbe (intervallumba), akkor az események elhelyezkedése Poisson-eloszlást követ. Ilyen események lehetnek az emberek születésének időpontja, egy telefonközpontba adott idő alatt befutó hívások száma, egy forgalmas útkereszteződésben egy adott idő alatt áthaladó járművek száma, egy adott V térfogathban lévő kolloid részecskék száma, egy tenger adottírészében található halak száma, a Bürker kamra látómezőjében megfigyelt vörös vagy fehérvérsejtek száma, egy adott radioaktiv preparátumban egy adott időegység alatt elbomló atommagok száma, stb.

Ezeknél az esetéknél a Poisson-eloszlás praktikusságát az adja meg, hogy csak egy paramétert kell ismerni, a ?-t, és ez meg is határozza az eloszlás alakját is.

Példa:

Vizsgáljuk meg, hogy a hemocitométer egyes négyzetrácsai ban megszámlálható élesztő sejtek száma Poisson-eloszlást követ-e. A hemocitométerre kicseppentették az élesztő szuszpenziót és megszámlálták azoknak a négyzeteknek a számát, amelyek 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 darab élesztőt tartalmaznak. A 4.2. táblázat - Élesztő sejtek eloszlása a hemocitométerben. (Student adatok) A 4.2. táblázatban tüntettük fel az élesztő sejtek eloszlását. (Az adatok Student adatai). A táblázat adataiból kiszámítható, hogy várható érték ? = 4,68, ezzel a paraméterrel meghatározott az egyes valószínűségi változó előfordulási valószínűségét, valamint kiszámoltuk, mennyi lenne a 0, 1, 2, 4, stb. élesztőt tartalmazó négyzetek száma. A 4.2. táblázat - Élesztő sejtek eloszlása a hemocitométerben. (Student adatok) Ezeket az adatokat is tartalmazza, a valószinűségek eloszlását ugyanakkor a...
4.2. ábrán is bemutatjuk. A számítás alapján látható, a kísérleti eloszlás jó közelítéssel Poisson-eloszlásnak tekinthető. (Az illeszkedés pontosságának megállapítására ?² próbát lehetne használni, lásd később.)

4.2. táblázat - Élesztő sejtek eloszlása a hemocitométerben. (Student adatai)

<table>
<thead>
<tr>
<th>A négyzetben megfigyelt sejtek száma ()</th>
<th>Négyzetek (megfigyelt) száma</th>
<th>Pk valószínűség (?=4,68)</th>
<th>Négyzetek (számított) száma</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0,00928</td>
<td>3,7</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>0,04343</td>
<td>17,4</td>
</tr>
<tr>
<td>2</td>
<td>43</td>
<td>0,10162</td>
<td>40,6</td>
</tr>
<tr>
<td>3</td>
<td>53</td>
<td>0,15852</td>
<td>63,4</td>
</tr>
<tr>
<td>4</td>
<td>86</td>
<td>0,18547</td>
<td>74,2</td>
</tr>
<tr>
<td>5</td>
<td>70</td>
<td>0,17360</td>
<td>69,4</td>
</tr>
<tr>
<td>6</td>
<td>54</td>
<td>0,13541</td>
<td>54,2</td>
</tr>
<tr>
<td>7</td>
<td>37</td>
<td>0,09053</td>
<td>36,2</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
<td>0,05296</td>
<td>21,2</td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>0,02754</td>
<td>11,0</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>0,01289</td>
<td>5,2</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>0,00548</td>
<td>2,2</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>0,00214</td>
<td>0,9</td>
</tr>
<tr>
<td>12-nél több</td>
<td>0</td>
<td>0,00114</td>
<td>0,4</td>
</tr>
<tr>
<td></td>
<td>400</td>
<td></td>
<td>400,0</td>
</tr>
</tbody>
</table>

2. Folytonos eloszlások

2.1. Egyenletes eloszlás

Eddig diszkrét változók eloszlásait elemeztük. Most tekintsük a folytonos változók eloszlásai közül a legegyszerűbbet, a folytonos egyenletes eloszlást. Egy változót egyenletes eloszlásúnak nevezzük, ha sűrűségfüggvénye a következő:

\[ f(x) = \begin{cases} 
0, & x \leq a \\
\frac{1}{b-a}, & a < x \leq b \\
0, & x > b 
\end{cases} \]

Tehát \( f(x) \) konstans (\( c \)) az \( (a,b) \) intervallumban, egyébként pedig 0. A konstans értéke (\( c \)) \( 1/(b-a) \) kell, hogy legyen, hogy a függvény alatti terület kiadja a biztos esemény bekövetkezési valószínűségét, azaz 1-et. Megemlíthetjük, hogy például az óramutató helyzete egyenletes eloszlású valószínűség változó.

Példa:

Egy kórházi telefonközpont telefonhívásainál azt tapasztalják, hogy a tárcsázást követő időtartam 5 és 85 másodpercig terjed. Az eltelt idő legyen a egyenletes eloszlású valószínűségi változó. Határozzuk meg a valószínűségi változó sűrűség- és eloszlásfüggvényét.

Feltettük, hogy a valószínűségi változó egyenletes eloszlású a 5 másodperctől 85 másodpercig terjedő intervallumban, ezért a sűrűségfüggvénye:
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szőllősi János dr.)

Eloszlás. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások

\[
f(x) = \begin{cases} 
0, & x \
1, & x \\ 
\frac{1}{80}, & 5 < x \\
0, & x > 85 
\end{cases}
\]

Az valószínűségi változó eloszlásfüggvénye pedig az alábbiak alapján adható meg:

\[
F(x) = P(X < x) = \begin{cases} 
0, & x \
\frac{x-85}{80}, & 5 < x \\
1, & x > 85 
\end{cases}
\]

A

4.3. ábrán tüntettük fel ennek az egyenletes eloszlásnak a sűrűség- és eloszlásfüggvényét.

2.2. Exponenciális eloszlás

A valószínűségi változót exponenciális eloszlású valószínűségi változónak nevezzük, ha az eloszlás \( f(x) \) sűrűségfüggvényére érvényes a következő egyenlet,

\[
f(x) = \lambda e^{-\lambda x},
\]

ha \( x \geq 0 \).

A \( \lambda \) állandót az eloszlás paraméterének tekintjük. A sűrűségfüggvény és az eloszlásfüggvény kapcsolata alapján az exponenciális eloszlású valószínűségi változó eloszlásfüggvénye:

\[
F(x) = F(\xi < x) = 1 - e^{-\lambda x},
\]

ha \( x \geq 0 \).

Exponenciális eloszlást követnek a különféle várakozási idők, például a radioaktív bomlás során az egyes atomok élettartama. Ugyancsak exponenciális eloszlásúak a használati tárgyak vagy azok különböző alkatrészeinek élettartamai. Az exponenciális eloszlású valószínűségi változó várható értéke és varianciáját a következő összefüggések adják meg:

\[
\overline{M}(\xi) = \mu = \frac{1}{\lambda}
\]
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szőllősi János dr.)

\[ D^2(\xi) = \sigma^2 = \frac{1}{\lambda^2} \]

Példa:

Egy CT berendezés működési ideje a meghibásodásig exponenciális eloszlású. A folyamatot leíró valószínűségi változó várható értéke legyen 500 óra. Határozzuk meg a exponenciális eloszlású valószínűségi változó sűrűség- és eloszlásfüggvényét. A fentiek alapján a \( \lambda \) paraméter értéke kiszámítható:

\[ \lambda = \frac{1}{M(\xi)} = \frac{1}{500} = 0,002 \]

A \( \lambda \) paraméter segítségével a CT-berendezés meghibásodásának sűrűségfüggvénye és eloszlásfüggvénye kiszámítható. A számított értékek ábrázolását a
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szőllösi János dr.)

4.5. ábrán mutatjuk be.

2.3. Normális eloszlás

A statisztikában a legfontosabb és leggyakrabban alkalmazott eloszlás a normális eloszlás. A normális eloszlás göbét először egy francia matematikus, Abraham de Moivre fedezte fel és közölte le 1733-ban. A normális eloszlást tudományosan két matematikus-csillagász, a francia Pierre-Simon Laplace és a német Carl Friedrich Gauss alapozta meg. Többen úgy vélik, hogy Laplace hozzájárulása a normális eloszlás tulajdonságainak tisztázásához jelentősebb volt, mint Gaussé, mégis Gauss után nevezték el a normális eloszlást Gauss eloszlásnak, miután Gauss volt az első, aki a normális eloszlást égitek mozgására alkalmazta. A természetben, az orvostudományban nagyon sok mért paraméter normális eloszlással írható le, mint például az egyének magassága, vérnyomása, súlya, stb. A normális elnevezés is arra utal, hogy a mért adatainktól ezt várjuk, mert ez a természetes viselkedésük.

Egy folytonos valószínűségi változót normális eloszlásúnak nevezünk µ és ? paraméterekkel, ha a sűrűségfüggvénye a következő képlettel adható meg:

\[ f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]

A megfelelő eloszlásfüggvény:

\[ F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \]
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szöllősi János dr.)

Mint említettük, ennek az eloszlásnak nagy jelentősége van a matematikai statisztikában, ezért tulajdonságait részletesen tárgyaljuk. Külön szépsége a normális eloszlásnak, hogy az eloszlást jellemző paraméterek a μ és a σ egyből kiolvashatók az eloszlás sűrűség vagy eloszlásfüggvényéből.

\[ M(\xi) = \mu \]

\[ D(\xi) = \sigma \]

A normális eloszlás sűrűségfüggvényét a

4.6. ábra, az eloszlás görbéjét pedig a
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szöllősi János dr.)

4.7. ábra

A Gauss görbe harang alakú és szimmetrikus a \( \mu \) várható értékre, és ez a pont egyúttal a függvény egyetlen maximumhelye. Miután a normális eloszlás szimmetrikus, a várható érték egyben az eloszlás mediánja és módusa is. Differenciálással meggyőződhetünk rá, hogy az \( f(x) \) függvénynek két inflexiós pontja van, mégpedig a \( \mu - \sigma \) és \( \mu + \sigma \) helyeken. Gyakori feltevés, hogy a mérési hibák eloszlása a \( \mu \) átlag körüli normális eloszlás, \( \mu \) t így szokás az eloszlás átlagának is nevezni. Az eloszlás \( \sigma \) paramétere az eloszlás standard deviációja, melyet a minta standard deviációjával közelíthetünk. A két paraméternek speciális jelentése van: annak valószínűsége, hogy egy egyedi megfigyelés a valódi értéktől (az eloszlás átlagától) egyszeres standard deviációira tér el, 0,682. Ez elég alacsonynak tűnik, ezért a kutatók a standard deviáció 2- vagy 3-szorosát szokták venni, amellyel ez a valószínűség 0,954-re illetve 0,998-re emelkedik. Tehát annak valószínűsége, hogy egy egyedi megfigyelés a valódi értéktől (az eloszlás átlagától) kétszeres standard deviációira tér el, 0,954. A
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások (Szöllősi János dr.)

4.8. ábra bemutatja a paraméterek jelentését az $f(x)$ sűrűségfüggvényen.

Normális eloszláscsaládba tartozó függvények alakja hasonló, egyik a másikba átszámolható, az $x$ tengely menti elhelyezkedésüket a $\mu$ (\(\sigma\))
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szöllősi János dr.)

4.9. ábra), a szélességét pedig a $\sigma$ paraméter határozza meg (}
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások (Szöllősi János dr.)

4.10. ábra). Látható, hogy míg a $\mu$ változtatása a Gauss görbe eltolását jelenti az x tengely mentén, addig $\sigma$ megváltoztatása a görbe laposságát befolyásolja, minél nagyobb a $\sigma$, annál laposabb és szélesebb a görbe. Vegyük észre, hogy minden esetben, (így a $\sigma$ megváltoztatásánál is) a görbe alatti terület egyforma, 1-el egyenlő, a biztos esemény valószínűségét adja meg.

A normális eloszlások családjának legjelentősebb tagja a standard normális eloszlás, aminek várható értéke $\mu = 0$, és szórása $\sigma = 1$. Ebben az esetben a sűrűségfüggvényt $\varphi(x)$-nel, az eloszlásfüggvényt pedig $\Phi(x)$ jelöljük, a függvényeket leíró képletek pedig tovább egyszerűsödnek:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$
$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$$

A standard normális eloszlás sűrűség illetve eloszlásfüggvényét mutatja a
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szöllősi János dr.)

\[
\phi(x) \approx 0.4 \quad \text{és} \quad \frac{1}{\sqrt{2\pi}} \approx 0.4
\]
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szőllősi János dr.)

Bármilyen normális eloszlás átalakítható standard normális eloszlássá. Ha a válószínűségi változó \( \mu \) várható értékű és \( \sigma \) szórású, az alábbi transzformáció

\[
Z = \frac{X - \mu}{\sigma}
\]

után kapott \( z \) változó 0 várhatóértékű, és 1 szórású standard normális eloszlású. Ezért, ha az \( x_1, x_2, x_3, \ldots, x_n \) minta egy \( \mu \) várható értékű és \( \sigma \) szórású eloszlású populációból származik, a minta \( z \) értékei, azaz a standardizált mintaelemek, standard normális eloszlásúak lesznek.

\( \Phi(x) \) függvény értékei táblázatos formában általában minden statisztikával foglalkozó könyvben megtalálhatók. Ezen a táblázatok egy része minden egyes \( x \)-hez megadják a sűrűségfüggvény alatti területet az \( x \)-től balra. Más táblázatok az eloszlás szimmetriája miatt csak pozitív \( x \)-ekre közlik a fenti értékeket, vagy nem az \( x \)-től balra, hanem jobbra eső területet vagy félterületet tartalmazzák. Ugyanakkor olyan táblázatok is állnak rendelkezésre, amelyek nem az eloszlás, hanem a sűrűségfüggvény értékeit tartalmazzák. A standard normális eloszlás táblázatok segítségével bármilyen normális eloszlás adatait kiszámolhatjuk a transzformációs képlet segítségével. Így elegendő egy táblázat készlet, a standard normális eloszlás táblázat készlet, nem kell minden különböző \( \mu \)-vel és \( \sigma \)-val rendelkező normális eloszlásra külön táblázatokat szerkeszteni.

1. példa. Adjuk meg a standard normális eloszlás alatti területet az \( x = -1,65 \) és \( x = 1 \) helyek között.Megoldás: \( \Phi(-1,65) = 0,0495 \), \( \Phi(1) = 0,8413 \). A keresett területet kivonással kapjuk: \( 0,8413 - 0,0495 = 0,7918 \)
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások

(Szöllősi János dr.)

2. példa. Bizonyos laboratóriumban a kísérleti patkányok testsúlyait normális eloszlásúnak találták \( \mu = 42 \text{ dkg} \) átlaggal és \( \sigma = 6 \text{ dkg} \) szórással. Egy ilyen populációban mi annak a valószínűsége, hogy a patkányok testsúlya 30 és 45 dkg közé esik? 

Megoldás: A 4.13. ábra vázlatosan mutatja ezt az eloszlást. Standardizálás után a \( \mu = 42 \)-nek megfelel a \( z = 0 \), a \( \sigma = 6 \) standard deviációknak pedig a \( z = 1 \). Alkalmazzuk a z transzformációt a 45-re és a 30-ra, kapjuk a következő standardizált értékeket:

\[
\begin{align*}
Z_{45} &= \frac{45 - 42}{6} = 0.5 \\
Z_{30} &= \frac{30 - 42}{6} = -2
\end{align*}
\]

\( z(0.5) = 0.6915 \) és \( z(-2) = 0.0228 \). Kivonás után \( 0.6915 - 0.0228 = 0.6687 \). Tehát várhatóan a populáció 67%-ának a testsúlya fog 30 és 45 dkg közé esni.

3. példa. Frankenstein professzor vámpír denevéreket telepít a laboratóriumbára. A denevérek tépőfogainak a hossza normális eloszlást követ \( \mu = 28 \text{ mm} \) átlaggal és \( \sigma = 4 \text{ mm} \) szórással. Frankenstein tudja, hogy azoknak az
állatoknak a harapása halálos, akiknek a tépőfogmérete a populáció felső 5%-ába esik. Számítsuk ki, hogy ez hány mm-es fogmértet jelent.

Megoldás. A szituációt a

4.14. ábra szemlélteti. Mivel az eloszlás táblázatban az \( x \)-től balra eső területek vannak meg, az \( x \)-től jobbra eső terület nagyságát kivonással kapjuk: 1-0,05=0,95. Mivel 0,95 nincs pontosan benne a táblázatban, 0,9505-öt fogunk használni. Az ehhez tartozó \( x \) érték 1,65 standard normális eloszlás esetén 1,65, amelyet most vissza kell számolnunk a \( \mu = 28 \) várható érték és \( \sigma = 4 \) szórású normális eloszlású valószínűségi változóvá.

\[
X = \mu + x \cdot \sigma = 28 + 1,65 \cdot 4 = 34,6
\]

Tehát, a 35 mm-nél hosszabb fogú denevérek esnek a populáció felső 5%-ába, akiknek halálos a harapása.

2.4. Lognormális eloszlás

Egy valószínűségi változó lognormális eloszlású, ha a változó logaritmusá

\[ \varphi = \ln \xi \]

normális eloszlású. A lognormális eloszlás sűrűségfüggvénye:
Az eloszlás várható értéke és variánciája

\[ M(\xi) = \mu = e^{m + \frac{\sigma^2}{2}} \]

\[ D^2(\xi) = \sigma^2 = e^{2m + \sigma^2} (e^{\sigma^2} - 1) \]

A statisztikában gyakran alkalmazott eljárás, ha a minta eloszlása nem normális, akkor vesszük a mintaelemek logaritmusát. Ezzel a művelettel gyakran sikerül az adatokat normális eloszlásúvá transzformálni. A

![Diagram](image)

4.15. ábra mutatja egy lognormális eloszlás görbéjét.

Hogyan is alakulhat ki a lognormális eloszlás? Az orvostudományban például sok valószínűségi változó csak pozitív értékeket vehet fel (pl. vércukor szint, koleszterin koncentráció stb.). Ilyenkor, ha a várható érték nem túl magas, közel van a 0-hoz, és ehhez képest a szórás nagy (vagy másféle a variációs kockációs nagy) az eloszlás öhatatlanul aszimmetrikussá válik, az eloszlás görbe a pozitív irányba eltorkul, elnyúlik. Azt is felhozhatjuk érvként, hogy pozitív valószínűségi változó nem is lehet normális eloszlású, hiszen a normális eloszlású valószínűségi változó egyaránt vehet fel negatív és pozitív értékeket. Ez a nézet azonban szükségtelenül korlátozó jellegű, hiszen ha az adott normális eloszlásnál elhanyagolható annak a valószínűsége, hogy nullánál kisebb értékeket vegyen fel, a normális eloszlást teljesen jól lehet leírni olyan valószínűségi változóval, amely csak pozitív értékeket vehet fel. Ha a variációs kockációs nél kisebb, az utóbbi megközelítés érvényes és normális eloszlást kapunk. Ha a variációs kockációs 1 vagy annál nagyobb, a csak pozitív értéket felvehető valószínűségi változó eloszlása torzult lesz, nagyon sok esetben lognormális eloszlással lesz megközelíthető. Az élő sejtek felszínén kifejeződő különböző féhérjék eloszlása gyakran követi a lognormális eloszlást. Ezekben az esetekben gyakran előfordul, hogy az áramlási citométerrel mért
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások (Szöllősi János dr.)

Fluoreszcencia intenzitások (amelyek az kifejeződési szintekkel arányosak) logaritmusát használjuk fel az eloszlási hisztogramok szerkesztésénél. Az így kapott hisztogramok normális eloszlást követnek, rajtuk a statisztikai elemzések könnyen elvégezhetőek. A statisztikai vizsgálat végén azonban nem szabad elfelejteni a számol paraméterek (átlag, szórás stb.) visszaalakítását.

2.5. $\chi^2$ eloszlás

A $\chi^2$ (khi-négyzet) eloszlást a normális eloszlásból származtatjuk. Tételezzük fel, hogy $\mu$ várható értékű és $\sigma$ szórású normális eloszlású $\xi$ valószínűségi változóból veszünk egy $n$ elemű mintát. A mintaelemeket standard normális eloszlású valószínűségi változóvá transzformálunk az ismert képlet alapján:

$$z_i = \frac{X_i - \mu}{\sigma}$$

Mindemellett $z$ értéket négyzetre emelünk és a négyzeteket összeadjuk. A $z^2$-ek összege $n$ szabadságfokú khi-négyzet eloszlást fog követni. Általánosságban tehát a

$$\chi^2_n = z_1^2 + z_2^2 + z_3^2 + \ldots + z_n^2$$

$n$ szabadságfokú khi-négyzet eloszlással írható le, amelynek sűrűségfüggvénye az alábbi összefüggéssel írható le:

$$f_n(x) = \left(\frac{n}{2}\right)^{-\frac{n}{2}} \frac{1}{2^n} e^{-\frac{x}{2}}, \quad x \geq 0$$

Az $n$ szabadságfokú khi-négyzet eloszlás várható értéke és varianciája:

$$\bar{M}(\chi^2) = \mu = n$$

$$D^2(\chi^2) = \sigma^2 = 2n$$

A
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások (Szöllősi János dr.)

4.16. ábrán látható, hogy a khi-négyzet eloszlások szabadsági fokoktól függő eloszlások én nem szimmetrikusak. Az $n = 1$ és az $n = 2$ szabadságfokú khi-négyzet eloszlás képe teljesen eltér a $n \geq 2$ szabadságfokú eloszlásokétől. Az $n \geq 2$ szabadságfokú khi-négyzet eloszlások módusa $(n-2)$ értékké egyezik meg.

A khi-négyzet csak 0 és +? között vehet fel értékeket, hiszen négyzetek összegéről van szó. Érdemes megjegyezni, hogy két vagy több khi-négyzet valószínűségi változó összegei szintén khi-négyzet eloszlást követnek. Az $n \geq 30$ szabadságfokú khi-négyzet eloszlások jól megközelíthetők a normális eloszlással. A későbbiek során látjuk fogyni, hogy a khi-négyzet eloszlást fél lehet használni illeszkedés-, homogenitás- és függetlenségvizsgálatnál.

2.6. A t-eloszlás

Egy ismert várhatóértékű ($\mu$) és szórású ($\sigma$) normális eloszlásból vett véletlenszerű minta empirikus várható értékével számított $u$ paraméter

$$ u = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} $$

standard normális eloszlást követ. Mi történik akkor, ha a szórást nem ismerjük és a mintából becsüljük meg a korrigált empirikus szórás ($s$) segítségével. Az így számított statisztika milyen eloszlást követ? Ezt a problémát oldotta meg W. S. Gossett statisztikus és „Student“ álnéven közölte az eredményeket 1908-ban. Az alábbi összefüggés alapján számolja ki a t paramétert.
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szöllősi János dr.)

\[ t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \]

Ezt a valószínűségi változót Student t-eloszlásnak hívjuk. Gossett kimutatta, hogy a \( t \)-eloszlás hasonlíthat a standard normális eloszláshoz, de egy kissé szélesebb eloszlást mutat, azaz kevésbé „csúcsos”, és az eloszlás alakja függ a minta méretétől, egészen pontosan \((n-1)\)-től, a minta szabadsági fokától. A \( t \)-eloszlás szimmetrikus és a szabadsági fok növelésével egyre inkább megközelíti a standard normális eloszlást.

A \( t \)-eloszlás sűrűségfüggvénye \( f = (n-1) \) szabadságfok esetén:

\[ f(t) = \frac{K}{\left(1 + \frac{t^2}{n-1}\right)^{\frac{f+1}{2}}} \]

ahol \( K \) a minta elemszámától \((n)\) függő konstans. A \( t \)-eloszlás várható értéke csak esetén létezik:

\( M(t) = \mu = 0 \)

A \( t \)-eloszlás varianciája esetén:

\[ D^2(t) = \sigma^2 = \frac{f}{f-2} \]

Az
Eloszlások. A biológiában és az orvostudományban gyakran előforduló nevezetes eloszlások
(Szöllősi János dr.)

4.17. ábra mutatja a t-eloszlást különböző szabadsági fokok esetén, valamint összehasonlításképpen a t-eloszlás végterületen vett határértékét a standard normális eloszlást. A statisztikai próbák esetén a t-eloszlás alkalmazásakor a különböző szabadsági fokokhoz, különböző valószínűség értékekhez tartalmazó kritikus értékeket táblázatokban tüntetik fel. Lásd például ebben a könyvben a megfelelő táblázatot.

2.7. Az F-eloszlás

Gyakran előfordul, hogy két varianciát kívánunk összehasonlítani, és ennek egyik lehetséges módja, hogy a hányadosukat képezzük: \( \frac{s_1^2}{s_2^2} \). Természetesen ha a két variancia megegyezik, hányadosuk 1 lesz. Rendszerint nem tudjuk az összehasonlítani kívánt populáció varianciáját, mert nem ismerjük azokat, ezért a varianciákat mintavétel alapján becsülik meg. Ilyenkor a variancia-hányados becsélése a mintaváltozókat használunk (\( s_1^2, s_2^2 \)). Ha az \( s_1^2 \) és \( s_2^2 \) korrugált empirikus szórásnögyzetek \( n_1 \) és \( n_2 \) elemű, normális eloszlású populációból származó mintákból számoltuk ki, a \( \frac{s_1^2}{s_2^2} \) hányados egy szabadsági fokoktól függő eloszlást, \( F \) eloszlást követ. (Az \( F \) eloszlást Sir R. A. Fisher emlékére nevezték el.) A hányados képzésekor rendszeresen a nagyobb empirikus szórásnögyzetet jelöljük \( s_1^2 \)-el és azt osztjuk el az alacsonyabb értékű \( s_2^2 \)-vel.
Eloszlások. A biológiában és az
orvostudományban gyakran
előforduló nevezetes eloszlások
(Szőllősi János dr.)

\[
F = \frac{S_1^2}{S_2^2}
\]

Valójában az \(F\)-eloszlás nem is egy, hanem két szabadsági foktól függ, \(f_1 = (n_1-1)\)-től ami a számláló, illetve \(f_2 = (n_2-1)\)-től ami a nevező szabadsági foka. Az \(F\)-eloszlás sűrűségfüggvénye:

\[
f_{f_1, f_2} = \frac{K \times F^{\frac{f_1-1}{2}}}{(f_1 \times F + f_2)^{\frac{f_1+f_2}{2}}}
\]

ahol \(K\) az \(f_1\) és \(f_2\) szabadsági fokoktól függő konstans. Az \(F\)-eloszlás várható értéke esetén:

\[
\bar{M}(F) = \mu = \frac{n_2}{n_2 - 2}
\]

Az \(F\)-eloszlás varianciája esetén:

\[
\sigma^2 = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}
\]

Néhány \(F\)-eloszlási göbét mutat be a
Eloszlások. A biológiában és az orvostudományban gyakran előfordul nevezetes eloszlások (Szöllősi János dr.)

4.18. ábra. A görbék a szabadsági fokoktól függően aszimmetrikusak a magasabb értékek felé elnyúltak. Ha mind a két szabadsági fok nő, az eloszlás görbék egyre szimmetrikusabb lesznek.

Amennyiben az F-eloszlást a statisztikai próbáknál a populációk varianciájának összehasonlítására kívánjuk felhasználni, a kiindulási nullhipotézisünk rendszeresen az, hogy a két variancia megegyezik, azaz $\sigma_1^2 = \sigma_2^2$. Ilyenkor az F-eloszlás képlete tovább egyszerűsödik:

$$F = \frac{s_1^2}{s_2^2}$$

A statisztikai döntések meghozatalakor F táblázatokat használnunk, amelyek megadják az egyes valószínűségi szintekhez tartozó F értékeket különböző szabadsági fokok esetén. Ilyenkor, mint említettük, figyelembe kell venni mind a nevező és mind a számláló szabadsági fokat. Az F próba gyakorlati alkalmazására majd a későbbiek során térünk ki.
5. fejezet - Becslések, referencia értékek, megbízhatósági tartományok

1. A statisztikai következtetésről

Problémafelvetés

Emlékezzük arra, hogy a biostatisztika egyik fő kérdése az, hogy az összegyűjtött adatainkból hogyan következtethetünk a megismernő ismeretlenre.

A bevezetőben bemutattuk, hogy milyen hasonlóság van a biostatisztika alapfeladata(i) és az orvos munkája között:

Néhány („kevés”) megfigyelés/tünet/mért adat alapján kell(ene) a páciens teljes egészségi állapotát megítélnie: felfedni mindazon elváltozásokat és kóros folyamatokat, melyek az egészségromlás okai/jelzői/következményei („diagnózis”). Majd kiválasztani egy/több olyan beavatkozást, mely ezen állapoton a kívánt (lehetséges legnagyobb) változást/javulást előidézni képes („terápiás cél és eszközök”). Ezután megkezdi a beavatkozást, miközben folyamatosan megfigyeli, hogy a kívánt irányban és mértékben halad-e a folyamat (újabb „adatgyűjtés, elemzés, diagnózis”, a terápia hatásának értékelése) – és, ha kell, módosít. Végül értékelni az eredményt (újabb adat-, tünet-gyűjtés, diagnózis”), amivel zárol illetve újra indul a ciklus. Mindaddig, amíg a „terápiás cél” el nem tűnik: csak itt lehet kilépni a folyamatból.

Tehát: kevés (a vizsgált élő szervezet állapotának teljes leírásához képest borzasztóan kevés) tényből kell következtetni a nemcsak nem eléggé ismert, hanem egy gyakorlatilag is megismerhetetlen dologra/állapotra.

A biostatisztika fogalomrendszerében ez: a minta (kevés adat-tény) alapján a sokaságra (ami ismeretlen, sőt gyakran megismerhetetlen – idea) kell következtetni, döntést hozni. Vagyis, azok a módszerek, amelyeket a biostatisztika az ilyen típusú kérdésekben döntéshozásra kifejleszt, közvetlenül alkalmazhatóak az orvosi gyakorlatban, a mindennapok döntéshozásaiában.

Nézzünk példaként néhány tipikus kérdést, mit is kívánunk megismerni:

Egy adott betegségben egy bizonyos laboradat eltér-e az egészségesekétől?

Példaü: vizeletből megmérünk 25 beteg (cukorbetegek esetleg adott fajtájú vesebetegek) kreatinin adatát – ez 25 számérték – ez tehát tény. És tudjuk valahonnan (tankönyvi adat vagy korábbi sok-sok labormérés alapján), hogy egészségeseknél milyen kreatinin értékek fordulnak elő: Mi az a középérték (= várható érték), ami körül az egészségesek adatai szóródnak. (Ez utóbbi természetesen teljes mértékben megismerhetetlen, vagyis egy idea.)

A kérdés: A betegcsoportunk kreatinin adatai ezen „egészséges közép” körül szóródnak-e, úgy, hogy az ettől való „kis” eltérés betudható a véletlen hatásának (pl. biológiai variabilitás), avagy az eltérés „jelentős”, tehát nem valószínű, hogy „ekkora nagy” eltérést csak a véletlen okozza. Az utóbbi esetben nyilván azt mondjuk, hogy a betegek kreatinin értékei érdemben eltérnek az egészségesekétől. Vagyis – talán – az eltérő kreatinin használható az adott betegségre jellemző, a diagnózist megalapozó/támogató tünetként. Igen vagy nem? Sok múlhat a döntésen!

Egy bemérő automata, pipetta, tablettázó gépsor töltőgép stb. elég pontosan a beállított adagot adagolja-e?

Egy méréssorozat értékei, pl. vérnyomásmérés adatai, vércukor-adatai, légsszennyezés-adatai, vízsszennyezés-adatai stb. meghaladnak-e egy bizonyos határértéket?

Mindegyik kérdésnél az a feladat, hogy hogyan következtethetünk a minta (ismert) adataiból az ismeretlen sokaságra.

Mindezen kérdések annyiban közösek, hogy egy ismert mintából, adatsorból számítható átlag és szórás mennyire van távol egy sokaságra feltételezett, vagy valódi várható értékétől illetve szórástól, tehát egy ismert minta jellemező alapján kellene valamilyen ítéletet hoznunk a sokaság várható értékére, szórására. Általánosabban fogalmazva, a mintánk milyen paraméterű sokaságból származik.

A fenti kérdéseknek konkrétán: A minta átlagából (és szórásából) hogyan lehet következtetni a sokaság várható értékére (és szórására). Erre a feladatra két különböző módszerrel is tudunk válaszolni. Az egyik a becslés,
2. A statisztikai becslés

2.1. A fejezet tanulmányozásához szükséges előismeretek

- Változók típusai, a folytonos változók
- A minta és az alapsokaság
- Hisztogram, relatív gyakoriság sűrűség, sűrűségfüggvény
  - A valószínűség mérőszáma: görbe alatti terület $p(a,b)$ (Ellenőrző kérdések)
- A minta jellemzése, a sokaság paraméterei
  - Átlag és szórás, a várható érték és a szórás (Ellenőrző kérdések)
- A normális eloszlás és jellemzői (egyebek mellett: 68-95-99,7%)
  - Mintapélda: hallgatók testmagassága
  - A minta mérete és várható terjedelme. (a szórás „működése”, a „3 szabály”)
  - A std normális eloszlás és értelmezése (Ellenőrző kérdések)
- A normális eloszlás kiemelt jelentősége az orvostanhallgatók számára:
  - A binomiális eloszlás és határeseti közelítése normálissal (Ellenőrző kérdések)
  - Az átlag eloszlása
  - Mintapélda: 16 fős csoportok átlagmagasságának eloszlása (Ellenőrző kérdések)

2.2. A pontbecslés

A statisztikai becslés egyszerű esete a pontbecslés. Az ismeretlen egy hozzá hasonlóval közelíthető (becsüljük). A „problémafelvetés” indító kérdéseinél a minta átlagát úgy tekinthetjük, mint a várható érték egy becslését, illetve a minta szórását úgy, mint a sokaság szórásának egy becslését.

Tulajdonképpen ugyanazt alkalmaztuk akkor, amikor az 50 hallgató testmagasság-adatából következtettünk (a normális eloszlásnál és az átlag eloszlásánál is alkalmazott példára utalok vissza): feltételeztük, hogy az orvostanhallgatók sokaságának várható értéke 170 cm pusztán azért, mert a testmagasság adatok átlagára ezt az értéket kaptuk. A – nek az egy pontbecslése.

További alkalmazás e fejezetben, amikor az ismeretlen -t a mintából vett szórással, vagyis az sx-szel helyettesítjük: A -nak az sx egy pontbecslése.

2.3. Az intervallumbecslés

 Második ilyen lehetőség az intervallumbecslés.

A probléma az előbbivel, a pontbecsléssel az volt, hogy nem igazán illeszkedik a célul kitűzött valószínűségi gondolkodásmóddhoz: Nem rendelhetünk „valószínűséget”, „megbizhatóságot” a pontbecsléshez, pedig tudjuk, hogy ez csak egy többé-kevésbé bizonytalan közelítés. Az ok: folytonos eloszlásnál minden „éles” értéke tartozó görbe alatti terület (= az adott érték bekövetkezési valószínűsége) nulla, ettől különböző valószínűség csak intervallumhoz tartozhat. (lásd: folytonos eloszlások, valószínűség sűrűség függvény)

Ezért felé meg az alább ismertetett (intervallummal becslő) gondolatmenet a valószínűségre alapozott döntéshozásnak (szemben a pontbecsléssel).
Az ismeretlen várható érték becslésénél az átlag és a szórás alapján két számot választunk – legyen ez a két szám „a” és „b” –, és az mondjuk, hogy az ismeretlen várható érték bizonyos valószínűséggel az (a,b) intervallumban – tehát „a” és „b” között – található. Ez a „bizonyos valószínűség” a becslés megbízhatóságára, ami nyilván kisebb, mint 100%, de azért lehetőleg nem sokkal kisebb: Többnyire 99%, 95% vagy 90%-ot szoktunk választani.

Vagyis, arra a kérdésre, hogy „Mennyi (lehet) a értéke?”, ahelyett, hogy „a várható érték = (kb.)” (ez lenne a pontbecslés), azt mondjuk: „a várható érték valahol az körül, egy (a,b) intervallumban – tehát „a” és „b” között – található” „p” valószínűség (ez pedig az intervallumbecslés). Ez az (a,b) a várható érték p%-os megbízhatóságával (idegen szóval: „konfidencia”) intervalluma.

Emlékezzünk, már korábban is használtunk ilyen nevezetes intervallumokat, melyekhez különböző valószínűség, megbízhatóság tartozott:Pl. az 5.1. táblázat - = 170 cm = 8 paraméterű normális eloszlás esetén adott intervallumra esés valószínűsége, és adott mintaelenszámánál az intervallumon belül és kívül eső elemek várható aránya. 5.1. táblázatban a testmagasság adatokra: Ha = 170 cm és = 8 cm (mint azt korábban feltételeztük – ez tehát pontbecslés!), akkor a ±2 intervallumban, tehát 154-186 cm-ig található az adataink kb. 95%-a. Ez úgy értelmezendő, hogy 1 adat esetén 95% a megbízhatóság, a valószínűség annak, hogy az 1 adat ebben az adott tartományban található. Sok adat esetén pedig az adatok 95%-a – pl 100 adatból 95 adat – található az adott intervallumban. Hasonlóan a ±3 tartomány az adatok 99,7%-át tartalmazza, stb.

<table>
<thead>
<tr>
<th>Interv</th>
<th>Arány (belül/kívül)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>162-178</td>
</tr>
<tr>
<td></td>
<td>154-186</td>
</tr>
<tr>
<td></td>
<td>146-194</td>
</tr>
<tr>
<td></td>
<td>Vsz.</td>
</tr>
</tbody>
</table>

5.2. táblázat - = 170 cm = 8 paraméterű normális eloszlás esetén adott intervallumra esés valószínűsége, és adott csoportszámnál az intervallumon belül és kívül eső elemek várható aránya.

<table>
<thead>
<tr>
<th>Arány (belül/kívül)</th>
</tr>
</thead>
<tbody>
<tr>
<td>164-176</td>
</tr>
<tr>
<td>168-172</td>
</tr>
</tbody>
</table>

Vagy, a 16 elemű minták átlagainak eloszlásánál (5.2. táblázat - = 170 cm = 8 paraméterű normális eloszlás esetén 16 elemű minták átlagainak adott intervallumra esésének valószínűsége, és adott csoportszámnál az intervallumon belül és kívül eső elemek várható aránya. 5.2. táblázat.) – amire = 170 cm és /n = 2 cm volt –, a ±2/n intervallum (a 166-174 cm közti tartomány), a csoportátlagok 95%-át tartalmazta (a sárga cella), és így tovább.

Ellenőrző kérdés: Az 5.2. táblázat szerint (a sárga cella alatt) a ±3/n tartományon belül – 164-176 cm között – viszont már a 100 átlagból az összes belül található. Miért van ez így?

Váltás: Tudjuk, hogy a ±3/n tartományon belül 99,7%-a van az átlagoknak. Miért mondhatjuk, hogy ez gyakorlatilag az összes átlagot jelenti? Hogy értjük ezt a 99,7%-ot? Úgy, hogy 1000-ból 3 (100-ból 0,3) van kívül! Tehát átlagosan kb. minden 3 db ilyen 100-as csoportból – vagyis 3szor 100 ilyen, 16 elemű minta-átlagból – 1 olyan 100-as csoport lesz, amelyből 1 adat (átlag) kívül esik ezen a ±3/n tartományon. Ezért mondhatjuk, hogy, ha csak 1 ilyen 100-as csoportunk van, akkor abban várhatóan minden átlag belül lesz ezen a ±3/n tartományon. (Várhatóan, de nem feltétlenül)

**2.3.1. Intervallumbecslés -ra és alapján**
A statisztikai intervallum-becslés gondolatmenetének megértéséhez induljunk ki a csoportátlagok már ismert eloszlásából. A testmagasság adatok példáján (5.2. táblázat - = 170 cm = 8 paraméterű normális eloszlás esetén 16 elemű minták átlagainak adott intervallumban esésének valószínűsége, és adott csoportszámnál az intervallumon belül és kívül eső elemek várható aránya. 5.2. táblázat) azt látjuk, hogy a körül ±1/n távolságban, tehát 168-172 cm-ig található az átlagok 68%-a. ±2/n távolságon belül, tehát 166-174 cm közé pedig a 95%-a. Ez azt jelenti, hogy 20 csoportátlag közül (sárga cella a táblázatban) várhatóan 19 ezen az intervallumon belül van, s csak 1 van ezen kívül. Az

![Diagram](image)

5.1. ábra mutat egy ilyen lehetséges helyzetet, 20 különböző csoportátlagot külön (fekete) pontonként ábrázolva: 19 belül van a tartományon, 1 pedig kívül.

100 csoportátlag (kék pontok) közül viszont várhatóan 95 lesz belül a tartományon, s csak 5 kívül, ahogy itt az ábra alsó sorában a 100 pont mutatja.

2.3.2. **Intervallumbecslés -re az és alapján**

Vegyük ezt a 100 különálló mintaátlagot, amit ez a 100 pont jelképez, (amit tehát úgy kaptunk, hogy 100, egyenként 20 fős csoport testmagasságait megmértük, és kiszámítottuk az átlagaikat), és rajzoljunk mindegyik köré egy ±2/n hosszúságú intervallumot. Tudjuk, hogy a körül egy ekkora intervallumon belül van 100 átlag közül kb. 95 (5.2. táblázat - = 170 cm = 8 paraméterű normális eloszlás esetén 16 elemű minták átlagainak adott intervallumba esésének valószínűsége, és adott csoportszámnál az intervallumon belül és kívül eső elemek várható aránya. 5.2. táblázat). A kérdés az, hogy ezen 100 különböző helyzetű, de egyforma hosszúságú intervallum közül hány tartalmazza a -t.

Ez jobban megérthető az 5.4. ábráról, amely ezeket az intervallumokat mutatja. Viszonyítási tartománként látszik a = 170 cm középpontú ±2/n = 8 cm széles tartomány (szaggatott vonallal). Látható, hogy:

Azon intervallumok, melyek közepje belül van ezen a ±2/n tartományon, azok tartalmazzák a -t. Hiszen az intervallumok középe és a távolsága kisebb, mint a 2/n. Azt is tudjuk, hogy azon intervallumok belüli – , vagyis 16-os csoportátlag, 95 volt a 100-ból.

A viszonyítási intervallumon kívüli 5 csoportátlag köré húzott ±2/n hosszúságú intervallumok pedig nem tartalmazzák a -t. Hiszen a -től ezen csoportátlagok távolabbik, mint a 2/n.

Vagyis a 100 közül van 95 olyan átlagunk, ami köré húzott ±2/n intervallum tartalmazza, és van 5 olyan átlagunk, ami köré húzott ±2/n intervallum nem tartalmazza a -t.

Vegyük észre viszont, hogy a való életben csak egy átlagunk van, s nem tudjuk, hogy az iménti 100 közül valójában melyik az: a 95 közül, vagy az 5 közül egy? Ezt értjük azon, hogy: Ha 100 átlagunk lenne ugyanabbnál az eloszlásból, akkor azok köré húzott 95 intervallum tartalmazná a -t, 5 viszont nem. Ezt az 95-öt tekinjük az – ±2/n – becslés megbízhatóságának, az 5-öt pedig a hiba kockázat-ának.

Ez kulcsfontosságú: Megérteti a megbízhatóság fogalmát: „Ha 100 átlagunk lenne…”

Egy opcionális megjegyzés: Az imént ismertetett becslés, amikor -ra és alapján készíttetünk intervallumot, csak „elméleti jelentőségű”: Mivel a számunkra fontos helyzetekben sem a sem és nem ismert (sőt meg sem ismerhető!), így ezzel a módszerrel csak ismeretlen jellemzők alapján lehetne becslőlni ténylegesen ismert (megmérhető) értékeket. Ennek tehát „gyakorlati jelentősége” nincsen, csak az a szerepe, hogy az intervallumbecslés „konstrukcióját” bemutassa. A statisztikai becslés, mint döntéshozó

---

Becslések, referencia értékek, megbízhatósági tartományok

59

Created by XMLmind XSL-FO Converter.
eljárás célja épp fordított: ismert adatok/értékek alapján kell becsülni a nem megmérhető ismeretlent. Alább bemutatjuk, hogyan lehet a megfordítást, a gondolatmenet „talpára állítását” megtenni.

2.3.3. Az ismeretlen helyett használjuk az ismert sx-et! Hogy változik a megbízhatóság?

Most jön a fő probléma: Nem tudjuk létrehozni a ±2/n hosszú intervallumokat, mert nem ismerjük a -t! És itt jön a második kulcsgondolat ebben a gondolatmenetben: Helyettesítsük a -t a mintából vett szórással, vagyis az sx-szel. Erre az ad alapot, hogy a -nak az sx egy pontbecslése. Van azonban ennek az ötletnek egy „apró” problémája: Az sx lehet kisebb is és nagyobb is, mint a ! Így az sx/n is lehet kisebb is és nagyobb is – mint a /n volt – egy konkrét minta és a belőle számolt sx esetén. Hogy befolyásolja ez az imént megértett megbízhatóságot?

Az 5.3. ábrán: Nézzük először azt a 95 intervallumot, amely ±2/n hosszú intervallumként korábban még tartalmazta a -t. Ezek közül azok, amelyek hosszabbak lesznnek (amelyénél az sx nagyobb lett, mint volt: pl. az ábrán az első intervallum), azok továbbra is tartalmazni fogják. Azonban lesznek, amelyek ebből a 95-ből (a -nál rövidebb sx miatt) rövidebb lesznek, s ezek egy része akár olymértekben is rövidülhet, hogy ezután már nem fogja tartalmazni. Mint itt az ábrán láthatjuk a második intervallumnál. Vajon ily módon hány eshet ki a 95-ből?

Egyéb esetek számoljuk ki azt, hogy hány lehet rövidebb! Feltételezhetjük, hogy az sx egyenlő (50%-50%) valószínűséggel lesz vagy kisebb, vagy nagyobb, mint a .

Ellenőrző kérdés: Hogyan becsülhető ez alapján, hogy a 95-ből hány intervallum rövidülhet?

Válasz: A rövidülő intervallumok lehetséges száma tehát becsülhető egy olyan binomiális eloszlás alapján, ahol az n = 95, a p = 0,5. Vagyis – a számolást elhagyom – 95% megbízhatósággal 40-60 azon intervallumok száma, melyek rövidebbek lesznek. Tehát ezek közül fordulhat elő néhánynál, hogy a továbbiakban már nem tartalmazza a -t.

Emlékezzünk, ez a becslés tehát az -ra és alapján tett becslésnél említett „elméleti jelentőségű” eljárás!

Nézzük, mi lehet az 5, eredetileg kívül levő közepű intervallummal? Néhány meghosszabbodhat, s ezek közül lehet olyan, amely ezután már tartalmazni fogja a -t, holott korábban nem tartalmazta. (Mint az ábra harmadik intervalluma.) Vajon az 5-ből hány lehet hosszabb? Ügye, legfeljebb 5!

Összegezve: 95-ből, ami eredetileg tartalmazta a -t legalább 40-nél van rá esély, hogy a rövidülés miatt ezután nem tartalmazza, míg az 5-ből, ami eredetileg nem tartalmazta, legfeljebb 5 tartalmazhatja ezután. A kérdés az, hogy: vajon mi a 95-ből kiesés és az 5-ből hozzájövő egyenlege?

Nyilván ez attól függ, mennyire tér el az Sx/n (a standard hiba), a /n-től. Ez az eltérés pedig annál nagyobb lehet, minél kisebb az adatszám amiből kiszámoljuk Sx-et, hiszen annál bizonytalanabb a pontbecslés.

5.3. táblázat -

<table>
<thead>
<tr>
<th></th>
<th>N = 16</th>
<th>adat</th>
<th>N = 6</th>
<th>adat</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>1,39</td>
<td>2,58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95%</td>
<td>1,29</td>
<td>2,71</td>
<td>2,82</td>
<td>3,20</td>
</tr>
<tr>
<td>99%</td>
<td>1,11</td>
<td>2,96</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Vegyük erre egy példát: Ha a \( /n = 2 \) cm lenne, akkor 16 adatnál az Sx/n lehetne a 2 helyett akár 1,29 vagy 2,71, tehát, mintegy 35%-kal eltérő a \(/n = 2\) cm-től bármelyik irányban, lefelé és fölfelé is. Viszont, ha csak 6 adatból számoljuk ki az átlagot és a szórást, akkor a növelt bizonytalanság – szemben az előbbi 16 adatból kiszámoltattal –, látszik abból, hogy ekkor az Sx értéke lehetne akár csak 0,8 vagy 3,2 (tehátmintegy 60%-kal lefelé is és fölfelé is eltérő) az eredeti 2 cm helyett. Ez látszik a mellékelten 5.3. táblázatból, ahol továbbá a biztonság vagy megbízhatóság hatása is jól követethető: Hogyan változnak ezen határok, ha a 95% helyett 90 vagy 90% megbízhatósággal kívánjuk ezeket az eltéréseket megadni.

Összegezve: annál kisebb a megbízhatóság, minél kisebb a mintákat mérete, vagyis az n, annál nagyobb mértékű rövidülés lehet és így annál több intervallum eshet ki a „tartalmazza a -t” csoportból: Az eredeti 100-ból 95 helyett mondjuk csak 92, 90, 85, stb. A kérdés tehát:

**5.4. táblázat -**

<table>
<thead>
<tr>
<th>p% megb. Interv.: ± t*sx/n. – a t értékei:</th>
<th>2</th>
<th>4,30</th>
<th>9,92</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2,57</td>
<td>4,03</td>
<td>5,51</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2,31</td>
<td>3,36</td>
<td>4,28</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>2,23</td>
<td>3,17</td>
<td>3,96</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>2,13</td>
<td>2,95</td>
<td>3,59</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2,09</td>
<td>2,85</td>
<td>3,42</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>2,01</td>
<td>2,68</td>
<td>3,16</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>1,96</td>
<td>2,58</td>
<td>3,00</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>p% megb. Interv.: ± z*/n. . – a z értékei</th>
</tr>
</thead>
<tbody>
<tr>
<td>p = 95%</td>
</tr>
<tr>
<td>1,96</td>
</tr>
<tr>
<td>2,58</td>
</tr>
<tr>
<td>3,00</td>
</tr>
<tr>
<td>p = 99%</td>
</tr>
<tr>
<td>p = 99,7%</td>
</tr>
<tr>
<td>p = 99%</td>
</tr>
<tr>
<td>2,58</td>
</tr>
<tr>
<td>3,00</td>
</tr>
<tr>
<td>3,00</td>
</tr>
</tbody>
</table>

Az 5.4. táblázat - 5.4. táblázat alján látjuk, hogy egy p% megbízhatóságu intervallum (a normális eloszlás alapján) n-től függetlenül az \( \pm z*/n \) adva. Az \( z \) konstans pár jellegzetes értéke (ezeket mutatja a táblázat): 95%-höz 1,96, 99,7%-höz 3.

Emlékezzünk, továbbá, hogy a \( z = 1 \)-szeres távolsághoz 68% megbízhatóság tartozott.

A megbízhatóság szokásos mértéke még a 99%, és ehhez tartozna egy közébső 2,58-as szorzófaktor, ami ugyancsak szerepel a táblázatban.
Vegyük azt is észre, hogy ezen p értékekhez tartozik egy-egy hibakockázati érték is (ezt itt most %-val jelölünk): a 95%-hoz 5%, a 99-hez 1%, a 99,7-hez 0,3%.

### 2.3.3.1. A t eloszlás: az intervallumok n-től függő megnyújtása

Ugyancsak vegyük észre, hogy a táblázatban közölt 3 szám (tehát a megfelelő megbízhatósághoz tartozó távolság adatok), valójában a normál görbét határozzák meg:

Hiszen abból származtak. Tehát egy-egy ilyen szánhármas meghatároz egy-egy megfelelő lefutású valószínűség-sűrűség függvényt: megadja, hogy mely intervallumon belül mekkora a valószínűség (görbe alatti terület).

Ezen pár fontos észrevétel után visszatérve a táblázat első részét vegyük szemügyre: hogyan is növeljük a szorzófaktort?

Maradjunk csak a 95% megbízhatóságnál: Így a kérdés az, hogyan is növeljük n-től függő mértékben az 1,96-os faktort? Az 5.7. táblázat felső része mutatja a megoldást: Ha nagyon nagy az n értéke (mondjuk 1000), akkor ez a növelés nagyon kicsi (észrehetetlen), hiszen nagyon nagy n esetén az sx valószínűleg nagyon közel lesz a -höz.

Tehát az Sx-cserenagy n esetén nem okoz érdemi intervallumhossz (s ezáltal megbízhatóság-) változást.

Azonban, ha csökken az n értéke (pl. ahogy a táblázatból látszik, n értéke 50 vagy az alatti lesz), akkor az eltérés az 1,96-hoz képest már jelentős: 2,0 – 2,1 – 2,2 – 2,3… (sőt, n = 5 körül felmehet 2,5-re), ahogy a táblázat mutatja.

Természetesen ugyanilyen növekedésre van szükség, ha 99%-os megbízhatóságot kérünk: akkor a 2,58 növekszik akár 4-ig vagy afölé, illetve a 99,7%-nál a 3 növekszik 4 … 5 … sőt eftélé.

Vagyis, ahogy a táblázatból látjuk, minden n-hez tartozik egy-egy újabb számsor, szemben a normális eloszlásnál talált n-től független z értékekkel. Ezeket a számokat itt a táblázatban t értékeken hívjuk. Minden egyes n-hez tartozó ilyen t értékek – ugyanúgy, ahogy az imént az adott z értékek a normális eloszlást –, egy-egy „t-eloszlás”-t határoznak meg: Minden n-hez egy-egy újabb „t-görbe” tartozik, amelyet az ún. szabadságfoka (n-1) határoz meg.

Így minden n-hez a megbízhatósági intervallum az ±t*sx/n módon lesz számolható. Magyarul: az intervallumot – a kívánt mértékben – egy növelt („t”) érték felhasználásával nyújtottuk meg. Konkrétan 16 adatnál – a táblázatból látjuk, ez megfelel n-1 = 15-ös szabadsági foknak –, tehát 16 adatnál az intervallum ±2,13*sx/n lesz, amely 95% megbízhatósággal tartalmazza a -t a korábban vett értélemben. (Tehát ezen azt értjük, hogy: „ha 100 átlagunk lenne, … stb.”)

### 2.3.3.2. A t-eloszlás jellegzetességei

Most itt közbevetőleg két további ábrán a t-eloszlással fogunk foglalkozni, s utána térünk vissza újra a konfidencia intervallumhoz. A kérdés az, hogy milyenek is ezek, a normális-tól némileg különböző t eloszlások? Hasonlítuk össze az

5.4. ábra alapján a std normálissal.
Emlékezzünk, hogy a folytonos eloszlásokat a valószínűség-sűrűség függvényvel ábrázoltuk: Például a normális eloszlás esetén ez egy „harang alakú” görbe (Gauss görbe). A standard normális eloszlás esetén a vízszintes tengelyen mérjük a várható értéktől való eltérést std hiba egységekben. Egy adott (a,b) intervallumhoz tartozó görbe alatti terület pedig azt mutatja, hogy milyen valószínűséggel lehet ekkora (a és b közti mértékű) eltérés. A görbe alatti teljes terület 1 (100%), így bármely véges intervallumhoz 1-nél kisebb valószínűség tartozik.

Tudjuk, hogy t-eloszlás esetén a 95%-nyi görbe alatti terület a (normálisra jellemző ±2-szoros, pontosan 1,96 szoros helyett) a 2,1-, 2,2-, 2,3-, ... stb. szoros std hibányi – vagyis az n-től függő mértékben egyre szélesebb – intervallumhoz tartozik. Tehát minden n-hez egy újabb t görbe tartozik mégpedig úgy, hogy középen kicsit alacsonyabbnak kell lenni (így lesz adott intervallumhoz tartozó terület/megbízhatóság kisebb). A széleken viszont kicsit magasabbnak, mint a std normális sűrűségfüggvénye, hiszen a teljes görbe alatti terület mindigek görbénél 1 kell legyen.

Az ábrán összevetünk egy ilyen t görbét – ez inkább csak ez eltérés jellegét mutatja – a std normális eloszlás sűrűségfüggvényével. A normális görbén (színezve az ábra jobb felén) a 95% megbízhatósághoz tartozó területfél (narancsszínre satírozva), középtől az 1,96 std hibányi távolságig tart. Az ábra bal oldalán a T görbénél – ugye, ez a görbe középen kicsit alacsonyabb, ennek megfelelően a széleken kicsit magasabb –, ugyanakkora távolságnál levágya (a sárga terület) mindig kisebb, mint a normál görbe alatti (a narancs) volt – a kék terület pedig nagyobb, mint a zöld. Hogy ugyanakkora (95%) legyen ez a középső (sárga) terület, ahhoz kijebb, a középtől távolabb kellene levágunk. (Tehát amikor a levágott terület a két oldalon 5, egy oldalon 2,5%). Konkrétan pl. a T5-ös görbe esetén 2,57 std hibánál (lásd 5.7. táblázat adatait).

Tehát ez a fontos különbség a t görbék és a normális eloszlás sűrűségfüggvénye között a „farokrészen”, a görbék elnyúló, szélső szakaszán érzékelhető. Ezért nézzük meg néhány különböző t görbére, valamint a normális sűrűségfüggvényre összehasonlításként ezeket a külső széleket.

Az 5.5. ábrán jól látszik, hogy a normális (itt sárgával satírozott) görbe 2,5%-nyi területe (ami itt most a jobb oldalon a 95%-on kívül van), 1,96-tól indul ki felé.

A normális fölött haladó T10-es görbénél, ha ugyanitt vágánk el, akkor a levágott terület jóval nagyobb lenne, mint a 2,5%, hiszen a T végig fölött halad a normálisnak innen kifelé, ezen a külső szakaszon. Ezért, hogy a levágott terület ugyanakkora legyen, mint a normálisnál, hát kijebb, így konkrétan a 2,23-nál kell levagnunk. (Ezt jelenti az, hogy a T10-es görbénél a 95% valószínűséghoz egy nagyobb távolság (nagyobb t érték), éspedig a 2,23 hibával.)

Ugyanez a T5-ös görbén, amely még följebb fut: Ennek megfelelően még kijebb kell levágni (éspedig 2,57-es távolságnál), hogy a levágott terület 95% legyen középen, és 2,5-2,5% a két végén.

Vagyis ezen „kritikus számok” (melyeket az 5.4. táblázat - 5.4. táblázatban láttunk) jelentése: Hol kell elvágni a görbét, hogy a középső terület 95%, a levágott két kis farokterület együtt 5% legyen.

2.3.4. Az ±t*sx/n intervallumbecslés a-re
Alkalmazzuk ezt most – visszatérve – a 100 mintaátlag köré húzott intervallumokra. Mit is alkalmazunk? Azt, hogy az n-től függő mértékben megöveljük a 2-es szorzófaktort. Ha 16 adatból számítottuk a mintaátlagokat, akkor konkrétan 2,13-ra növeljük a szorzófaktort.

Magyarul, (a 100 különböző minta 100 átlaga és 100 szórása alapján) a különböző középpontú (különböző átlagú) és különböző hosszúságú (eltérő szórású) minták intervallumait úgy számoljuk ki, hogy a std hibát 2,13-al szorozzuk, szemben a normál eloszlásnál korábban használt 2-es (1,96-os) szorzófaktorral.

Nézzük meg, mi lesz ennek a hatása az eredetileg 95 belül levő átlag esetén. (Az Sx- csere miatt egy részük rövidebb, egy részük hosszabb lett, mint az eredeti, s most ezt mind egységesen – „t” mértékében – megnyújtjük!)

Az Sx- csere miatt változott hosszúságú intervallumok egy része változatalanul tartalmazta -t. Ezek a t-vel történt megnyújtás után – ezt mutatja az

5.6. ábrán az az 1., 2., és 3. bejelölt mintát jelző intervallum – változatalanul tartalmazzák a -t. Viszont voltak olyan, eredetileg belül levő átlagú intervallumok, amely az Sx- csere („jelentős” rövidülés) miatt a továbbiakban már nem tartalmazzák: Ezek egy része a t-vel történt nyújtás után újra tartalmazhatja, más része azonban továbbra sem (az utóbbira példa az ábra 96-ként jelölt intervalluma).

Viszont az eredetileg kívül levő közepű 5 intervallumból (akár az Sx- csere miatti hossznövekedés, akár a t-vel történt megnyújtás miatt) lehet olyan, ami már ezután tartalmazni fogja a -t. Így kompenzálva az imént látott „eredetileg belső közepű, de az Sx- csere miatt -t nem tartalmazó” intervallumokat (lásd az ábra 95. jelű intervallumát).

E két hatás (vagyis a belül levő közepű, de túlzottan megrövidült és a kívül levő közepű, de eléggé megnyúlt intervallumok száma) épik kompenzálni fogja egymást.

A t értéke azért épp akkora, mert ennél a növekedésnél egyenlítődik ki éppen a kieső és a belépő intervallumok száma, vagyis áll helyre a megbízhatóság eredeti (Sx- csere előtti) értéke! Vagyis a 2-es faktor növelése miatt újra épp 95 intervallum fogja tartalmazni a -t. Továbbá a 95 belső közepű, de sokat rövidült, s így ?-t már nem tartalmazó intervallum(ok) plusz mindazon külső középüek, amik nem növekedtek meg eléggé az 5 közül, hogy ezután már tartalmazzák, azok együttvéve újra csak 5 intervallumot adnak a 100 közül, melyek továbbra sem tartalmazza a -t.

Tehát: Az

-\frac{t*sx}{n} módon készített 100 különböző középpontú (amelyek a minta-átlagainkat jelentik) és különböző hosszúságú (amelyek a minták szórásainkból adódnak) intervallumból – az n-től függő („t érték”) megnyújtás miatt – újra 95 fogja tartalmazni a -t. Vagyis az \frac{t*sx}{n} intervallum a -nek egy p% megbízhatóságú intervallumbecslése, ami abban az esetben, ha az n = 16 és p = 95%, akkor éppen t = 2,13-al számolandó.

Ez az a t-érték, amit a 5.5. táblázat - 5.5. táblázat tartalmazott: Mutatva, hogy függ n-től és p-től ez a t-érték. Amely mögött, ugye, az egyes t eloszlásokhoz tartozó görbék húzódtak. Minden statisztika könyvben részletes táblázatot találhatunk a n-hez tartozó (tehát különböző szabadsági fokú) t-görbék különböző megbízhatósághoz tartozó t értékeirel. Javasoljuk, hogy ennek ellenére érdemes megjegyezni legalább a leggyakrabban használt 95%-os megbízhatósághoz tartozó alábbi értékeket:

5.5. táblázat -

2.3.5. Mintafeladatok a -re vonatkozó ±t*sx/n intervallumbbecslésre

1. mintapélda

Egy 16 fős hallgatói csoport testmagasság adataiból az átlag 175 cm, a szórás 10 cm. Kér dés, lehet-e a sokaság várható értéke a 170 cm? Magyarul: tulajdoníthatjuk-e a 170 és a kapott 175 cm közti eltérést csak a véletlennek?

Ha a 16 fős csoport hallgatói külföldiek, eltér-e lényegesen a külföldi hallgatók testmagassága a magyar hallgatókra jellemző 170 cm-es várható értékétől, vagy az eltérés tulajdonítható csak a véletlennek? (Megjegyzés: Emlékezzünk, a 170 cm-es várható értéket pontbecsléssel nyertük egy nagyobb létszámú magyar hallgatói minta átlagából!)

A számolás rendkívül egyszerű: Mivel t értéke 95% megbízhatóság és n-1 = 15 szabadsági fok esetén: 2,13, a -re vonatkozó 95%-os megbízhatósági intervallum: 175±2,13*10/16 cm, vagyis 175±2,13*2,5 cm. Tehát az intervallum: 169,67 – 180,33 cm. A 170 cm ezen belül van, tehát egy lehetséges , egy lehetséges várható érték,(Tekintve, hogy a 95% megbízhatóságot használtuk, ezért ez a becslésünk 5% hibakockázattal jár.) Tehát a 175 cm nem tér el lényegesen a 170 cm-től!

2. mintapélda

Egy 9 fős beteg csoport szérum albumin adatára 3,9 mg/100 ml átlag és 0,6 mg/100 ml szórás adódott. Lehet-e vajon a betegek várható értéke az egészségesekre jellemző 4,2 mg/100 ml? Másként fogalmazva: lényegesen eltér-e a kapott 3,9 a várt 4,2-től? Még másként fogalmazva: lényegesen eltér-e a betegeknél kapott érték az egészségeseketől?

A számolás: A 95% megbízhatósági intervallum a -re: 3,9±2,3*0,6/9, vagyis 3,9±2,3*0,2. Vagyis 3,43 – 4,33 mg/100 ml. A 4,2 ezen belül van, tehát egy lehetséges várható érték – ismét 5% hibakockázattal. Tehát a 3,9 nem tér el lényegesen a 4,2-től!

3. Összefoglalás

3.1. A gondolatmenet összefoglalása

Ez a fejezet arról szólt, hogyan lehet ismeretlen mennyiségre becslésként (ismert adatokból) egy intervallumot megadni, amelyben az ismeretlen mennyiség megkívánt valószínűséggel megtalálható.

A leggyakoribb példája ennek, amikor a (mintából ismert) és sx segítségével egy (a,b) intervallumot hozunk létre, amely valamely p% (100-nál kisebb, de ahhoz közeli) megbízhatósággal tartalmazza a (-z ismeretlen) -t. Ezt hívjuk a várható értékre vonatkozó – p% – megbízhatósági intervallumnak. Beláttuk (megértettük!), hogy ennek kiszámolása: ±t*sx/n, ahol az és az sx a mintából számított jellemzők, a n a minta méretéből adódik, a t pedig az n-től és a kívánt megbízhatóságtól függő állandó (az n-hez tartozó t görbéből származó adat).

Maga a számolás egyszerű, de kérdés, akkor miért kellett ez a bonyolult gondolatmenet? Mi is volt ez a gondolatmenet?
1. A és a ismeretében tudunk egy – kivánt megbízhatóságú – intervallumot megadni a minta átlagra a normális eloszlás alapján. („az átlag eloszlása” xxx oldal – ebből indultunk ki)

A és a ismeretében meg tudjuk mondani, hogy egy adott intervallumban milyen valószínűséggel szerepel egy mintaátlag. Vagy a és a ismeretében meg tudjuk mondani, hogy 95% valószínűséggel milyen intervallumba kell esnie az adott mintaátlagnak?

2. A mintaátlag és a ismeretében tudunk adni a -re egy intervallumot – ugyancsak a normális eloszlás alapján. Ez az előbbieknak a megfordítása.

3. Azonban probléma, hogy nem ismert! Viszont ismerjük helyette a minta szórását! Beláttuk, hogy a helyett az sx-et használjuk, akkor csökken a megbízhatóság (nő a bizonytalanság).

4. Ezt kompenzálja az intervallumnak az n-töl függő mértékű megnyújtása. Itt alkalmaztuk a t értékeket. Ezek háttérében a t eloszlás van.

5. Végül tehát:

\[ \pm t \times \frac{sx}{\sqrt{n}} \]

Az orvos munkájában alapvető mozzanat, hogy ismeretlen (sőt teljességében megismerhetetlen) dologról kell valószínűleg és megbízhatólag döntést hozni. Ez az illetőséget és bizonytalanságot biztosít. A gondolatmenet kíváló példája a valószínűségek alapján történő döntéshozásnak, így teljes megértése érhetetlen. Ha bővíteni vagy továbbfejleszteni akarjuk gondolkodási eszközeinket, készüljünk a sikerrel.

Az intervallumbecslés további alkalmazásaival a tananyag későbbi fejezeteiben találkozunk majd.

3.2. A várható érték megbízhatósági intervalluma: az intervallumbecslés felidézése

Az 5. fejezetben megismertük a statisztikai becslés módszerével, mégpedig egy konkrét eljárás alapján. Ez az intervallumbecslés a várható értékre, szokásos elnevezéssel: a várható érték megbízhatósági intervalluma. Ebben az esetben az a kérdés, hogy hol található a várható érték, a ? Az a válasz rá, hogy valahol az átlagérték, a mintabeli átlagérték körül. Egy tartományt határoztunk meg. Nézzük ezt a segédábrát ehhez!

Az körül húztunk egy tartományt tôle bizonyos távolságra, és azt állítottuk, hogy ebben található egy bizonyos p valószínűséggel, megbízhatósággal a várható érték. Konkrétan, ez a távolság t*sx/n – magyarul t-szer a standard hiba – volt, ahol ez a t érték természetesen függött attól, hogy hány adatból határoztuk meg az átlagértékét és a szórást.

Minél kevesebb adatból határozzuk meg az átlagot (és az Sx-et), annál kevésbé állíthatjuk, hogy az jól közelíti a várható értéket. Ugye emlékszünk erre korábbról: „átlag szórása”.

Továbbá a t értéke függ attól, hogy milyen megbízhatóságot kívánunk. Ugye erre is emlékszünk korábbról: „hányszor szigma” a norm eloszlásnál.…”

Annál szélesebbnek kell lenni az intervallumnak – magyarul, ebben az esetben annál nagyobb a t értéknek –, minél nagyobb megbízhatósággal szeretnénk a μ-re vonatkozó kijelentéshoz megtenni, minél nagyobb valószínűséggel állítani azt, hogy benne van a μ az intervallumban.

A
Képzeljük el, hogy száz különböző mintát veszünk. Mindegyikből kiszámítjuk az átlagértéket és a szórást. Így az:

\[ \pm \frac{s\times\sqrt{n}}{n} \]

ez száz különböző intervallumot jelent! Ebből a 100-ból átlagosan p számú intervallum tartalmazni fogja a \( \mu \)-t, 1-p pedig nem. Ez a \( p \) megbízhatóság jelentése.

A leggyakoribb alkalmazása ennek a kilencven öt százalékos konfidencia intervallum, amelynél ezen száz különbözőből kilencvenöt tartalmazni fogja a \( \mu \)-t, 1-p pedig nem! Ahogy a 5.4. táblázatban látjuk, ha tizenhat adatból számítottuk a mintajellemzőket, és \( p = 95\% \), akkor ennek a t-nek az értéke 2,13 lesz. Vegyük észre, hogy a 100-ből 5 intervallum viszont nem tartalmazza a \( \mu \)-t! Ezt a megbízhatósággal szemben hibakockatnak hívjuk.

Fontos észrevétel, hogy rendszerint csak egy mintánk van, egy mérési adatsorunk, és nem tudjuk, hogy ez vajon a 95, vagy az 5 közül való: Tehát, ez az egyetlen vajon tartalmazza-e vagy nem a várható értéket? Nem tudjuk! 

Ezt a bizonytalanságot jelenti a 95%-5% értékpár! Láttuk, hogy a számítás maga nagyon egyszerű! Láttunk rá mintapéldát:

Ráadásul, ha statisztikai programot alkalmazunk, akkor ezt a számolást maga a program elvégzi. A nehéz az interpretáció, az értelmezés: Azt érteni, hogy: „ha száz adatsorunk lenne, akkor ebből…”.

### 4. Ismeretellenőrző és gyakorló kérdések és feladatok

Egy gyógyszergyári gépsor pontosságát félóránként kivett egy-egy tabletta, tehát a nyolcórás műszak alatt tehát összesen 16 tabletta hatóanyag tartalmának ellenőrzésével végezzük. (A 16 adatból a hatóanyag-tartalom átlaga 102,1 mg a szórás 4 mg.) Ez alapján:

Milyen becslés, ha azt mondjuk, hogy a várható érték 102,1 mg?

És, ha azt mondjuk, hogy 101 mg?

Valaki azt állítja: a 102,1 mg nagyobb valószínűséggel igaz, mint a 101 mg. Igaz-e ez?

Egy másik vélemény: 101,5 és 102,5 mg között nagyobb valószínűséggel található a várható érték, mint 100,5 és 101,5 között. Igaz-e ez?

Tudnánk-e hozzávetőleges becslést adni, hogy milyen valószínűséggel található a várható érték 101-103 mg között? (A 0-100% skálán – mégis – hol helyezkedhet el ez az érték?)

Megadhatunk-e egy olyan tartományt, amire azt mondhatjuk, hogy ebben aztán biztosan benne van az ismeretlen várható érték?

Mi a – fejezet alapján adható – válasz: 95% megbízhatósággal mely tartományban található a v-é?

Mekkora az előbbi döntés hiba-kockázata?

Mít értünk az előbbi 2 kérdésben a 95% megbízhatóságon, illetve a hibakockázaton?

Ha csak 90% megbízhatóságú döntést kivánunk, akkor hogyan alakulna az intervallum: szélesebb vagy keskenyebben lenne (esetleg: ugyanakkora maradna vagy: ugyanolyan széles, de eltérő helyzetű lenne - s ez utóbbi esetben: kisebb vagy nagyobb lenne-e a közepe)?

És, ha 99% megbízhatósággal teszünk becslést az intervallumra (az hogyan viszonyulna a 95%-oshoz)?

Ha a fenti példa adatai (mintatárgy 102,1 mg a szórás 4 mg) a 16 helyett 50 tabletta adatából kaptuk volna, akkor hogyan alakulna a 95% megbízhatóságú intervallum: szélesebb vagy keskenyebben (stb. az előző felsorolás szerint…) lenne-e mint az előbbi (16 adatból) esetben?

5.6. ábrán láttuk ennek a grafikus értelmezését: Ez az értelmezés az egész gondolatkör kulcskérdése! (itt hivatkozni az ábrára és a magyarázó szövegére… vagy az alábbi szövegre:).
Hogyan alakulna ez esetben (50 adathóból) a valószínűség becslése előbbi 101-103 mg tartománynra (szemben az ottani, 16 adattal becsült valószínűséggel)? Nagyobb, kisebb vagy kb. ugyanakkora lenne? Ha nagyobb vagy kisebb, akkor kb. milyen mértékű lenne a változás? Ha kb. ugyanakkora, úgy az estleges (kis) eltéréstre mit mondanánk: mely értéket nem éri el biztosan a változás: pl. „kb. ugyanakkora, de az ellérésem nem több, mint 1-2%”? A fenti kérdéseket meg lehet ismételni (esetleg részben szétszóttva) 6-8 különböző számpélda között. Lehet, persze, hagyományosabb módon is kezdeni: a 6-8 számpéldán kiszámoltatni példánként különböző (a többségre 95%-os, de 1-2-nél 90 illetve 99%-os) megbízhatóságú megbízhatósági intervallumokat. Így a hallgató először begyakorolja a számolás technikáját. Aztán jöhet az iménti megértést ellenőrző/elmélyítő kérdéssor!

További (megértést és gondolkodást ellenőrző) kérdések:

- Hallgatók testmagasság adataiból két eredményünk van (minta mérete, átlaga, szórása): 1, (16, 168,2 cm, 8 cm) és 2, (100, 170,5 cm, 7 cm). Mit mondanánk a várató értékre becslésként, ha
- 1, minta alapján pontbecslésként?
- 2, minta alapján pontbecslésként?
- 1, és 2, minta alapján pontbecslésként? (hogy tekintsük a két minta megbízhatóságát e becslés szempontjából?)
- 1, minta alapján intervallumbecslésként?
- 2, minta alapján intervallumbecslésként?
- 1, és 2, minta alapján intervallumbecslésként?

- Idézzük fel a megbízhatóság jelentését! Például 90% megbízhatóságnál 100-ból 90 intervallum tartalmazná a ?-i.
- De, ha nekünk csak 1 ilyen kísérletünk (1 adatsorunk, s a belőle számított 1 intervallumunk van, akkor az vajon tartalmazza-e vagy nem?
- És, ha tényleg elvégeznénk a 100 kísérletet (100 adatsor, 100 különböző intervallum), akkor abból valóban 90 tartalmazná? Nem lehet-e, hogy csak 89? Esetleg 92?
- Mekkora értékre mondanánk, hogy na, az azért már nem lehet…
- … és mi jut erről akkor eszünkbe? Mit mondhatunk, 100 ténylegesen kiszámolt intervallumból legalább és legfeljebb mennyi tartalmazza?
- Mennyiben hasonló és/vagy mennyiben különböző ez a kérdés ahhoz, amikor azt vizsgáltuk, hogy 95 intervallumból hány rövidül meg (ténylegesen, legalább és legfeljebb) az Sx-? csere miatt?
- Fontos utasítás: A válasz mellé fontos indoklást is adni: „ez a válasz, és azért, mert…”
- és a válasz értékelésében fontosabb az indoklásból kiderülő gondolkodásmód, mint önmagában a „végeredmény” pontossága vagy pontatlansága)
6. fejezet - A hipotézisvizsgálat

1. A fejezet tanulmányozásához szükséges előismeretek

- Változók típusai, a diszkrét és a folytonos változók.
- A binomiális eloszlás (Ellenőrző kérdések).
- A minta és az alapsokaság.
- Hisztogram, relatív gyakoriság sűrűség, sűrűségfüggvény.
  - A valószínűség mérőszáma: görbe alatti terület p(a,b) (Ellenőrző kérdések).
  - A minta jellemzése, a sokaság paraméterei.
  - Átlag és szórás, a várható érték és a szórás (Ellenőrző kérdések).
- A normális eloszlás és jellemzői (egyebek mellett: 68-95-99,7%).
  - Mintapélda: hallgatók testmagassága.
    - minta mérete és várható terjedelme. (a szórás „működése”, a „3 szabály”)
  - A std normális eloszlás és értelmezése (Ellenőrző kérdések)
- A normális eloszlás kiemelt jelentősége az orvostanhallgatók számára
  - A binomiális eloszlás és határeseti közelítése normálissal (Ellenőrző kérdések)
- Az átlag eloszlása
  - Mintapélda: 16 fős csoportok átlagmagasságának eloszlása (Ellenőrző kérdések)
- A megbízhatósági intervallum a várható értékre
  - A statisztikai becslés és fajtáit: pont és intervallumbecslés
  - A megbízhatóság értelmezése az intervallumbecslésnél (Ellenőrző kérdések)
- A -re vonatkozó ±t*sx/n intervallumbecslés
  - A t-ecélzások és eltérésük a std normálistól
- Mintapélda: 16 tabletta hatóanyagtartalma – mennyi lehet a ? (Ellenőrző kérdések)

2. A hipotézisvizsgálatról

Az előző fejezetben megismertük a két valószínűségre alapozott döntéshozó módszer közül az intervallumbecslés eljárását. Erre példaként megismertük a (folytonos, normális eloszlású) adatokból a várható értékre adható becslést. Arra a kérdésre, hogy „Mennyi (lehet) a értéke?”; azt mondjuk: „A várható érték valahol az körül, – éspedig az ±t*sx/n intervallumban – található ‘p’ valószínűséggel.

Ez a várható érték p%-os megbízhatósági („konfidenca”) intervalluma.

A hipotézisvizsgálatra ugyanezt a mintapéldát tekintve: Azt kérdezzük, hogy „lehet-e a értéke egy konkrét „a” érték – és a válasz „igen” vagy „nem” lehet.
Természetesen, a válaszunkban ismét csak egy bizonyos (100%-nál kisebb, de ahhoz közelű) mértékben bízhatunk, és egy bizonyos (vállalt!) hibakockázattal döntünk!


Nézzük a második fajta döntéshozó módszer, a hipotézis vizsgálat gondolatmenetét. Azt állítsuk, hogy ezt a legtöbb ember naponta sokszor gyakorolja a hétköznap élet döntési helyzeteiben!

Hadd vegyünk ezért egy hétköznap életből vett példát, azt demonstrálva, hogy ez a gondolatmenet nem egy mesterséges találmány, ezt valóban eredményesen használjuk a mindennapokban – és egyúttal fedezzük is fel a módszer sablon-lépései és a gondolkodás jellegzetességeit!

Tételezzük fel, hogy éjszaka – talán félállomban – mintha locsogó eső hangját hallottam volna. Hogy döntöm el reggel, hogy jajon esett-e éjjel az eső, vagy csak álmodtam az egészét?

1. „Valószínűleg csak álmodtam ezt az esőhangot, mert nem szoktam éjszaka felébredni” – gondolom –, tehát: Nem esett!” – feltételezem első válaszként a kérdésemre.

2. Tudom, hogy számomra mit jelent a határ a „valószínű” és a „nem valószínű” között.)

A második lépést a gondolatmenetben egyelőre zárójelben írtam le: A fenti mondatot nem szögezem le magamban, egyszerűen tudom, hogy így van, ez a számomra nem igényel komolyabb megfontolást, mérellegést: Bármiről el tudom dönteni egy adott helyzetben, hogy én azt „valószínűnek” vagy „nem valószínűnek” tekintem. E második lépés lényege ugyanis: Mi a határ a kis és a nagy valószínűség között?

A következő fejezetben ennek a kiemelt jelentőségéről – miért és miért itt? – bővebben beszélünk

3. Odamegyek az ablakhoz, kinézek és hatalmas tócsákat látok az úttesten. „Tévedtem” – gondolom – ha nem esett az éjjel, mitől lennének ezek a tócsák az utcán? Talán csőtörés volt? Esetleg az öntözőautó itt locsolta ki a megmaradt vizét az úttestre a házunk előtt? Nem valószínű!”…

4. ...”Szóval tévedtem!”…

Tehát a negyedik pont egy egyszerű szabály: A valószínűség alapján döntünk a hipotézist illetően: Ha a harmadik lépésben a tényt (bármi is az: itt most a tócsákat): „nem valószínű” következményének kell tekintenem az 1., pontban megfogalmazott kezdő feltételezésemől (a „hipotézis”-tól, mely szerint „nem esett” és csak álmodtam)? És állást foglalok, hogy ekkora elérést kicsi vagy nagy valószínűségűnek tekintek („valószínű” vagy „nem valószínű”) – nyilván korábbi tapasztalataim alapján – akkor, ha az 1. pontbeli kiinduló gondolatomban helyes lenne.

5. „Tehát: Esettt!”

Az ötödik pont a következtetés: Fogalmazzuk meg, hogy mit jelent ez az előbbi döntés hétköznap nyelven.

A példánkban elvetettük a „Nem esett, csak álmodtam” hipotézist, tehát azt mondjuk: „Esettt!”

Az imért feltárt 5 gondolkodási/logikai lépés alkalmaz arra, hogy ismert tények alapján valamely ismeretlen helyzetről (hipotézisről) döntést hozzunk. Kér dés az, hogy ez az ő lépés vajon csak most működött, vagy egy univerzálisan alkalmazható, tudományos igényt kielégítő, ismételhető módszer. Erre egy egyszerű próba: Vajon...
mennyire függ a kezdő feltevéstől a gondolatmenet működése? Tételezzük fel az 1. pontban az előbbi hipotézisünk ellenkezőjét:

1. Tegyük fel, hogy „Esett!”. …
2. Változatlan (tapasztalatból jól tudjuk a határt a „valószínű” és a „nem valószínű” közt).
3. Ugyanaz, a tény az tény!; „tócsákat látok”. Viszont, ha éjjel esett, akkor ennek „valószínű” következménye, hogy reggel tócsákat látok. Tehát itt, a harmadik pontban a valószínűség megváltozik!
4. Tehát fordítva kell dönteni: A tócsák léte („tény”) valószínű következménye egy estleges éjjeli esőnek („hipotézis”). Így nincs ok elvetni a hipotézist!
5. Esset az eső!

Az ellentétes hipotézisből kiindulva ugyanazon következtetésre jutottunk!

Vegyük észre, hogy ennek egy oka van: mindkét gondolatmenetben ugyanarra a „tényre” (tócsák!) alapoztuk a döntést! Beláttuk tehát, hogy a döntés e gondolatmenet szerint nem függ a hipotézistől! De vajon függ-e a ténytől, amit tapasztalunk, amire alapozunk? Ehhez a fejezet végén ellenőrző feladatsor található!

Tehát az eredmény nem függ attól, hogy mi volt az eredeti feltevésünk (hipotézisünk), az ötlépéses gondolatmenet jól működik, valóban kielégíti a tudományos követelményeket.

A hipotézisvizsgálat a biometriában

4.1. Az egymintás t próba

Alkalmazzuk az előbbi öt lépést egy adatértékelési helyzetre, egy konkret számpéldára! Használjuk ugyanazt a feladatot, amit a megbízhatósági intervallumnál harmadik példaként használtunk. (16 tabletta hatóanyagtartalma – az átlag – 102,1 mg, a szórás 4 mg, lehet-e a várható érték a dobozra írt névleges 100 mg?):

1. Tegyük fel, hogy a = 100 mg igaz.

Magyarul nincs lényeges eltérés a (hipotézis szerinti) várható érték 100, és a mintaátlag 102,1 között: A 2,1 mg eltérés csak a véletlen műve. Ezt kezdő, vagy „nullhipotézisnek” hívjuk, a jele H0.

2. A kis- és nagy valószínűség (a „nem valószínű” és a „valószínű”) közti határt kell megmondanunk itt: Ezt döntési szintnek hívjuk és p-val jelöljük. Legyen ez most öt százalék: p = 0,05.

Ez azt jelenti, hogy az öt százaléknél nagyobb valószínűséget „nagy valószínűség”-nek fogjuk tekinteni, az ötnél kisebbet pedig „kicsi”-nek. (A következő fejezetben e határ megválasztásáról még részletesen beszélünk!)

3. Milyen „távol van” a tény a hipotézistől? Mekkora is az a 2,1 mg eltérés? És: Milyen valószínűségű a „legalább ekkora (itt 2,1 mg) eltérés” akkor, ha a H0 igaz (= 100 mg)?

A „féltőségi hipotézisvizsgálatban" felismert öt lépést harmadik lépésében ez utóbbi „B” lépés volt az alapvető. De ennek meghatározásához most tudunk majd, mekkora is az eltérés (hisz „kis” eltérést nagy valószínűséggel okozhat csak a véletlen, míg „nagy” eltérés kevésbé valószínű), vagyis szükségünk van előbb az „A” pontra.

Ebben támaszkodhatunk az előző fejezetbeli számolásunkra:

Beláttuk, hogy a μ és az standard hiba (sx/n) egységekben kifejezett különbségéhez tudunk valószínűséget („megbízhatóságot”) meghatároznia: mégpedig a t valószínűség-sűrűség függvények alapján. A példánkban:

A „távolság” (= tény – hipotézis) 102,1 – 100 = 2,1 mg, a std hiba 4 mg/16 = 1 mg, tehát a távolság 2,1 mg/1 mg = 2,1 std hibányi. A „valószínűség” pedig: Az n-1 = 15 szabadsági fokú t-eloszlásnál a t = 2,13-nál nagyobb különbséghez tartozna 5%-nyi görbe alatti terület (6.7. táblázat és 6.9. ábra alapján). A „legalább 2,1-szer std hibányi
különbséghez” (ami példánkban a tény és a hipotézis között van: (102,1-100)/(sx/n) = 2,1/(4/4) = 2,1-szeres különbséghez) vajon milyen valószínűség tartozik? Az ábráról jól látszik, hogy a 2,1 az „beljebb” van, mint a 2,13 tehát a hozzá tartozó (ettől „kifelé” elhelyezkedő) terület nagyobb, mint 5%.

A „valószínűség” (ezt ’p’-vel jelöljük) nagyobb, mint 0,05 (5%, vagyis az ?), tehát pg ?.

4. Emlékezzünk, a döntési szabály az, hogy ha:

kicsi a valószínűség (pg ?), tehát „nem valószínű”, hogy a véletlen okozzon legalább ekkora („nagy”) különbséget a tény és a hipotézis között. Ekkor elfeljük a hipotézist és azt mondjuk, hogy az eltérés szignifikáns ? szinten.

nagy a valószínűség (pg ?), tehát „valószínű”, hogy a véletlen okozzon ilyen („kicsi”) különbséget a tény és a hipotézis között. Ekkor nincs okunk elfelteni a hipotézist hipotézist és azt mondjuk, hogy az eltérés nem szignifikáns ? szinten.

A p = ? esetben akármerre dönthetünk. Példánkban ez utóbbi a helyzet: pg0,05, tehát megtartjuk a hipotézist

5. Következtetés: a mintáink átlaga nem tér el szignifikánsan a hipotetikus várható értéktől, tehát a μ lehet 100 mg. Másként fogalmazva: legalább ekkora eltérés (nagy valószínűséggel) lehet a véletlen miatt.

Az eljárás átgondolása:

Megvizsgáltuk, hogy a mintha átlaga milyen távol van egy feltételezett (H0 szerinti) várható értéktől. Ezt a távolságot standard hiba egységekben kifejezve mérjük, és a mértéke a „t” érték. Kiszámoltuk, hogy milyen valószínűséggel fordulhat elő legalább a tapasztalt mértékű eltérés, ha a H0 igaz.

Itt van „elrejtve” a módszer alkalmazásának egy fontos feltétele: a valószínűséget úgy tudjuk kiszámolni, ha feltételezünk, hogy az adataink a normális eloszlást követik.

Ezt az ábrát én szétszedném két külön ábrára, melyeken a levágott részek szimmetrikusak. Ez a távolság lehet nagy – nézzük itt az 6.13. ábrán: Nagy (t2) távolságnál a „legalább ekkora különbség”-hez tartozó valószínűség (a görbe alatti piros farok-terület) kicsi lesz (p2), ha a H0 igaz.

Emlékezzünk: A görbe alatti pl. a és b közti –pl. a és b közti – terület azt jelenti, hogy milyen valószínűséggel fordul elő a mintatálfad adott mértékű – a és b közti – eltérése a várható értéktől.

A „nagy távolság – kicsi valószínűség” az azt jelenti, hogy nagy különbségeket csak kis valószínűséggel okozhat a véletlen. Pedig a nagy különbség (ez tény) előállít! Ez valószínűtlen lenne, ha a H0 (ez csak hipotézis volt) igaz. Ez az ellentmondás feloldása: a H0 valószínűleg téves! Ezért elfeljük.

Ezzel szemben – az ábra bal oldalán ezt rajzoltuk fel –, kis (t1) távolsághoz nagy (p1)valószínűség, nagy farok-terület tartozik. Ennek jelentése, hogy:

Kis eltérés-t nagy valószínűséggel okozhat a véletlen, ha a H0 igaz. Tehát a kis eltérés (ami a tény) nem mond ellent a H0-nak (ami a hipotézis volt), így nincs ok elvetni a nullhipotézist.

Ez a terület, a valószínűség – a t és az n ismeretében – a sűrűségfüggvényből kiszámolható. Az SPSS a p értékeket pontosan kijárja. Vagy, ahogy az előbb számoltunk a 16 tabletta hatóanyagára: Csak becsüljük, hogy a „p” valószínűséget a döntési határhöz képest (a példában ? = 5% volt) mekkora. Elég volt tudni, hogy a p nagyobb, mint 5% Ez alapján már meghozhattuk döntést.

A módszer a neve: egymintás t próba. Egy adatsorból hozzunk döntést a értékére nézve.

4.2. Három fontos megjegyzés

4.2.1. A nullhipotézisről

Épp az imént, a hétköznapi hipotézisvizsgálati gondolatmenetnél azt bizonyítottuk, hogy mindegy, mit fogalmazunk meg induló hipotézisnek: Az eljárás működik! Csakhoz, amikor tényleges számértékek (adatok) alapján hozzunk döntést, akkor pontos valószínűségkalkulációt kell végezni. (Nem éleg csak „valószínű” vagy „nem valószínű” hétköznapi különbségtétel.) Valószínűségeket akkor tudunk pontosan számolni, ha az értékeinket csak a véletlen befolyásolja. Ezért tényleges számolások esetén nem mindegy, hogy mi a H0: Ennek
a „nincs hatás” helyzetnek kell lennie, ekkor ugyanis a mért értékeinket csak a véletlen (mérési hiba, biológiai variabilitás, stb.) téríti el a vár értéktől.

4.2.2. A döntési szintről

A második megjegyzés: A valószínűségek közötti határ („kicsi” vagy „nagy”) igen élesen értelmezendő. Ha például ez a határ a szokásos 5 %, akkor az 5,01%, az már „nagy” valószínűség, a 4,99% pedig „kicsi” valószínűség, holott a kettő között hajszállnyi különbség van csak. Ezért nagyon fontos, hogy ez a határ micsoda: ezen műlök, hogy hogyan fogunk dönteni. A határ (az döntési szint) megválasztása nem lehet önkényes, továbbá nem lehet mindig ugyanaz (pl. 5 %) érték sem! A következő fejezetben átgondoljuk, hogy milyen esetekben kell – nemcsak lehet, hanem kell! – ettől lefelé, avagy fölfelé elterünk.

Még nem értem odáig, de valahol meg kell említeni, hogy nem szabad „mereven” ragaszkodni a szignifikáns - nem szignifikáns döntéshez, mindig értelmezni kell a kapott p-értéket, mert p = 5,01 nem szignifikáns ugyan, de csak 5%-os szinten nem az.

4.2.3. A döntésről

A harmadik megjegyzés: Vegyük észre, hogy a „H₀ megartása” döntés nem azért született, mert bizonyítottuk, hogy H₀ igaz! Hanem „csak” azért, mert nem volt elég (nyomós) ok az elvetésre.

Ellenőrző kérdés: Mi lehet ennek a (két, alapvetően különböző) oka?

Válasz: Vagy nincs különbség (tényleg igaz a H₀), vagy csak nem elég érzékeny a vizsgálatunk, hogy a létező különbséget igazolni tudja.

5. A hipotézisvizsgálata további két alkalmazása

5.1. Páros t próba

Nézzünk egy újabb példát, ami igen gyakori típus az orvosi gyakorlatban. A feladat, hogy egy diétával kombinált tréning hatását vizsgáljuk: Csökken-e a páciensek koleszterin szintje? A tréning előtt és után is megnéztük a koleszterin értékeket. A 12 páciens laboreredménye két adatsor:

<table>
<thead>
<tr>
<th>Sorszám</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Után</td>
<td>200</td>
<td>236</td>
<td>216</td>
<td>233</td>
<td>224</td>
<td>216</td>
<td>296</td>
<td>195</td>
<td>207</td>
<td>247</td>
<td>210</td>
<td>209</td>
</tr>
<tr>
<td>Változás</td>
<td>-1</td>
<td>5</td>
<td>-27</td>
<td>-4</td>
<td>-21</td>
<td>-30</td>
<td>-40</td>
<td>-33</td>
<td>-20</td>
<td>-74</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Előtt</td>
<td>201</td>
<td>231</td>
<td>221</td>
<td>260</td>
<td>228</td>
<td>237</td>
<td>326</td>
<td>235</td>
<td>240</td>
<td>267</td>
<td>284</td>
<td>201</td>
</tr>
</tbody>
</table>

Azonban ez a két adatsor páronként összetartozó adatokból áll! Valójában az a kérdés, hogy a különbség a két adat, az ‘előtt’ és az ‘után’ között, tehát a különbség, a változás érdemben eltér-e 0-tól. Tehát nézzük csak a változás adatsort, ami a harmadik sorban látható. A változás (egyetlen adatsor), átlaga 20,17 a szórása 23,13 a standard hiba 6,68. Nézzük most a hipotézis-vizsgálat 5 lépését: (Ez tehát a már ismert egymintás t-próba, melyet a különbségre alkalmaztunk.)

1. H₀: az, hogy a kezelés nem hatékony. Tehát a változás várható értéke 0.

2. A döntési szint (az ?) 5% legyen.

3. Megnézzük, hogy a tapasztalt eltérés (a tény 21,17 a hipotézis szerint várt 0-hoz képest) hányszorosan a standard hibának (a 6,68-nak). Ez a távolság (a t értéke) 3,02.

Mi a valószínűsége annak, hogy legalább ekkora távolság csak a véletlen miatt álljon elő. Erre az egyszerűbb módszer, hogy a számítógép ezt kiszámolja: p = 0,0117 (1,17%).
4. Mivel p? (= 0,05), tehát elvetjük a Hₐ-t!

5. K övetkeztetés: a diéta és a tréning hatékony volt, a változás érdemi, szignifikáns.

Ennek a módszernek a neve páros t próba.

Mint beláttuk, ez valójában egyeztethető t-próbára a „változás”-ra. Azért van külön neve, mert a változás adatokat két – rendszerint egy „előtt” és egy „után” – adat különbségeként kaptuk meg.

Ellenőrző kérdés: Hogyan tudnánk becsülni a p valószínűségét? (Azt kellene tehető belátni, hogy a p érték 5%-nál kisebb. (És, há t esetleg azt, hogy egy százaléknál viszont nagyobb!))

Összefoglalva a két példa (egyenmérték és páros) alapján az egyenmér tek t próbáért: Ha folytonos és normális eloszlású a változó (az adataink), és a kérdés a minta átlaga és egy várható érték eltérése, tehát, hogy van-e különbség, hatás, vagy pedig az eltérés csak a véletlen műveknek tulajdonítható (ez utóbbi a kezdő hipotézis – emlékszünk?), akkor egyenmérték t próbával válaszolhatunk a kérdésre. Mégpedig úgy, hogy kiszámoljuk, milyen valószínűséggel okozhat legalább ekkora (a standard hiba t-szerese) eltérést csak a véletlen, ha a H₀ igaz. Ez a p valószínűség a (-t, t) intervallumon kívüli terület az adatoknak megfelelő t eloszlás sűrűségfüggvénye alatt, amit számítógéppel, vagy táblázattal meg tudunk határozni.

Ha ez különbség nagy, akkor az azt jelenti, hogy a p valószínűség kisebb, mint egy előre meghatározott döntési szint (az ?). Vagyis legalább ekkora különbség előfordulása kis valószínűségű, ha a nullhipotézis igaz, ezért elvetjük a nullhipotézist.

Ha ez a különbség nem nagy, tehát a p valószínűség nem kisebb, mint a döntési szint (vagyis p „elég nagy”), akkor viszont nincs okunk elvetni a nullhipotézist.

Speciális eset volt, amikor két, páronként összetartozó adatsorból a különbségre végeztük el ugyanezt, ezt hívíjuk páros t próbának. Vegyük észre, hogy ekkor a különbség adatsora (egyetlen minta), mindig a = 0 nullhipotézissel vizsgálandó, hiszen az a kérdés, hogy a különbség nulla-e? Vagyis: volt-e hatás, vagy nem?

Kérdés: Mint láttuk, a t próba feltétele az adatok normális eloszlása. A 6.14. táblázat különbség adatsairól vajon jogosan tételezhetjük-e fel a normalitást? (Egy képzeletben felrajzolt hisztogramon milyen képet mutatnak ezek az adatok?)

5.2. Az előjél-próba

Nézzünk most újra az előbbi példát (6.14. táblázat.)! Az előbb elemeztük a változás adatsorát. Vegyük észre, hogy a változás a tizenkét páciensből kettőnél volt pozitív, és tíznél negatív. Magyarul tíznél csökkent a koleszterinszint és csak kettőnél nőtt. Hogyan tudjuk ezt az eddigihez hasonló („hipotézisvizsgálat”) gondolatmenettel kiértékelni?

1. A nullhipotézis ugyanaz, mint az előbb, tehát, hogy a kezelés nem hatékony!

Ekkor mit várnunk – s most vizsgáljuk csak a változás negatív vagy pozitív voltát, tehát –az előjeléket illetően? Ha nem lenne hatékony a kezelés, és a változás csak a véletlen miatt térne el a nullától: akkor azt várnánk, hogy a tizenkét páciensből hatszor pozitív, hatnál negatív, hatnál negatív legyen a különbség. Mi is ez? Ez a binomiális eloszlás esete! Hiszen, hogy hány pozitív és negatív változás van (ha ez a két kimenetet egyenlő valószínűségű), akkor ezzel ugyanaz a helyzet, mint amikor tizenkét csücs dobók egy érmén, s azt kérdésem, hány szor késett a legjobb, és hány szor írás. Véletlenszerű, hogy fej, vagy írás lesz-e egy-egy feldobás eredménye. Ha nincs hatása a döntésünknél a tréningnek, akkor véletlenszerű az, hogy egy páciens koleszterinszintje két mérsz között csökken, vagy nő. Tehát ez esetben (nullhipotézis) 6 negatív és 6 pozitív változást várnunk. Ez tehát a várható érték.

2. A döntési szint maradjon 5%.

3. A kérdés, hogy mikkora a H₀ esetén várt és a tényleges adat eltérése, és milyen valószínűségű a (k = 6) várható értéktől legalább ekkora (esetünkben legalább 4) vagy ennél nagyobb eltérést?

Példánkban 2 esetben kaptunk növekedést (ami 4-gyel töl el a várt 6-tól), és 10 esetben csökkenést. Ha felrajzoljuk ennek a binomiális eloszlásnak a B (12, 0.5) paraméterű binomiális eloszlásnak a valószínűségi értékeit, akkor ez úgy fognak nézni, hogy picivel több, mint 20% a valószínűsége, hogy 6 fej, 6 írás lesz. Az 5
illetve a 7 fej eredmény már picivel 20% alatti valószínűségű. A 4 vagy 8 fej az 12% körüli, a 3 illetve 9 fej az 5% körüli valószínűségű. A számmunka most érdekes 2 illetve 10 fej valószínűsége 1,6%, az 1 illetve 11 fej valószínűsége 0,3%, végül a 0 illetve 12 fej mindössze 0,02% valószínűségű. Tehát a valószínűség ennek a három pici valószínűség összegének kétszerese:

Ez a valószínűség mintegy 3,8%, ami kisebb, mint 5%.


5. Tehát a diéta és a tréning hatékony volt. A változás érdemi, lényeges, szignifikáns.

Ennek a módszernek a neve előjel próba.

Vegyük észre azt, hogy itt nem használtunk az értékelésnél a változás adatok számértékeit. Tehát az a korábban (a t próbánál) alkalmazott feltétel, hogy az adatsor normális eloszlású, itt értelmetlen. Nincs ilyen feltétel, csak az előjeleket vizsgáltuk. Épp ez a módszer alkalmazható olyankor, hogyha nem tudjuk bizonyítani, nem tudjuk feltételezni, hogy az adataink normális eloszlásúak. Ekkor az előjel próbá meg nyugodtan alkalmazható, hiszen csak az előjeleket vizsgálja.

6. Összefoglalás

6.1. A hipotézisvizsgálat gondolatmenete

Miről is volt szó e fejezetben? Megismertük a hipotézisvizsgálat 5 lépéses gondolatmenetét egy hétköznapi példán. (Ez volt az „esett-e az éjjel?” kérdés példa.) Az öt lépés a következő:

1. A vizsgálandó hipotézis (H₀) megfogalmazása.

2. Mi a határ a kis és a nagy valószínűség között? Ez a döntési szint, az ?.

3. Milyen mértékű az eltérés és milyen valószínűségű a legalább a tapasztalt mértékű eltérés a H₀ esetén várhatóhoz képest (tehát, ha a H₀ igaz)? Ezt a valószínűséget jelöljük a p.


5. A következtetés: megfogalmazzuk, hogy mit is jelent ez az előbbi döntés a kérdést illetőe

Továbbá megismertünk három konkrét módszert: Az egymintás- illetve a páros t próbát, valamint az előjel próbát. Mind a három hipotézisvizsgálat, ugyanezt az 5 lépéses gondolatmenetet tartalmazza.

6.1.1. A hipotézisvizsgálat működése: az egymintás t próba

Nézzük most meg ezt a módszert elvi megvilágításban, az egymintás t-próba mintáján.

Az első fontos különbség a hétköznapi gondolatmenethez képest, hogy a hipotézisvizsgálat esetén a H₀ mindig a „nincs hatás” helyzet, amikor csak a véletlennek van szerepe az adatok változásában.

Ekkor tudjuk ugyanis kiszámolni pontosan azt a valószínűséget, amely máj a döntésünknek az alapja lesz.

A „tabletták hatóanyagtartalma” példán azt vizsgáltuk, hogy egy névleges, feltételezett várható értéktől mennyire tér el a kapott mintáltalag. (Az alapkérés: okozhatja-e ezt az eltérést pusztá véletlens? A válasz az eltérés mértékétől függ: hogy mekkora a „t” értéke.)

Vegyük észre, hogy ez a t érték, egy pusztá szám – mint „távolság” –, önmagában nem jelentene semmit: Gondoljunk arra, ha például azt kérdezzem, hogy milyen messze vagyunk a városközponttól, akkor válaszolhatnám azt rá, hogy 3 a távolság, de esetleg válaszolhatnám azt is, hogy kb. 2 a távolság. Az egyik esetben azt értem rajta, hogy ez kb. 3 km, a másik esetben az, hogy ez kb. 2 mérföld. Tehát nem mindegy, hogy mi a mértékegység.

Itt a t mindig „standard hiba” (tehát Sx/n) egységen méri ezt a távolságot, ezért két különböző esetben ugyanazon t érték (pl. a „kettő”) csak azt jelenti, hogy „kétszeres standard hiba”.
A hipotézisvizsgálat

A standard hiba – persze – esetről esetre eltér, kisebb vagy nagyobb: így pl. a \( t = 2 \) esetről esetre más különbség. Nézzük az 6.15. ábrát!

Ha ez a \( t \) (értékű) távolság nagy, azt jelenti, hogy a feltételezett \( \mu \) az \(-t\)ól nagy (\( t_2 \)) távolságból van. Akkor a legalább ekkora \( \mu \) különbség (ezt jelöli a piros nyíl) valószínűsége (\( p_2 \)) – vagyis az ettől nagyobb eltérésekhez tartozó görbe alatti terület – kicsi. (Az a kicsi piros farok-terület!)

Miért is? Emlékezzünk, az \( \alpha \) fejezetben: A lehetséges eltérések valószínűségét a megfelelő \( t \) valószínűség sűrűség függvény görbe alatti területe adjja. Tehát csak egy-egy intervallumra vonatkozóan kapunk nullától eltérő valószínűséget, egy-egy \( t \)-értékhöz tartozó terület (valószínűség) az mindig nulla.

Tehát, ha a \( t \) értéke nagy, a \( p \) értéke pici, vagyis kicsi a valószínűsége annak, hogy legalább ilyen nagy eltérést csak a véletlen okozza. Ezt az alapján, ha nullvalószínűsítünk.

Ellentét esetben, ha a \( t \) távolság kicsi: az ábra túloldalán zölddel ábrázolva (közelről induló zöld nyíl), akkor a hozzá tartozó valószínűség nagy lesz (nagy \( t \)-zől terület). Tehát, nagy valószínűséggel okozhat ekkora vagy ennél nagyobb eltérést csak a véletlen, ha nullaváltoztatok a hipotézisünk.

A kettő között természetesen van egy határ (a kicsi és a nagy távolság határa). Ezen \( t \) határértékehez tartozó görbe alatti terület éppen az \( \alpha \) érték, aminek a szokásos értéke az 5%. (erre térünk még vissza később, hogy ettől néha el kell térni vagy lefelé, vagy fölfelé.

Megjegyzés: számítógépes értékelés esetén a program pontosan kiszámolja az adott \( t \) távolsághoz tartozó \( p \) értéket. Így könnyen dönthetünk az alapján, hogy ez kisebb vagy nagyobb, mint az \( \alpha \) határérték.

Ha nincs számítógép, s csak a \( t \) távolságot ismerjük, akkor viszont a döntéshez elég azt tudni, hogy az \( \alpha \)-hoz tartozó \( t \) értéknel („kék határvonal” az \( \alpha \)) ez kisebb-e vagy nagyobb? (Ha kisebb, akkor a \( p_\alpha \), tehát nincs ok a hipotézis elvetésére. Ha nagyobb, akkor a \( p_\alpha \), tehát elvetjük a \( H_0 \)-t.) Ezek a \( t \) határértékek a \( \alpha \)-hoz és a szabadsági fokhoz) táblázatban megtalálhatók.

7. Ismeret-ellenőrző és gyakorló kérdések és feladatok

Amikor a hipotézisvizsgálatban döntünk (igen vagy nem) a kiinduló feltevésünk sorsáról, akkor ez az igen-nem döntés milyen megbízhatóságú?

Valaki azt mondja: Az igen/nem döntés egy éles választás, tehát a döntés teljes bizonyosságot! Vajon igaza-e? Ha igen/ha nem, miért?

Játsszuk végig gondolatban a hipotézisvizsgálat fenti 5 lépését az alábbi szituációkra:

- Hipotézis: esett, tény: nagy tócsák!
- Hipotézis: esett, tény: minden porszáraz!
- Hipotézis: nem esett, tény: június vége, délelőtt 9óra, szikrázó napsütés élénk széllel, nincs tócsa az úton!
- Hipotézis: nem esett, tény: június vége, délelőtt 9óra, szikrázó napsütés élénk széllel, nincs tócsa az úton!
- Hipotézis: esett, tény: június vége, délelőtt 9óra, szikrázó napsütés élénk széllel, rohanó felhők, nincs tócsa az úton, a tetőcsatorna kiömlőnyílása előtt nedves hordalékkupac látható!
- Hipotézis: nem esett, tény: június vége, délelőtt 9óra, szikrázó napsütés élénk széllel, rohanó felhők, nincs tócsa az úton, a tetőcsatorna kiömlőnyílása előtt nedves hordalékkupac látható!

Találjuk ki további megfigyelési – (tény) – helyzeteket az előbbi hipotézisekhez, s játsszuk ezekkel is végig a döntéshez vezető 5 lépéses gondolatmenetet!

Keressünk olyan (lehetséges) jelenségeket, amik egy éjszakai eső után több órával az eső következményei lehetnek, a korábbi esőt mutathatók/igazolhatják.

Keressünk olyan (lehetséges) jelenségeket, amik egy éjszakai eső ellen szóló jelenségek lehetnek több órával később, ezzel kizárva a korábbi eső lehetőségét.

Ezen „tünetekből” kombináljunk realiss vagy reálsnak tűnő „tényhelyzeteket” az előbbi kérdéshez.

Vegyük észre, gondolkodjunk el azon, hogy ez a „tüntetgyűjtő” feladat nagyon rokon azzal, amit egy orvosnak
A hipotézisvizsgálat

folyamatosan végeznie kell, ha a páciense tüneteit fél kívánja deríteni, ismerni: Észre kell vennie mindazt (még ha a páciensnek esetleg nem is tűnt fél, nem panaszodott), ami az ő egészségének állapotát jól jelezni képes: Az esetleges betegségeket és/vagy az egészséget! Mindkettőt, hisz miként az esőnek és a hiányának egyaránt vannak tünetei, úgy a betegségnél, de az egészségénél is!

Találjunk ki más alaphelyzetet és vizsgáljuk meg az 5 lépéses gondolatmenet működését eltérő hipotézisek és tények esetén az előbbi két kérdés mintájára!

Példa új alaphelyzetre: Épp munkába indulok, amikor a reggeli rádió-hírek egy fél mondata megüti a fülemet: … az X és Y utcák kereszteződésében, a baleseti helyszínél még tart, torlódás van! Tudom, hogy X és Y gyakori utcanevek, az én városomban is vannak, kereszteződnek is, de még legalább 2 másik városról tudok, ahol ilyen kereszteződés szintén van. Nem hallottam, milyen városban történt az eset.

Milyen tényhelyzetek lehetségesek, amiket a nap során később, amikor arra járok, láthatok? Játsszuk végig a lehetséges hipotézisekkel és tényekkel az 5 lépést a döntésig.

További hétköznapi próbahelyzetek a döntés gondolatmenetére:

Háttér ismeret az értékeléshez („eloszlás”): Pécsen júliusban a várható napi maximális hőmérséklet 26-27C. A napi maximumok 100 éves maximuma 40C minimuma 12C. (Az eloszlás értékelése: 1, mekkora a szórás? 2, mely értékek milyen gyakoriak? Ennek ismeretében:
Vizsgáljuk meg ugyanezt a fenti többi tényhelyzetre is (külön-külön)! Ugyanezt (módosított számértékekkel) végezzük el januári hőmérsékletekkel (megadandó!) is.
Ha a feladatot öszi vagy tavaszi adatokkal értékeljük, mi lesz az eltérés az iménti téli vagy nyári helyzethez képest?

Mit tapasztaltunk: Hány helyzetben próbáltuk ki a módszert? Hányban működött és hányban nem?
Az előbbi fejezet előjelpróbájánál használt B (12; 0,5) eloszlás valószínűségeit számoljuk ki!

Milyen ez az eloszlás? Milyen az eloszlásfüggvény alakja?
Miért volt jogos a 3 kis valószínűség kétszeresét venni p értékeként?
Mennyi az eloszlás várható értéke és szórása?
A „4”, mint eltérés-érték, hányszoros szórásnál nagyobb illetve kisebb?
Ez alapján tudunk-e „fejbol” döntést hozni a feladat H₀-ét illetően – nem tekintve a kiszámolt valószínűségeket
Lehetnek-e eldönthető H₀-ek az alábbi kijelentések?

A különbség nulla.
Az eltérés nem jelentős.
Az eltérés jelentős.
Az eltérés csak a véletlen műve.
A változás mértéke egy adott érték (pl. 15 g/l)
Mind egyenlő, a tapasztalt eltéréseket csak a véletlen okozta.

Az alábbi kijelentések közül melyik nem lehet egy hipotézisvizsgálat eredménye (a döntés)?

A különbség nem nulla.
Az eltérés nem jelentős.
Az eltérés jelentős.
Az eltérés csak a véletlen műve.
A változás mértéke egy adott érték (pl. 15 g/l)
Mind egyenlő, a tapasztalt eltéréseket csak a véletlen okozta.

Az alábbi kijelentések közül melyik nem lehet egy hipotézisvizsgálat eredménye (a döntés)?

A különbség nem nulla.
Az eltérés jelentős.
Az eltérés nem jelentős.
Az eltérés nem csak a véletlen műve.
A változás mértéke egy adott érték (pl. 15 g/l)
A változás mértéke eltér egy adott értéktől (pl. 15 g/l-től).
Mind egyenlő, a tapasztalt eltéréseket csak a véletlen okozta.
Mind eltérő, a tapasztalt eltéréseket nem a véletlen okozta.
Mind egyenlő, a tapasztalt eltéréseket csak a véletlen okozta.
Legálább egy eltér a többtől.
Az alábbi 1-4 kérdéseket oldjuk meg előbb becsülve az eredményt, majd SPSS-el számolással is. (A becslést illetve a „kb.” adatokat csak fejben számolással, kerekített adatokkal próbáljuk meghatározni!)

Az alapfeladat: Normál testsúlyúak és erősen elhízotta vérnyomását vizsgáltuk. Az adatok (n, átlag, szórás) (1 tizedesre kerekítve) a diasztolés vérnyomásra: 24; 77,9; 10,4 (normál csoport) és: 33; 90,2; 13,2 (kövér csoport) (SPSS-nél a pl. sav tartalmazza az adatokat)

1. Lehet-e a normál testalkatúak vérnyomása, a sokaság várható értéke (), az „irodalmai” 80 Hgmm?
   Mi a H₀? És mi legyen a döntési szint, az ??
   Mennyire tér el a mintaátlag a hipotetikus várható értékől? Mekkora a t érték? Milyen egységben mérve?
   (Mekkora a standard hiba?) továbbá: Mekkora a p és mi ennek a jelentése?
   Hogy döntünk és mely számadat(ok) alapján?
   Mit jelent ez a döntés a feladat eredeti kérdését illetően?

2. Lehet-e a kövérek vérnyomása, a sokaság várható értéke, a az egészséges „irodalmai” 80Hgmm?
   Mi a H₀? És mi legyen a döntési szint, az ??
   Mennyire tér el a mintaátlag a hipotetikus várható értékől? Mekkora a t érték? Milyen egységben mérve?
   (Mekkora a standard hiba?) továbbá: Mekkora a p és mi ennek a jelentése?
   Hogy döntünk és mely számadat(ok) alapján?
   Mit jelent ez a döntés a feladat eredeti kérdését illetően?

3. Eltér-e egymástól az elsőéves hallgatók jobb- és balkéz szorító ereje? (a „kezero.sav” adatait értékeljük az SPSS programmal!)
   Hány eredeti adatsorunk van? Mire vonatkozik valójában a kérdés? (különbség!) Milyen az eredeti (jobb és balkéz) adatok eloszlása? És milyen a különbség-adatok eloszlása? Milyen módszert választunk?
   Mi a H₀? És mi legyen a döntési szint, az ??
   Mennyire tér el a mintaátlag a hipotetikus várható értékől? Mekkora a t érték? Milyen egységben mérve?
   (Mekkora a standard hiba?) továbbá: Mekkora a p és mi ennek a jelentése?
   Hogy döntünk és mely számadat(ok) alapján?
   Mit jelent ez a döntés a feladat eredeti kérdését illetően?

4. Ismételjük meg a 3. feladat értékelését a különbség (’diff’) változó adatait használva!
   Hány adatsorunk van most? Mire vonatkozik a kérdés? Milyen módszert választunk?
   Mi a H₀? (1) És mi legyen a döntési szint, az ??
   Mennyire tér el a mintaátlag a hipotetikus várható értékől? Mekkora a t érték? Milyen egységben mérve?
   (Mekkora a standard hiba?) továbbá: Mekkora a p és mi ennek a jelentése?
   Hogy döntünk és mely számadat(ok) alapján?
   Mit jelent ez a döntés a feladat eredeti kérdését illetően?

5. Milyen feltétel teljesülése mellett jogos az 1-4 kérdésekben használt t próbák alkalmazása?
   Milyen meggondolás alapján tételezhetjük fel, hogy ez teljesül?
Mi van, ha mégsem teljesül? Milyen módszer jön szóba az értékelésre ekkor?

6. Ismételjük meg a 3. feladat értékelését az előjelpróbával!

Hány adatsorunk van most? Mire vonatkozik a kérdés? Miért és milyen módszert választunk?

Mi a \( H_0 \)? És mi legyen a döntési szint, az ??

Milyen eltérés vizsgálata alapján dönthetünk e módszernél? Mekkora az eltérés itt? Milyen eloszlás alapján értékelhető ez az eltérés? továbbá: Hogyan számolható és mekkora a \( p \) értéke? Mi ennek a jelentése?

Hogy dönthünk és mely számadat(ok) alapján?

Mit jelent ez a döntés a feladat eredeti kérdését illetően?
7. fejezet - A két módszer: hasonlóságok és különbözőségek

Potó lászló dr.

1. A fejezet tanulmányozásához szükséges előismeretek

- Változók típusai, a folytonos változók A minta és az alapsokaság
  - Hisztogram, relatív gyakoriság sűrűség, sűrűségfüggvény
    - A valószínűség mérőszáma: görbe alatti terület p(a,b) (Ellenőrző kérdések)
  - A minta jellemzése, a sokaság paraméterei
    - Átlag és szórás, a várható érték és a szórás (Ellenőrző kérdések)
- A normális eloszlás és jellemzői (egyebek mellett: 68-95-99,7%)
  - Mintapélda: hallgatók testmagassága
    - minta mérete és várható terjedelme. (a szórás „működése”, a „3 szabály”)
  - A std normális eloszlás és értelmezése (Ellenőrző kérdések)
  - A normális eloszlás kiemelt jelentősége az orvostan hallgatók számára
    - Az átlag eloszlása
    - Mintapélda: 16 fős csoportok átlagmagasságának eloszlása (Ellenőrző kérdések)
- A megbízhatósági intervallum a várható értékre
  - A statisztikai becslés és fajtái: pont és intervallumbecslés
    - A megbízhatóság értelmezése az intervallumbecslésnél (Ellenőrző kérdések)
  - A -revonatkozó
    - ±t*sx/n intervallumbecslés
    - A t eloszlások és eltérésük a std normáltól
    - Mintapélda: 16 tabletta hatóanyagtartalma – mennyi lehet a ? (Ellenőrző kérdések)
- A megbízhatósági intervallum a várható értékben
  - A statisztikai becslés és fajtái: pont és intervallumbecslés
    - A megbízhatóság értelmezése az intervallumbecslésnél (Ellenőrző kérdések)
  - A -re vonatkozó ±t*sx/n intervallumbecslés
    - A t eloszlások és eltérésük a std normáltól
    - Mintapélda: 16 tabletta hatóanyagtartalma – mennyi lehet a ? (Ellenőrző kérdések)
- A hipotézisvizsgálat
A két módszer: hasonlóságok és különbözőségek

1. A hipotézisvizsgálat alap gondolatmenete
   - „hétkőznapi” döntési helyzetekben
   - adatok értékelése esetén? (Ellenőrző kérdések)

2. A kivitelezés 5 alaplépése
   - A $H_0$ és megfogalmazásának alapelve
   - A döntési szint és szerepe a döntésben
   - A számolás (mit is kell kiszámolni?) és a döntés (mi a döntés logikája).
   - A következtetés (Ellenőrző kérdések)

3. Mintapélda az egymintás t próbára
4. Mintapélda a páros t próbára
5. Mintapélda az előjel próbára és: (Ellenőrző kérdések)

2. A várható érték megbízhatósági intervalluma: az intervallumbecslés felidézése

2.1. A hipotézisvizsgálat alapgondolatának felidézése

A hipotézisTesztelés gondolatmenete naponta többször is alkalmazott döntéshozó séma! Eszerint:

Van egy kérdés, s erre válaszként gondolatban (rendszerint öntudatlanul) megfogalmazunk egy hipotézist, egy hipotetikus választ („alaphelyzet”).

utána megnézzük a tényeket, s mérlegeljük, hogy a tények milyen valószínűséggel lehetnek a hipotetikus „alaphelyzet” következményei.

Ez a valószínűség lehet „nagy” avagy „kicsi”.

És ez alapján döntünk:

Ha a mérlegelésben a valószínűség nagynak adódott, akkor nincs okunk elvetni a hipotézist, hiszen a tények valószínű következményei lehetnek a hipotézisnek, tehát nincs ellentmondás. Ha viszont kicsi ez a valószínűség – vagyis: „kis valószínűségű/valószínűtlen”, hogy a tények a hipotézisünk következményeiként állhattak elő, tehát a tények ellentmondani látszanak a hipotézisünknek: ezért elvetjük a hipotézist.

Emlékezzünk, hogy az ennek kapcsán bemutatott példa – az „éjszaka esett-e az eső, vagy nem?” – arra is alkalmas volt, hogy ellenőrizzük, mi lesz, ha az ellentett hipotézisből indulunk ki. Azt találtuk, hogy ez az ötlépéses gondolatmenet – függetlenül a kiinduló hipotézistől – működőképes. Majd megnéztük ugyanezt a 16 tabletta hatóanyagtartalma példára. A kérdés az, hogy lehet-e a várható érték a névleges 100 mg? És az ismerős őt lépés erre is működött.

3. Az egymintás t próba és a megbízhatósági intervallum

3.1. Mi ezen két gondolatmenet egymás való viszonya?

A megbízhatósági intervallum esetén van egy tartomány, amelyen belül lévő értékek a „lehetséges” várható értékek, a kívül levők pedig a „nem lehetségesek”. (Természetesen adott hibakockázattal: ha az intervallum 95% volt, akkor ez 5%.)

Ezzel szemben hipotézis tesztesztesetén (az előbbi egymintás t próba példáján):
Ha nagy a távolság, vagyis t nagy, a hozzá tartozó valószínűség (a „p”) kicsi, akkor elvetjük a H₀-t. Ha a t nem nagy, a hozzá tartozó valószínűség nem kicsi, akkor nincs okunk elvetni a H₀-t. Mindkét előbbi esetben a p valószínűséget az ?-hoz, a döntési szintezhez viszonyítjuk.

Hozzuk össze ezt a két kritériumot: A „nagy távolság, kis valószínűség” és a „kis távolság nagy valószínűség” hogyan illeszkedik a „belül” vagy „kívül” fogalom-párral? Nézzük ezt a 7.16. ábrán

Tehát egy belső hipotetikus µ-re nincs okunk elvetni, egy külsőre pedig elvetjük a H₀-t.

4. Az első- és a másodfajú hiba

Vegyük azt észre, hogy egy külső hipotetikus µ esetén a nagy távolság nem jelenti, hogy kizárt egy (legalább ilyen) nagy eltérés előfordulása! Csak azt jelenti, hogy ennek a valószínűsége kisebb, mint az ?, konkrétan éppen p. Tehát, ha mi elutasítjuk a null hipotézist, akkor p valószínűségével bizony rossz döntést hozunk! Ez a p ugyan kicsi, de nem nulla.

Annak esélyét, hogy rossz döntést hozunk a H₀ elutasítása esetén, azt hívjuk az „elsőfajú hiba” kockázatának. Amelynek a valószínűsége tehát kicsi, de nem nulla, hanem éppen a p.

Egy belső hipotetikus µ-re, amely kis távolságot, tehát kis t értéket jelent, ennek megfelelően nagy p értékel jár, készítsünk egy következő ábrát: ahol a és b az -tól r*sx/n-re van (7.17. ábra).

Mi erre az épp megvizsgált közeli µ-ra mondjuk azt, hogy ez egy lehetséges µ. És, ha az ettől eltérő, de még a megbízhatósági intervallumon belüli másik hipotetikus µ-t nézneink? Arra is azt mondánk, hogy lehetséges érték. És egy harmadikra, a -t; Arra is! Egy negyedikre? Arra is! Tehát, ha körülf eltelezett µ-re az mondhatjuk, hogy lehetséges µ érték, miért pont az az egy, amit az imént megvizsgáltunk, a µ, lenne az igazi?

Tehát, ha nem vetjük el a H₀-t, ugyancsak hibás döntés lehet: Ezt hívjuk másodfajú hibának.

Mit vettünk észre tehát? Bárhogyan is döntünk, mindkét döntésünk esetén hibákockázattal jár a döntés. Ha elvetjük a H₀-t, az elsőfajú hibát kockázatjuk, ha megtartjuk, akkor másodfajú hibát kockázatjuk.

4.1. Hogyan csökkenthetjük ezt a kockázatot?

Az elsőfajú hibanál úgy, hogy csökkentjük az ?-t. Miért?

Az ? csökkentésével egyre növeljük a „megtartjuk a H₀-t” (= „belül”) intervallumot. Így egyre kevesebb H₀-at utasítunk el, hiszen az elutasításhoz vezető p értékek tartománya csökkenni fog. Mír értünk ezen? Ha pl. 5%-ról 1%-ra csökkentem az ?-t akkor p = 4% ... 3% ... 2% esetén is „nagy”-nak minősül ezentől a p valószínűség, tehát nem fogjuk elvetni a H₀-t.

Magyarul: csökkenthetjük ezt a kockázatot? Az? csökkentésével csökkenthetjük az elsőfajú hiba kockázatát, ami maga a p (és aminek a maximális értéke épp az ?). Ezzel azonban egyre több H₀-at megtartunk, hiszen a „megtartjuk a H₀-t” (= „belül”) intervallumot növeltük! De ezzel a megnövelt intervallummal megváltozik a másodfajú hiba valószínűsége! Magyarul, csökken a p értékei és a H₀-at megtartunk, a? csökken a másodfajú hiba valószínűség.

Ellentét esetben, a másodfajú hibanál próbáljuk meg növelni az ?-t! Mivel jár ez?

Egyre csökkentjük, szűkítjük a belső részt, egyre csökkentjük a „megtartjuk a H₀-t intervallumot”. Egyre kevesebb H₀-at tartunk meg, hiszen az ehhez tartozó p értékek tartománya csökken. Pl. ha 5%-ről 10%-ra növeljük az ?-t, akkor p = 6 ... 7 ... 8 ... 9% esetén is „kicsi”-nek minősül ezentől a p, hiszen a határ az a 10%! Magyarul elvetjük a H₀-t ezen értékekre is.

Vagyis a „megtartjuk a H₀-t” intervallum csökkentésével csökkenni fog a másodfajú hiba valószínűségét. Ezzel együtt azonban egyre több H₀-at utasítunk el, hisz a „külső” (elutasítom a H₀-t) tartomány nő! Vagyis ugyanezzel
egyre növeljük az elsőfajú hiba valószínűségét, ami maga a p (és az elutasításhoz vezető p értékek tartománya növekedett).

Összegezve: bármelyik hibakockázatot akarjuk csökkenteni, ez tehetséges az ? változtatásával! Azonban a másik hibafajta kockázata ezzel együtt növekedni fog. De, ha bármelyik hiba kockázatát csak a másik rovására tudjuk csökkenteni ? változtatásával, akkor: mi az optimális ??

4.2. Mi az optimális ? (döntési szint)?

4.2.1. Egy hétköznapi példa: a bírósági ítélet

Ha például bíróság valamely bűnügyben ítélt, akkor a jelenlegi gyakorlat szerinti alapállás az, hogy ártatlan a gyakorlásosított. (A nullhipotézis az ártatlanság vélelme!) Próbáljuk meg most erre a helyzetre alkalmazni az előbbi kétféle hibacskökkentő törekvést.

Az első törekvés legyen az, hogy „csak nehogy egy ártatlan börtönbe kerüljön”. Mit jelent ez? Ez azt jelenti, hogy egészsen addig megtartjuk az ártatlanság vélelmet, ameddig nincs valami nagyon nyomós okunk ennek elvetésére.

Ad abszurdum képzeljük el, hogy a bíró ezt a gyakorlatot folytatja: Csak akkor ítélik bűnügybe, ha azóta több bűnös is szabadon marad, és az azt jelenti, hogy megtartunk hibás nullhipotéziseket, magyarul nő a másodfajú hiba esélye. Erre egy közismert példa Amerikában a második világháború előtti időszak hírhedt maffia vezére, Al Capone esete. Mindenki tudta, hogy sok bűn terheli, de nem sikerült börtönbe zárni, csak egyszer, akkor is rövid időre: adócsalás miatt. Nem sikerült bebizonyítani, hogy a H0, vagyis az ártatlanság vélelme elvétendő.

A második lehetséges törekvés az lenne, hogy „csak nehogy egy bűnös szabadon maradjon”. Ez csak egy hipotetikus bíróságon működhet, hisz az ártatlanság vélelme a tényleges bírósági tárgyalások alapelve.

Ebben az esetben a nullhipotézis (az ártatlanság vélelme) elvetését preferáljuk! Ezt jelenti az ? növelése. Ez azonban azzal jár, hogy akire csak a gyanú árnyéka vetül azt elítéljük!

Ezzel csökkentjük a másodfajú hiba esélyét: egyre kevesebb bűnös üssza meg! Persze, így egyre több ártatlan is börtönbe kerül! Ez az ára annak, hogy nehogy véletlenül egy tényleges bűnös szabadon maradjon! Ezzel persze az elsőfajú hiba esélyét növeljük, vagyis, hogy elvetünk egy igaz nullhipotézist, tehát egy ártatlan mégis börtönbe kerül.

Nézzünk egy másik példát ugyanerre a két törekvéshöz, mely társ az orvosi gyakorlathoz közelebb áll.

4.2.2. Egy új gyógyszer bevezetése és mellékhatása. Az ? megválasztása

Tételezzük fel, hogy egy új gyógyszert vezetünk be. A nullhipotézis az, hogy ez a gyógyszer hatáslan. (Ugye, ez a negatív állítás. Emlékszünk-e még a 6. fejezetére?)

Az első törekvés az, hogy „csak nehogy egy hatástalan gyógyszert gyártásba vonjunk”. Ugyan miért? Mert nagyon drága egy új gyógyszer gyártását beindítani, drága azt megfelelő reklámmal a piacra bevezetni és így tovább. Pl., ha ez a gyógyszer már a harmadik oklevelen fokat fajta a piac, mert egy új vényomáscsökkentő vezetnénk be, amivel már így is tele vannak a gyógyszertárak: több különböző gyár több különböző készítménye kapható. Ekkor igen meggondolni, hogy ha csak ugyanolyan jó, de nem határozottan jobb gyógyszerünk van, mint a már piac levők, akkor miért vezessük be? Miért költsünk rá egy csomó pénzt, amikor nem jobb, csak épp ugyanolyan?

Ez esetben természetesen a „hatástalan” (= nem elég hatásos) ítéletet, a nullhipotézis megtartását kell preferálnunk! Magyarul, csökkenteni kell az ?.t: legyen mondjuk az ? = 0,01. Csak akkor mondjuk, hogy a gyógyszer hatásos, ha nagyobb megbízhatósággal jobb, mint a jelenleg piac levők. A kisebb ? miatt szélesebb lesz a C.I., nagyobb a megbízhatóság a „hol a ?” kérdésre.

93

De, egy gyógyszerrel lehet az is a kérdés, hogy egy bizonyos mellékhatás létezik-e vagy sem? Pl. ha a gyógyszerkísérletek során valaki arra panaszkozott, hogy fájt a hasa, akkor beirjuk-e azt, hogy ennek a gyógyszernek egy lehetséges mellékhatása a hasfajás?

A nullhipotézis az, hogy nincs meg ez a mellékhatás. Itt az a kérdés, hogy mi a célszerű preferencia? Természetesen az, hogy „csak nehogy egy mellékhatás rejthet maradjon”. ír? Mert, ha nem írjuk fel, akkor ez a fogyasztók bizalomvesztésével járhat. Esetleg valaki az egyébként meglévő mellékhatást tapasztalva bíróság elé viszi a gyártót és a gyógyszergyár egy csomó pénzt kell fizetnie kárterítésként, mert bebizonyosodhat, hogy ez a mellékhatás mégis létezett, de nem sorolták fel a tájékoztatóban. Ezzel szemben mi történik, ha ott apró betűvel az amúgy is meglévő négy, öt, hat, nyolc mellékhatás mellé még odakerül pluszként a hasfajás. Legtöbben rendszerint úgysem olvassák el tehát ebben az esetben inkább kell vállalni annak a kockázatát, hogy egy akár nem is létező mellékhatást mellékhatásként tüntetünk fel!

Magyarul, a nullhipotézis elvetését kell preferálnunk! Ezt tehetjük az ? növelésével. Legyen tehát az ? = 0,1. Ezzel persze az elsőfajú hiba esélye nő; annak kockázata, hogy elvetünk egy igaz nullhipotézist: tehát azt kockáztatjuk, hogy egy nem létező mellékhatást mellékhatásként tüntetünk fel.

Ezen két gondolatmenet érvényessége között természetesen van egy „határsáv”.

Nem azt mondjuk tehát, hogy mindig ? = 5%-ot kell választanunk! Hanem azt mondjuk, hogy ha egyik előző gondolatmenetünk sem indokolt (tehát nem indokolt sem a H₀-nak a védelme az elvetés kárára, sem a H₁ elvetése a H₀ védelme kárára – vagyis a határsában, a „senki földjén” érezzük magunkat), abban az esetben egy arany középütnak tekintjük az 5%-ot. Tehát nem mindig, hanem csak akkor, ha nincs határozott preferenciánk az előbbi két gondolatmenet között.

5. Összefoglalás

Ebben a fejezetben összevetettük az intervallum becslés módszerét a hipotézis vizsgálatával. Konkrétan a várható érték megbízhatósági intervalluma és az egymintás t próba összevetésével megállapítottuk:

1. Az intervallumon kívül elutasítunk egy hipotetikus µ-t belül pedig nem utasítunk el.

2. Mindkét döntés hibakockázattal jár, bevezettük az első, illetve másodfajú hiba fogalmát, illetve megismertük, hogy melyik döntés mely hibának a kockázatával jár. Ez egy nagyon fontos gondolat! Emlékezzünk arra, hogy az orvos mindig információhiányos helyzetben dönt, tudnia kell tehát, hogy nincs egyszer és mindenkorra jó döntés: Minden döntését hibakockázattal hozza. Az egyetlen tárgy a biometria, amely nemcsak hogy felhívja erre a figyelmet, de be is bizonyítja, hogy ez törvény szerű, hogy így van. Továbbá az egyetlen olyan tárgy, amely megmutatja, hogy vannak egzakt, mérhető, elemezhető eljárások, módszerek, amelyek egészéig az ilyen döntési helyzetek kezelhetők.


4. Ez a választás nem tetszőleges, a feladattól függően vagy növelni vagy csökkenteni kell, és az a szokásos 5% csak egy arany középű: Ha nincs határozott preferenciánk, akkor választjuk.

6. Ismeret-ellenőrző és gyakorló kérdések és feladatok

(A becslést illetve a „kb.” adatokat csak fejben számolással, kerekített adatokkal próbáljuk meghatározni!)

Az alapfeladat:

Normál testsúlyúak és erősen elhízottak vérnyomását hasonlítjuk össze. Az adatok (n, átlag, szórás) (1 tizedesre kerekítve) a szisztolés vérnyomásra: 24; 125,1; 17,1 (normál, 1-es csoport) 33; 147,2; 23,3 (kövér, 2-es csoport)
A két módszer: hasonlóságok és különbözőségek

A megbízhatósági tartomány (C.I.):

1. Kb. mekkora az 1-es standard hibája?
   
   A becslés egy lehetséges gondolatmenete:
   
   Ha \( n = 25 \) lenne, akkor a \( n = 5 \) lenne, tehát \( n = 24 \) -re a \( n \) kicsivel kisebb, mint 5.
   
   A \( 3 \times n \) tehát valamivel 15 alatt lenne, a \( 4 \times n \) valamivel 20 alatt. Így a 17,1-es szórás kb. 3,5\( \times n \), tehát SE3,5.

2. Kb. mekkora a 2-es standard hibája? (Becsüljük meg az előbbi példa mintájára!)

3. Az 1, vagy 2, 95%-os C.I-a szélesebb – vagy talán egyformák?
   
   Ha egyformák, miért? (mi egyforma, ami miatt a C.I. egyforma lesz?)
   
   Ha különbözők, miért? (mi különböző, ami miatt a C.I. különböző lesz?)

4. Kb. mekkora az 1-es 95% C.I. a v.é-re? (t értékét csak becsüljük meg 1 tizedes pontossággal)

5. Lehet-e a várható értéke (95%megebizh.) 120?
   
   A. És 119? Vagy 122, 124, 126, 128, 130?
   
   
   C. Na, és mi a helyzet a 118-al? (Vigyázat, még mindig csak a becslésnél, a fejben számolásnál tartunk!)

6. Ha 99% megbízhatóságot kívánunk, csökken vagy nő-e ez az intervallum?
   
   Lehet-e ekkor a várható érték 118? És 117? (Mi volt a helyzet a 117-el 95% megbízhatóságnál?)

7. A 4, 5, 6, feladatot végezzük el a 2-es mintára is.
   
   Ez esetben természetesen más számértékek lesznek „érdekesek” az 5, és 6, feladatban: 5/a. 140-154 közti számok 5, B. 138 és alatta, 156 és fölötte. 5, C. 139?

8. A várható érték 95% C.I. mennyi pontosan, táblázatból kikerestett t értékkel számolva?
   
   Mennyi az 1-es, és mennyi a 2-es minta esetén?
   
   Van-e olyan érték, ami közös várható érték lehet mindkét mintára?
   
   Mire következtetünk ebből: van-e különbség a normál testsúlyúak és a kövérek szisztolés vérnyomása között?

A hipotézisvizsgálat

Az alábbi kérdésekhez használjuk az SPSS programot, a pl. sav adathalmazt. A kérdéses csoportokat a Q4 változó 2 (normál) és 4 (kövér) értéke adja meg.

1. Készítsünk egy számegyenest, s jelöljük be rajta az előbb vizsgált értékeket. Kezdjük az 1-es mintával, pl. a 110-140 értéktartomány feltüntetésével.
   
   Vizsgáljuk meg, lehetnek-e a értékei sorban az egyes (fenti) értékek:
   
   Írjuk fel minden megvizsgált érték alá az egymintás t próba által kiírt t és p értékeket!
   
   Mely értékekre utasítjuk el, s melyekre nincs okunk elutasítani a H0-t, az ? választott értéke alapján?
   
   A számegyenesen hol található a p = ? érték, s ez milyen t értékhez tartozik?
   
   Hány ilyen értéket adó hely van a számegyenesen?
Hol találhatók a pg? értékei és hol a pl? értékei a számegyenesen?

Mely tartományban fogadjuk el az értékeket H₀-ként? És mely tartományban utasítkuj el?

2. A 2, feladat részfeladatait ismételjük meg a 2-es mintára vonatkozóan is.

Az első és másodfajú hibák


Biztos-e az, hogy jól döntünk, ha eszerint döntünk?

Van-e ebben valamekkora tévedési kockázat? Ha nem, miért nincs? Ha igen, miért van? És mekkora ez a kockázat? Mi a neve ennek a hibának?


Biztos-e az, hogy jól döntünk, ha eszerint döntünk?

Van-e ebben valamekkora tévedési kockázat? Ha nem, miért nincs? Ha igen, miért van? És mekkora ez a kockázat? Mi a neve ennek a hibának?


5. Ha az ? értékét 0,05-ről 0,01-re változtatjuk, változik-e ezzel a döntési hibakockázat?

Mekkora volt a hibakockázat a 116-os feltételezett várható érték elvetésekor?

Mi változik, ha az ? értékét 0,05-ről 0,01-re változtatjuk?

Mi a helyzet a 117-es érték esetén ugyanezzel?

6. Ha az ? értékét 0,05-ről 0,1 -re változtatjuk, változik-e ezzel a döntési hibakockázat?

Mekkora volt a hibakockázat a 120-as feltételezett várható érték megtartásakor?

Mi változik, ha az ? értékét 0,05-ről 0,1-re változtatjuk?

Mi a helyzet a 122-es érték esetén ugyanezzel?

7. Az 5, és 6, kérdések megoldása alapján magyarázzuk el:

A H₀ elvetése vagy megtartása irányába tolja el a mérlegelést az ? növelése?

A H₀ elvetése vagy megtartása irányába tolja el a mérlegelést az ? csökkentése?

8. Mi változik, ha az ? értékét 0,05-ről 0,1-re változtatjuk?

Mi változik, ha az ? értékét 0,1-ről 0,05-re változtatjuk?

Mi a helyzet a 122-es érték esetén ugyanezzel?

9. Mikor indokolt – a 16., kérdés válasza alapján – a H₀ megtartását preferálni?

Tudunk-e a 7. fejezet példái alapján újabb példákat adni, amikor egy döntés kétféle hibakockázata (ha hibás az igen válasz és, ha hibás a nem válasz) következményei nem egyforma súlyúak, s ezért szándékosan valamilyen javára „súlyoznunk” kell a mérlegelésben?

Pár hétköznapi példa ötletadónak: Ha átmegyek a befagyott víz jegén, nyerek egy fél órát (fél napot). Ha beszakad a jég, mit veszek…? Ha ittasan/fáradtán autóval én vezetek hazáig, nyerek pár órát, s megtakarítom a buszjegy árát. És, ha balesetet okozok…? Ha rohanok a buszmegállóig, elérhetem a következő buszt, s 1 órával hamarabb odaértek… (hova is??), de ha rohanok, megízzadok. Mitől függ itt, hogy mi a jó döntés? Kvízjátékban dupla vagy semmi szabály mellett választhatok 2 válasz közül, de fogalmaim sincs, melyik a jó válasz…, vagy befejezve a játékot elvihetem, amit eddig elértem. Mi a jó…
döntés, ha egy 32 millió Ft csúcsonyeremjegy tartó játékban ez a játék elején (pl. 4000 Ft-os) tétnel történik, s mi, ha 8 milliónál a következő duplázás lehetőségénél tartunk? Az ismerős játsszőr melletti forgalmazás útra ritkán (1000… 10000 autó közül 1 élé) kigurulhat a labda, s rendszerint követi egy rohanó kisgyerek. Sietnem kell, a „p” értéke nagyon kicsi. Lassítsuk, vagy bizszzák abban, hogy nem most realizálódik az a kis valószínűség?

10. Mikor indokolt – a 16, kérdés válasza alapján – a \( H_0 \) elvetését preferálni?

Gondoljunk meg válaszunkat és keressünk további példákat a 17, kérdésnél felsoroltak alapján.

11. Keressünk orvosi vonatkozású példákat a 9, és 10, kérdés esetére! Gondolatserkentő ötletek:

Ultrahanggal vizsgálunk vesetumor gyanúval vizsgálatra küldött pácienset. A lelet negatív. Két eset lehetséges: 1, Valóban nincs tumor, s jogosan küldjük haza „nincs probléma” eredménnyel a pácienset. 2, A vizsgálat bizonytalansága érvényesült, s a negatív lelet csak „másodfajú hiba”: Van tumor, de a vizsgálat nem adott elég bizonyítékot, hogy elveszítük a nullhipotézist. Ez utóbbi esetben nyilvánvaló feladatunk lenne további vizsgálatokat végezni. De ez drága, s a páciens is további tortúrákön (esetleg igen kellemetlen, sőt fájdalmas) vizsgálatoknak teszi ki. Mi a súlyosabb hiba:

Ha elfogadjuk a negatív leletet, de az hibás, s a tumort később már csak egy későt, nem kezelhető állapotban ismerjük fel? Vagy: Ha kételkedünk a vizsgálatban, pedig valójában az helyes, s mi mégis feleslegesen további kellemetlen és drága vizsgálatokat végezünk el?

Ennek alapján melyik irányban kell a döntési szintünket módosítani az eredeti \( UH \)-vizsgálat értékelésekor?

Influenzás időszakban tipikus influenzás tünetekkel jelentkezik egy páciens a rendelőnkben. E helyzetben tekintsünk két mérlegelhető kérdést:

1. Influenzája van-e a betegnek? (tények: influenzás tünetek, influenzás időszak)

2. Valamilyen ’x’ veszélyes trópusi betegsége van, ami ugyanezen tünetekkel (is) járhat, de nálunk csak olyanoknál fordulhat elő, akik a trópusokról térek haza (tehát: tünetek vannak, de szinte kizárt a betegség).

Ha csak az első kérdést mérlegeljük, hazaküldjük a beteget „ágynyugalom, vitamin, folyadék a láz megszűnése után még 2 napig” terápiás javaslattal. De, ha tévedünk, s mégis ’x’ (vagy ’y’, esetleg ’z’, stb.) betegsége van, amit nem is vizsgáltunk azon az alapon, hogy ezek mind ritkák…?

Gondoljuk végig ezt a problémakört az első- és másodfajú hiba, és az ? változtatása szempontjából!

Az iménti példák igazolják, hogy az orvos mindennapi munkájában alapvető döntési helyzetek háttérében a hipotézisvizsgálat elemi lépései és gondolatmenete áll.

Továbbá, hogy a döntési hibák és kockázatuk mérlegelése alapvető feltétele, hogy a helyzetnek megfelelő legjobb döntést tudja az orvos meghozni.

Ez már önmagában elégőséges ok, hogy bizonyítsa: az orvosképzésben alapvető jelentőségű a statisztikai döntéshozás módszereinek megismerése, a gondolkodási készségek ezirányú fejlesztése.
8. fejezet - Csoportok összehasonlítása folytonos eloszlásból származó adatok esetén

Boda Krisztina dr.

1. A fejezet tanulmányozásához szükséges előismeretek

Folytonos változó, folytonos eloszlás

Statisztikai minta

Normális eloszlás
t-eloszlás

Nullhipotézis

Szignifikancia

2. Bevezetés

2.1. Összetartozó és független minták

Két folytonos eloszlásból származó minta összehasonlításához a megfelelő módszer megválasztása függ a kísérleti elrendezéstől és az adatok eloszlásától.

A kísérleti elrendezéstől függően kétféle – összetartozó vagy független – mintákat kapunk. Ha ugyanazokról a kísérleti egyedekről gyűjtünk információt kétszer egymás után, (általában valamely beavatkozás, kezelés előtt és után), az ilyen módon végzett kísérletet önkontrollós kíséreltnek nevezzük, a kapott minták összetartozó vagy párosított adatok. Ugyancsak összetartozó minták keletkezhetnek akkor is, ha nem ugyanazok az egyedek, de a kísérlet szempontjából párosíthatók az adataik. Az így kapott adatokat illesztett pároknak (matched pairs) nevezzük. Egy másik kísérleti terv szerint a kísérleti egyedeket két kezelés mellett megfigyeljük, és az így kapott minták függetlenek, mivel az egyes csoportokban különböző egyedek szerepelnek. Egy adott adathalmazban független csoportokat képeznek a vizsgált tényezőnek egy másik kategorikus változó (pl. nem) szerinti csoportjai. A független minták elemszáma nem szükségségesen egyenlő.

2.2. Paraméteres és nemparaméteres próbák

Az adatok eloszlása is meghatározza a módszerválasztást. Az ún. paraméteres módszerek alkalmazásának feltétele a normalitás: önkontrollós adatok esetén a különbségek, független minták esetén a csoportok normális eloszlású sokaságból származnak. Mivel tudjuk, hogy a normális eloszlást a két paramétere (, ) egyértelműen meghatározza, ezért a két minta eloszlásának összehasonlítása a paraméterek normális eloszlásával vezethető vissza, innen az elnevezés.

Ha a normalitás feltétele nem teljesül, gyakran egy megfelelően választott transzformációval normalizálhatjuk adatainkat, és a transzformált adatokon végezzük el az összehasonlítást. Másik lehetőség az ún. nemparaméteres próbák alkalmazása, amelyeknek nem feltétele a normalitás. Ezek a próbák a paraméteres próbákkal szemben nem feltételezik az eloszlás paramétereinek megismerhetőségét, sőt gyakran azok létezését is.

A nemparaméteres próbák nagy része a rangsorolásán alapul. Ezek a próbák az adatok tényleges számtétele helyett csak azok egymáshoz viszonyított nagyságrendjét veszik figyelembe. A rangsorolásán alapuló próbák még olyan típusú adatokra is alkalmazhatóak, amelyeknek az értékeit nem lehet pontos számmal jellemezni, csak...
Csoportok összehasonlítása
folytonos eloszlásból származó
datok esetén


A választás a paraméteres és nemparaméteres módszerek között nem mindig egyértelmű. A normalitást sokszor
vagy nem tudjuk ellenőrizni (pl. kis elemszám miatt), vagy nem is akarjuk. Ugyanakkor, az itt tárgyalt
paraméteres próbák robusztusak – azaz, feltételeik kisebb-nagyobb megsértése mellett az eredményeik
érvényesek maradnak.

3. Két összetartozó minta összehasonlítása

3.1. Példa

Egy speciális diéta hatásosságát szeretnénk megvizsgálni önkontrollos kísérlettel. Minden vizsgálati személy
testsúlyát megmérjük a diéta előtt és után. A hipotetikus kísérlet eredménye 10 kísérleti személyen a 8.1. táblázatban látható.

8.1. táblázat - Egy diéta-kísérlet adatai

<table>
<thead>
<tr>
<th>Vizsgált személy</th>
<th>Testsúly a diéta előtt</th>
<th>Testsúly a diéta után</th>
<th>Különbség</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>95</td>
<td>90</td>
<td>5</td>
</tr>
<tr>
<td>3.</td>
<td>75</td>
<td>72</td>
<td>3</td>
</tr>
<tr>
<td>4.</td>
<td>110</td>
<td>100</td>
<td>10</td>
</tr>
<tr>
<td>5.</td>
<td>81</td>
<td>75</td>
<td>6</td>
</tr>
<tr>
<td>6.</td>
<td>92</td>
<td>88</td>
<td>4</td>
</tr>
<tr>
<td>7.</td>
<td>83</td>
<td>83</td>
<td>0</td>
</tr>
<tr>
<td>8.</td>
<td>94</td>
<td>93</td>
<td>1</td>
</tr>
<tr>
<td>9.</td>
<td>88</td>
<td>82</td>
<td>6</td>
</tr>
<tr>
<td>10.</td>
<td>105</td>
<td>99</td>
<td>6</td>
</tr>
<tr>
<td>Átlag</td>
<td>90,8</td>
<td>86,8</td>
<td>4</td>
</tr>
<tr>
<td>SD</td>
<td>10,79</td>
<td>9,25</td>
<td>333,3</td>
</tr>
</tbody>
</table>

Az átlagok között különbség eléggé nagyknak látszik. Igen ám, de viszonylag kevés számú adatunk van egy elvileg végötem nagy sokaságból. Lehetséges, hogy pusztán véletlenül kaptunk ekkora a különbséget? Megismételve a kísérletet másik 10 emberen, jaj, ha azok hasonló lesz-e az átlagos változás, vagy esetleg sokkal kisebb, akár 0? Következtetéseinket nem a vizsgált 10 emberre, hanem a sokaságra, például általában a diétát alkalmazó emberekre szeretnénk lehittni. Ha a diéta nem hatásos, akkor ez azt jelenti, hogy a második mérés nagyon hasonlít az elsőhöz, vagyis csak véletlenszerű különbségek vannak a mérések között, azaz – mindkét minta ugyanazon sokaságból vett véletlen mintának tekinthető. Ekkor az átlagos különbség is kicsi, a sokaságot tekintve az átlagos különbség nullával egyenlő. Ha a diéta hatásos, akkor sokaságot megkísérjük az átlagos különbség nullától eltérő számmal jellemezhető. Ez útóból két monddattal két állítást fogalmaztunk meg, melyek közül az egyik mellett döntünk a statisztikai próba során: (1) a sokaság különbség-átlaga nulla – ezt nevezzük nullhipotézisnek, és (2) a sokaság különbség-átlaga nem nulla – ezt nevezzük alternatív hipotézisnek. A döntéshez még meg kell adni, hogy mennyi ingadozást engedünk a véletlennek: ez tetszőleges valószínűség lehet, mégis leggyakrabban 5%-ot szoktak megadni, (szignifikanciaszint, ).

3.2. Paraméteres módszer – páros t-próba

A paraméteres módszerek feltételezik, hogy adataink normális eloszlású sokaságból származnak. Két normális
eloszlás összehasonlítása az átlagaik és a szórások összehasonlítását jelentik, azonos szórások esetén csupán az átlagokat kell összehasonlítani. A két csoportot összehasonlító paraméteres próbák (t-próbák) az átlagos
változást vizsgálják, nullhipotézisünk az, hogy a két vizsgált sokaság átlaga azonos (\(=\)). Önkontrollos kísérletek esetén az adatpárok különbségét véve egyetlen adatsort kapunk, a normalitás feltétele erre a különbség-adatomsorra vonatkozik. A módszer neve páros t-próba vagy régebbi nevére egymintás t-próba a különbségre. A módszert az előző fejezetben részletesen ismertettük, itt most csak röviden foglaljuk össze.
A nullhipotézis szerint a különbség-sokaság átlaga nulla (\(\mu_{\text{különbség}}=0\)), az alternatív hipotézis szerint nullától eltérő (\(\mu_{\text{különbség}}\neq 0\)). A próbát elvégezhetjük a különbség-átlagra vonatkozó adott megbízhatóságú konfencia-intervallum segítségével: ha a 0 beleszik a konfencia-intervallumban, az eltérés nem szignifikáns adott szinten, ha kívül esik, akkor az eltérés szignifikáns. A statisztikai szignifikancia megállapítását mégis leggyakrabban a különbség-minta átlagából képzett, \(t\)-vel jelölt mennyiség segítségével vizsgáljuk, amely még az elemcsömtől és a különbség szórásától is függ:

\[
t = \frac{\bar{d}}{s_d} \sqrt{n}
\]

ahol

- a különbségek átlaga,
- \(s_d\) a különbségek standard deviációja.

Bebízonyítható, hogy ha a normalitás feltétele teljesül és igaz a nullhipotézis, akkor a fenti mennyiség Student-\(t\)-eloszlást követ, \(n-1\) szabadságfokkal. Ha valóban nincs különbség az összehasonlítandó két csoport között, akkor a

- és így a \(t\) is kicsi lesz, közel nullához. A \(t\)-eloszlás táblázatából adott szignifikanciaszint és szabadságfok mellett meghatározható, hogy mi az a határ, ameddig „elfogadhatóan kicsi” a \(t\)-érték, az ún. kritikus \(t\)-érték (\(t_\alpha, n-1\)). Ha az általunk számolt \(t\)-érték abszolút értéke nagyobb, mint a kritikus érték,

\[
|t| > t_{\alpha/2, n-1}
\]

akkor az alternatív hipotézis mellett döntünk, azaz elvetjük a nullhipotézist, és azt mondjuk, hogy a különbség szignifikáns szinten, jelölése pl. Ha

\[
|t| < t_{\alpha/2, n-1}
\]

akkor a nullhipotézis mellett döntünk, és azt mondjuk, hogy a különbség nem szignifikáns szinten, pg.

Ma már a kritikus értékek táblázatból való kikeresése helyett inkább az ún. \(p\)-értéket használják a döntéshez, mivel ezt a legtöbb statisztikai rendszer kiszámítja. Ekkor a \(p\)-érték és összehasonlításával döntünk. A \(p\)-érték annak valószínűsége, hogy ha igaz a nullhipotézis, akkor legalább a most kapott vagy még nagyobb abszolút értékű \(t\)-értéket kapjunk.

\(Tárolt\) \(p\)-próba egyoldalas alternatív hipotézis esetén. Ha a nullhipotézissel szemben csak azt vizsgáljuk, hogy a változás pozitív-e, (vagy, más probléma esetén, negatív-e), akkor a kritikus értéket a \(t\)-eloszlás egyik szélén keressük, így adott esetén a kritikus értéket 2 „oszlopában” kell keresni (\(t_2, n-1\)), a \(p\)-érték pedig a kétoldalas \(p\)-értékének fele lesz. Ritkán alkalmazunk egyoldalas próbát, a hipotéziseket ugyanis a kísérlet elvégzése előtt állítjuk fel, amikor legtöbbször nem ismerjük a változás irányát.

### 3.2.1. A példafeladat megoldása páros \(t\)-próbával

\(H_0: \mu = 0\) (a sokaság különbség-átlaga 0)

\(H_1: \mu \neq 0\) (a sokaság különbség-átlaga 0-tól eltér, kétoldalas alternatív hipotézis)

Legyen \(\alpha = 0,05\).

Ekkor

\[
t = \frac{\bar{d}}{s_d} \sqrt{\frac{n}{4}} = \frac{3,333}{\sqrt{10}} = 3,795
\]

- a szabadságfok \(df = 10-1 = 9\), a kritikus érték \(t_{0.05,9} = 2,262\). Mivel 3,795<2,262, elvetjük a nullhipotézist, és azt mondjuk, hogy a különbség szignifikáns 5%-os szinten.

Statisztikai program használatával a 9-es szabadságfok esetén \(t = 3,795\)-höz tartozó kétoldalas \(p\)-érték \(p = 0,00425\) jóval kisebb, mint 0,05, tehát ugyanúgy szignifikáns a különbség. A
8.1. ábra a példa adatainak megfelelő t-eloszlást, kritikus értéket és p-értéket (fekete színű terület) mutatja be.

Egyoldalas alternatív hipotézis esetén csak azt vizsgáljuk, hogy a testsúly csökkent-e, azaz, a változást jelző különbség átlag a populációban pozitív-e. Ekkor 5%-os szignifikancia szint esetén a kritikus t-érték $t_{0.1,9} = 1.66$, ehhez hasonlítva a 3.795-öt, azt kapjuk, hogy a hatás szignifikáns. Az egyoldalas p-érték $p = 0.002125$. Az ábrán a kétoldali p-érték a két kis fekete terület nagysága együttesen az eloszlás két szélén.

A kétoldalas próbát a 95%-os konfidencia-intervallum alapján is elvégezhetjük. AZ intervallum számításához ugyanazt a táblázatbeli kritikus t-értéket használjuk, azaz, az átlaghoz hozzáadjuk és levonjuk a

$$ t_{a,\alpha/2} = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{3.3333}{\sqrt{10}} = 2.384 $$

mennyiséget. Az intervallum két végpontja 4,2384, a különbség-átlagra vonatkozó 95%-os megbízhatósági intervallum tehát (1,616, 6,384). Annak valószínűsége, hogy az átlagos fogyást jellemző sokaság-átlag ebben az intervallumban esik, 95%. Az intervallum azonban a nullhipotézis szerinti „0” átlagot nem tartalmazza. Ezért úgy döntünk, hogy az eltérés szignifikáns 5%-os szinten. Megjegyezzük, hogy a konfidencia-intervallumot – amellett, hogy döntésre is alkalmazható - elsősorban a „hatás” nagyságának becslésére használják.

### 3.3. Nemparaméteres módszerek, előjelpróba, Wilcoxon-féle előjeles rangpróba

A nemparaméteres módszerek esetén az adatpárok különbségeinek eloszlásáról nem feltételezzük az eloszlás normalitását, mivel vagy nem tudjuk ellenőrizni, vagy nincs értelme az ellenőrzésnek, pl. nyilvánvalóan ferde vagy ordinális adatok esetén.

Ebben a fejezetben a ragsoroláson alapuló nemparaméteres módszereket mutatjuk be, önkontrollos kísérlet esetén ilyen próbához az előjelpróba és a Wilcoxon-féle előjeles rangpróba. A nullhipotézise mindkettőnek az, hogy a két sokaság eloszlása megegyezik.

Az előjelpróba (lásd az előző fejezet) elvégzéséhez először képezzük a két minta különbségét, majd megszámoljuk a negatív és a pozitív különbségek számát (a nullákat kihagyjuk). Ha az eredeti két változó azonos eloszlású, akkor körülbelül azonos számú negatív és pozitív különbséget kapunk. A próbához létezik egy statisztikai táblázat, amelyben megtalálható, hogy adott mintaelmszám és esetén hány eltérés tekinthető szignifikánsnak (ennek kiszámítása a binomiális eloszlás alapján). Nagy mintaelmszám esetén $(g \geq 30)$, alkalmazható az a formula, amelyre már a normális eloszlás táblázata használható a p-érték megkeresésére. Az előjelpróbát egyszerűsége miatt általában gyors tájékozódás céljára használják.
A Wilcoxon-féle előjelű rangpróba nem csak az előjeleket, hanem a különbségek közötti nagyságrendeket is figyelembe veszi, így nagyobb erejű, mint az előjelpróba. Az előjelpróba helyett a következő: a mintaelemek közötti különbségeket rangsoroljuk az előjelktől függetlenül, az esetleges nullákat kihagyjuk.

A rangsorolást a következőképpen végezzük: az adatsort nagyság szerint sorba rendezzük, és a legkisebbnek adjuk az 1-es rangszámot, a következőnek a 2-est, stb. Összesen \( n \) rangszámot osztunk ki. Egyenlő adatok esetén is egyre növekvő rangszámot adunk, majd az egyenlő adatokhoz tartozó rangszámokat utólag korrigáljuk a megfelelő rangszámok átlagával (a korrigált rangszámokat kapcsolt rangszámoknak nevezzük). A rangsorolás helyességét úgy ellenőrizhetjük, hogy összeadjuk a kapott rangszámokat, ennek az összegnek meg kell egyeznie \( n(n+1)/2 \) (az első „n” egész szám összegével).

A próba során ezután külön összeadjuk a pozitív vagy a negatív különbségeket tartozó rangszámokat (sőt, elég csak az egyiket). Ha igaz a nullhipotézis és a két sokaság azonos eloszlású, akkor a pozitív és a negatív különbségekhez tartozó rangszámösszegek körülbelül egyformák lesznek. Minél nagyobb az eltérés valamelyik a két sokaság között, annál nagyobb lesz az eltérés a két rangszámösszeg között is. Megkora eltérést tekinthetünk még véletlenszerűnek? Kis mintaelemszám esetén (nl30 vagy nl50) táblázatok állnak rendelkezésre, amelyek adott -hoz megadják azt az intervallumot, amekkora rangszámösszeg még véletlen eltérésnek tekinthető. Nagy mintaelemszám esetére, vagy ha nagyon sok a kapcsolt rang, egy közelítően normális eloszlású statisztika segítségével a normális eloszlás alapján vizsgálható a szignifikancia:

\[
Z = \frac{\sum_{i=1}^{n} R_i}{\sqrt{\sum_{i=1}^{n} R_i^2}}
\]

Itt a számlálóban az összes előjelű rangsázm összege szerepel, a nevezőben pedig ezek négyzetösszegéből vont négyzetgyök. A számítógépes programrendszerek általában csak ezt a normális közelítésből származó \( p \)-értéket számítják ki még kis mintaelemszám esetén is, amikor pedig a közelítés nem túl jó.

### 3.3.1. A példafeladat kiértékelése előjelű rangpróbával

A különbségek rangsorolását a különbségek abszolút értékén végezzük (tehát az előjelktől függetlenül). A legkisebb szám az 1-es, ez kapja az 1-es rangszámot. Mivel két darab 1-es különbség van, a 1-es rangszámokat utólag korrigáljuk úgy, hogy minden két 1-esnek a 2-es rangot adjunk el. A kapott rangsorokat következőképpen elkészítjük:

### 8.2. táblázat - Az előjelű rangpróba elvégzése a 8.1 táblázat adatain.

<table>
<thead>
<tr>
<th>Vízsgált személy</th>
<th>A testsúlyok különbsége</th>
<th>Rangszámok előjeltől függetlenül</th>
<th>Kapcsolt rangok</th>
<th>Előjelű rangok négyzetei</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>3.</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>4.</td>
<td>10</td>
<td>9</td>
<td>9</td>
<td>81</td>
</tr>
<tr>
<td>5.</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>49</td>
</tr>
<tr>
<td>6.</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>7.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8.</td>
<td>1</td>
<td>2</td>
<td>1,5</td>
<td>1,5</td>
</tr>
<tr>
<td>9.</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>49</td>
</tr>
<tr>
<td>10.</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>49</td>
</tr>
<tr>
<td>Összeg</td>
<td></td>
<td></td>
<td></td>
<td>282,5</td>
</tr>
<tr>
<td>1.</td>
<td>-1</td>
<td>1</td>
<td>1,5</td>
<td>-1,5</td>
</tr>
</tbody>
</table>
Az egyenlő előjelhez tartozó rangokat összeadjuk. A pozitív rangok összege $R_+ = 43.5$, a negatívoké $R_- = 1.5$. A táblázathoz $p = 0.05$ és $n = 9$ esetén 5-40 intervallumba eső rangszámösszegek tekinthetők véletlennek. Mindkét rangszámösszeg ezen az intervallumon kívül esik, tehát szignifikáns különbséget kapunk 5%-os szinten. Az elemszám most kicsi, és csak 1 kapcsolt rang szerepel, mégis számítsuk ki a (8.2) képlettel a $z$-értéket: $z = 42/282.5 = 2.499$. Az ehhez tartozó $p$-érték a standard normális eloszlásból $p = 0.012$. Tehát a változás szignifikáns 5%-os szinten a nemparaméteres módszer alapján is.

4. Két független csoport összehasonlítása

4.1. Példa

A diéta hatóerejét most a következő kísérleti terv szerint vizsgáljuk. Az önkényt jelentkező pacienteket véletlenszerűen két csoportba sorolunk, az egyik csoport hagyományos étrendet követ (kontroll csoport), a másik csoport diétázik. A kúra hatóerejét a két csoportban mért testsúlyváltozást összehasonlítva lehet teszteni: mondhatjuk-e, hogy a diétázó csoport tagjai átlagosan többet fogynak, mint a hagyományos étrendet követők? A változásokat tartalmazó oszlop független mintát jelent, minden személy vagy az egyik, vagy a másik csoportba tartozik. Két csoport összehasonlításakor kedvező, ha az elemszámok megegyeznek, ám ez nem feltétele a próbta elvégzésének. A 8.3. táblázat - Diéta kísérlet adatai 8.3. táblázat tartalmazza egy elképzelt kísérlet során kapott testsúlyváltozásokat, amelyben a diétás csoportban 10, a kontroll csoportban 11 személyről vannak adataink.

8.3. táblázat - Diéta kísérlet adatai

<table>
<thead>
<tr>
<th>Vizsgálati csoport</th>
<th>Vizsgált személy</th>
<th>Testsúlyváltozás</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Átlag</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>3.333</td>
<td></td>
</tr>
<tr>
<td>Kontroll</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>-2</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>-2</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Átlag</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>2.145</td>
<td></td>
</tr>
<tr>
<td>Diéta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>
4.2. Paraméteres módszer – kétmintás t-próba

Azt vizsgáljuk, hogy a két független minta ugyanazon vagy két különböző sokaságból származó két véletlen minta. Paraméteres módszerek esetén feltételezzük, hogy mindkét minta normális eloszlású sokaságból származik. Mivel a normális eloszlást két paraméterre meghatározza, az eloszlás két paraméterét, az átlagokat vagy a szórásokat lehet összehasonlítani a két csoportban.

4.2.1. Átlagok összehasonlítása, kétmintás t-próba

A kétmintás t-próba a két átlagot hasonlítja össze, a kapott mintaátlagok segítségével a sokaság-átlagokra következtetünk. A nullhipotézise az, hogy a két sokaságnak, amelyekből a mintákat vettük, azonos az átlaga (kontroll = diéta), kétoldalas ellenhipotézise pedig az, hogy különbözők az átlagok (kontrolldiéta), egyoldalas ellenhipotézis szerint a kontroll sokaság átlaga nagyobb, mint a kezelt csoport átlaga (kontrollg2diéta).

A kétmintás t-próbának két „változata” van attól függően, hogy a varianciák egyenlők-e vagy sem. Először tegyük fel, hogy a két populációban a varianciák is azonosak. Jelölje m és n a két minta elemszámát,

\[ x \text{ és } y \]

a mintaátlagokat, és s, és s, a minta standard deviációkat. Ha a feltételek teljesülnek és a nullhipotézis igaz, akkor a következő képlettel kiszámítható mennyiség

\[ t = \frac{x - y}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \]

n+m-2 szabadságfokú t-eloszlást követ, ahol

\[ s_p^2 = \frac{(n-1) \cdot s^2_x + (m-1) \cdot s^2_y}{n + m - 2} \]

a közös variancia összevont (pooled) becslése. Ekkor a t-eloszlás táblázatából adott szignifikanciaszint és szabadságfok mellett meghatározható az a kritikus t-érték (t, n+m-2), amelyet az általunk számított t-értékkel összehasonlítva döntünk. Ha az általunk számolt t-érték abszolút értéke nagyobb, mint a táblabeli kritikus érték,

\[ |t| > t_{\alpha/2, n+m-2} \]

akkor az alternatív hipotézis mellett döntünk, azaz elvetjük a nullhipotézist, és azt mondjuk, hogy a különbség szignifikáns szinten, jelölése pl. Ha

\[ |t| < t_{\alpha/2, n+m-2} \]

akkor a nullhipotézis mellett döntünk, és azt mondjuk, hogy a különbség nem szignifikáns szinten, pg.

Különböző varianciák esetén más képletet és szabadságfokot kell számítani (Welch próba). Ekkor

\[ t = \frac{x - y}{\sqrt{s^2_x + s^2_y \left(1/n + 1/m\right)}} \]

a szabadságfok

\[ g^2 = \frac{(n-1) \cdot (m-1)}{s^2 \cdot (m-1) + (1-g^2) \cdot (n-1)} \]

ahol
Csoportok összehasonlítása  
Folytonos eloszlásból származó  
adatok esetén

4.2.2. A varianciák összehasonlítása

Annak eldöntéséhez, hogy a (8.3) és (8.5) képletek közül melyiket alkalmazzuk, össze kell hasonlítani a varianciákat. Ez egy újabb próba, amelyben azt a nullhipotézist teszteljük, hogy a két normális eloszlású sokaság azonos variancijáú (\(H_0: \sigma_1^2 = \sigma_2^2\)). Két variancia összehasonlítását kézi számolással egyszerűen elvégezhetjük. Ehhez a nagyobb varianciát osztjuk a kisebb varianciával, a kapott mennyiség \(F\)-eloszlású lesz: 

\[
F = \frac{v_1 s_1^2}{v_2 s_2^2}
\]

melynek két szabadságfoka van: (a nagyobb szórásnégyzetű minta elemszáma-1) és (a kisebb szórásnégyzetű minta elemszáma-1). Az \(F\)-próba táblazatai azonban az eloszlás egyik oldalát tartalmazzák, az 1-nél nagyobb értékeket, ezért kellett a nagyobb varianciát osztanunk a kisebbel. Mivel az \(F\)-táblázatok egyoldalasak, a próbánk pedig kétoldalas, ezért \(5\%\)-os kétoldali próbához a 2,5%-os \(F\)-táblázatot kell használnunk. Egy \(5\%\)-os \(F\)-táblazattal tehát 10%-os kétoldalas szinten teszteljük a varianciák különbségességét. A számítógépes programok többnyire az ún. Levene próbával végzik a varianciák összehasonlítását, amelynek nem feltételezi az adatok normális eloszlása, a számítását itt nem részletezzük.

4.3. A példafeladat kiértékelése kétmintás \(t\)-próbával

\(H_0: \mu_1 = \mu_2\) (a két sokaság átlaga megegyezik)  
\(H_a: \mu_1 \neq \mu_2\) (a két sokaság átlaga különbözne – kétoldalás próba).

Tegyük fel, hogy az adataink normális eloszlású sokaságból származnak. Az adatok eloszlását megnézve, nincs okunk kételkedni a normálisban, nincs nagyon kiugró érték és viszonylag szimmetrikus is mindkét adat sor. Ilyen kis elemszám esetén nem is tudjuk biztonságosan ellenőrizni a normálisitást (lásd normalitásvizsgálat).

Akkor, hogy el tudjuk dönteni, melyik „\(t\)-próba képletet” alkalmazzuk, először a varianciákat hasonlítjuk össze, kézi számítással. „Szemre” a két standard deviációt összehasonlítva látható, hogy az egyik egy kicsit nagyobb, mint a másik. Kérjeds, hogy ez a nagyságbeli különbség véletlennek tekinthető-e. Elvégezve a próbát, \(F\)-re a következő értéket kapjuk:

\[
F = \frac{10,79^2}{9,25^2} = \frac{116,4241}{85,5625} = 1,36069
\]

A szabadságfokok kiszámításánál a nagyobb varianciájú csoport a 10 elemű diétás csoport, a kisebb varianciájú a 11 elemű kontroll csoport. Így tehát az \(F\)-eloszlás táblázatából a kritikus értéket a számítható 9-es szabadságfokú és oszlopa és a nevező 10-es szabadságfokú sorában kell keresni. A 9,10 szabadságfokú (egyoldalas) \(F\)-eloszlás táblázatából \(p = 0,05\) esetén \(F_{0,05; 9; 10} = 3,02\). Mivel az általunk számított \(F\)-érték ennél kisebb, a varianciák azonosagára vonatkozó nullhipotézist elfogadjuk, a különbség nem szignifikáns 10%-os szinten, \(p \geq 0,1\). Nyilván a különbség 5%-os szinten sem szignifikáns, tehát \(p \geq 0,05\). A programrendszer által számított Levene próba eredménye hasonló: \(p = 0,189\geq 0,05\), nem szignifikáns 5%-os szinten.

Miután úgy döntöttünk a sokaságok varianciáról, hogy azok egyenlőek, az átlagok összehasonlítására a \(t\)-érték számításához az azonos varianciákhoz tartozó (8.3) képletbe helyettesítünk:
A szabadságfok $= 10 + 10 = 20$-ra. A táblalétra kritikus érték $t_{0.05, 20} = 2.109$. Ennél a mi számított $t$-értéknél abszolút értéke nagyobb, tehát az eltérés szignifikáns $5\%$-os szinten. A statisztikai rendszer által számított $p$-érték $p = 0.023$, szignifikáns. A $p = 0.023$ azt jelenti, hogy ha igaz lenne, hogy a dióta nem hatásos, akkor $2.3\%$ az esély arra, hogy pusztán a véletlen folytán kapjunk 3 kg-os vagy ennél nagyobb átlagos különbséget. Ez az esély pedig nagyon kicsi, ezért tekintünk ezt a különbséget jelentősnek. Az átlagosan 3 kg-os különbség a két csoport közötti testsúlycsökkentésen statisztikailag szignifikáns különbség, más dolog viszont annak megítélése, hogy ekkora különbség jelentős-e a „szakmailag”.

A $t$-próba másik, Welch-féle módosított képletet ($8.5$) kellene alkalmaznunk, ha a variációk különbözők lennének. Annak ellenére, hogy elfogadottuk a variációk egyenlőségét, csak a számítás bemutatására, végezzük el a számítást különböző variáció esetére is: $t = 2.426$, szabadságfok $= 15,122$. Ekkor a kritikus érték $t_{0.05, 15} = 2,1315$, ennél nagyobb a számított $t$-érték abszolút értéke ($2,426g2,1315$), tehát a különbség ebben az esetben is szignifikáns. A $p$-érték ebben az esetben $p = 0.028$ lenne.

4.4. Nemparaméteres módszer – Mann-Whitney próba

Ha a normálitást nem tudjuk, vagy nem akarjuk ellenőrizni, vagy ha – bár az eredeti eloszlás folytonos, de az adatokat ordinális skálán mérjük –, a két független csoportot nemparaméteres módszerekkel hasonlíthatjuk össze. Sok eljárás van, melyek mindegyike azt a nullhipotézist teszti, hogy a két minta azonos eloszlásból származik. A leggyakrabban mégis egy rangsorolásos eljárást alkalmaznak a kétmintás próbára nemparaméteres megfelelőjeként. Ezt a próbát szokták Wilcoxon próbának is nevezni, mivel eredetileg Wilcoxon dolgozta ki. Röviddel utána Mann és Whitney közölte ennek egy másik értelmezését. Mégis, megkülönböztetésül az összetartozó adatok kiértékelésére szolgáló Wilcoxon-féle előjelű rangpróbától, a független minták összehasonlítására szolgáló eljárást Mann-Whitney próbának nevezik általában.

A módszer igen egyszerű: a két mintát együtt rangsoroljuk, vagyis csoporttól függetlenül készítjük el a számítást különböző varianciák esetére is: $t = 2,426$, szabadságfok $= 15,122$. Ekkor a kritikus érték $t_{0.05, 15} = 2,1315$, ennél nagyobb a számított $t$-érték abszolút értéke ($2,426g2,1315$), tehát a különbség ebben az esetben is szignifikáns. A $p$-érték ebben az esetben $p = 0.028$ lenne.

Egy kicsit bonyolultabb a Mann-Whitney-U statisztika számítása, amely két csoport elemeinek a pára állításán alapul. Az egyik csoport minden egyes elemét ($x_i$) pára állítjuk a másik csoport minden egyes elemével ($y_i$), az így keletkezett párok száma $n1n2$. Megvizsgáljuk, hogy a párok között hány olyan van, ahol az első szám kisebb, mint a másik ($x_i < y_i$). Ezeknek a pároknak a száma a Mann-Whitney-U-val jelölt statisztika (pontosabban, ha vannak a párok között egyenlők is, akkor az egyenlő párok számának a felét még hozzávezessük $U$-hoz). Ha a két populáció között nincs különbség, körülbelül egyforma számú olyan pár lesz, amelyekben $x_i < y_i$, mint amelyekben fordított a helyzet. Ha nagyon sok vagy nagyon kevés ilyen pár van, az arra utal, hogy a két populációban lévő számok nem egyformák egymáshoz viszonyítva. Az $U/n1n2$- hányados annak a valószínűségnének a becslése, hogy egy, az első populációból véletlenszerűen választott új egyed értéke kisebb lesz, mint a másik populációból választott új egyed. $U$ különben a $T$ ismeretében is kiszámítható a következő képletten: $U = n1n2 + 1n1\text{ord}(n2\text{ord} + 1) - T$.
Megjegyzés: A probléma szimmetriája miatt ezt a próbát sokféleképpen lehet tárgyalni, az egyes szakkönyvek el is térnek egymástól a kisebb elemekből származó adatok esetén.

4.4.1. A példafeladat kiértékelése Mann-Whitney U-próbával

A példafeladat adatait először nagyság szerint rendeljük, majd az együttes mintára a legkisebb 1-esével kiosztjuk a rangszámokat. Az egyenlő adatokhoz tartozó rangszámokat korrigáljuk (kapcsolt rangok). Ezeket az adatokat a 8.5. táblázat tartalmazza. Azt elemzhetjük, hogy a kisebb rangszám összege a kisebb belőle tekintet tekintetek esetén, vagy pedig bármelyiket. Eszerint természetesen a táblazataik is eltérnek egymástól. A számítógépes szoftverek között is vannak eltérések, pl. bizonyos szoftverek Wilcoxon-éle statisztikaként a két rangszámösszeg közül a kisebbet tekintik; $U$ értéket az első csoportra számítják ki, és ha ez nagyobb, mint $n_1n_2/2$, akkor $U' = n_1n_2 - U$ értéket írják ki.

8.4. táblázat - A 8.3. táblázat adatai a testsúlyváltozás (diéta előtt és után különbsége) szerint rendezve, rangszámokkal.

<table>
<thead>
<tr>
<th>Vízsgált személy</th>
<th>Testsúlyváltozás</th>
<th>Csoport</th>
<th>Rangszám</th>
<th>Kapcsolt rangok</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>-2</td>
<td>Kontroll</td>
<td>2</td>
<td>1,5</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>Diéta</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>Diéta</td>
<td>4</td>
<td>5,5</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>Kontroll</td>
<td>5</td>
<td>5,5</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>Kontroll</td>
<td>6</td>
<td>5,5</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>Kontroll</td>
<td>7</td>
<td>5,5</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>Diéta</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>Kontroll</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>Kontroll</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>Kontroll</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>Diéta</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>Kontroll</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>21</td>
<td>3</td>
<td>Kontroll</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>Diéta</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>Diéta</td>
<td>16</td>
<td>16,5</td>
</tr>
<tr>
<td>17</td>
<td>5</td>
<td>Kontroll</td>
<td>17</td>
<td>16,5</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>Diéta</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>Diéta</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>Diéta</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>Diéta</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>19</td>
<td>-2</td>
<td>Kontroll</td>
<td>1</td>
<td>1,5</td>
</tr>
</tbody>
</table>

8.5. táblázat - A 8.3. táblázat adatai az eredeti sorrendben, rangszámokkal.

<table>
<thead>
<tr>
<th>Vízsgált személy</th>
<th>Testsúlyváltozás</th>
<th>Csoport</th>
<th>Rangszám</th>
<th>Kapcsolt rangok</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>5</td>
<td>Diéta</td>
<td>16</td>
<td>16,5</td>
</tr>
<tr>
<td>3.</td>
<td>3</td>
<td>Diéta</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>4.</td>
<td>10</td>
<td>Diéta</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>5.</td>
<td>6</td>
<td>Diéta</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>6.</td>
<td>4</td>
<td>Diéta</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>7.</td>
<td>0</td>
<td>Diéta</td>
<td>4</td>
<td>5,5</td>
</tr>
</tbody>
</table>
### A kezelt csoport rangszámösszege

A kezelt csoport rangszámösszege $R_1=140$, a kontroll csoporté $R_2=91$. Mivel a kisebb elemszámú csoport a kezelt, így a próbastatisztikánk értéke $T = R_2 = 140$. Ellenőrizzük számolásunk helyességét: a két rangszámösszeg együttens $91 + 140 = 231$, ez egyenlő az első „$n$” egész szám összegével, az $n(n+1)/2$ képlet szerint $21 \cdot 22/2 = 231$.

A Mann–Whitney U próba táblázata szerint 10-11 elemszámhoz tartozó kritikus értékek = 0,05 esetén 81-139, mivel a $T = 140$ ezen az intervallumon kívül esik, a különbség szignifikáns 5%-os szinten. A $z$-érték kiszámítására nincs szükség, mert az elemszám kicsi, mégis mintaképpen kiszámítjuk:

\[
z = \frac{140 - 10(10 + 11 + 1)/2}{\sqrt{10 \cdot 11(10 + 11 + 1)/12}} = \frac{-30}{14,2} = -2,12
\]

Standard normális eloszlás és = 0,05 esetén a kritikus érték $z = 1,96$. Mivel a $|z| = 2,12 > 1,96$, a különbség szignifikáns 5%-os szinten. A statisztikai program által számolt $p$-érték $p = 0,033$.


## 5. Két csoport összehasonlítására vonatkozó feladatok és kérdések

Bizonyos betegek vér PH értékét mérték sóoldat beadásakor és 20 perccel később ($n=18$). Változott-e az átlagos PH érték? Az alapstatisztikák a következők:

<table>
<thead>
<tr>
<th>Vizsgált személy</th>
<th>Testülyváltozás</th>
<th>Csoport</th>
<th>Rangszám</th>
<th>Kapcsolt rangok</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.</td>
<td>1</td>
<td>Diéta</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>9.</td>
<td>6</td>
<td>Diéta</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>10.</td>
<td>6</td>
<td>Diéta</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>Rangszámösszeg, $R_1$</td>
<td></td>
<td></td>
<td></td>
<td>140</td>
</tr>
<tr>
<td>11.</td>
<td>2</td>
<td>Kontroll</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>12.</td>
<td>0</td>
<td>Kontroll</td>
<td>5</td>
<td>5,5</td>
</tr>
<tr>
<td>13.</td>
<td>1</td>
<td>Kontroll</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>14.</td>
<td>0</td>
<td>Kontroll</td>
<td>6</td>
<td>5,5</td>
</tr>
<tr>
<td>15.</td>
<td>3</td>
<td>Kontroll</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>16.</td>
<td>1</td>
<td>Kontroll</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>17.</td>
<td>5</td>
<td>Kontroll</td>
<td>17</td>
<td>16,5</td>
</tr>
<tr>
<td>18.</td>
<td>0</td>
<td>Kontroll</td>
<td>7</td>
<td>5,5</td>
</tr>
<tr>
<td>19.</td>
<td>-2</td>
<td>Kontroll</td>
<td>1</td>
<td>1,5</td>
</tr>
<tr>
<td>20.</td>
<td>-2</td>
<td>Kontroll</td>
<td>2</td>
<td>1,5</td>
</tr>
<tr>
<td>21.</td>
<td>3</td>
<td>Kontroll</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Rangszámösszeg, $R_2$</td>
<td></td>
<td></td>
<td>91</td>
<td></td>
</tr>
<tr>
<td>I.</td>
<td>-1</td>
<td>Diéta</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
8.6. táblázat -

<table>
<thead>
<tr>
<th>Beadáskor</th>
<th>20 perc múlva</th>
<th>Különbség</th>
</tr>
</thead>
<tbody>
<tr>
<td>Átlag</td>
<td>7,3821</td>
<td>7,3911</td>
</tr>
<tr>
<td>SD</td>
<td>0,0331</td>
<td>0,03336</td>
</tr>
</tbody>
</table>

1. Milyen típusú a két összehasonlítandó minta?
2. A különbség eloszlásról normális eloszlást feltételezve, milyen próbával teszteli a kérdést?
3. A páros t-próbának fogalmazza meg a nullhipotézisét és az alternatív hipotézisét kétoldalas próba esetén a feladatnak megfelelően!
4. Mennyi a szabadságfokok?
5. Mennyi a táblabeli t-érték = 0,05 esetén?
6. Számítsa ki a t-értéket és válaszoljon arra a kérdésre, hogy szignifikáns-e az átlagos PH-érték változása?

Hasonlitsuk össze a páros t-próbát és a kétmintás t-próbát, végezzük el a 8.1. táblázat adatain a kétmintás t-próbát, nem véve figyelembe, hogy összetartozó mintákról van szó! Vajon mi lehet az oka, hogy ugyanaz a két csoport közötti különbség most nem szignifikáns?

Útmutatás: Ugyanazon adatok esetén teljesen eltérő következtetésre jutottunk, különböző kísérleti elrendezéseket feltételezve. Független csoportok esetén nem elegendő az információ ahhoz, hogy az átlagosan 4 kg súlyveszteséget jelentősnek tekintsük (t = 0,89, df = 18, p = 0,385). 10-10 független mérés alapján és az adott szórások mellett a 4 kg súlyveszteség véletlennek tekinthető. Az önkonszroltos kísérlet előnye abban van, hogy a saját kontroll (adat-párok) figyelembe vételével csökkenti a hibavarianciát.

A diéta kísérlet adatai részletesen. A 8.7. táblázat - 8.8 táblázatban megadjuk az elképzelt kísérlet adatait. Milyen t-próbákat végezne az alábbi adatokon?

Útmutatás: Bármely kezelés esetén a változás vizsgálata csak akkor meggyőző, ha azt egy kontroll (kezelést nem kapott) csoporttal hasonlítjuk össze. Elvileg tehát elvégezhető a változás vizsgálata csoportonként, és a csoportok összehasonlítása a diéta kezdetén és végén. Azonban a kiértékeléskor nem helyes a t-próbák halmozott alkalmazása. Általánban helytelen bármely olyan módszer ismértelt alkalmazása egyazon kísérleten belül, amelyet két adatsor összehasonlítására „találtak ki”, a 4 adatsort egyszerre varianciaanalíziszel lehet összehasonlítani. Egyetlen t-próbát végezhetünk mégis: a változások független adatsorait kétmintás t-próbával összehasonlítva megkapjuk, hogy nagyobb volt-e az átlagos testsúlyváltozás a diétás csoportban, mint a kontrollban.

8.7. táblázat -

<table>
<thead>
<tr>
<th>Vizsgálati csoport</th>
<th>Vizsgált személy</th>
<th>Testsúly a diéta előtt</th>
<th>Testsúly a diéta után</th>
<th>Különbség</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>95</td>
<td>90</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>75</td>
<td>72</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>110</td>
<td>100</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>81</td>
<td>75</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>92</td>
<td>75</td>
<td>88</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>83</td>
<td>82</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>94</td>
<td>93</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>88</td>
<td>82</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>105</td>
<td>99</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>
Kétféle tanítási módszer tesztelésére 34 hallgatót véletlenszerűen 2 csoportba soroltak. A maximum 100 pontos teszteredmények a következők voltak:

A: 89,89,91,92,92,92,92,93,95,95,96,97,97,97,98
B: 50,66,73,74,84,88,89,90,91,91,91,93,93,93,96,98

Hasonlitsa össze a két csoport adatait a megfelelő módszerrel.

Útmutatás: a második csoportban eloszlása nyilvánvalóan nem normális, ezért nemparaméteres próbára alkalmazása javasolt.

<table>
<thead>
<tr>
<th>Vizsgálati csoport</th>
<th>Vizsgált személy</th>
<th>Testsúly a diéta előtt</th>
<th>Testsúly a diéta után</th>
<th>Különbség</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vizsgált személy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Átlag</td>
<td>90,8</td>
<td>86,8</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>10,79</td>
<td>9,25</td>
<td>3,333</td>
<td></td>
</tr>
<tr>
<td>Kontroll</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>93</td>
<td>91</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>84</td>
<td>84</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>83</td>
<td>82</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>92</td>
<td>92</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>101</td>
<td>98</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>83</td>
<td>82</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>95</td>
<td>90</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>93</td>
<td>93</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>84</td>
<td>86</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>98</td>
<td>100</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>90</td>
<td>87</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Átlag</td>
<td>99,5</td>
<td>89,55</td>
<td>1,00</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>6,314</td>
<td>6,039</td>
<td>2,145</td>
<td></td>
</tr>
<tr>
<td>Diéta</td>
<td><strong>1</strong></td>
<td><strong>85</strong></td>
<td><strong>86</strong></td>
<td><strong>-1</strong></td>
</tr>
</tbody>
</table>
9. fejezet - Kettőnél több csoport összehasonlítása folytonos adatok esetén

Boda Kriszta dr.

Több csoport összehasonlításakor automatikusan adódk az az ötlet, hogy alkalmazzuk a két csoport összehasonlítására már megismert módszereket, mindig két-két csoportot használva össze, pl. végezzünk több t-próbát páronként. Ez az eljárás helytelen – bár sajnos elég gyakori az orvosi irodalomban. Ha ugyanis olyan eljárásokat használnánk párosával, egyenként 0,05 szinten, amelyeket két csoportra (vagy változóra) fejlesztettek ki, akkor az első fajta hiba valószínűsége halmazódhat, azaz, a megengedhetőnél nagyobb valószínűséggel ítélhetünk – tévessz és különbözőnek valójában azonos populációkat. Emlatt hibás pl. több csoport esetén az átlagok összehasonlítására páronként kétminta t-próbákat végezni. Nem tudhatjuk ugyanis, hogy a szignifikáns eredmények közül melyek tulajdoníthatók a véletlennek, és melyek tükröznék valódi különbséget.

Ezt a hibánövekedést könnyen megérthetjük, ha belegondolunk az első fajta hiba jelentésébe: az 0,05 szint azt jelenti, hogy ha a nullhipotézis igaz, (pl. az összehasonlítandó populációk között nincs különbség), az első fajta hiba elkövetésének valószínűsége 0,05 = 1/20. Ha több, azonos populációból vett mintát páronként hasonlítunk össze, 20 közül átlagosan 1 összehasonlítás szignifikáns eredményre vezet, pusztán a véletlen műveként.

Az első fajta hibanövekedést speciális esetekben ki is lehet számítani. Több csoport összehasonlítása esetén nem egy, hanem több hipotézist tesztelünk, adott, mondvuk 0,05 szinten. Mindegyik tesztnél elkövetjük az első fajta hibát, aminek a valószínűsége 0,05 – egyetlen hipotézisre. Ez azt jelenti, hogy annak valószínűsége, hogy elvetjük az igaz nullhipotézist 0,05. Ebből annak valószínűsége, hogy elfogadjuk az igaz nullhipotézist: 1 - 0,05 = 0,95. Ha még azt is feltesszük, hogy a hipotézisek függőek, akkor annak valószínűsége, hogy két teszt esetén mindkettőt elfogadjuk, 0,95 x 0,95 = 0,9025, három teszt esetén ugyanez 0,953 = 0,8574. Annak valószínűsége, hogy a vizsgált 3 hipotézis közül legalább egyet tévessz és elutasítsünk: 1 - 0,8574 = 0,1426, ez lényegesen nagyobb, mint 0,05. Általában k számú függő hipotézis összehasonlítása esetén 1- (1-0,05) annak valószínűsége, hogy az egész döntésnél elfogadjuk az első fajta hibát, vagyis valamelyik összehasonlításnál tévessz szignifikáns különbséget állapítunk meg. Ezt a valószínűséget a kísérletre vonatkozó szignifikanciaszintnek nevezzük, szemben -val, ami az összehasonlításonkénti szignifikanciaszint. Ha a hipotézisek nem függőek, akkor a kísérletenkénti szignifikanciaszintre felső határt adhatunk: ez a határ k.


Ha az eloszlás nem normális, vagy a csoportok variánciái nem azonosak, gyakran célra vezető módszer a mintaadatok transzfórmálása. Elég gyakori példa eloszlás a lognormális eloszlás, ebben az esetben az adatok logaritmusát véve, nem csak az eloszlás normálisátható, de a variánciák közötti különbség is megszüntethető. Ha a normalitást nem tudjuk ellenőrizni, vagy az adatok folytonos eloszlásból származnak ugyan, de ordinális skálán mérik (pl. a fájdalom egy 10 fokozatú skálán mérik), akkor a csoportok eloszlásainak összehasonlítása nemparaméteres módszerekké történhet. A leggyakrabban alkalmazott nemparaméteres módszerek több független minta esetén a Kruskal-Wallis próba, és több (egy szempont szerint) összetartozó minta esetén a Friedman próba.

1. Paraméteres eljárások, variánsanalízis

A variánsanalízis több, azonos szórású, normális eloszlású populáció átlagának az összehasonlítására szolgáló módszer, amelyet ANOVA néven is emlegetnek az angol elnevezés betűinek rövidítésésként (Analysis of Variance). A variánsanalízis a t-próbák általánosítása több csoport esetére. Azért hivjuk variánsanalízisnek,
mert az átlagokat hasonlítja ugyan, de ezt többféle módon definiált variációk segítségével teszi. A varianciaanalízis a teljes adathalmaz össz-szóródását (pontosabban összvarianciáját) vizsgálja abból a szempontból, hogy azt csupán a véletlen ingadozás okozza-e, vagy ahhoz valamilyen más tényező, pl. a csoportok átlagai közötti különbség is hozzájárul.

Többféle varianciaanalízis van a kíséreti elrendezéstől függően. Amennyiben a csoportok függetlenek, és csak egyetlen szempont szerint különböznek (pl. többféle kezelést vagy többféle betegcsoportot hasonlítunk össze), akkor egyszempontos varianciaanalízis a módszer neve. Ha a csoportok függetlenek, de többféle szempont szerint is vizsgálhatók (pl. nemek szerint és kezelések szerint is), akkor két- vagy többszempontos varianciaanalízissel hasonlítjuk össze az átlagokat. Ha a csoportok összetartozó minták csoportjai, (pl. ugyanazokon az egyedeken több mérést végeznek több időpontban vagy különböző kíséreti körülmények között), akkor az ún. ismérelt méréses varianciaanalízist kell alkalmaznunk.

1.1. Egyszempontos varianciaanalízis

1.1.1. Példa

Egy kísérletben (Farkas és mtsai, 2003.) lokális iszkémiának alávetett, izolált patkányszívben a szívfrekvencia és a QT-szakasz hosszának változását vizsgálták három antiaritmiás gyógyszer hatására. 5 Mm K+ káliumion koncentráció esetén, 25 perccel a lokális iszkémia után a QT-szakasz hosszát a 9.1. táblázat - 9.1. táblázatban látható értékeket kapták. Vizsgáljuk meg, hogy a 4 csoportban van-e különbség a QT-szakasz átlagos hosszából!

9.1. táblázat -

<table>
<thead>
<tr>
<th>Kontroll</th>
<th>Quinidine</th>
<th>Lidocaine</th>
<th>Flecainide</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>84</td>
<td>56</td>
<td>65</td>
</tr>
<tr>
<td>68</td>
<td>89</td>
<td>76</td>
<td>73</td>
</tr>
<tr>
<td>66</td>
<td>78</td>
<td>72</td>
<td>71</td>
</tr>
<tr>
<td>54</td>
<td>81</td>
<td>66</td>
<td>61</td>
</tr>
<tr>
<td>89</td>
<td>69</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>Átlag</td>
<td>60,4</td>
<td>67,3</td>
<td>68,0</td>
</tr>
<tr>
<td>SD</td>
<td>6,80</td>
<td>5,49</td>
<td>4,34</td>
</tr>
<tr>
<td>61</td>
<td>76</td>
<td>65</td>
<td>69</td>
</tr>
</tbody>
</table>

1.2. Az egyszempontos varianciaanalízis elve

Több csoport összehasonlítása lényegében a csoportok eloszlásának összehasonlítását jelenti. Minden mérés hibával jár, a mintaadatok csoportonként pusztán a véletlen miatt is különböznek. A kérdés éppen ez: annak elődöntése, hogy az egyes minták ugyanabbl a sokaságból származnak-e, vagy nem. Az egyszempontos varianciaanalízisben azonos szórású normális eloszlásból származó mintákkal dolgozunk, így a kérdés az átlagok összehasonlítására korlátozódik.

Mielőtt a példafeladatot megoldanánk, gondolkodjunk el a varianciaanalízis elvén. Játsszuk azt, hogy ismerjük a valóságot, vagyis használjunk ún. véletlenszám-generátort. A

112
9.1.a.,b. ábrán látható, normális eloszlásból származó 6 elemű mintákat az Excel program segítségével készítettük. Az 9.1.a. ábra három, N(2,1) eloszlásból származó 6 elemű mintát ábrázol – ezek átlagai és szórásai tehát csak a véletlen miatt különböznek, az átlagok és a szórások közelítően 2 illetve 1 egységnyiek. A teljes adathalmaz – azaz, az összes adat szóródása lényegében ugyanakkora, mint az egyes csoportoké (azaz, közelítően 1 egységnyi). A
9.1.b. ábrán a második és a harmadik csoportot 2, illetve 1 egységgel feljebb „toltuk”. Az egyes csoportok szórásai nem változtak tehát, de az átlagok közötti különbség mesterséges, nem a véletlen műve. Ettől az eltolástól – az átlagok közötti különbségtől - a teljes adathalmaz össz-szórása megnőtt, sokkal nagyobb lett, mint az egyes csoportok szórása.

Konkrét adatok kiértékelésekor nem ismerjük a valóságot, a mintaadatok alapján azonban megvizsgálhatjuk az adatok szóródását. Többféle szóródást is vizsgálhatunk:

- Az összesített adathalmaz szóródását csoporttól függetlenül.
- A csoportok szóródását csoportonként külön-külön, az ún. csoportokon belüli szóródást.
- Az átlagok közötti különbségből adódó – ún. csoportok közötti szóródást.

Ha igaz az, hogy a sokaság-átlagok azonosak, akkor az átlagok eltérése ugyanaz a véletlen ingadozás okozza, mint a csoportokon belüli ismétlések különbözőségét: a csoportok közötti és a csoportokon belüli variancia ugyanannak a (hiba)varianciának a becslése (

1.3. Többszörös összehasonlítások

Ha a variancia-analízis eredménye nem szignifikáns, akkor az analízisnek vége, és azt mondjuk, hogy a csoport-átlagok között nincs szignifikáns különbség. Ha a variancia-analízis eredménye szignifikáns, ez azt jelenti, hogy a populációk átlagai között van különbség. Azt, hogy ez a különbség hol van, mely csoportok átlagai egyformák és melyek közöttik különbözik, az ún. többszörös összehasonlításokkal tudjuk felérteni. Ezek között leggyakrabban az összes lehetséges páronkénti összehasonlítás elvégzése, de vannak olyan módszerek is, amelyek az átlagokat „hasonlítják”, „homogen” osztályokra csoportosítják, tehát egy ilyen osztályon belüli csoport szignifikánsan különbözik bármely másik osztálybeli csoporttól. A többszörös összehasonlító módszerek többed-kevésbé az egész kísérletre vonatkozóan „garantálják”, hogy az első faja hiba valószínűsége -nál ne legyen nagyobb, tehát adott szignifikanciaszinten tartják az első faja hiba valószínűségét a teljes kísérletre nézve.


Előfordul, hogy a variancianalízis eredménye szignifikáns, mégsem találunk páronként szignifikáns különbséget. Ilyenkor is kell, hogy legyen „valahol” különbség, esetleg a kontroll csoport különbözik az összes többi csoporttól, azokat egy „nagy” csoportnak véve. Az olyan összehasonlításokat, ahol csoport-együttesek átlagait hasonlítjuk más csoportokhoz vagy csoport-együttesekhez, kontraszt-analízisnek nevezzük. Scheffe
módshoz az átlagok egy vagy több súlyozott kombinációját teszteli az első fajta hiba növekedése nélkül. Ezeket az összehasonlításokat ebben a részben nem tárgyaljuk.

A „post-hoc” összehasonlítás veszélyei.

Ideális esetben a kísérlet előtt már tudnunk kell, hogy mely csoportok közötti különbség érdekel bennünket. A gyakorlatban mégis gyakori, hogy a kísérlet elvégzése után kapott eredmények ismeretében előre nem tervezett összehasonlításokat is elvégezünk. Ha pl. két csoport között nem vártunk eltérést, mégis nagyobb különböző átlagokat kaptunk, késztetést érezhetünk arra, hogy ezt a különbséget is teszteszük. Ha csak a legnagyobb eltérést mutató csoportok közötti különbséget hasonlíthatjuk össze, annak valószínűsége, hogy szignifikáns különbséget kapunk, nagyobb lesz ahhoz képest, mint ha a vizsgálat megkezdése előtt véletlenszerűen választott különbségszintet tesztenélünk. Ilyen esetben a kísérletre vonatkozó első fajta hiba megnő, nagyobb lesz az előre deklarált hibájából. A konzervatívabb többszörös összehasonlítási módszereket megkereshetünk, mely átlagok különböznek melyekből, minden lehetséges párt összehasonlítva. Ha a csoportok páronként összehasonlítását aztán végezzük, miután a kísérletet elvégeztük, ezeket a konzervatív módszereket kell alkalmaznunk. Egyedül Scheffé módszere enged meg ilyen utólagos „kutakodást”.

1.4. Az egyszempontos varianciaanalízis feltételei

A varianciaanalízis feltétele, hogy minták normális eloszlású, azonos variánciátú populációból származnak. Az azonos variánciák feltétel azt jelenti, hogy minden minta szóródási körületbeli ugyanakkora, vagyis ugyanannak a (közös) variánciának a becslése. Ez utóbbi feltételt sokan nem figyelik, vagy nem veszik figyelembe, pedig megsértése hibás következtetésekre vezethet.


A minták függetlensége alapvető feltétel. Gyakori hiba, hogy orvosi jegyzőkönyvekből, kórlapokból úgy állítanak össze csoportokat, hogy egyes személyek több csoportba is bekerülnek, pl. a betegség előrehaladottsága miatt, vagy mivel többször jelentek meg vizsgálaton. Az ilyen csoportok nem tekinthetők függetlennek.

1.5. A varianciaanalízis elvégzéséhez, a számításokhoz szükséges képletek

Legyen adott: \( t \) számú független minta, jelölje \( n_i \) az \( i \)-edik minta elemszámát, \( x_{ij} \) a \( i \)-edik minta \( j \)-edik elemét. Az első minta elemei tehát \( x_{11}, x_{12}, \ldots x_{1n} \), a második \( x_{21}, x_{22}, \ldots x_{2n} \), az utolsó \( x_{t1}, x_{t2}, \ldots x_{tn} \). Feltesszük, hogy mindegyik minta normális eloszlású, azonos szórású populációból származik, azaz az \( i \)-edik minta \( N(n_i) \) eloszlásból. A tesztelendő nullhypothesjez az, hogy a vizsgált populációk átlagai megegyeznek, a minták egy és ugyanazon populációból származnak, az alternatív hipotézis szerint legalább egy minta más populációból származik: \( H_0: \mu_1 = \mu_2 = \ldots = \mu_t \).

Jelölje \( N \) az össz-elemzésom, \( n \) az \( i \)-edik minta átlagát, *** az összes adat átlagát.

A minták közötti variáncia becsleése az egyes mintavariánciák „egyesítése”:
Kettőnél több csoport összehasonlítása folytonos adatok esetén

\[
S_{\text{belé}}^2 = \frac{\sum_{i=1}^{t} \sum_{j=1}^{n_i} (x_{i,j} - \overline{x}_i)^2}{(n_1 - 1) + (n_2 - 1) + \ldots + (n_t - 1)}
\]

A minták közötti variancia becslésének számlálója az egyes minta-átlagok és a teljes adathalmaz átlaga közötti, az elemszámokkal szorzott négyzetösszeg:

\[
S_{\text{készt}}^2 = \frac{\sum_{i=1}^{t} n_i (\overline{x}_i - \overline{x})^2}{t - 1}
\]

Az összvariancia becslése pedig

\[
S^2 = \frac{\sum_{i=1}^{t} \sum_{j=1}^{n_i} (x_{i,j} - \overline{x})^2}{N - 1}
\]

A nullhipotézis teszteléséhez a két varianciát kell összehasonlítanunk: a minták közötti variancia nagyobb-e, mint a mintákon belüli. (Itt egyoldalas próbát végzünk, ha a minták közötti variancia kisebb, mint a mintákon belüli, máris megtartjuk H₀-t). Két varianciát F-próbával hasonlítunk össze. Most

\[
F = \frac{S_{\text{készt}}^2}{S_{\text{belé}}^2}
\]

t-1, N-t szabadságfokú F-eloszlású valószínűségi változó. Adott esetén az F-eloszlás táblázatbeli értékét jelölje \( F_{t-1,N-t} \). Ha \( F_{t-1,N-t} \) elvetjük H₀-t és azt mondjuk, hogy a minták közötti nagy varianciát a mintaátlagok közötti eltérés okozta, van tehát az átlagok között a többidő különböző. Ha \( F_{t-1,N-t} \) akkor megtartjuk a nullhipotézist.

A varianciaanalízis számításait általában táblázatba szokták foglalni, amint a 9.2. táblázat - A varianciaanalízis táblázata 9.2. táblázatban is látható:

9.2. táblázat - A varianciaanalízis táblázata

<table>
<thead>
<tr>
<th>A szóródás oka</th>
<th>Négyzetösszeg</th>
<th>Szabadságfok</th>
<th>Variancia</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Csoportok között</td>
<td>( Q_k = \sum_{i=1}^{t} )</td>
<td>t-1</td>
<td>( S_k^2 = \frac{Q_k}{t-1} )</td>
<td>( F = \frac{S_k^2}{S_b^2} )</td>
</tr>
<tr>
<td>Csoportokon belül</td>
<td>( Q_b = \sum_{j=1}^{n} )</td>
<td>N-t</td>
<td>( S_b^2 = \frac{Q_b}{N-t} )</td>
<td></td>
</tr>
<tr>
<td>Teljes</td>
<td>( Q = \sum_{i=1}^{t} \sum_{j=1}^{n_i} )</td>
<td>N-1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A teljes adathalmaz össz-varianciájának számlálója, a teljes eltérés-négyzetösszeg két komponensre, a minták közötti és a mintákon belüli(hiba) eltérés-négyzetösszegekre bonthatók: \( Q = Q_k + Q_b \). A szabadságfokok hasonló módon összeadódnak: N-1=t+1+N-t. A megfelelő négyzetösszegeket osztva a megfelelő szabadságfokokkal kapjuk a minták közötti és a mintákon belüli variánciákat, melyek a nullhipotézis fennállása esetén ugyanannak.
1.5.1. A példafeladat megoldása

Vizsgáljuk a nullhipotézist = 0,05 szinten! A tesztelendő nullhipotézis szerint mindegyik minta ugyanolyan átlagú sokaságból származik, vagyis a négy mintaátlag közötti ingadozás csakis a véletlen miatt van. A

Mielőtt nekikezdenénk a varianciaanalízisnek, vizsgáljuk meg a feltételeket. A normalitást 6 elemre nem érdemes ellenőrizni, de az adatokból látszik, hogy nincs okunk kételkedni a normalitásban. A szóródásokat „szemre” ellenőrizve azt látjuk, hogy nincs túl nagy különbség az egyes csoportok szóródásaiban. A Levene-próba eredménye $F = 0,537, 3$ és $19$ szabadságfokkal, $p = 0,663$, azaz jóval nagyobb, mint a 0,05. Tehát a variációk azonosságát elfogadjuk, a varianciaanalízis feltétele teljesül.

9.3. táblázat - A varianciaanalízis táblázata a mintapélda adataira

<table>
<thead>
<tr>
<th>A szóródás oka</th>
<th>Négyzetösszeg</th>
<th>Szabadságfok</th>
<th>Variancia</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Csoportokon belül</td>
<td>665,367</td>
<td>19</td>
<td>35,019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teljes</td>
<td>2180,957</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Csoportok között</td>
<td>1515,590</td>
<td>3</td>
<td>505,197</td>
<td>14,426</td>
<td>0,000</td>
</tr>
</tbody>
</table>

A varianciaanalízis számításainak eredménye a 4.10. táblázatban látható. A számunkra leglényegesebb eredmény a jobboldalon található $p$-érték, amely olyan kicsi, hogy az első három tizedesjegye csupa 0, és a számítógép gyakran így írja ki. Tehát $p$ 10,05, a csoport-átlagok között szignifikáns különbség van.

Mivel a varianciaanalízis eredménye szignifikáns, meg kell vizsgálnunk, mely csoportok között van különbség. A kísérlet célja itt a kontrolltól való eltérés vizsgálata volt, tehát a Dunnett próbát alkalmazzuk. Eszerint csak a Quinidine-nél kezelt csoport átlaga különbözik szignifikánsan a kontrolltól (9.4. táblázat - A Dunnett próba eredménye a mintapélda adataira 9.4. táblázat)

9.4. táblázat - A Dunnett próba eredménye a mintapélda adataira

<table>
<thead>
<tr>
<th></th>
<th>Az átlagok különbsége</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kontroll – Lidocaine</td>
<td>6,9333</td>
<td>.158</td>
</tr>
<tr>
<td>Kontroll – flecainide</td>
<td>7,6000</td>
<td>.113</td>
</tr>
<tr>
<td>Kontroll – Quinidine</td>
<td>22,4333</td>
<td>.000</td>
</tr>
</tbody>
</table>

Vizsgáljuk meg a többszörös összehasonlításokat LSD és Bonferroni módszerrel is! A kontrollhoz képest az LSD próbával a Flecainidine-átlag is szignifikánsnak adódott (9.5. táblázat - Többszörös összehasonlítások LSD módszerrel a mintapélda adataira 9.5. táblázat). Tudunk kell azonban az LSD módszerről, hogy kissé „engedékeny”, azaz, előfordulhat, hogy növeli a tényleges első fajta hibát. A számítógépes programok a Bonferroni módszert úgy hajtják végre, mintha minden csoportot minden csoporttal össze akarnánk hasonlítani (9.6. táblázat - Többszörös összehasonlítások Bonferroni módszerrel a mintapélda adataira 9.6. táblázat). Mivel 4 csoport esetén 6 lehetséges páronkénti összehasonlítás van, az “LSD” p-értékeihez képest a “Bonferroni” módszer p-értékei éppen 6-szorosak – kivéve az utolsót, ahol már az 1-et is meghaladta volna a 6-szoros szorzó, ott természetesen a maximális 1 érték szerepel. A Dunnett próbával összehasonlítva a p-értékeket, látható, hogy a kontrollhoz képest kisebb p-értékeket kaptunk, mint Bonferroni módszerrel. Ez is mutatja, hogy a Bonferroni módszer sok összehasonlítás esetén túlságosan konzervatív lehet, tehát bizonyos valódi különbségeket nem tud kimutatni.

9.5. táblázat - Többszörös összehasonlítások LSD módszerrel a mintapélda adataira

<table>
<thead>
<tr>
<th></th>
<th>Az átlagok különbsége</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kontroll – Lidocaine</td>
<td>-6,9333</td>
<td>0,068</td>
</tr>
<tr>
<td>Kontroll – Flecainide</td>
<td>-7,6000</td>
<td>0,047</td>
</tr>
<tr>
<td>Kontroll – Lidocaine</td>
<td>15,5000</td>
<td>0,000</td>
</tr>
<tr>
<td>Kontroll – Flecainide</td>
<td>14,8333</td>
<td>0,000</td>
</tr>
<tr>
<td>Lidocaine– Flecainide</td>
<td>-.6667</td>
<td>0,847</td>
</tr>
<tr>
<td>LSD Kontroll – Quinidine</td>
<td>-22,4333</td>
<td>0,000</td>
</tr>
</tbody>
</table>

9.6. táblázat - Többszörös összehasonlítások Bonferroni módszerrel a mintapélda adataira

<table>
<thead>
<tr>
<th></th>
<th>Az átlagok különbsége</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kontroll – Lidocaine</td>
<td>-6,9333</td>
<td>0,408</td>
</tr>
<tr>
<td>Kontroll – Flecainide</td>
<td>-7,6000</td>
<td>0,284</td>
</tr>
<tr>
<td>Quinidine – Flecainide</td>
<td>15,5000</td>
<td>0,001</td>
</tr>
<tr>
<td>Quinidine – Quinidine</td>
<td>14,8333</td>
<td>0,002</td>
</tr>
<tr>
<td>Lidocaine – Flecainide</td>
<td>-.6667</td>
<td>1,000</td>
</tr>
<tr>
<td>Bonferroni Kontroll – Quinidine</td>
<td>-22,4333</td>
<td>0,000</td>
</tr>
</tbody>
</table>

Megjegyzés. A teljes kísérlet a fentnél bonyolultabb volt: a QT-szakasz hosszát 3 Mm K⁺ és 5 Mm K⁺ káliumion-koncentráció mellett, az antiarritmias szerek alacsony és magas koncentrációján, a lokális iszkémia után több időpontban vizsgáltuk. A teljes kísérletből egy részproblémát emeltünk ki, amely egyszemponos varianciaanalízissel oldható meg. Helytelen lenne azonban az egyszempontos ANOVA-k ismételt alkalmazása minden egyes kálium koncentráció, koncentráció és időpont esetén, mert ezzel ugyanúgy megnövelhető az első fajta hibaelláttszínleg a teljes kísérletre nézve. A teljes kísérlet kiértékelése négszempontos, ismételt méréses varianciaanalízissel lehetséges 3 független szemponittal (gyógyszerhatás, a K⁺ hatása és a koncentráció hatása) és egy ismételt méréses szemponittal (időhatás).
2. Nem normális eloszlású sokaságok összehasonlítása, nemparaméteres próbák

Ha az egyes csoportokra a normalitás nem feltételezhető vagy nyilvánvalóan nem teljesül, akkor nemparaméteres módszereket alkalmazhatunk. Ebben az esetben is igaz, hogy a két csoportra kidolgozott próbák ismételt alkalmazása páronként nem helyes, a teljes kísérletre vonatkozó első fajta hiba ugyanúgy növekedhet. Önkritikus kísérlet esetén ilyen próba az ún. Friedman próba. Studiunk a két csoportra kidolgozott próbák ismételt alkalmazása páronként nem helyes, a teljes kísérletre vonatkozó első fajta hiba ugyanúgy növekedhet.

2.1. Több összetartozó minta nemparaméteres összehasonlítása: Friedman próba

Több csoport is képezhet összetartozó mintát: ha ugyanazokról az egyedekről több mérést készíthetünk különböző időpontban vagy különböző körülmények között. Gyakran előfordul, hogy a vizsgálatba bevont pacienseket nem csak a kezelés előtt és után figyeljük meg, hanem a kezelés után több időpontban is vizsgáljuk a kezelés hatását. Ilyenkor nem helyes az önkritikus kísérletre megismert Wilcoxon próbája páronkénti alkalmazása. Helyette az ún. Friedman próbát kell alkalmazni. Szokták nemparaméteres ismételt méréses ANOVA-nak is nevezni. A nullhipotézise: több, összetartozó minta ugyanazon populációból származik.

Jelölje \( t \) az ismételt mérések számát, \( n \) az egyedek számát. Az eljárás a következő: az adatokat egyenként rangsoroljuk, 1-től t-ig, majd a kapott rangszámokat mérésenként összeadjuk; az összegeket jelölje \( R_i \), \( i = 1, \ldots, t \). Ha egyes ismételt mérések között nincs különbség, akkor csak a véletlenség miatt, melyik rangszám melyik ismétléshez „kerül”, következéképpen a rangsánszögek is körülbelül egyformák. A Friedman próba ettől az egyformaságtól való eltérést teszti fel a következő képlet, amely közvetlőleg \( t-1 \) szabadsággfokú \( \chi^2 \)-eloszlást követ

\[
\chi^2 = \frac{12}{tn(t+1)} \sum_{i=1}^{t} R_i^2 - 3n(t+1) \sum_{i=1}^{n} (e_i^3 - e) \quad 1 - \frac{nt(t^2-1)}{2}
\]

A nevező csak kapcsolt rangok esetén lesz 1-től különböző. Az „\( e \)-k” jelentése az egyes sorokban az egyenlő elemek száma. Ha a \( p \)-érték alapján nem szignifikáns a különbség, az analízisnek vége. Szignifikáns esetben a különbség pontosítására, a páronkénti összehasonlításra a kétminta próbákat kell alkalmaznunk (pld. Wilcoxon-féle eljeges rangpróba) korrekcióval, pl. Bonferroni korrekcióval.

2.1.1. Példa

A lokális iszkémia vizsgálat kísérletben (Farkas és mtsai, 2003) 5 Mm K⁺ káliumion koncentráció esetén, a lokális iszkémia előtt 6 perccel (lap) és 1 perccel, utána pedig 5 és 25 perccel a QT-szakasz hosszára a 9.7. táblázat - Mintaadatok a Friedman próbahez 9.7. táblázatban található értékeket kapták.

<table>
<thead>
<tr>
<th>Állat sorszáma</th>
<th>Alap</th>
<th>Lokális iszkémia előtt 1 perc</th>
<th>Lokális iszkémia után 5 perc</th>
<th>Lokális iszkémia után 25 perc</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>53</td>
<td>51</td>
<td>66</td>
<td>53</td>
</tr>
<tr>
<td>3.</td>
<td>71</td>
<td>7</td>
<td>82</td>
<td></td>
</tr>
</tbody>
</table>

9.7. táblázat - Mintaadatok a Friedman próbahez
Az összehasonlításból a 3. állat adatait a hiányzó érték miatt törölni kell. Most az ismétlések száma, \( r=4 \) és a vizsgált egyederék száma, \( n=5 \). A 9.6. táblázat - Többszörös összehasonlítások Bonferroni módszerrel a mintapélda adataira 9.6. táblázat adatait soronként rangsorolva a 9.8. táblázat tartalmazza. Egy sor tartalmaz kapcsolt rangokat, mégpedig 2 elem egyenlő, tehát ebben a sorban \( e=2 \). 5 adat alapján a Friedman-próba eredménye szignifikáns: \( \chi^2 = 10,346 \), szabadságfok = 3. A \( \chi^2 \) eloszlás táblázatában 3-as szabadságfokhoz és \( \alpha=0,05 \) esetében tartozó kritikus érték 7,815. Mivel 10,346 ennél nagyobb, a különbség szignifikáns 5%-os szinten. A \( p \)-érték \( p=0,017 \).

9.8. táblázat - A próbaadatok ragsorolása

<table>
<thead>
<tr>
<th>Állat sorszáma</th>
<th>Alap</th>
<th>Lokális iszkémia előtt 1 perc</th>
<th>Lokális iszkémia után 5 perc</th>
<th>Lokális iszkémia után 25 perc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>2.</td>
<td>2,5</td>
<td>1</td>
<td>4</td>
<td>2,5</td>
</tr>
<tr>
<td>3.</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>4.</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5.</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>6.</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td><strong>Ranszműsszeg</strong></td>
<td><strong>R_i=6,5</strong></td>
<td><strong>R_i=14</strong></td>
<td><strong>R_i=19</strong></td>
<td><strong>R_i=10,5</strong></td>
</tr>
</tbody>
</table>

2.2. Több független minta nemparaméteres összehasonlítása: Kruskal-Wallis próba

Az egyszempe tos varianciaanalízis nemparaméteres megfelelő módszereként általában a Kruskal-Wallis próbát szokták emlegetni, holott több nemparaméteres módszer is van annak a nullhipotézisnek a tesztelésére, hogy több, független minta ugyanazon populációból származik.

A Kruskal-Wallis próba végrehajtása hasonló a Mann-Whitney U próbához. Sőt, 2 független csoport esetén mindkét módszer azonos eredményt ad. A mintákat együtt rangsoroljuk, vagyis csoporttól függetlenül készítjük el a rangsza mokat. Egyenlő adatok esetén koriggálunk a rangszámok átlagával. Végül csoportonként külön-külön összeadjuk a rangszámokat. A rangszámoszszeggelenképp egy képlet segítségével kiszámíthatjuk a \( H \)-val jelölt próbastatisztika értékét. Ez a képlet, ha nincsenek kapcsolt rangok vagy kevés számú kapcsolt rang esetén a következő: jelölje ismét \( t \) az összehasonlítandó csoportok számát, jelölje \( R_i \) az \( i \)-edik csoport rangszámösszegét. Ekkor

\[
H = \left( \frac{12}{N(N+1)} \sum_{i=1}^{t} \frac{R_i^2}{n_i} \right) - 3(N+1)
\]

szerint számolt próbastatisztika alapján dönthetünk, kis elemszám esetén táblázatból kereshető ki a szignifikancia, nagy elemszám esetén pedig a \( H \)-statsztika közelítően \( \chi^2 \)-eloszlású \( t-1 \) szabadságfokkal, tehát \( \chi^2 \) -eloszlás táblázata használható.

Az eredményképpen kapott \( p \)-érték alapján csak annyit tudunk megállapítni, van-e a csoportok között a többitől eltérő. Nemszignifikáns esetben az analízist befejezőük. Ha a \( p \)-érték alapján szignifikáns a különbség, (de csak akkor!) a páronkénti összehasonlításra a következő eljárást lehet alkalmazni: az \( i \)-edik és a \( j \)-edik csoport különbözõk, ha a következő egyenlőtlenség teljesül:
Kettőnél több csoport összehasonlítása folytonos adatok esetén

\[
\frac{R_i - R_j}{n_i - n_j} > t_{\alpha, N-t} \left( \frac{N(N+1)N-1-H}{12} \right)^{1/2} \left( \frac{1}{n_i} + \frac{1}{n_j} \right)^{1/2}
\]

ahol \( t_{\alpha, N-t} \) az \( N-t \) szabadságfokú \( t \)-eloszláshoz tartozó kritikus érték. Itt tehát tulajdonképpen a rangszám-átlagok különbségeinek abszolút értékét hasonlítjuk egy, a jobb oldalán szereplő kifejezéshez.


Az egyszempontos varianciaanalízis mintapéldájának adatain (9.1. táblázat - 9.1. táblázat) most végezzük el a Kruskal-Wallis próbát!

A rangszámokat és a rangszámösszegeket a 9.9. táblázat tartalmazza. Ezekből az adatokból a Kruskal-Wallis \( H \) értéke

\[
H = \frac{12}{23(23+1)} \left( \frac{26^2}{5} + \frac{122.5^2}{6} + \frac{62.5^2}{6} \right) - 3(23+1) = \frac{1}{46} \left( 135.2 + \frac{23137.5}{6} \right) - 72 = 14.77
\]

a szabadságfok = 3. A \( z \)-eloszlás táblázatából \( \alpha = 0.05 \) és 3-as szabadságfok esetén a kritikus érték 7,815. Mivel 14,77 > 7,815, a különbség szignifikáns 5%-os szinten. A számítógépes program a kapcsolt rangokat is figyelembe vette a \( H \) számításánál, s ezért \( H = 14,837, df = 3, p = 0.002 \). A többszörös összehasonlításokhoz számítsuk ki a kifejezés jobboldalát! Most \( N=23, t_{0.05,20} = 2.093 \), a kifejezés jobb oldalához két számítást kell végezni az elemzéseknek megfelelően:

\[
2,093 \left( \frac{23(23+1)}{23-4} \right)^{1/2} \left( \frac{1}{5} + \frac{1}{6} \right)^{1/2} = 2,093 \cdot \left( \frac{46 \cdot 0.3747 \cdot 11}{30} \right)^{1/2} = 5,26
\]

9.9. táblázat - Rangszámok, rangszámösszegek és -átlagok a mintapélda adataira.

<table>
<thead>
<tr>
<th>Csoport</th>
<th>Adat</th>
<th>Rangszám</th>
<th>Rangszámösszegek</th>
<th>Rangszámátlagok</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>53</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>68</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>66</td>
<td>8,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>54</td>
<td>2</td>
<td>26</td>
<td>5,2</td>
</tr>
<tr>
<td>Quinidine</td>
<td>76</td>
<td>17,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quinidine</td>
<td>84</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quinidine</td>
<td>89</td>
<td>22,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quinidine</td>
<td>78</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quinidine</td>
<td>81</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quinidine</td>
<td>89</td>
<td>22,5</td>
<td>122,5</td>
<td>20,42</td>
</tr>
<tr>
<td>Lidocaine</td>
<td>65</td>
<td>6,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lidocaine</td>
<td>56</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lidocaine</td>
<td>76</td>
<td>17,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lidocaine</td>
<td>72</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lidocaine</td>
<td>66</td>
<td>8,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lidocaine</td>
<td>69</td>
<td>12</td>
<td>62,5</td>
<td>10,42</td>
</tr>
<tr>
<td>Flecainide</td>
<td>69</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flecainide</td>
<td>65</td>
<td>6,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flecainide</td>
<td>73</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flecainide</td>
<td>71</td>
<td>14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Képezzük az összes lehetséges párt a rangszám-átlagokból és hasonlítsuk az átlagok különbségeit a most kiszámított mennyiségekhez. A 9.10. táblázat szerint a Quinidine különbözik a többi csoporttól. Az 1-4 összehasonlítás „határeset, mint ahogyan az a paraméteres esetben is történt az LSD összehasonlításnál (9.5. táblázat - Többszörös összehasonlítások LSD módszerrel a mintapélda adataira 9.5. táblázat). Ezek a páronkénti összehasonlítások ugyanúgy korrekcióra szorulnak tehát, mint az LSD-módszer. Emiatt az 1-4 összehasonlítás korrekció után biztosan nem lesz szignifikáns.

9.10. táblázat -

<table>
<thead>
<tr>
<th>Csoport</th>
<th>Adat</th>
<th>Rangszám</th>
<th>Rangszámösszegek</th>
<th>Rangszámátlagok</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flecainide</td>
<td>61</td>
<td>4,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flecainide</td>
<td>69</td>
<td>12</td>
<td>65</td>
<td>10,83</td>
</tr>
<tr>
<td>Control</td>
<td>61</td>
<td>4,5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Képezzük az összes lehetséges párt a rangszám-átlagokból és hasonlítsuk az átlagok különbségeit a most kiszámított mennyiségekhez. A 9.10. táblázat szerint a Quinidine különbözik a többi csoporttól. Az 1-4 összehasonlítás „határeset, mint ahogyan az a paraméteres esetben is történt az LSD összehasonlításnál (9.5. táblázat - Többszörös összehasonlítások LSD módszerrel a mintapélda adataira 9.5. táblázat). Ezek a páronkénti összehasonlítások ugyanúgy korrekcióra szorulnak tehát, mint az LSD-módszer. Emiatt az 1-4 összehasonlítás korrekció után biztosan nem lesz szignifikáns.

A számítógépes programok nem végzik el a fenti számítást, ezért végezzünk páronként Mann-Whitney U-próbával is és alkalmazzuk a Bonferroni korrekiót! Mivel a kontrollhoz 3 csoportot kell hasonlítnunk, Mann-Whitney U-próba által adott p-értékeket 3-mal kell megszoroznunk. Az eredmény a 9.11. táblázat - Többszörös összehasonlítás a mintapélda adataira. 9.11. táblázatban látható.

9.11. táblázat - Többszörös összehasonlítás a mintapélda adataira.

<table>
<thead>
<tr>
<th>Populációk</th>
<th>j</th>
<th>Rangszám-átlagok i</th>
<th>j</th>
<th>Különbség abszolút értéke</th>
<th>kifejezés jobb oldala</th>
<th>Páronkénti hasonlítás eredménye</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>5,32</td>
<td>10,42</td>
<td>5,22</td>
<td>5,262 (p&lt;0,05)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>5,2</td>
<td>10,83</td>
<td>5,63</td>
<td>5,262 (p&lt;0,05)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>20,42</td>
<td>10,42</td>
<td>10</td>
<td>5,017 (p&lt;0,05)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>20,42</td>
<td>10,83</td>
<td>9,59</td>
<td>5,017 (p&lt;0,05)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>10,42</td>
<td>10,83</td>
<td>0,41</td>
<td>5,017</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>5,2</td>
<td>20,42</td>
<td>15,22</td>
<td>5,262 (p&lt;0,05)</td>
<td></td>
</tr>
</tbody>
</table>

A számítógépes programok nem végzik el a fenti számítást, ezért végezzünk páronként Mann-Whitney U-próbával is és alkalmazzuk a Bonferroni korrekiót! Mivel a kontrollhoz 3 csoportot kell hasonlítnunk, Mann-Whitney U-próba által adott p-értékeket 3-mal kell megszoroznunk. Az eredmény a 9.11. táblázat - Többszörös összehasonlítás a mintapélda adataira. 9.11. táblázatban látható.

9.11. táblázat - Többszörös összehasonlítás a mintapélda adataira.

<table>
<thead>
<tr>
<th>Kontroll – Lidocaine</th>
<th>Mann-Whitney U-próba p-értéke</th>
<th>p-érték Bonferroni korrekcióval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kontroll – Flecainide</td>
<td>0,052</td>
<td>0,156</td>
</tr>
<tr>
<td>Kontroll – Quinidine</td>
<td>0,004</td>
<td>0,012</td>
</tr>
</tbody>
</table>

3. Több csoport összehasonlítására vonatkozó feladatok és kérdések

3.1. Feladat, számítógépes „játék” az Excel programmal.

Generáljunk 100 olyan mintát, amely ugyanolyan eloszlásból származik, és hasonlítuk össze őket t-próbákkal! Például: generáljunk 100 darab, N(20,2) eloszlásból származó, 10 elemű mintát!

Növeljük a minta elemszámát és nézzük meg a változást!

Változtassuk a szórást és figyeljük az eredményt!

Mivel magyarázzuk, hogy – bár csak véletlen ingadozás lehet a minták között, mégis kapunk szignifikáns különbséget?

Átlagosan hány szignifikáns különbséget várunk 5%-os szinten, 100 összehasonlítás esetén?
Útmutatás Az Eszközök menü Adatalemzés moduljában van egy véletlenszámgenerátor, aminek segítségével könnyen létrehozhatunk adott eloszlásból származó véletlen számokat. A 9.3. ábra tartalmaz egy ilyen próbálkozás részletét. „Véletlenül” a 100-ből éppen 5 lett szignifikáns, annyi, amennyit vártunk. Újabb próbálkozásnál lehet, hogy 4-at, 6-ot kapunk, tehát nem pontosan 5-öt, mivel 100-ből átlagosan 5 szignifikáns különbség várható, nem pontosan 5.

<table>
<thead>
<tr>
<th></th>
<th>Csoportokon belül</th>
<th>Teljes</th>
<th>Csoportok között</th>
</tr>
</thead>
<tbody>
<tr>
<td>Négyzetösszeg</td>
<td>395562,7</td>
<td>661356,4</td>
<td>265793,6</td>
</tr>
<tr>
<td>Szabadságfok</td>
<td>332</td>
<td>338</td>
<td>6</td>
</tr>
<tr>
<td>Variancia</td>
<td>1191,454</td>
<td>19,90794</td>
<td>44298,935</td>
</tr>
<tr>
<td>F</td>
<td>37,181</td>
<td>19,4932</td>
<td>37,181</td>
</tr>
<tr>
<td>p</td>
<td>0,0001</td>
<td>0,0001</td>
<td>0,0001</td>
</tr>
</tbody>
</table>

9.12. táblázat - Egy varianciaanalízis eredménye

<table>
<thead>
<tr>
<th>A szóródás oka</th>
<th>Négyzetösszeg</th>
<th>Szabadságfok</th>
<th>Variancia</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Csoportokon belül</td>
<td>395562,7</td>
<td>332</td>
<td>1191,454</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teljes</td>
<td>661356,4</td>
<td>338</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Csoportok között</td>
<td>265793,6</td>
<td>6</td>
<td>44298,935</td>
<td>37,181</td>
<td>0,0001</td>
</tr>
</tbody>
</table>

Hány csoportot hasonlítottak össze ebben a vizsgálatban?

Összesen hány adat szerepelt a vizsgálatban?
Mennyi a csoportok közötti variancia? Hányszorosza ez a csoportokon belüli varianciának?

Mennyi az F két szabadságfoka?

Van-e különbség az összehasonlított csoportok átlagai között?

3.3. Három féle tanítási módszer tesztelésére a hallgatókat véletlenszerűen 3 csoportba sorolták. A maximum 100 pontos teszteredmények a következők voltak:

1. csoport: 74, 84, 86, 87, 89, 90, 95, 97, 99
2. csoport: 81, 84, 86, 87, 87, 87, 91, 92, 93, 95, 98
3. csoport: 66, 73, 74, 84, 88, 90, 91, 96, 100

Van-e különbség a 3 csoport átlagos teljesítményében? Hasonlítsa össze a 3 csoport adatait a megfelelő módszerrel.

Útmutatás 3 független csoportot az eloszlásból származó adatok összehasonlításával lehet összehasonlítani. A pontoszámok 0 és 100 között helyezkednek el, az itt látható pontoszámok az intervallumnak inkább a felső felében, ezért gondolhatunk arra, hogy az eloszlás kissé ferde és elvileg nem lehet normális. Mégis, az eloszlást jobban megnevezve, nem igazán találunk kiugró értéket. Feltűnő, hogy a 3. csoportban nagyobb a szóródás, mint az első kettőben. Paraméteres módszert alkalmazva a Levene-teszt eredménye p = 0,061, „éppen” nem szignifikáns 5%-os szinten. A szórások azonosságát tehát kénytelenek vagyunk elfogadni. A varianciaanalízis p-értéke p = 0,492, a különbség nem szignifikáns a 3 csoport átlagában.

Mivel kétségeink támadhatnak a normalitás miatt és a szórások azonossága miatt is, Végezzünk inkább nemparaméteres analízist! A Kruskal-Wallis próba szerint p = 0,878, tehát a 3 csoport eloszlása azonosnak mondható, nincs szignifikáns különbség a tanítási módszerekben.

4. Összefoglalás
Ebben a fejezetben folytonos eloszlásból származó adatok összehasonlításával foglalkoztunk. Láttuk, hogy a módszer függ az eloszlástól: normális eloszlást feltételezve ún. paraméteres eljárásokat alkalmaztunk, egyébként pedig nemparaméteres próbákat. A módszer attól is függ, hogy összetartozó vagy független csoportokat hasonlítunk össze. Végül megállapítottuk, hogy kettőnél több csoport esetén nem alkalmazhatók a két csoportot összehasonlító módszerek, mivel ilyenkor megnőhet az összes csoportra (az egész kísérletre) vonatkozó első fajta hiba valószínűsége.

A módszerek alkalmazhatósága tehát nagyban függ bizonyos feltételektől. Ha ezek a feltételek nem teljesülnek, akkor is el tudjuk végezni a próbát, de téves eredményt kaphatunk. Pl. ki tudjuk számítani a t-értéket és a hozzá tartozó p-értéket nyilvánvalóan nemnormális eloszlás esetén is, ennek alapján azonban hibás következtetésre juthatunk. A feltételek teljesülésének a megállapítása nem is mindig egyszerű.

Természetesen vagy paraméteres, vagy nemparaméteres módszert kell alkalmaznunk, nem pedig mindkettőt. Mégis, ha kétségeink vannak a feltételek teljesülése felől, végezzük el a paraméteres és a nemparaméteres próbát is. Ha a feltételek teljesülnek, a két módszer nagyon hasonló eredményt fog adni. Ha a két módszer eredménye nagyon eltérő, akkor a nemparaméteres próba alkalmazása tűnik megbízhatóbbnak.

Ha a feltételek teljesülnek vagy csak kevésbé nem teljesülnek, akkor a paraméteres módszerek alkalmazása javasolt több okból is: jobban alkalmazhatók becslések és a confidencia-intervallumok kiszámítására és könnyen általánosíthatók bonyolultabb kísérelti elrendezések esetére.

5. Irodalomjegyzék 9. fejezethez
Kettőnél több csoport összehasonlítása folytonos adatok esetén


10. fejezet - KORRELÁCIÓ ÉS REGRESSZIÓ ANALIZIS

Dinya Elek dr.

1. Korreláció számítás

1.1. Bevezetés

A statisztikai alkalmazások gyakori problémáját képezik azok a vizsgálatok, melyek során azt nézzük, hogy egy vagy több független változó milyen hatással van a függő változóra, milyen erős a kapcsolat közöttük illetve hogyan írható le, fejezhető ki ez a kapcsolatot. A kapcsolat elemzésnek az elsőfajtját korreláció-, az utóbbit regressziószámításnak nevezzük. A regressziószámítás a változók közötti – véletlenszerű – kapcsolatban lévő törvény szerűségeket, tendenciát igyekszik kifejezni függvények (egyenletek) formájában. Mindegyik elemzési módszer sajátos feladattal bír, de szoros kapcsolatban is állnak egymással.

A korrelációszámítás a változók közötti kapcsolat erősségét vizsgálja. Nyilvánvaló, hogy a két vizsgálati módszer egymást kiegészíti: a változók közötti erős korreláció azt jelenti, hogy a regressziószámítással nyert függvényt használhatjuk a változók közötti kapcsolat jellemzésére, míg a gyenge korreláció épp az ellenkezőjére utal.


A vizsgálatokat aszerint csoportosítjuk, hogy a változók közötti kapcsolat lineáris vagy attól eltérő. Általában lineáris kapcsolatra törekszünk, mert ez a kapcsolat a legjobban érthető, ugyanakkor matematikailag is a legjobban kezelhető. Ha a probléma nem lineáris, akkor transzformációk (pl. logaritmus) segítségével megpróbáljuk azt lineárisra változtatni.

A továbbiakban a korrelációval, annak egyszerűbb formájával, a kétváltozós lineáris korrelációval ismerkedünk meg. A megismert tulajdonságok egyszerűen általánosíthatók a többszörös korrelációra is.

1.2. Lineáris korreláció (parametrikus korreláció)

A fogalom megértéséhez tekintsük a
10.1. pontfelhő ábrákat:

Az

10.1.a. és
10.1.b. ábrán az összetartozó \((x, y)\) pontpárok szorosan egy képzeletbeli egyenes mentén helyezkednek el. Az \(x\) értékek növekedésével a \(y\) értékek is nőnek, illetve az \(x\) értékek csökkenésével az \(y\) értékek csökkennek. Ilyen esetekben az \(x\) és \(y\) változó közötti korrelációt lineáris pozitív (10.1.a.) illetve lineáris negatív korrelációk (
10.1.b. ábrán) nevezzük. Ha a pontok egy egyenes mentén vagy szorosan mellette helyezkednek el, akkor a korreláció értéke maximális illetve maximum közeli érték. A

10.1.c. és
10.1.d. ábra lazább pozitív és negatív korrelációs kapcsolatot mutat. Az

10.1.e. ábra a változók közötti korrelálatlanságot mutatja, hiszen az egyik változó eloszlása teljesen független a másik eloszlásától. Az
10.1. ábra esetén a két változó közötti kapcsolatot nemlineárisnak nevezzük, mivel a pontpárok egy görbe vonal mentén helyezkednek el.

A korreláció szimmetrikus fogalom, nincs értelme a kapcsolatban kiemelni az egyik vagy a másik változót, hiszen a fordítottja is igaz. Az a tény, hogy a változók korrelátlanok még nem jelenti azt, hogy a változók függetlenek is egymástól. Korrelátlan változók között is lehet kapcsolat csak ezt a kapcsolatot nem tudjuk számszerűsíteni. Fordítva azonban igaz, hogy a független változók egyben korrelátlanok is, vagyis semmilyen kapcsolat nincs a két változó között. Az ilyen változók értékei egymástól függetlenül alakulnak. A korreláció számszerű értéke mellett érdemes a változók pontfelhő (scatter plot) diagramját is megérinteni, mert a két információ együttesen ad teljes felvilágosítást a változók viselkedéséről.

A lineáris korrelációs (vagy Pearson–féle) együttátható értékét a következő módon határozzuk meg:

$$ r = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2 \sum_{i=1}^{N} (y_i - \bar{y})^2}} $$

ahol
az $x_i$ értékek,
az $y_i$ értékek átlagait jelöli.

Az $r$ értéke dimenzió nélküli szám és a $[-1, 1]$ zárt intervallumban helyezkedik el: $-1$ esetén (maximális) negatív, $+1$ érték esetén (maximális) pozitív kapcsolációval beszélünk. Ha az $r = 0$, akkor a vizsgált két változó kapcsolatát korreláltanak (de nem függetlennek) nevezzük (lázd
10.1.e. ábra). A korrelációs értékeket $r \geq 0.7$ felett mondjuk erős kapcsolatnak, de az $r$ értékek valós tartalmat a szakmai megfontolások adják meg. Lehetnek olyan vizsgálatok, amelyben kisebb értékeknek is jelentősége lehet. A képlettben, a számíthatónál álló kifejezést kovarianciának (közös szóródásnak) nevezzük, amelynek szabadságfoka $(N-1)$. Tulajdonképpen a változók közötti kapcsolat jellemzésére a kovariancia értéke is megfelelő lenne, de nagyságát a változók értékei befolyásolják. Így a korrelációs együttthatók nem váltnak összehasonlíthatóvá. Ezért szerepel a nevezőben standardizáló tényezőként a két szórás szorzata. Ennek eredményeként lesz az $r$ értéke standardizált érték és változók összehasonlíthatóvá. Az $r^2$ értéket determinációs együttthatónak nevezzük, amely két variancia hányadosaként írható fel

$$r^2 = \frac{s_y^2}{s_y^2}$$

ahol $s_y^2$: $Y$ varianciának az a része, amit az $x$ független változó megmagyaráz, $s_x^2$: $Y$ teljes variációja

Hasonlóan írható fel a korreláció szimmetriája miatt $x$-re is

$$r^2 = \frac{s_x^2}{s_x^2}$$

Az $r^2$ értéke tehát azt fejezi ki, hogy az $X$ változó a $Y$ variációjának hány %-át magyarázza (hány százalékéért a felelős).

Minél magasabb az értéke, annál szorosabb a két változó között a lineáris kapcsolat. Az $r^2$ értéke 0 és 1 közötti szám és $r^2 l r$. A 10.1. táblázat az $r$ és $r^2$ összefüggését mutatja

### 10.1. táblázat -

<table>
<thead>
<tr>
<th>$r^2$ (%)</th>
<th>1</th>
<th>4</th>
<th>9</th>
<th>16</th>
<th>25</th>
<th>36</th>
<th>49</th>
<th>64</th>
<th>81</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>0,</td>
<td>0,2</td>
<td>0,3</td>
<td>0,4</td>
<td>0,5</td>
<td>0,6</td>
<td>0,7</td>
<td>0,8</td>
<td>0,9</td>
<td>1,0</td>
</tr>
</tbody>
</table>
KORRELÁCIÓ ÉS REGRESSZIÓ
ANALIZIS

A táblázat alapján pl. \( r = 0.9 \) esetében \( r^2 = 81\% \) ami azt jelenti, hogy az \( Y \) változó öszvarganciájának 81%-át magyarázza meg az \( X \) változó, ennyi %-ért a „felelős” és 19% varianciaszint más tényező (pl. mérési hiba) okozza.

Mik a lineáris korrelációsánitás feltételei?

Az \( Y \) és \( X \) változók mennyiségi (folytonos) változók legyenek és mindegyikük eloszlása normális eloszlást kövezen vagy, ha \( X \) értékei rögzítettek (pl. időpontok), akkor \( Y \) eloszlás \( X \) minden rögzített értéke esetén normális eloszlást kövenesen

Az összes kovariancia legyen lineáris.

Minden \( X \) értékehez \( Y \) érték is tartozzon.

Az \( X \) és \( Y \) értékeket egymástól függetlenül mérjük.

A minták választása legyen véletlen mintavétel.

Kiugró (outliers) értékek esetén nagy gondossággal kell eljárni, mivel ezek az adatok erőteljesen befolyásolják az \( r \) értékét. Már egy kiugró érték hatása is jelentős lehet. Ilyen esetekben még nagyobb hangsúlyt kap a pontfelhő ábra tanulmányozása.

1.2.1. Korrelációs együttható szignifikancia

A korrelációsánitatekésnél is tartsuk szem előtt, hogy a számítást a populációnból vett mintákkal végezzük, de az eredményt az egész populációra akarjuk érvényesíteni. Ha vesszük az \( X \) és \( Y \) változók összes populációbeli \( N \) számú mintáját, akkor az így kapott sokaságot kétváltozós sokaságnak nevezzük, amelyről feltételezzük a normális eloszlást. E kétdimenziós normális eloszlás korrelációját az elméleti korrelációs együttható méri, amit \( \rho \) jelölünk. A mintából meghatározott \( r \) ennek a elméleti korrelációs együtthatónak a becslése. A értéke a \([-1, 1]\) intervallum. Az \( r \) eloszlása nem szimmetrikus eloszlás, a \(-1, 0, +1\) értékek kivételével csak jól közelíti. Az \( r \) eloszlása épí a végpontok miatt ferde eloszlás, ami \( = 0 \) esetén válik szimmetrikussá. Ahhoz, hogy tényleg nyugodt legyünk az \( r \) értékét illetően, ellenőrzésére szignifikancia vizsgálatot kell végezni egy \( N-2 \) szabadágtörő t-statisztika felhasználásával:

\[
t = r \sqrt{\frac{N-2}{1-r^2}}
\]

Szignifikáns eltérés esetén a \( H_0: \rho = 0 \) hipotézist elvetjük és az \( r \) értékét valós lineáris kapcsolatnak minősítjük.

A \( 0 \) vagy \( \rho = 1 \) (0 populációs korrelációs érték) hipotézisek tesztelésénél az \( r \) eloszlása aszimmetrikus, de az ún. Fisher–féle Z transzformációval normális eloszlást kapunk

\[
Z = \frac{1}{2} \ln \left( \frac{1+r}{1-r} \right)
\]

10–es alapú logaritmus használata esetén

\[
z = 1,1513 \sum \ln \left( \frac{1+r}{1-r} \right)
\]

\( z = 1,1513 \) ? szorozva \( \ln (1 + r/1 - r) \)

Az eloszlás átlaga és szórása

\[
\mu_z = \frac{1}{2} \ln \left( \frac{1+\rho_1}{1-\rho_2} \right)
\]

és

\[
\sigma_z = \frac{1}{\sqrt{N-3}}
\]

Created by XMLmind XSL-FO Converter.
A $z$ értékét a korrelációs együthatható konfidencia intervallumának a meghatározására is felhasználhatjuk, amely 5%-os szignifikanciaszinten: alsó érték:

$$Z_A = z - \frac{1.96}{\sqrt{N-3}}$$

felső érték:

$$Z_F = z + \frac{1.96}{\sqrt{N-3}}$$

Az adatokat visszatranszformálva kapjuk az $r_A$ és $r_F$ értékeket és

$$r_F = \frac{e^{2Z_F} - 1}{e^{2Z_F} + 1}$$

A Z-transzformáció segítségével két korrelációs együthatható ($r_1$, $r_2$) eltérésének szignifikanciáját is teszthetjük a

$$Z = \frac{z_1 - z_2 - |\mu_{z_1} - \mu_{z_2}|}{\sqrt{\frac{1}{N_1-3} + \frac{1}{N_2-3}}}$$

képlet alapján, ahol $\mu_{z_1}$, $\mu_{z_2}$, az $r_1$ és $r_2$ együthathatók z eloszlásbeli átlagai, $\sigma_{z_1}$, $\sigma_{z_2}$: az $r_1$ és $r_2$ együthathatók z eloszlásbeli szórásainak különbsége:

$$\sigma_{z_1-z_2} = \sqrt{\frac{1}{N_1-3} + \frac{1}{N_2-3}}$$

Példa

A 10.2. táblázat - Egy tantárgyra fordított felkészülési idő és az elért eredmény 10.2. táblázat egy tantárgyra fordított felkészülési időt és az elért pontszámot mutatja. Határozzuk meg a változók közötti $r$ értéket.

**Megoldás:** Egészítsük ki az adatokat az alábbi számolásokkal

### 10.2. táblázat - Egy tantárgyra fordított felkészülési idő és az elért eredmény

<table>
<thead>
<tr>
<th>Ráfordítási idő ($x_i$)</th>
<th>Elérte pontszám ($y_i$)</th>
<th>($x_i$)²</th>
<th>($y_i$)²</th>
<th>($x_i$) ($y_i$) szorzva ($x_i$) ($y_i$)²</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>22</td>
<td>28.09</td>
<td>1672.81</td>
<td>216.77</td>
</tr>
<tr>
<td>1</td>
<td>38</td>
<td>18.49</td>
<td>620.01</td>
<td>107.07</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>18.49</td>
<td>723.61</td>
<td>115.67</td>
</tr>
<tr>
<td>3</td>
<td>61</td>
<td>5.29</td>
<td>118.81</td>
<td>25.07</td>
</tr>
<tr>
<td>3</td>
<td>61</td>
<td>5.29</td>
<td>3.61</td>
<td>4.37</td>
</tr>
<tr>
<td>4</td>
<td>57</td>
<td>1.69</td>
<td>34.81</td>
<td>7.67</td>
</tr>
<tr>
<td>4</td>
<td>61</td>
<td>1.69</td>
<td>3.61</td>
<td>4.37</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>1.69</td>
<td>8.41</td>
<td>3.77</td>
</tr>
<tr>
<td>5</td>
<td>67</td>
<td>0.09</td>
<td>16.81</td>
<td>-1.23</td>
</tr>
<tr>
<td>5</td>
<td>73</td>
<td>0.09</td>
<td>102.01</td>
<td>-3.03</td>
</tr>
<tr>
<td>6</td>
<td>65</td>
<td>0.49</td>
<td>4.41</td>
<td>1.47</td>
</tr>
</tbody>
</table>
KORRELÁCIÓ ÉS REGRESSZIÓ
ANALIZIS

<table>
<thead>
<tr>
<th>Ráfordítási idő (xᵢ)</th>
<th>Elért pontszám (yᵢ)</th>
<th>(xᵢ - x)²</th>
<th>(yᵢ - y)²</th>
<th>(xᵢ - x)² szorzva (yᵢ - y)²</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>72</td>
<td>0.49</td>
<td>82.81</td>
<td>6,37</td>
</tr>
<tr>
<td>7</td>
<td>73</td>
<td>2.89</td>
<td>102.01</td>
<td>17.17</td>
</tr>
<tr>
<td>7</td>
<td>71</td>
<td>2.89</td>
<td>65.61</td>
<td>13.77</td>
</tr>
<tr>
<td>8</td>
<td>79</td>
<td>7.29</td>
<td>259.21</td>
<td>43.47</td>
</tr>
<tr>
<td>8</td>
<td>83</td>
<td>7.29</td>
<td>404.01</td>
<td>54.27</td>
</tr>
<tr>
<td>10</td>
<td>84</td>
<td>22.09</td>
<td>445.21</td>
<td>99.17</td>
</tr>
<tr>
<td>12</td>
<td>85</td>
<td>44.89</td>
<td>488.41</td>
<td>148.07</td>
</tr>
<tr>
<td>12</td>
<td>93</td>
<td>44.98</td>
<td>906.01</td>
<td>201.67</td>
</tr>
<tr>
<td><strong>Átlag</strong></td>
<td>62.9</td>
<td><strong>12.11</strong></td>
<td><strong>371.19</strong></td>
<td><strong>62.98</strong></td>
</tr>
<tr>
<td><strong>Összeg</strong></td>
<td>1258</td>
<td>242.2</td>
<td>7423.8</td>
<td>1259.6</td>
</tr>
</tbody>
</table>

$$\sum_{i=1}^{12}(x_i - \bar{x})^2 = 242.2$$
$$\sum_{i=1}^{12}(y_i - \bar{y})^2 = 7423.8$$

Az r értéke a képletnek megfelelően

$$r = \frac{1259.6}{\sqrt{242.2 \cdot 7423.8}} = 0.9393$$

Az r érték szignifikanciájához használjuk a t-eloszlású statisztika képletét

$$t = 0.9393 \cdot \sqrt{\frac{20 - 2}{1 - 0.9393^2}} = 11.6151$$

Az eloszlás szabadságfoka df = \(N - 2 = 20 - 2 = 18\). Az 5%–os szinten a hozzátartozó kritikusérték 2,101. Mivel a \(t\)-érték nagyobb a kritikus értékénél, ezért a korreláció szignifikáns, értékét elfogadjuk, azaz, nem a véletlen miatt kaptunk ilyen nagy értéket. A korrelációs együttható 95%–os konfidencia intervalluma a megadott formulákkal könnyen meghatározható.

Tegyük fel, hogy a vizsgálatot \(N = 22\) fővel is elvégeztük és \(r_i = 0.85\) korrelációs értéket kaptunk. Vizsgáljuk meg, hogy a két korrelációs együttható között 5%–os szignifikanciaérték mellett van-e szignifikáns eltérés?

Először mindkét értékre határozzuk meg a Fisher–féle z értéket:

$$z_1 = \frac{1}{2} \ln \left( \frac{1 + r_1}{1 - r_1} \right) = \frac{1}{2} \ln \left( \frac{1 + 0.85}{1 - 0.85} \right) = 1.256$$

$$z_2 = \frac{1}{2} \ln \left( \frac{1 + r_2}{1 - r_2} \right) = \frac{1}{2} \ln \left( \frac{1 + 0.9393}{1 - 0.9393} \right) = 1.7321$$

A szórások eltérése

Created by XMLmind XSL-FO Converter.
Most azt tesztekjük, hogy a két minta átlagai nem térnek el szignifikánsan egymástól, vagyis \( \bar{z} \neq 0 \), így a standard normális eloszlás z értéke

\[
\begin{align*}
z &= \frac{z_1 - z_2 - 0}{\sigma_z} = \frac{1,256 - 1,7321}{0,3338} = -1,4263
\end{align*}
\]

Mivel a z-re iga a −1,96 z 1,96 reláció, ezért a \( H_0 \) hipotézist megtartjuk, tehát az \( r_1 = 0,85 \) és \( r_2 = 0,9393 \) korrelációs együttáthatók szignifikánsan nem térnek el egymástól.

### 1.3. Nemparaméteres korreláció

Ezeket a korrelációkat rangkorrelációknak is hivják.

#### 1.3.1. Spearman–féle rangkorreláció

A módszer a linéris korrelációs együttátható speciális esetének tekinthető. A kapcsolat szorosságának mérésére a két változó rangszámainak különbségét használjuk fel:

\[
r_s = 1 - \frac{6 \cdot \sum d_i^2}{N^2 - N}
\]

ahol \( d_i = x_i - y_i \), az x és y rangjainak különbsége; \( N \) = a mintaszám.

Az együttátható értékei a \( -1 \) és 1 intervallumba esnek: minél közelebb vannak ezek az értékek a \( -1 \)–hez vagy +1–hez, annál szorosabb a kapcsolat a két változó között. A \( -r_s \) értékei a kapcsolatot úgyis értelmezhető, hogy a két ismérve szerinti rangsor fordított sorrendben van.

Kapcsolt rangok estén az rs kiszámítása a következőképpen módosul

\[
r_s = \frac{1}{6} \left( N^3 - N \right) - \left( T_x + T_y \right) - \sum d_i^2
\]

ahol

\[
T = \sum_{j=1}^{i} \frac{1}{12} \left( t_j^3 - t_j \right)
\]

\( t \) : a kapcsolt rangok száma, \( j = 1, 2, 3, ... \), i az azonos rangsázmú csoportok száma.

A Pearson féle korrelációs együttáthatóhoz hasonlóan \( r_s \)-re is ellenőrizhető az a hipotézis (\( H_0 \)), hogy a populációbeli korrelációs együttátható 0, az alábbi t–statisztikával:

\[
t_s = r_s \sqrt{\frac{N - 2}{1 - r_s^2}}
\]

amely \( N \)–2 szabadságfokú t–eloszlást követ. Ha az így kiszámított \( t \) g a táblázatbeli kritikus értéknél, akkor az rs értéket a két változó kapcsolatának a jellemzésére használhatjuk. Ellenkező esetben nincs valós kapcsolat a két változó között.
Az r és a lineáris korrelációs együtható (r) eloszlása nagy mintaszám esetén azonos (hiszen a t-statistika is megegyezik). Ennek ellenére a két korrelációs együthatót nem szabad egymással helyettesíteni, mert jelentésük egészen más.

1.3.2. Kendall–féle rangkorreláció

A két változó kapcsolatát mérő együtható a Spearman–féle korrelációs együtható alternatívája. A számításhoz az egyes változók rangszámainak természetes sorrendjét vizsgáljuk, pl. X: 1, 2, 3, 4, 5; Y: 1, 5, 3, 2, 4.

Az X rangjai természetes sorrendben szerepelnek, míg az Y rangjai nem. Az Y változóban a rangok eltéréseinek a súlyát, az S értéket úgy határozzuk meg, hogy minden különböző Y rangpárhoz vagy a (+1) vagy a (–1) súlyt rendeljük annak megfelelően, hogy a párok adatai természetes sorrendben vannak-e vagy sem. Pl. az Y változó esetén az (1, 5) pár (+1) az (5, 3) pár (–1) súlyt kap. Ennek megfelelően a súlyok:

+1, +1, +1, +1, –1, –1, –1, –1, +1, +1

S a súlyok összege, így az S = 2. A súlyoknak megfelelően

$$S_{\text{max}} = \frac{1}{2} N(N + 1) = 1$$

, ha minden pár súlya (+1) és

$$S_{\text{min}} = -\frac{1}{2} N(N + 1) = -1$$

ha minden pár súlya (–1).

A értékét a következő formula határozza meg

$$\tau = \frac{S}{\frac{N(N-1)}{2}}$$

A értek a [–1, +1] intervallumban helyezkedik el: +1 érték jelenti, hogy a rangpárok sorrendje természetes és –1 jelenti a fordított sorrendet.

Kapcsolt rangok esetén az olyan Y pár, amelyben azonos kapcsolt rangok szerepelnek 0 súlyt kapnak, de ha két Y pár felett az X értékei kapcsoltak, szintén 0 súlyt kap az ilyen pár pl. X: 1, 2, 5, 2, 5, 4, 5, 5, 5; Y: 3, 2, 3, 5, 3, 5, 1, 6.

Rendre a súlyok: –1, +1, +1, –1, +1, 0, +1, –1, +1, 0, –1, +1, –1, +1, 0 = 2 = S

Kapcsolt rang esetén a

$$\tau = \frac{S}{\sqrt{\frac{1}{2} N(N - 1) - T_X} \sqrt{\frac{1}{2} N(N - 1) - T_Y}}$$

képlet alapján számoljuk, ahol TX az X változó, TY az Y változóinak kapcsolt rangjainak a számát jelenti:

$$T_X = \frac{1}{2} \sum t_i(t_i - 1)$$

és

$$T_X = \frac{1}{2} \sum t_j(t_j - 1)$$
A szignifikancia érékét a
\[
z = \frac{|S| - 1}{\sqrt{\frac{N(N-1)(2N+5)}{18}}}
\]
formula alapján határozzuk meg, amely standard normális eloszlást követ: 5%-os szignifikancia szinten −1,96 z +1,96 reláció estén a H₀ hipotézist megtartjuk, ellenkező esetben elvetjük. A H₀ hipotézis az, hogy a változók között nincs valós kapcsolat.

A Spearmann rs és a Kendall –féle korrelációs együthatók noha azonos feladatot látnak el, mégis különböznak. Ha ugyanazon az adathalmazon számítjuk ki őket, az r, értéke nagyobb lesz mint a értéke. A számítása bonyolultabb, különösen kapcsolt rangok esetén, ezért az ilyen problémák megoldását számítógéppel végezzük. A két értéket nem lehet összehasonlítani, mert más értelemmel bírnak.

2. Regresszióanalízis

2.1. Bevezetés

A biostatisztikai vizsgálatok során gyakran szükséges, hogy a változók közötti kapcsolatot függvény formájában fejezzük ki. Az egyik kitüntetett változót (függő változót) két vagy több változó (független változó) függvényeként (egyenleteként) akarjuk kifejezni. Az így meghatározott egyenlet azt fejezi ki, hogy a függő változót hogyan magyarázza, a többi változó, vagyis milyen hatással vannak rá. Pl. a szívinfarktus hogyan magyarázható a testsúly, magasvérnyomás, dohányzás stb. rizikófaktorokkal, milyen befolyást gyakorolnak a betegségre. A változók közötti kapcsolat egyenletszerű megismerése azért fontos, mert az egyenlet (a modell) használata általánosítja és függetleníti a probléma megoldást a mintaválasztástól.

A keresett egyenlet alakja a legkülönfélébb lehet. Az egyenlet kereséshez nagy segítséget nyújt a változók megoszlását mutató pontfelhő, mint azt a korrelációszámításnál látjuk. Ha a pontfelhő alakja egy elnyújtott ellipszishez hasonló, akkor a változók közötti kapcsolat egy lineáris egyenlettel kifejezhető. Ebben az esetben a pontok közé egy, a pontokra jól illeszkedő egyenes húzható. Az Excel programmal kényelmesen tudunk ilyen jellegű vizsgálatot végezni.

2.2. Kétváltozós lineáris regresszió

Kétváltozó közötti kapcsolatot leíró egyenlet alakja (}
10.2. ábra) \( y = a + bx \), ahol \( y \): a függő változó, \( x \): a független változó, \( a \): az \( y \) tengely metszete, \( b \): az egyenlet meredeksége (az szög tangense).

Az egyenletet az \( y \) változó \( x \) változóra vonatkozó regressziós egyenletének nevezzük. Hasonlóképpen lehet beszélni az \( x \) változó \( y \)-ra vonatkozó regresszójáról is, mert ekkor az \( x \)-t fejezzük ki az \( y \) függvényében (\( x = a + by \)). A

10.2. ábrával kapcsolatos megjegyzés, hogy ha a kapcsolat negatív, akkor az egyenes is fordított irányú lesz, és a \( b \) értéke negatív lesz. A pontokra legjobban illeszkedő egyenes megkeresésekor azt az egyenest kell választani a számos egyenes közül, amely esetében a megfigyelési adatoknak a regressziós egyenestől mért átlagos eltérése a legkisebb (\( Y_{D_i} \), a rezidum)

Az egyenes megkeresésére a négyzetes átlagot használjuk és magát a görbeillesztési eljárást leggyakrabban a legkisebb négyzetek módszerével végezzük. Legyen a keresett regressziósfüggvény alakja (a jelenti az egyenlet által becsült \( y \) értékeket)
A feladat úgy meghatározni az egyenlet \((a, b)\) paramétereit, hogy az \(Y\) értékek eltérésének négyzetösszege minimális legyen:

\[
y_D = \sum_{i=1}^{N} (y_i - \hat{y}_i)^2
\]

minimális

Helyettesítsük be az egyenletbe a regressziós függvény általános alakját

\[
y_D = \sum_{i=1}^{N} (y_i - (a + bx_i))^2
\]

minimális

kifejezést kapjuk. A feltételnek eleget tevő \(a\) és \(b\) értékét szélsőérték számítással kapjuk meg

\[
b = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N} (x_i - \bar{x})^2}
\]

és

\[
a = \bar{y} - b\bar{x}
\]

A \(b\) paraméter jelentése: az \(X\) független változó egységnyi változása milyen nagyságú változást okoz az \(Y\) függő változóban. Az \(a\) értéke a tengelymetszet magasságát adja.

A regressziószámítás feltétele, hogy az \(Y\) változó eloszlása legyen normális és a minta legyen random módon kiválasztva. Az \(X\) változóra egyedül a hibamentes adatfelvétel a kritérium.

### 2.2.1. Együttthatók konfidencia intervalluma

A korrelációszámításnál is láttuk, hogy az \(r\) értéke mintáról mintára változik. Hasonló a helyzet a regressziós együttthatókkal is, amelyek szintén mintáról mintára változnak. A populáció paraméterei legyenek \(A\) és \(B\), ekkor a regressziós együttthatók változó értékeire igaz, hogy \(M(a) = A; M(b) = B\).

Az együttthatók standard hibája:

\[
s_a = R_e \sqrt{\frac{\sum_{i=1}^{N} x_i^2}{N\sum_{i=1}^{N} (x_i - \bar{x})^2}}
\]

és

\[
s_b = \frac{R_e}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2}}
\]

, ahol
A regressziós együthatók normális eloszlásúak.

Az együthatók konfidencia intervallumai: \( b \pm t_{N-2} \cdot s_b \) és a \( t_{N-2} \cdot s_a \) szorozva és a \( t_{N-2} \cdot s_b \) szorozva.

A regressziós együtható b értékének szignifikanciájára vonatkozó statisztika:

\[
 t = \frac{b}{s_b},
\]

amely \( N - 2 \) szabadságfokú t-elozlást követ. A t statisztikával azt teszteljük, hogy \( H_0: b = 0 \) (\( H_1: b \neq 0 \))

### 2.2.2. A regresszió ANOVA táblázata

Az \( Y \) függő változó variációját az alábbi módon bonthatjuk fel

\[
 \sum_{i=1}^{N} (y_i - \bar{y})^2 = \sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{N} (y_i - \hat{y}_i)^2
\]

ahol a második tag a reziduális négyzetösszeg, amely az \( y_i \) értékek regressziósegényes körüli szóródását fejezi ki. Nyilván, minél nagyobb ez a rész, annál rosszabb a regresszió hatásfoka, annál kevésbé érvényesül az \( X \) hatása.

A regressziós összefüggés szignifikanciáját az ANOVA táblázat (10.3. táblázat - Regressziós ANOVA tábla 10.3. táblázat) alapján vizsgálhatjuk. Minden regressziós vizsgálat eredményéhez ez a táblázat is szorosan hozzájárul: a táblázat p értéke alapján döntünk arról, hogy a regressziós vizsgálat érvényes-e vagy sem.

A regresszió eredményének tanulmányozását is ezzel a táblázattal kell kezdeni, ugyanis a \( H_0: \) nincs kapcsolat \( X \) és \( Y \) változók között; \( H_1: \) van kapcsolat X és Y változók között.

Ha az eredmény szignifikáns (p \( \leq 0,05 \)), akkor fogadhatjuk csak el a valósnak a változók közötti kapcsolatot.

### 10.3. táblázat - Regressziós ANOVA tábla

<table>
<thead>
<tr>
<th>Forrás</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regresszió</td>
<td>( \sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2 )</td>
<td>1</td>
<td>( SS_R )</td>
<td>( \frac{SS_R}{SS_H} )</td>
</tr>
<tr>
<td>Reziduális (hiba)</td>
<td>( \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 )</td>
<td>( N - 2 )</td>
<td>( SS_H )</td>
<td>( \frac{SS_H}{N - 2} )</td>
</tr>
<tr>
<td>Total</td>
<td>( \sum_{i=1}^{N} (y_i - \bar{y})^2 )</td>
<td>( N - 1 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Az ANOVA táblázat felépítése

Az F értékét a df = 1, N – 2 szabadságfoknál keressük ki a táblázatból.
Példa: Határozzuk meg a korrelációs fejezetben megadott példa változóinak lineáris regressziós egyenletét, ANOVA táblázatát és a paraméterek 95%-os konfiddenciaintervallumát.

Megoldás: Először vizsgáljuk meg a változók által meghatározott pontfelhő diagrammot (10.4. ábra)

A pontok elhelyezkedése az

10.1.a. ábrához nagyon hasonló, így bátran alkalmazhatjuk a regressziós analízist.

A táblázat számolt értékeit behelyettesítve kapjuk
KORRELÁCIÓ ÉS REGRESSZIÓ
ANALIZIS

\[
b = \frac{\sum_{i=1}^{12} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{12} (x_i - \bar{x})^2} = \frac{1259,6}{242,2} = 5,2006
\]

\[
a = \bar{y} - b \cdot \bar{x} = 62,9 - 5,2006 \cdot 5 = 35,3365
\]

A keresett regressziós függvény tehát: Elért pontszám = \( y = a + bx = 35,3365 + 5,2006 \cdot x \)

Az ANOVA táblázat elkészítéséhez az alábbi adatokra (10.4. táblázat - Adatok az ANOVA táblázathoz 10.4. táblázat) van szükségünk:

### 10.4. táblázat - Adatok az ANOVA táblázathoz

<table>
<thead>
<tr>
<th>xi</th>
<th>yi</th>
<th>( y^* = 35,34 + \frac{5,2006x}{5,2006} )</th>
<th>((\bar{y}_i - \bar{y})^2)</th>
<th>((y_i - \bar{y}_i)^2)</th>
<th>(y_i - \bar{y})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>26</td>
<td>35,3365</td>
<td>759,74653</td>
<td>87,170232</td>
<td>1361,61</td>
</tr>
<tr>
<td>0</td>
<td>22</td>
<td>35,3365</td>
<td>759,74653</td>
<td>177,86553</td>
<td>1672,81</td>
</tr>
<tr>
<td>1</td>
<td>38</td>
<td>40,5371</td>
<td>500,0993</td>
<td>6,4368764</td>
<td>620,01</td>
</tr>
<tr>
<td>1</td>
<td>36</td>
<td>40,5371</td>
<td>500,0993</td>
<td>20,585276</td>
<td>723,61</td>
</tr>
<tr>
<td>3</td>
<td>52</td>
<td>50,9383</td>
<td>143,08227</td>
<td>1,1272069</td>
<td>118,81</td>
</tr>
<tr>
<td>3</td>
<td>61</td>
<td>50,9383</td>
<td>143,08227</td>
<td>101,23781</td>
<td>3,61</td>
</tr>
<tr>
<td>4</td>
<td>57</td>
<td>56,1389</td>
<td>45,712473</td>
<td>0,7414932</td>
<td>34,81</td>
</tr>
<tr>
<td>4</td>
<td>61</td>
<td>56,1389</td>
<td>45,712473</td>
<td>23,630293</td>
<td>3,61</td>
</tr>
<tr>
<td>4</td>
<td>60</td>
<td>56,1389</td>
<td>45,712473</td>
<td>14,908093</td>
<td>8,41</td>
</tr>
<tr>
<td>5</td>
<td>67</td>
<td>61,3395</td>
<td>2,4351602</td>
<td>32,04126</td>
<td>16,81</td>
</tr>
<tr>
<td>5</td>
<td>73</td>
<td>61,3395</td>
<td>2,4351602</td>
<td>135,9672</td>
<td>102,01</td>
</tr>
<tr>
<td>6</td>
<td>65</td>
<td>66,5401</td>
<td>13,250328</td>
<td>2,371908</td>
<td>4,41</td>
</tr>
<tr>
<td>6</td>
<td>72</td>
<td>66,5401</td>
<td>13,250328</td>
<td>29,810508</td>
<td>82,81</td>
</tr>
<tr>
<td>7</td>
<td>73</td>
<td>71,7407</td>
<td>78,157976</td>
<td>1,5858365</td>
<td>102,01</td>
</tr>
<tr>
<td>7</td>
<td>71</td>
<td>71,7407</td>
<td>78,157976</td>
<td>0,5486365</td>
<td>65,61</td>
</tr>
<tr>
<td>8</td>
<td>79</td>
<td>76,9413</td>
<td>197,15811</td>
<td>4,2382457</td>
<td>259,21</td>
</tr>
<tr>
<td>8</td>
<td>83</td>
<td>76,9413</td>
<td>197,15811</td>
<td>36,707846</td>
<td>404,01</td>
</tr>
<tr>
<td>10</td>
<td>84</td>
<td>87,3425</td>
<td>579,43581</td>
<td>11,172306</td>
<td>445,21</td>
</tr>
<tr>
<td>12</td>
<td>85</td>
<td>97,7437</td>
<td>1214,0834</td>
<td>162,40189</td>
<td>488,41</td>
</tr>
<tr>
<td>12</td>
<td>93</td>
<td>97,7437</td>
<td>1214,0834</td>
<td>22,50269</td>
<td>906,01</td>
</tr>
<tr>
<td>Összeg</td>
<td></td>
<td>6550,5994</td>
<td>873,0479</td>
<td>7423,8</td>
<td></td>
</tr>
</tbody>
</table>

Az adatok alapján az ANOVA tábla

### 10.5. táblázat - ANOVA tábla

<table>
<thead>
<tr>
<th>Forrás</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
</table>

Created by XMLmind XSL-FO Converter.
Az $F_{1,18} = 4.4139$ kritikusérték $1 F = 135.06$ ami erős szignifikanciát jelez, tehát a változók között a kapcsolat valós.
11. fejezet - Döntések előfordulási gyakoriságokról. Kontingencia táblázatok

Kaposi András dr.

1. Nominális változók és a kontingencia táblázat

Az adataink, a változók osztályozásáról a 2. fejezetben volt szó. A legegyszerűbb csoportot a kategóriás és azon belül pedig a nominális adatok képviselik. Az egyes adatok között mindössze az egyezésről illetve a nem egyezésről kell tudni dönteni. Ilyen nominális adat pl. a nem. Extrém esetekből eltekintve bárkirel megállapítható, hogy a nők vagy a férfiak csoportjába tartozik. Ha változónak mindössze két értéke van (mint a nem esetén), akkor azt kétértékűnek, dichotomikusnak vagy binárisnak nevezzük. Ha valamilyen módszert kidolgozunk az egyszerűbb adattípusra, akkor azt a módszert használhatjuk vagy általánosíthatjuk az összetettebbre is.


Gyakorisági táblázatok. A nominális (vagy akként kezzelt) adatok esetén a statisztikai vizsgálódás első lépése többnyire szintén ezen „hétköznapi” táblázat összeállítása. Második lépés (illetve első, ha az előbbi lépést átugorjuk) egy speciális táblázat létrehozata, amely nem az adatok pusztá felsorolása, hanem azok előfordulási gyakoriságainak a táblázatba rendezése. Nézzük a legegyszerűbb ilyen példát! Egy felnőtteket gyógyító háziorvos kigyűjt, hogy egy adott napon az őt felkereső betegek neme milyen volt (nő, nő, nő, férfi, nő, férfi, nő, nő, nő, férfi, nő, férfi, nő, nő, nő, férfi, nő, férfi, nő, férfi, nő, férfi, nő, férfi, nő, férfi, nő), majd ebből készíts „hétköznapi” (11.1. táblázat) és gyakorisági (11.2. táblázat - Gyakorisági táblázat 11.2. táblázat) táblázatokat.

11.1. táblázat - "Hétköznapi” táblázat

<table>
<thead>
<tr>
<th>n:</th>
<th>nő: f</th>
<th>férfi</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>n</td>
<td>f</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>f</td>
</tr>
<tr>
<td>n</td>
<td>f</td>
<td>n</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>f</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>f</td>
</tr>
<tr>
<td>n</td>
<td>f</td>
<td>n</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>f</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>f</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>f</td>
</tr>
</tbody>
</table>

11.2. táblázat - Gyakorisági táblázat

<table>
<thead>
<tr>
<th>nem</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>nő</td>
<td>férfi</td>
</tr>
</tbody>
</table>

Átláthatóbb a „hétköznapi” táblázat az egyszerű felsorolásnál (pl. egyforma hosszúak a cellák, nincs új sorban folytatódás), de sokkal tömörébb és talán az is mondható, hogy szebb a gyakorisági táblázat, mint akár a felsorolás, akár pedig a 11.1. táblázat - "Hétköznapi” táblázat 11.1. táblázat. Minden esetben a statisztikai gondolkodás számára ez a 11.2. táblázat - Gyakorisági táblázat 11.2. táblázat egy egyszerősített táblázat. A nem egy nominális valósizinüségi változó, amely két értéket vehet fel (kétértékű, dichotom, bináris változó), ezért kettő az oszlokok száma. A táblázatnak van még fejléje (vagy „fej”-sora) ebben van a nő és a férfi, de ezek
Döntések előfordulási gyakoriságokról. Kontingencia táblázatok

A táblázatok egyszerű és felismerhető, rendszerint rendelkeznek több oszloppal és sorral. A táblázatok az egyes változók összefüggéseit illusztrálnak, és azaz a válaszadók változásai és a változók összefüggéseit írják le. A táblázatok segítségével készen állnak a felszámolás és a feltételes valószínűségek megadására.

11.3. táblázat - Normális valószínűségi változó

<table>
<thead>
<tr>
<th>a kockafeldobás eredménye</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>gyakoriságok</td>
<td>14</td>
<td>16</td>
<td>19</td>
<td>14</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>

Ez a felírás nemcsak tömöríti az adatokat, hanem kérdéseink feltevésére ösztönöz. Mielőtt azonban a kérdéseinket feltennénk, mutassuk be az egyszerű vagy más néven egydimenzióis gyakorisági táblázatot utáni következő legegyszerűbb esetet is, amely nyilván a kétdimenziós gyakorisági táblázat lesz. A kétdimenziós táblázatoknál nemcsak az oszlopok száma nagyobb egynél, hanem a sorok száma is. A kétdimenziós (two-dimensional) táblázatot a kétszemélyes (two way) táblázatoknak is nevezik. Ezek között a legegyszerűbb táblázatnak két sora és két oszlopa van. Használjunk egy legegyszerűbb, modellt (11.4. táblázat - Kétdimenziós gyakorisági táblázat 11.4. táblázat) Tegyük fel, hogy a példánkbeli háziorvos körzetében – most nem tárgyalt szociológiai okok miatt – ha egy orvos felkereső beteg nem kap gyógyszert, akkor úgy érzi, hogy nem törődnek vele. Mivel a háziorvosunk úgy gondolja, hogy a beteg a vele való törődéstől is gyógyul, ezért minden hozzá forduló betegnek „gyógyszert” ír fel. Akin – az orvos legjobb meggyőződése szerint – nincs szüksége hatóanyagot tartalmazó „gyógyszert”, annak hatóanyagot nem tartalmazó „gyógyszert” (azaz placebót) ír fel, akinek pedig szüksége van rá, az hatóanyagot tartalmazó gyógyszert (most röviden hatóanyagosat) kap. Az egyéb speciális esetektől most tekintsünk el!

11.4. táblázat - Kétdimenziós gyakorisági táblázat

<table>
<thead>
<tr>
<th>a „gyógyszer” fajtája</th>
<th>nem</th>
</tr>
</thead>
<tbody>
<tr>
<td>placebo</td>
<td>12</td>
</tr>
<tr>
<td>hatóanyagos</td>
<td>7</td>
</tr>
<tr>
<td>nó</td>
<td>9</td>
</tr>
<tr>
<td>férfi</td>
<td></td>
</tr>
</tbody>
</table>

A táblázatot 2-szer 2-esnek (kétszerszettesnek) hívjuk, vagy a szorzást kiértékelve négymezősnek (1. megvastagított keret nélkül). Mivel két különböző szempont alapján csoportosítottuk a gyakoriságotokat, ezúttal nemcsak fejléc (féjsor) van a nő illetve férfi kategóriákkal, hanem fejoszlop is placebó illetve hatóanyagos kategóriákkal. Sokszor még a szempontok (vélemények) neveit is szerepeljük a féjsor alatt és a fejoszlop mellett (nem illetve a „gyógyszer” fajtája). Ismét megjegyezzük, hogy egyenértékű táblázatot nyernénk, ha az oszlop és sorkategóriákat felcserélnénk.

Kontingencia táblázat. A kétdimenziós gyakorisági táblázatot (frequency table) leggyakrabban kontingencia táblázatnak (contingency table, ct) hívják. Az érintést, érintkezést (com/con + tangere) jelentő kontingencia szó arra utal, hogy az egyes cellák gyakorisága a két kategóriá „érintkezése”, együttes fennállása esetén érvényes gyakoriság. A 2. fejezetben tárgyalt feltételes valószínűség fogalmának felhasználásával lehet egzaktabbal megfoglalni, hogy hogyan kapjuk a cellák gyakoriságait. A kontingencia átvitt értelemben feltételességet, eshetőséget, lehetőséget, véletlenséget is jelent. Szigorú értelemben kontingencia táblázat a legalább kétdimenziós gyakorisági táblázat, tágabb értelemben az irodalomban az egydimenziós gyakorisági táblát is annak hívják. A későbbiekben mi is ez utóbbi szóhasználatot követjük.

Szokásos kérdések. A gyakorisági táblázatok segítségével többfajta kérdésre kaphatunk választ. Ezek közül a példánk alapján a következőket tárghyújuk meg.
Döntések előfordulási gyakoriságokról. Kontingencia táblázatok

A kétdimenziós kontingencia táblázatot meghatározó változók (pl. a nem illetve a kapott „gyógyszer” fajtája, 11.4. táblázat - Kétdimenziós gyakorisági táblázat 11.4. táblázat) között van-e kapcsolat, vagy pedig függetlenek egymástól? A módszer neve ilyenkor kapcsolatvizsgálat vagy függetlenségvizsgálat.


2. Függetlenségvizsgálat khi-négyzet próbával 2-szer 2-es táblázat esetén

A két nominális (vagy akként kezelt) valószínűségi változó függetlenségvizsgálatát a 11.4. táblázat - Kétdimenziós gyakorisági táblázat 11.4. táblázatban összefoglalt példa alapján beszélünk meg. Kicsit kibővített formában az 11.5. táblázat - Kontingencia táblázat a megfigyelt (observed) gyakorisági értékek alapján. A kibővítés a cellák és a tulajdonságok betűkkel való jelzésére illetve a marginális gyakoriságok kiszámítására vonatkozik. Ez utóbbiak a sor és oszlopösszegeket jelentik (v.ö. peremvalószínűségek a 2. fejezetben).

<table>
<thead>
<tr>
<th>X:</th>
<th>Y:</th>
<th>a</th>
<th>b</th>
<th>a + b</th>
<th>c</th>
<th>d</th>
<th>c + d</th>
<th>a + c</th>
<th>b + d</th>
<th>a + b + c + d = n =</th>
</tr>
</thead>
<tbody>
<tr>
<td>X:</td>
<td>nem</td>
<td>12</td>
<td>3</td>
<td>15</td>
<td>7</td>
<td>9</td>
<td>16</td>
<td>19</td>
<td>12</td>
<td>31</td>
</tr>
</tbody>
</table>

Az orvoshoz forduló betegek alapsokaságán két tulajdonságot (valószínűségi változót) értelmeztünk. Ezek pedig a nem (X-szel jelölt változó) és a kapott „gyógyszer” fajtája (Y-nal jelölt változó). Elfordított-e az a feltételezésünk, hogy az X és Y valószínűségi változók függetlenek? Más szavakkal: elfordított-e az a feltételezés, hogy a placebót illetve hatóanyagos gyógyszert kapott páciensek csoportja és a páciensek neme függetlenek egymástól, azaz nincs kapcsolat közöttük? Ha a placebó illetve hatóanyagos gyógyszer kapása független a nemtől, akkor a nők és férfiak között azonos arányban kell lenni a placebót illetve hatóanyagos gyógyszert kapott pácienseknnek. A 6. fejezetben tárgyalt hipotézisvizsgálatok megfelelően itt a nullhipotézis (H₀) az, hogy a kapott gyógyszer fajtája független a nemtől. Az alternatív hipotézis (Hₐ) pedig az, hogy a kapott gyógyszer fajtája nem független a nemtől. Azaz: Hₐ: a kapott gyógyszer fajtája függ a nemtől. Sejtésünk az, hogy a háziorvos megítélése alapján a nők nagyobb arányban kapnak placebót, mint a férfiak.

Vizsgáljuk meg, hogy milyen gyakoriságokat kellene kapunk, ha igaz a nullhipotézis!

Mekkora lenne pl. a várt gyakoriság (expected frequency) a bal felső cellában (azaz a placebót kapott nők száma), ha a H₀ igaz?


Mivel a H₀ fennállása esetén független eseményekről van szó, ezért annak a valószínűsége, hogy egy nő placebót kapott, megegyezik a mintán belül női arány és a mintán belüli placebót kapott arányának szorzatával, azaz független valószínűségeket össze kell szorzni (v.ö. 2. fejezet). A számmunka fontos gyakoriság-értékeket, azt a gyakoriságot, amelyet függetlenség esetén kapnánk – az ún. várható gyakoriságot - pedig úgy kapjuk meg, hogy a szorzatékként megkapott valószínűséget megszorozzuk a minta elemeinek számával. Tehát ha a H₀ igaz, akkor:
Döntések előfordulási gyakoriságokról. Kontingencia táblázatok

A bal felső cella várható gyakoriság a: \[ a' = \frac{a + c}{n} \cdot \frac{a + b}{n} \cdot n = \frac{(a + c) \cdot (a + b)}{n} \]

A jobb felső cella várható gyakoriság a: \[ b' = \frac{b + d}{n} \cdot \frac{a + b}{n} \cdot n = \frac{(b + d) \cdot (a + b)}{n} \]

A bal alsó cella várható gyakoriság a: \[ c' = \frac{a + c}{n} \cdot \frac{c + d}{n} \cdot n = \frac{(a + c) \cdot (c + d)}{n} \]

A jobb alsó cella várható gyakoriság a: \[ d' = \frac{b + d}{n} \cdot \frac{c + d}{n} \cdot n = \frac{(b + d) \cdot (c + d)}{n} \]

Konkréttan: a bal felső cella: 19x15/31 = 9,19, a bal alsó cella: 19x16/31 = 9,81, a jobb felső cella: 12x15/31 = 5,81, a jobb alsó cella: 12x16/31 = 6,19. Az így kiszámított számértékekből egy segéd-kontingencia táblázatot készítünk (11.6. táblázat - Segéd-kontingencia táblázat a várható (expected) gyakoriság értékekkel 11.6. táblázat).

11.6. táblázat - Segéd-kontingencia táblázat a várható (expected) gyakoriság értékekkel

<table>
<thead>
<tr>
<th></th>
<th>X: nő</th>
<th>X: férfi</th>
<th>összes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>„gyógyszer”fajtája</td>
<td>(a' = 9,19)</td>
<td>(b' = 5,81)</td>
<td>(a' + b' = 15)</td>
</tr>
<tr>
<td>Y: 2</td>
<td>placebó</td>
<td>(c' = 9,81)</td>
<td>(d' = 6,19)</td>
</tr>
<tr>
<td>összes</td>
<td>(a' + c' = 19)</td>
<td>(b' + d' = 12)</td>
<td>(a' + b' + c' + d' = n = 31)</td>
</tr>
</tbody>
</table>

Vegyük észre, hogy a várható gyakoriságokat a következő módon kaptuk meg: várható gyakoriság = oszlopösszeg szorozva sorösszeg/a minta elemszáma

A segéd-kontingencia táblázat peremgyakoriságai természetesen megegyeznek az eredeti kontingencia táblázat peremgyakoriságaival.

Ha igaz a nullhipotézis, akkor a megfigyelt és a várható gyakoriságok nem nagyon térnek el egymástól. Hogyan konstruálható meg egy olyan szám (statisztika), amellyel ezt lehet teszteni? Megmutatható, hogy ha a megfigyelt és várt gyakoriságok különbség-négyzetéit osztjuk a várt gyakoriságokkal, majd ezeket összegezzük, akkor egy olyan statisztikát kapunk, amely khi-négyzet eloszlást követ (lásd 3. fejezet) és rendelkezik a fenti tulajdonságokkal. Formulával kifejezve:

\[ \chi^2 = \sum_{i} \frac{(O_i - E_i)^2}{E_i}, \]

ahol \(O\) a megfigyelt (observed), \(E\) pedig a várt (expected) gyakoriságot jelenti az i-dik cellában. Az összegzést minden cellára el kell végezni (a példánkban i = 1-től 4-ig). A szabadsági fok 2x2-es táblázat esetén 1, általában pedig úgy kapjuk meg, hogy kiszámítsuk az (oszlopok száma –1)x(sorok száma –1) szorzatot. (Ellenőrzés: 2-szer 2-es táblázattal: \((2 - 1) \cdot (2 - 1) = 1\)

Mielőtt a kiértékelést elvégeznénk, beszéljük meg a teszt alkalmazási feltételeit!

Az alkalmazhatóság feltételei. Szükséges, hogy eléggé nagy legyen a mintaelemszám. Ezt a feltételt a várható gyakoriságok segítségével fogalmazzuk meg: a segéd-kontingencia táblázatban levő várható gyakoriságok között nem lehet olyan cella, amelynek értéke kisebb mint 1, valamint 1 és 5 közötti értékből is maximum az összecellaszám 20%-a lehet. (Négyezősz táblázatra ez azt jelenti, hogy mindegyik várható gyakoriságnak 5-nél nagyobbaknak kell lenni.)

149
Döntések előfordulási gyakoriságokról. Kontingencia táblázatok

Azt is megjegyezzük, hogy a khi-négyzet teszt alkalmazási köre azért is olyan széles, mert a változók eloszlására nincs semmiféle felétel, valamint a legegyszerűbb adatok, azaz nominális adatok esetén is alkalmazható. Ha a változó folytonos lenne, akkor osztályok bevezetésével kategorizálhatunk. A módszer sem nem igényli, sem nem használja ki az adatok közötti esetleges sorrendiséget vagy azt, hogy akár kivonás és az osztás is értelmezett közöttük.

A 11.6. táblázat - Segéd-kontingencia táblázat a várható (expected) gyakoriság értékekkel 11.6. táblázat alapján a legkisebb várható gyakoriság (férfi és placebót kapott): 5,81 g 5, azaz megfelel az alkalmazhatósági előírásnak.

Az első fajta hiba megadása. A maximálisan méltányolható első fajta hibának vegyük a szokásos 5%-ot (? = 0,05)! Ez azt jelenti, hogy elvethetjük a nullhipotézist, ha kiszámolt khi^2 érték olyan nagy, hogy annak a valószínűsége, hogy ez az érték az adott szabadsági fokú khi^2 eloszláshoz tartozik, kisebb vagy egyenlő mint 5%. Az

![f(χ^2)](image)

11.1. ábrán ez annak felel meg, hogy a ?^2 számértékre g ?^2 ; ?

Az ? = 5% első fajta hibához az egy szabadsági fokú khi-négyzet eloszlás esetén a 3,841-es számérték tartozik. Lássuk, hogy mekkora a példánkbeli számított khi-négyzet érték!

A próbastatisztika kiszámítása. A 11.5. táblázat - Kontingencia táblázat a megfigyelt (observed) gyakorisági értékekkel 11.5. és 11.6. táblázat - Segéd-kontingencia táblázat a várható (expected) gyakoriság értékekkel 11.6. táblázat gyakorisága alapján végezzük el a kijelölt műveleteket! khi^2 = (12–9,19)/9,19+(3–5,81)/5,81 +(7–9,81)/9,81+(9–6,19)/6,19 = 4,288

Mivel 4,288 g 3,841, a
Döntések előfordulási gyakoriságokról. Kontingencia táblázatok

11.1. ábrán vázolt eset áll fenn, azaz elvetjük a nullhipotézist. Így nem fogadható el az a feltételezésünk, hogy az X és Y valószínűségi változók függetlenek. Más szavakkal: nem fogadható el az a feltételezés, hogy a placebót illetve hatóanyagos gyógyszert kapott páciensek csoportjai és a páciensek neme függetlenek egymástól. Elfogadjuk viszont az alternatív hipotézist (HA-t): a kapott gyógyszer fajtája nem független a nemtől, más szavakkal a kapott gyógyszer fajtája függ a nemtől. Egészen biztosak azonban nem lehetünk dolgunkban. A

![Graph](image1)

11.1. ábra: jobbra vonalkázott terület mutatja meg, hogy mekkora annak a valószínűsége, hogy elvetettük a nullhipotézist, mégis igaz (ez következik az első fajta hiba definíciójából, v. HIV). Egyelőre csak annyit tudunk, hogy ez a valószínűség kisebb, mint 5%, ami az általunk választott maximálisan méltányolható hibaszázalék. Ha a khi-négyzet táblázatunk tartalmaz további valószínűségeket, akkor pontosíthatjuk a tényleges első fajta hiba számértékét. Az egyik rendelkezésre álló táblázat szerint (Hajtman jegyzet) a 6.635-ös khi-négyzet értékhez 0,01 = 1 % hiba valószínűség tartozik. Mivel a mintánkból számított khi-négyzet érték (4,288) kisebb, mint ez az érték, 1%-esnek választott maximálisan méltányolható hiba esetén nem vethettük volna el a nullhipotézist. Ennek ismeretében annyit mondhatunk, hogy a tényleges első fajta hiba kisebb ugyan 5%-nél, de nagyobb, mint 1%.

A teszt elvégzése Excel táblázatkezelővel. A táblázatkezelő programok sikeréhez az is hozzájárult, hogy viszonylag könnyű velük számos statisztikai vizsgálatot elvégezni. A most tárgyalás alatt levő khi-négyzet teszt elvégzésére az Excel táblázatkezelő programban egy egyszerű beépített függvény áll rendelkezésre

Created by XMLmind XSL-FO Converter.
A felső és alsó része ugyanannak az Excel táblázatnak két különböző verziójú megjelenítése. A B3 és C4 közötti területre írótak be a megfigyelt gyakorisági adatokat. Tehát ez a kontingencia táblázat helye. Egyaránt látjuk a felső és az alsó táblázatban is, a számok cellákon belüli balra illetve jobbra igazításával ne törődjünk! A felső táblázatban azt látjuk, amit beírtunk: tőle vagy számokat, akkor azokat, ha végrehajtandó utasítást (képletet), akkor a képletet magát. Pl. a D3-as cellába az a képlet került, hogy „=SUM(B3:C3)”, azaz szummázzuk (adjuk össze) a B3 és C3 közötti cellatartalmakat. Ezzel az utasítással számolatuk ki a sorok peremgyakoriságait. Az alsó Excel táblázatnál a szokásos megjelenítést használtuk: minden cellában a kiértékelés cellatartalmak vannak, ha szöveg vagy szám a beírá, akkor a cellatartalom kiértékelése maga a szöveg vagy szám, ha pedig képlet, akkor a kiszámolt érték. Így a D3-as cella értéke: (12 + 3 =) 15. Ha a D3 tartalmát az Excel másoló funkcióval lemosoljuk a D4 és D5 cellába, akkor őda beíródnak a további peremgyakoriságok. A D3 cellába beírt „=SUM(B3:C3)” egy utasítás, pontososan egy relatív hivatkozás, ami azt jelenti, hogy adja össze a tőle balra levő két cellában levő értéket. Ezért elegendő az Excel másoló funkcióját használni. Beírándó még a B5 cellába a „=SUM(B3:B4)” függvény (amely utasítás pontososan azt jelenti, hogy adja össze a tőle felfele levő két cellában levő értéket) és lemosolódnak a C5 cellába is. Ezzel elkészítettük az oszlopok peremgyakoriságait. Mindegy, hogy a D5 cellába az összegzés a sorok peremgyakoriságainál történik-e vagy pedig az oszlopok peremgyakoriságainál. Mindegyik esetben az érték a minta elemszámát adja meg (ami most 31).


A D13 cellába van beírva a CHITEST(megfigyelt gyakoriságok helye, várható gyakoriságok helye) függvény, konkrétan: „=CHITEST(B3:C4,B9:C10)”. A beépített Excel függvény ilyenkor kiszámítja azt a szignifikanciaszintet, amely mellett a nullhipotézist esetlegesen el lehet vetni. Ha ez az érték nagyobb, mint 0,05 = 5%, akkor mégsem sokás a nullhipotézist elvetni. Ha pedig kisebb, mint 5%, akkor elvetjük, és az 5% helyett használható az a hibavalószínűség, amelyet az Excel kiszámított. A példánkban ez a szignifikanciaszint 0,0384 = 3,84%. Míg a táblázatok alapján csak annyit állíthattunk, hogy a tényleges első fajta hiba kisebb, mint 5% (továbbá esetleg azt, hogy 1% és 5% közé esik).

A nullhipotézis elvetésének megjegyzése. Arra a következtetésre jutottunk, hogy a nem és a kapott gyógyszer között kapcsolat van. Pontosabban: a példánkbeli háziorvosok körzetében lakó hölgyek szignifikánsan különböző (konkrétan nagyobb) arányban fordulnak orvoshoz, azaz igénylik az orvosi végrehajtást, mint a férfi társaik. Ezen állításunk megbízhatósági szintje (1-től nagyobb) arányban fordulnak orvoshoz, azaz igénylik az orvosi törődést, hogy a tényleges első fajta hiba kisebb, mint 5% (továbbá esetleg azt, hogy 1% és 5% közé esik).

A nullhipotézis elvetésének megjegyzése. Arra a következtetésre jutottunk, hogy a nem és a kapott gyógyszer között kapcsolat van. Pontosabban: a példánkbeli háziorvosok körzetében lakó hölgyek szignifikánsan különböző (konkrétan nagyobb) arányban fordulnak orvoshoz, azaz igénylik az orvosi végrehajtást, mint a férfi társaik. Ezen állításunk megbízhatósági szintje (1-től nagyobb) arányban fordulnak orvoshoz, azaz igénylik az orvosi törődést, hogy a tényleges első fajta hiba kisebb, mint 5% (továbbá esetleg azt, hogy 1% és 5% közé esik).

A nullhipotézis elvetésének megjegyzése. Arra a következtetésre jutottunk, hogy a nem és a kapott gyógyszer között kapcsolat van. Pontosabban: a példánkbeli háziorvosok körzetében lakó hölgyek szignifikánsan különböző (konkrétan nagyobb) arányban fordulnak orvoshoz, azaz igénylik az orvosi végrehajtást, mint a férfi társaik. Ezen állításunk megbízhatósági szintje (1-től nagyobb) arányban fordulnak orvoshoz, azaz igénylik az orvosi törődést, hogy a tényleges első fajta hiba kisebb, mint 5% (továbbá esetleg azt, hogy 1% és 5% közé esik).

**Újabb példa.** Családorvosok generalizált szorongásos betegek esetén hasonlították össze Grandaxin, illetve Anxiron kezelés hatékonyságát. Mindkettő gyógyszer (szedatívum) hatékonynak bizonyult. A következő kérdés: van-e a – teljesen tünetmentessé tevő – hatékonyságuk között különbség? Azaz van-e kapcsolat a tünetmentessé válók aránya és a kezelésre használt gyógyszer között? Nullhipotézis: nincs kapcsolat a tünetmentessé válók aránya és a terápiás szer fajtája között, azaz a kezelés hatására tünetmentessé válók aránya az Anxiron szedők között ugyanakkora, mint a Grandaxinnal kezelteknél. (A 6/13 és a 8/16 közötti különbség pusztán a véletlen ingadozás miatt nem teljesen egyforma.) Alternatív hipotézis: Van kapcsolat a tünetmentessé válók aránya és a terápiás szer fajtája között. A maximálisan méltánolható első fajta hibának vegyük a szokásos 5%-ot (?. 0,05!)

11.7. táblázat - Kontingencia táblázat a megfigyelt gyakorisági értékekkel
Döntések előfordulási gyakoriságokról. Kontingencia táblázatok

<table>
<thead>
<tr>
<th></th>
<th>Anxiron</th>
<th>Grandaxin</th>
<th>összes</th>
</tr>
</thead>
<tbody>
<tr>
<td>tünetmentes</td>
<td>6</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>nem-tünetmentes</td>
<td>7</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>összes</td>
<td>13</td>
<td>16</td>
<td>29</td>
</tr>
</tbody>
</table>

Készítsük el a segéd-kontingencia táblázatot a peremgyakoriságok alapján!

11.8. táblázat - Segéd-kontingencia táblázat a várható gyakorisági értékekkel

<table>
<thead>
<tr>
<th></th>
<th>Anxiron</th>
<th>Grandaxin</th>
<th>összes</th>
</tr>
</thead>
<tbody>
<tr>
<td>tünetmentes</td>
<td>6,28</td>
<td>7,72</td>
<td>14</td>
</tr>
<tr>
<td>nem-tünetmentes</td>
<td>6,72</td>
<td>8,28</td>
<td>15</td>
</tr>
<tr>
<td>összes</td>
<td>13</td>
<td>16</td>
<td>29</td>
</tr>
</tbody>
</table>

A segéd-kontingencia táblázatban minden várható gyakoriság nagyobb, mint 5, tehát a teszt alkalmazhatósága szempontjából további vizsgálóádásra nincs szükség. Az $\alpha = 5\%$ első fajta hibához az egy szabadsági fokú kihénymény először esetén a 3,841-es számérték tartozik.

A 11.8. táblázat - Segéd-kontingencia táblázat a várható gyakorisági értékekkel 11.7. és 11.8. táblázat - Segéd-kontingencia táblázat a várható gyakorisági értékekkel 11.8. táblázat gyakoriságai alapján végezzük el a kijelölt műveleteket! $\chi^2 = (6–6,28)^2/6,28+(8–7,72)^2/7,72+(7–6,72)^2/6,72+(8–8,28)^2/8,28 = 0,042$

Mivel 0,042 < 3,841, ezért nem vehetjük el a nullhipotézist. Azaz a mintánk alapján nincs ok, hogy a két gyógyszer hatásossága különböző legyen. Jelen esetben a CHITEST függvény értéke 0,837 = 83,7%. Ez sokkal magasabb, mint a maximálisan méltányolható 5%-os első fajta hiba. Másként megfogalmazva: 100% – 83,7% = 16,3% megbízhatósággal „állíthatnánk”, hogy van különbség a kétféle szedatívum között a tünetmentesség elérése szempontjából. Ha táblázatkezelővel végezzük el a tesztet, azt a valószínűséget kapjuk meg, amely mellett a nullhipotézist elvethetjük. Jelen esetben a CHITEST függvény értéke 0,837 = 83,7%.

3. Függetlenségvizsgálat khi-négyzet próbával nagyobb méretű gyakorisági táblázatok esetén

Ahogy az előző szakaszban megbeszéltem, a kétdimenziós gyakorisági táblázatok között a 2-szer 2-es a legegyszerűbb. Ezt a faja elrendezést akkor kapjuk, ha az alapsokaságunk mindkét felhasznált tulajdonság kétértékű valószínűségi változóval írható le. A kiszámított példában a nem és a kapott gyógyszer fajtája (placebő illetve hatóanyagos) volt. Számos esetben azonban a változók kétértékűek, a kétből többéértékűek. Ha a korábbi példa nem ennyire leegyszerűsített (minden orvoshoz forduló kap „gyógyszer"-t), akkor a terápia leírására kettőnél többértékű változó szükséges. Így a sorok száma kettő helyett $r$-re ($r$ g 2) nő. Ha az orvoshoz fordulók közé a nők és férfiak mellé a hermafroditákat is bevesszük, akkor az oszlopok száma is kettő fölé nő.

Általánosan: ha egy alapsokasághoz tartozó két tulajdonság közül az egyik egy $r$-értékű, a másik pedig egy $s$-értékű diszkrét valószínűségi változóval jellemzhető, akkor az együttet viselkedésüket egy $r$ számú sorból és $s$ számú oszlopból álló gyakorisági táblázattal ($r$-szer $s$ mezős kontingencia táblázattal) írhatjuk le (11.9. táblázat - $r$-szer $s$ mezős kontingencia táblázat megszegyelt gyakorisákokkal a vastagított részben, valamint fejsorral, fejoszloppal és peremgyakoriságokkal 11.9. táblázat). A gyakoriságokat céliszerű kétindexes mennyiségként ($k_{ij}$) felirni: az első index a soruk a sorszáma ($k_{i}$-nél i), a második index pedig az oszlop ($k_{j}$-nél j).
Döntések előfordulás
gyakoriságokról. Kontingencia
táblázatok

<table>
<thead>
<tr>
<th>X</th>
<th>X₁</th>
<th>X₂</th>
<th>...</th>
<th>Xᵣ</th>
<th>összesen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y₁</td>
<td>k₁,₁</td>
<td>k₁,₂</td>
<td>...</td>
<td>k₁,ᵣ</td>
<td>k₁,₋</td>
</tr>
<tr>
<td>Y₂</td>
<td>k₂,₁</td>
<td>k₂,₂</td>
<td>...</td>
<td>k₂,ᵣ</td>
<td>k₂,₋</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yᵣ</td>
<td>kᵣ,₁</td>
<td>kᵣ,₂</td>
<td>...</td>
<td>kᵣ,ᵣ</td>
<td>kᵣ,₋</td>
</tr>
<tr>
<td>összesen</td>
<td>k₋,₁</td>
<td>k₋,₂</td>
<td>...</td>
<td>k₋,ᵣ</td>
<td>n</td>
</tr>
</tbody>
</table>

11.10. táblázat - r-szer s mezős segéd-kontingencia táblázat a várt gyakoriságokkal

<table>
<thead>
<tr>
<th>X</th>
<th>X₁</th>
<th>X₂</th>
<th>...</th>
<th>Xᵣ</th>
<th>összesen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y₁</td>
<td>k₁,₁/n</td>
<td>k₁,₂/n</td>
<td>...</td>
<td>k₁,ᵣ/n</td>
<td>k₁,₋</td>
</tr>
<tr>
<td>Y₂</td>
<td>k₂,₁/n</td>
<td>k₂,₂/n</td>
<td>...</td>
<td>k₂,ᵣ/n</td>
<td>k₂,₋</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yᵣ</td>
<td>kᵣ,₁/n</td>
<td>kᵣ,₂/n</td>
<td>...</td>
<td>kᵣ,ᵣ/n</td>
<td>kᵣ,₋</td>
</tr>
<tr>
<td>összesen</td>
<td>k₋,₁</td>
<td>k₋,₂</td>
<td>...</td>
<td>k₋,ᵣ</td>
<td>n</td>
</tr>
</tbody>
</table>

A függetlenségvizsgálat alapvető kérdése: elfogadható-e az a feltételezés, hogy az X és Y változók függetlenek?

A nullhipotézis: \( H_0 : X \text{ és } Y \) függetlenek.

Az alternatív hipotézis: \( H_A : X \text{ és } Y \) nem függetlenek.

A khi-négyzet teszt alkalmazhatóságának feltételei:

A segéd-kontingencia táblázatban levő várható gyakoriságok között nem lehet olyan cella, amelynek értéke kisebb mint 1.

A segéd-kontingencia táblázatban levő várható gyakoriságok értékei között az 5-nél kisebbek száma maximum az összcellaszám 20%-a lehet.

A vizsgálandó próbastatisztika:
Döntések előfordulás

Számbásági fok: DF = (r – 1) szorzva (s – 1), azaz (a sorok száma – 1) szorzva (az oszlopok száma – 1). Kritikus

tartomány: azon $p^2$ értékek tartománya, amelyekre $p^2 g 2^2$.

A maximálisan méltányolható első fajta hiba szokásos értéke: $? = 0,05 = 5%$

Döntés: Ha a khi-négyzet számított értéke a kritikus tartományba esik, azaz $p^2$ nincsen $p^2$ akkor a

nullhipotézist elvetjük. Ha pedig a $p^2$ nincsen $p^2$ akkor nem vetjük el a nullhipotézist. Az eljárás végén mindkét

esetben értékeljük, hogy a vizsgált tulajdonságok összefüggése (a nullhipotézis elvetésekor) illetve a

függetlensége (a nullhipotézis elfogadásakor) milyen szakmai következményekkel jár.

Példa. Egy bizonyos szembeneteg (nem artériás típusú ischaemias opticus neuropathia) sikeres műtéti

korrekciónál jelent meg 1989-ben egy közlemény (11.11. táblázat).

Minthogy e betegségben korábban semmiféle hatással kezdési módszer

visszegyorsult, ezt a műtétet sok helyen alkalmazták. Rövidesen azonban ez a gondolatban

beavatkozásokról is megjelentek beszámolók, ezért számolva a 25 klinikai centrum 244 ilyen betegét, akik

közül 119 főn elvégezték a műtétet, 125 betegen nem. A felmérés eredménye:

11.11. táblázat - Kontingencia táblázat a megfigyelt gyakorisági értékekkel

<table>
<thead>
<tr>
<th></th>
<th>műtött</th>
<th>nem műtött</th>
<th>összes</th>
</tr>
</thead>
<tbody>
<tr>
<td>javult</td>
<td>39</td>
<td>53</td>
<td>92</td>
</tr>
<tr>
<td>nem javult</td>
<td>52</td>
<td>56</td>
<td>108</td>
</tr>
<tr>
<td>romlott</td>
<td>28</td>
<td>16</td>
<td>44</td>
</tr>
<tr>
<td>összes</td>
<td>119</td>
<td>125</td>
<td>244</td>
</tr>
</tbody>
</table>

Készítsük el a segéd-kontingencia táblázatot a peremgyakoriságok alapján (11.10. táblázat - r-szer s mezős

segéd-kontingencia táblázat a várt gyakoriságokkal 11.10. táblázat)!

Nullhipotézis: nincs kapcsolat a gyógyulás mértéke és az alkalmazott terápiá között, azaz a műtéten átesettek

között ugyanolyan arányban vannak jelen a javultak, nem javultak illetve azok akiknek az állapota romlott, mint

akkienél műtéttel nem alkalmaztak. A maximálisan méltányolható első fajta hiba 5%-ot (?) = 0,05!

A segéd-kontingencia táblázatban minden várható gyakoriság nagyobb, mint 5, tehát a teszt alkalmazhatósága

szempontjából további vizsgálódásra nincs szükség. A szabadsági fokok száma: (a sorok száma – 1) szorzva (az

oszlopok száma – 1) = 1-szer 2 = 2. Az ? = 5% első fajta hibához a két szabadsági fokú khi-négyzet eloszlás

esetén a 5,991-es számérték tartozik.

A 11.11. táblázat - Kontingencia táblázat a megfigyelt gyakorisági értékekkel 11.11. és 11.12. táblázat - Segéd-

kontingencia táblázat a várható gyakorisági értékekkel 11.12. táblázat gyakoriságai alapján végezzük el a

kijelölt műveleteket!

11.12. táblázat - Segéd-kontingencia táblázat a várható gyakorisági értékekkel

<table>
<thead>
<tr>
<th></th>
<th>műtött</th>
<th>nem műtött</th>
<th>összes</th>
</tr>
</thead>
<tbody>
<tr>
<td>javult</td>
<td>44,87</td>
<td>47,13</td>
<td>92</td>
</tr>
<tr>
<td>nem javult</td>
<td>52,67</td>
<td>55,33</td>
<td>108</td>
</tr>
<tr>
<td>romlott</td>
<td>21,46</td>
<td>22,54</td>
<td>44</td>
</tr>
<tr>
<td>összes</td>
<td>119</td>
<td>125</td>
<td>244</td>
</tr>
</tbody>
</table>
Mivel 5,407 l 5,991, ezért nem vethetjük el a nullhipotézist. Azaz a mintánk alapján nincs okunk feltételezni különbséget a két módszer (műtét illetve nem műtét) hatásossága között. Ha táblázatkezelővel végezzük el a tesztet, azt a valószínűséget kapjuk meg, amely mellett a nullhipotézist elvethetjük. Jelen esetben a CHITEST függvény értéke 0,067 = 6,7%. Ez valamivel magasabb, mint a maximálisan méltányolható 5%-os első fajta hiba. Másként megfogalmazva: 100% – 6,7% = 93,3% megbizhatósággal „állíthatnánk”, hogy van különbség a kétfele szedatívum között, de mivel – előre elődöntöttük, hogy a maximálisan méltányolható első fajta 5%-ez, ezért az ahhoz tartozó megbízhatósági szintnek legalább 95%-osnak kellene lennie. Látszik, hogy a jelen esetben nem vagyunk távol ettől a szinttől. Mindenesetre a nullhipotézis elfogadása is arra a döntésre ösztönözik, hogy ne végezzük tovább a kérdés műtétet, mert nem jogtól annak a betegcsoportnak az állapota, ahol a beavatkozást elvégeztetik, mint ahol nem. Tehát kerülendő a minden bizonynal kockázatos és költséges beavatkozás (a műtét).


11.13. táblázat - Kontingencia táblázat a megfigyelt gyakorisági értékekkel

<table>
<thead>
<tr>
<th></th>
<th>műtött</th>
<th>nem műtött</th>
<th>összes</th>
</tr>
</thead>
<tbody>
<tr>
<td>nem romlott</td>
<td>91</td>
<td>109</td>
<td>200</td>
</tr>
<tr>
<td>romlott</td>
<td>28</td>
<td>16</td>
<td>44</td>
</tr>
<tr>
<td>összes</td>
<td>119</td>
<td>125</td>
<td>244</td>
</tr>
</tbody>
</table>

(Eredmény: 5%-nál kisebb [konkrétan 2,93%-os] szignifikanciaszint mellett elvethető az a nullhipotézis, hogy nincs összefüggés a kétfele alkalmazott eljárás és a beteg állapotának súlyosbodása között. Tehát: 95%-nál nagyobb [konkrétan 97,7%-os] megbízhatósággal állíthatjuk, hogy a műtét ront a beteg állapotán.)

4. Az összefüggés erősségének jellemzésére használt mérőszámok

Ha a függetlenségvizsgálata a nullhipotézis elvetésével zárul, akkor természetesen merülhet fel a kérdés: mennyire szoros a megállapított összefüggés? Ennek a jellemzésére számos mérőszámot lehet használni, amelyek közül néhányat megemlítünk. Mindegyik esetben a lehetséges értékek 0 és +1 közé esnek. Tökéletes függetlenség esetén az értékük 0, maximális összefüggés esetén pedig +1.

**Kontingencia együttható.** Néha a khi-négyzetes függetlenségvizsgálatot bevezető Pearson nevét is hozzátesszük az elnevezéshez (Pearson's contingency coefficient)

$$C = \sqrt{\frac{X^2_{számított}}{X^2_{számított} + n}}$$

Csuprov-féle kontingencia együttható:

$$T = \sqrt{\frac{X^2_{számított}}{n \cdot \sqrt{(r-1)(s-1)}}}$$

Cramer-féle kontingencia együttható:
Döntések előfordulási gyakoriságokról. Kontingencia táblázatok

\[ V = \sqrt{\frac{X^2_{számított}}{n \cdot (\min(r, s) - 1)}} \]
12. fejezet - Egy adatsor illeszkedése eloszláshoz

Kaposs Károly dr.

1. A statisztikai illeszkedésvizsgálat célja és fajtái

A 11. fejezetben megfogalmaztunk olyan kérdéseket, amelyeket egydimenziós kontingencia táblázattal kapcsolatban tehtünk fel: pl. a nem a 11.2. táblázatban vagy a kockafeldobás eredménye a 11.3. táblázatban megegyezik-e egy feltételezett eloszláshoz? Ha a gyakoriságokat ismert valószínűségekből kapott gyakoriságokkal hasonlítjuk össze, akkor *tiszta illeszkedésvizsgálatot* végzünk. Ennek speciális esete ha a feltételezett eloszlás az egyenletes eloszlás. Pl. ugyanannyi nő fordul orvoshoz, mint férfi (11.2 táblázat), illetve a kockafeldobás eredményeként ugyanannyisor kapjuk a hat lehetséges esetet, azaz szabályos a dobókocka (11.3. táblázat). Ilyenkor *egyenletes eloszlásra történő illeszkedésvizsgálat* az eljárás neve. Ha pedig nem ismerjük annak az eloszlásnak a paramétereit, amelyre a megfigyelt értékeket illeszteni szeretnénk, pusztán a típusát, akkor *becsléses illeszkedésvizsgálatot* végzünk. Az eloszlás típusa alapján a megfigyelt gyakoriságokból becsüljük az eloszlás paramétereit. Sokszor ez az eloszlás a normális eloszlás, amely esetben két paraméter becsülésére van szükség és az eljárás neve: *normalitásvizsgálat*.

O: megfigyelt gyakoriságok; E: ismert gyakoriságok

Eösszehasonlítása: tiszta illeszkedésvizsgálat

Eösszehasonlítása: *egyenletes eloszlásra történő illeszkedésvizsgálat*

P: eloszlás fajtája ismert csak, a paraméterek becslése A-ból

Eösszehasonlítása: becsleses illeszkedésvizsgálat

Az illeszkedésvizsgálatok áttekintése. A dőlt nagybetűk egydimenziós gyakorisági táblázatokat jelentenek

2. Egyenletes eloszlásra történő illeszkedésvizsgálat

2.1. 1. példa.

A 12.1. táblázat - 12.1. táblázat a 11.2 megfigyelt gyakoriságokat tartalmazó kontingencia táblázat (O) kibővítése lefelé a feltételezett gyakoriságokat tartalmazó segéd-kontingencia táblázattal (E; ), illetve fejoszloppal és peremgyakoriságokkal. Mivel a kétértékű (nominális) valószínűségi változóról azt feltételezzük, hogy egyenletes eloszlású, azaz ugyanannyi nő fordul orvoshoz, mint férfi, ezért a feltételezett gyakoriság minden cellában megegyezik a peremgyakoriság és a lehetséges kimenetelek számának hányadosával.

<table>
<thead>
<tr>
<th></th>
<th>Nő</th>
<th>férfi</th>
<th>összesen</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>31/2 = 15,5</td>
<td>31/2 = 15,5</td>
<td>31</td>
</tr>
<tr>
<td>O</td>
<td>19</td>
<td>12</td>
<td>31</td>
</tr>
</tbody>
</table>

A khi-négyzet teszt alkalmazhatósági feltétele illeszkedésvizsgálatoknál ugyanaz, mint a már megbeszélt függetlenségvizsgálatnál, azaz a segéd-kontingencia táblázatban levő változó gyakoriságok között pedig nem lehet olyan cella, amelynek értéke kisebb mint 1, valamint 1 és 5 közötti érték többé is maximum az összegszám 20%-a lehet. Mivel 15,5 g 5, az alkalmazhatósággal nincs gond. A szabadsági fokok száma a (cellák száma) -1, jelen esetben 2- 1 = 1. A maximalisan méltányolható első fajta hibának nagyobb és 5%-ot (? = 0,05)! Egy szabadsági fok esetén ehhez a szignifikanciaszint hez tartozó ? = ? = 0,05! A nullhipotézis (H0): a megfigyelt gyakoriságok (A) az adott (jelen esetben egyenletes) eloszlásnak felelnek meg.

A khi-négyzet statisztika értéke képletben megfelelően, azaz a megfigyelt és várt gyakoriságok különbség-négyzetéit osztjuk a várt gyakoriságokkal, majd ezeket összegezzük: khi = (19-15,5)²/15,5+(12-15,5)²/15,5 = 1,581. Mivel 1,581 l 3,841, a nullhipotézist nem vethetjük el. Táblázatkezelő programot használva az a signifikanciaszint, amely mellett elvethetnénk a nullhipotézist, 0,209 = 20,9%, de ez lényegesen magasabb, mint a maximalisan méltányolható 5%. Tehát megtartjuk a nullhipotézist: a példával való komparálás esetén
ugyanannyi nő fordul orvoshoz, mint férfi. A megfigyelt 19 illetve 12 fő közötti különbségben pusztán a véletlennek van szerepe. (Az ezen megállapításkor esetleg elkövetett második fajta hiba nagyságáról nincs információ.)

2.2. 2. példa.

Az előbbi péddankhoz hasonlóan a 12.2. táblázat - 12.2. táblázat a 11.3. megfigyelt gyakoriságokat tartalmazó kontingencia táblázat (A) kibővítése lefelé a feltételezett gyakoriságokat tartalmazó segéd-kontingencia táblázattal (B), illetve fejoszloppal és peremgyakoriságokkal. Mivel a hatértékű (nominális) valószínűség változóról azt feltételezzük, hogy egyenletes eloszlású, azaz a kockafeldobás eredményeként ugyanannyiszor kapjuk a hat lehetséges esetet, ezért a feltételezett gyakoriság minden cellában megegyezik a peremgyakoriság és a lehetséges kimenetelek számának hányadosával (100/6 = 16,67).

2.2. táblázat

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>összesen</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16,67</td>
<td>16,67</td>
<td>16,67</td>
<td>16,67</td>
<td>16,67</td>
<td>16,67</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td><strong>O</strong></td>
<td>9</td>
<td>14</td>
<td>16</td>
<td>19</td>
<td>14</td>
<td>28</td>
<td>100</td>
</tr>
</tbody>
</table>

Mivel 16,67 g 5, a teszt alkalmazható. A szabadsági fokok száma a jelen esetben 6-1 = 5. A maximálisan működő első fajta hiba végünk 5%-ot (? = 0,05)! Ót szabadsági fok esetén ehhez a szignifikanciaszinthez tartozó ?²-értéket: 11,070. A nullhipotézis (H₀): a megfigyelt gyakoriságok (A) az egyenletes eloszlásnak felelnek meg.

A khi-négyzet statisztika értéke képletnek megfelelően: khi² = (9-16,67)²/16,67+(14-16,67)²/16,67 +(16-16,67)²/16,67+(19-16,67)²/16,67+(14-16,67)²/16,67+(28-16,67)²/16,67 = 12,44. Mivel 12,44 g 11,070, a nullhipotézist legalább 5%-os szignifikanciaszinten elvethetjük. Táblázatkezelő programot használva az a szignifikanciaszint, amely mellett elvethetjük a nullhipotézist 0,029 = 2,9%.

3. Tiszta illeszkedés vizsgálata

Az emberi AB0 vércsoport fenotípusbeli gyakoriságait az allélek gyakoriságai határozzák meg. Egy modell szerint az európai gyakoriságokat a következő allél relatív gyakoriságok alapján kapjuk meg: p = 3/12, q = 1/12, r = 8/12 (p + q + r = 1). Egy 500 fős mintában a következő vércsoport fenotípusokat határozták meg: A: 215 fő, B: 65 fő, AB: 25 fő, 0: 195 fő. Kérdés: a megfigyelt gyakoriságok illeszkednek-e a modell által várható értékekhez?

A khi-négyzet statisztika értéke képletnek megfelelően: khi² = (215-197,92)²/197,92+(65-59,03)²/59,03 +(25-20,83)²/20,83 +(195-222,22)²/222,22 = 6,247. Mivel 6,247 g 7,815, a nullhipotézist nem vethetjük el. Táblázatkezelő programot használva az a szignifikanciaszint, amely mellett elvethetnél a nullhipotézist 0,100
= 10,0%, de ez magasabb, mint a maximálisan méltányolható 5%. Így megtartjuk a nullhipotézist: a minta adatsora kielégítően illeszkedik a modellhez. Az esetleg elkölvetett második fajta hiba nagyságát nem ismerjük.

4. Becsléses illeszkedésvizsgálat. Speciális eset: normalitásvizsgálat

Az eddig megbeszélt khi-négyzettes tesztek mind alkalmazhatók voltak nominális valószínűségi változó esetén. Az egyes kategóriák között még csak rendezés sem volt (amivel pedig a kategóriális változók másik típusa, az ordinális változók már rendelkeznek), illetve ha lett volna, akkor sem használtuk volna ki. A becsléses illeszkedésvizsgálatnál az eloszlás típusa alapján a megfigyelt gyakoriságokból becsüljük az eloszlás paramétereit. Itt csak azzal az esettel foglalkozunk, ha az eloszlás a normális eloszlás (Gauss eloszlás). A valószínűségi változók közül normális eloszlása a numerikus változóknak lehet. Tehát a becsléses illeszkedésvizsgálatot csak ezekre lehet elvégezni. A többi khi-négyzet teszthez hasonlóan itt is a megfigyelt és a várható gyakoriságokat hasonlítjuk össze. Ahhoz, hogy megfigyelt gyakoriságaink lehessenek, folytonos változó esetén is csoportokra (kategóriákra) kell bontanunk a változót. Azaz ha eleve nem volt diszkrét a változó, akkor azzá kell tennünk. Összefoglalva: becsléses illeszkedésvizsgálatot diszkrét numerikus változó esetén végezhetünk.

Az orvostanhallgatók egyik orvosi fizika gyakorlatán mikroszkóp segítségével mérőpárokban meghatározzák béka vörösvérsejtek hosszabbik átmérőjét. Egy mérőpár 50 db véletlenszerűen kiválasztott vörösvérsejt mér meg. Az alábbi kontingencia táblázat (12.4. táblázat - 12.4. táblázat) kilenc mérőpár összesen 450 mérési adata alapján készült.

<table>
<thead>
<tr>
<th>Kontingencia táblázat</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
</tr>
<tr>
<td>15</td>
</tr>
</tbody>
</table>

Kérdés, hogy a béka vörösvérszejt hosszabbik átmérője, mint valószínűségi változó, normális eloszlást követ-e? Nullhipotézis: a minta normális eloszlású populációból származik. Mivel a feltételezett normális eloszlás elméleti értékeit nem ismerünk, azokat a mintából számontapasztalati értékekkel becsüljük. Becslése a mintából számont számítani közép (? ), becslése pedig az adatok tapasztalati szórása (s ). Konkrétan = 24,76 ?m, s = 3,65 ?m. Ezen két paraméter alapján az illeszteni kivánt normális eloszlás sűrűségfüggvénye ismert: G(x; , s ). A sűrűségfüggvény definíciója szerint a görbe alatti terület 1 = 100%, a minta adatából készíthető gyakorisági eloszlásfüggvény görbe alatti területe pedig a minta elemszámával egyezik meg, ezért G(x; , s ) függvényt meg kell szorozni n = 450-nel. A
12.1. a. ábrán fekete lépcsős függvényként látható a megfigyelt gyakorisági sűrűség függvény (O) és folytonos pirossal pedig a várt folytonos ($E_{cont}$) függvény. Ez utóbbi tehát $n$-szer $G(x; s_1, s_2)$.

Mivel a várt függvény folytonos eloszlású, ezért végrehajtanunk az oszályokra bontást. Az osztályoknak meg kell egyezniek a megfigyelt eloszlásfüggvény osztályaival. Jelen esetben az osztályszélesség 1 ?m, a határok az egész ?m-eknél vannak és az osztályok alulról zártak, azaz pl. a 10 ?m-es adat a 10 és 11 ?m közötti tartományba számítódik. (A 2. fejezetben volt arról szó, hogy milyen a szerepe a választott osztályszélességnek illetve hogy mennyire nem lényeges a határok kezdetének elhelyezkedése és az osztályok alulról vagy felülről való zártsága.) A
12.1.b. ábrán csak a piros függvény (E) különbözik az előzőtől (Econt), mivel ezt már annak az osztályokra bontásával kaptuk. Szintén nem túlzottan lényeges, hogy az egyes osztályokba pl. a 10 és 11 ?m közötti tartománya, a folytonos függvény 10 vagy 11 ?m-es helyettesítési értékét írjuk. Javaslat: használjuk a középsőt (itt: 10,5 ?m-t)!
12.1.c. ábrán a megfigyelt (O) és a várt (E) lépcsős függvények különbségeit láthatjuk, amelyek súlyozottan ugyan, de meghatározzák a khi-négyzet értéket. A nullhipotézis feletti döntéshez kiszámolni kívánt khi-négyzet érték szemléletesen látható a
12.1.d. ábrán. Az ábrázolt lépcsős függvény minden egyes osztályban (a nominális változóknál szokásos szóhasználatban: kategóriában) tartalmazza a (megfigyelt – várható)/várható \([O-E]/E\) értékeket. Ezek összege, azaz a lépcsős függvény alatti területe megadja a kiszámolni kívánt khi négyzet értéket.

Vegyük észre, hogy míg a C ábrán a 25 és 26 \(\mu\)m közötti osztályban az O-E függvény meglehetősen nagy, lényegileg megegyezik a 17 és 18 \(\mu\)m közötti értékként (C ábra), addig ugyanezen osztályok járuléka a khi négyzet érték kiszámításában lényegesen különböző. Ennek oka az egyes osztályokban különböző súlyfaktor, mégpedig a várható érték reciproka (1/E) érték. A várható (E) érték a 25 és 26 \(\mu\)m közötti osztályban sokkal nagyobb, mint a 17 és 18 \(\mu\)m között, így a reciproka súlyfaktorként a leírt hatást okozza.

12.5. táblázat -
A szabadsági fokok száma a becsléses illeszkedésvizsgálatoknál $n-b-1$, ahol $n$ az osztályok száma, $b$ pedig az eloszlás paramétereihez számító száma. Normalitásvizsgálatnál $b = 2$. Jelen esetben a szabadsági fók: $22-2-1 = 19$. A maximálisan méltányolható első fajta hibának vegyünk 5%-ot ($\alpha = 0.05$)! Tizenkilenc szabadsági fők esetén ehhez a szignifikanciaszinthez tartozó $\chi^2$ érték: 30,144.

A khi-négyzet kiszámolásának teljes felírásától eltekintünk, csak az első három tagot adjuk meg: khi$^2 = (0-1.38)^2/1.38 + (4-2.76)^2/2.76 + (10-5.13)^2/5.13 + ... = 11.81$. Mivel 11.81 1 30,144, nincs okunk a nullhipotézis elvetésére. Táblázatkezelőt használva megkapjuk azt a szignifikanciaszintet, amely mellett „elvethetnénk” a nullhipotézist, ez pedig 0.898 = 89.8 %. Jó okunk van a nullhipotézis megtartására: a béka vörösvérsejtjeinek hosszabbik átmérője normális eloszlású.

<table>
<thead>
<tr>
<th>$E$</th>
<th>1.3</th>
<th>2.7</th>
<th>5.1</th>
<th>8.8</th>
<th>14.2</th>
<th>21</th>
<th>28</th>
<th>37</th>
<th>43</th>
<th>48.1</th>
<th>49.1</th>
<th>46.4</th>
<th>40.8</th>
<th>33.2</th>
<th>25.1</th>
<th>17.6</th>
<th>11.4</th>
<th>6.8</th>
<th>3.8</th>
<th>2</th>
<th>0.9</th>
<th>0.4</th>
<th>450</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>össz</td>
<td></td>
</tr>
</tbody>
</table>

| össz | 450 |

A kontingencia táblázat

<table>
<thead>
<tr>
<th>Kontingencia táblázat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$</td>
</tr>
<tr>
<td>15</td>
</tr>
</tbody>
</table>

| össz | 450 |
12.1. ábra. ).
II. rész - Második rész
Tartalom

13. Két faktor szerinti keresztszályozás ............................................................................................................. 169
   1. Példa .............................................................................................................................................. 169
14. Két faktor szerinti keresztszályozás ............................................................................................................... 180
   1. Rögzített és véletlen faktor .................................................................................................................. 182
   2. Egy véletlen faktor szerinti varianciaanalízis ..................................................................................... 182
   3. A példa megoldásának folytatása (5. táblázat) ................................................................................. 184
   4. Keresztszályozás két véletlen faktor szerint ...................................................................................... 184
15. Keresztszályozás egy rögzített és egy véletlen faktor szerint: véletlen blokk ..................................... 188
   1. Példa .............................................................................................................................................. 188
   2. A példa folytatása ............................................................................................................................ 190
   3. Hierarchikus osztályozás ................................................................................................................ 191
   4. Példa .............................................................................................................................................. 191
      4.1. A varianciakomponensek becsélése ...................................................................................... 193
5. Megjegyzés a kísérleti tervek és a modellek alkalmazásához .......................................................... 193
16. Többváltozós elemzések: korreláció és regresszió ................................................................................. 195
   1. Bevezetés ........................................................................................................................................ 195
   2. Többszörös korreláció ....................................................................................................................... 195
   3. Többváltozós regressziós számítás .................................................................................................. 196
      3.1. Regressziós eljárások ................................................................................................................ 197
      3.2. Logisztikus regresszió .............................................................................................................. 197
      3.3. Probit regresszió ........................................................................................................................ 199
      3.4. Poisson regresszió ..................................................................................................................... 199
   4. Nem-lineáris regresszió .................................................................................................................... 199
13. fejezet - Két faktor szerinti keresztszöztályozás

Kemény Sándor dr.

A jelenségek megismerése legtöbbször a változók közötti összefüggések megismerését jelenti. A változókat különböző skálákon mérhetjük, a „mérést” itt általánosabban értve:

névleges (nominal, categorical): eredendően nem szám, például férfi-nő, egyik vagy másik gyógyszer, melyik üzemben végezzük a gyártást, az A, B vagy C városban van egy kórház stb. Még ha kódszámot adunk is az egyes szinteknek, sorrendjük nem értelmezhető (pl. a férfiak személyi száma 1-essel kezdődik, a nőké 2-essel, de ez nem sorrend, mint tudjuk!)

sorrendi (ordinal, ordered categorical): eredendően ez sem szám, például hadnagy, főhadnagy, százados. Itt nincs értelme annak a kérdésnek, hogy a főhadnagy mennyivel magasabb rangú, mint a hadnagy. Illetve az a tény, hogy az alhadnagy és a főhadnagy között két rendfokozat van, míg a főhadnagy és a százados között csak egy, még nem mond semmit arról, hogy az alhadnagy és a főhadnagy között nagyobb-e a különbség, mint a százados és a főhadnagy között.

Ezt a két skálát együtt minőségi skálának nevezzük.

intervallum- (interval): eredendően szám, például a C-ban mért hőmérséklet. A 40 C-ös folyadék éppen 20 C-kal melegebb a 20 C-osnál, de nem kétszer olyan meleg.

arányos (proportional): eredendően szám, például a K-ben (Kelvinben) mért hőmérséklet, tömeg, koncentráció stb. Természetes zérus-pontja van, és 4 kg cukor tömege éppen kétszerese a 2 kg-énak.

Ezt a két skálát együtt mennyiségi skálának nevezzük.

A függő változót legtöbbször intervallum-skálán mérjük, tehát mennyiségi jellegű és folytonos (pl. túlélési idő, a vér koleszterin-szintje). A független változó lehet mennyiségi vagy minőségi. Ha a független változót minőségi (tipikusan névleges) skálán értelmezzük (pl. melyik kezelésben részesül a beteg, milyen fajta táptalajt használtunk a tényezőkhez, melyik baktériumtörzssel oltunk, férfi vagy nőt páciensről van szó), a függő és független változók közötti összefüggést varianciaanalízissel vizsgálhatjuk.

Minthogy a jelenségek a valóságban többváltozósak (a gyógyulás nemcsak az alkalmazott gyógyszer fajtájától függ, a testsúly-gyarapodás nem egyedül a csokoládé-fogyasztás mértéken múlik), kísérleteinknél is legtöbbször több változó hatását kívánjuk megismerni. Szokás ezeket a függő változókat faktoroknak is nevezni.

1. Példa

(Box - Hunter - Hunter: Statistics for experimenters, John Wiley a Sons, 1978, p. 228)

Egy kísérletosrozatban azt kívánták vizsgálni, hogy háromféle méreg (I, II, III.) hatásosságában van-e különbség (vagyis hogy a mérgek mindegyiké egyforma gyorsan pusztítja-e el a kártékony állatokat), másrészt viszont ha háziláttal is eszik a méregből, a rendelkezésre álló négy gyógyszer (A, B, C, D) egyformán hatásos-e a méregzés ellen.

A kísérleti eredményeket (y a túlélés ideje 10 órás egységekben) az 13.1. táblázat - Kísérleti eredmények I. táblázat mutatja.

13.1. táblázat - Kísérleti eredmények

<table>
<thead>
<tr>
<th>gyógyszer</th>
</tr>
</thead>
</table>

...
Két faktor szerinti keresztosztályozás

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0,31</td>
<td>0,82</td>
<td>0,43</td>
</tr>
<tr>
<td></td>
<td>0,45</td>
<td>1,10</td>
<td>0,45</td>
</tr>
<tr>
<td></td>
<td>0,46</td>
<td>0,88</td>
<td>0,63</td>
</tr>
<tr>
<td></td>
<td>0,43</td>
<td>0,72</td>
<td>0,76</td>
</tr>
<tr>
<td>méreg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A kísérleti eredmények természetesen nem véletlenül születtek így, a kísérletező tudatosan kombinálta a méreg mindhárom szintjét (I, II, III.) a gyógyszer mind a négy szintjével (A, B, C, D), és mindegyik ilyen kombinációnál (cellában) 4 állat szerepelt, vagyis a kísérletező egy kísérleti tervet hajtott végre. Ez a kísérleti terv keresztosztályozásnak nevezzük. Ha az ismétlések (itt állatok) száma minden cellában azonos, a terv kiegyensúlyozott tervnek nevezzük, az ilyen terv statisztikai szempontból előnyös (bár nem mindig sikerül így végrehajtani).

A kísérletező kérdései a következők lehetnek:

1. Azonos-e a túlélési idő a háromféle méreggel?

Természetesen látjuk az adatokból, hogy nemhogy a háromféle méreggel (azonos gyógyszer adása mellett), hanem még egy méreg-gyógyszer-kombinációjánál (tehát a terv egy cellájában) sem azonos a túlélési idő a négy állatra. Tehát pontosabban a kérdés a túlélési idő várható értékre vonatkozik, vagyis azt kérdezzük, hogy ha az ingadozást kivennénk a kísérleti eredményekből, a három sor egyforma lenne-e.

2. Azonos-e a túlélési idő a négyféle gyógyszerrel?

Itt is azt kérdezzük, hogy a túlélési idő várható értéke azonos-e, vagyis ha az ingadozást kivennénk a kísérleti eredményekből, a négy oszlop egyforma lenne-e.

Elképzelhető, hogy az egyes gyógyszerek nem egyformán gátolják a különböző mérgek hatását, ezt kölcsönhatásnak nevezik, és igen gyakori. Gondoljunk arra, hogy más a hatása egy pohár bor elfogyasztásának illetve egy szem altatóának külön-külön, mint ha az altatót egy pohár borral öblítjük le. Tehát a harmadik kérdés:

3. Van-e kölcsönhatás a mérgek és a gyógyszerek között?

A modell felépítése a kísérleti hiba jellegének tisztázásával kezdődik, itt ez a következő: \( y_{ijk} = \bar{\gamma}_i + \gamma_j + \delta_{ij} \quad (i=1, \ldots, r; j=1, \ldots, q; k=1, \ldots, p) \)

A képlet úgy olvasszuk, hogy az \( i \)-edik sor (\( i \)-edik méreg) és a \( j \)-edik oszlop (\( j \)-edik gyógyszer) kombinációjánál kapott \( k \)-adik eredmény egy \( \bar{\gamma}_i \) elméleti értékből és egy \( \gamma_j \) kísérleti hibából tevődik össze. Az \( \delta_{ij} \) kísérleti hiba az állatok különbözőségeből és az egyes állatok körülményeinek esetleges különbözőségeből (másik ketrec, más világítás, más hőmérséklet) adódik, ha maga a mérés is bizonytalanság, az is a kísérleti hibát növeli.

A mérgek száma (a méreg-faktor szintjeinek száma) \( r \) (itt 3), a gyógyszerek \( q \) (itt 4), az ismétlések \( p \) (itt 4), összesen
$$\text{rpm} = 3 \cdot 4 \cdot 4 = 48$$

kísérleti pontunk van.

Maga a $$\tau_i$$ elméleti érték az i-edik méreg, hatásának, a $$j$$-edik gyógyszer, hatásának és az $$?_i$$ kölcsönhatásnak az összege, és mindezeket egy közepes értékhez képest képzeljük el: $$y_{i,j} = \tau_i + \tau_{i,j} + ?_i + \tau_j$$

Az $$i$$-edik méreg hatása itt úgy értendő, hogy az $$i$$-edik méreg felértékelését az $$i$$-edik méregvel az $$i$$-edik méregtelui idővel különbözik.

Az adatok elemzésével nemcsak a föntebb (a., b., c.) kérdésekre adhatunk választ, hanem számszerűsíthetjük az egyes mérgéket, gyógyszerek hatását illetve kölcsönhatásaikat, vagyis becsülést adhatunk az $$i$$., $$j$$., és $$?_i$$ paraméterekre.

Az $$\omega$$ hibákra a következő feltételezésekkel élünk:

- várható értékük zérus;
- szórásnégyzetük (varianciájuk) $$\omega^2$$, konstans, vagyis mindegyik $$ij$$ (méreg-gyógyszer-kombinációval jellemzett) csoportra egyforma nagyságú (homoszcedaszticitás);
- az $$\omega$$ hibák csoportokon belül és csoportok között is függetlenek egymástól;
- normális eloszlásúak.

E feltételek megfogalmazása nem azt jelenti, hogy biztosak vagyunk teljesülésükben, hanem egészen pontosan azt, hogy az ezeken a feltételeken alapuló matematikai appratus csak akkor érvényes, ha a feltételek teljesülnek.

Az adatok elemzésével nemcsak a föntebb (a., b., c.) kérdésekre adhatunk választ, hanem számszerűsíthetjük az egyes mérgéket, gyógyszerek hatását illetve kölcsönhatásaikat, vagyis becsülést adhatunk az $$i$$., $$j$$., és $$?_i$$ paraméterekre.

A feltételek teljesülése vagy nem teljesülése a jelenség és a kísérletek természeti út fégg, és a lehetőségekhez képest ellenőrizzünk kell őket vagy az adatok földolgozása előtt, vagy annak során, de legkésőbb a végleges következtetések levonása előtt.

Például a szórásnégyzet nem szükségtelen azonos az egyes csoportokra. Lehet, hogy valamilyik méreggel szemben az állat-egyedek ellenálló képessége különböző, akkor ott nagyobb lesz a szórás, másik méreg jobban egyformán hat az egyes állatokra, akkor annál a méregnél kisebb a szórás.

A kísérleteket úgy végeznénk például, hogy az egyik gyógyszert kapó állatok az ablakhoz közelebbi ketrecben vannak, ahol több fényhez jutnak, a másik gyógyszert kapó állatok távolabb vannak az ablaktól, akkor nem tudhatnánk, hogy a két gyógyszer méreg-elleni hatása között tapasztalt különbögség nem a különböző fénymennyiségnek tulajdonítható-e. Hasonlóan, ha az egyik nap az I. méreg adjuk azoknak az állatoknak, amelyeket erre jelölünk ki, a következő napon a II. méreg következik stb., nem lehetünk biztosak abban, hogy az egyes napokon különböző időjárási viszonyok nem okoznak-e különbséget az állatok reakciójában, amit tévesen a mérgék közötti különbségeknek tulajdonítunk. Szintén hasonlóan, ha az egyik alomból kikerült állat-egyedek kapják az egyik mérgert, a másikból kikerültek a másikat, a méreg hatásának vélt különbögségben benne van az állat-egyedek genetikai különbözsége is. A hibák függetlenségét teljes randomizálással érjük el, vagyis hogy melyik állat melyik méreg és melyik gyógyszert kapja, és mikor kerül sorra, sorsolással kell eldönteni. Másképp fogalmazva minden állatnak egyforma esélye kell legyen arra, hogy valamilyik méreg-gyógyszer kombinációval jellemzett cellájába kerüljön a táblázatnak.

A normális eloszlásra az elvégzendő statisztikai próbák feltételeként van szükség.

Legelőször (mint mindig) érdemes az adatokat grafikusan szemléltetni. Az
1. ábra (dobozos ábra, angolul box-plot) a két faktor (a méreg és a gyógyszer) szintjei szerint csoportosítva mutatja a mérési eredményeket. Egy doboz a táblázat egy cellája, vagyis egy méreg-gyógyszer-kombináció. A dobozok közepén lévő négyszög a medián (a nagyság szerint sorbarendezett érték közül a középső, itt, minthogy 4 adat van, a középső kettő átlaga), a doboz határai a kvartilisek, a dobozból kiálló karok a legkisebb és legnagyobb értékeket mutatják.

Hasznos a következő ún. hatás-ábra is:
2. ábra)

Úgy tűnik az ábrából, hogy mind a gyógyszerek között, mind a mérgek között van különbség, tehát a hatásuk jelentős. A hatás-ábrából azt is láthatjuk, hogy a mérgek és a gyógyszerek hatása is jó közelítéssel additív, a vonalak majdnem párhuzamosak egymással, ez azt sejtheti, hogy nincs kölcsönhatás.

Következik a számszerű elemzés, a tulajdonképpeni varianciaanalízis.

A további tárgyalás érdekében jelöljük $y_{ijk}$-val az $i$-edik sorban ($i$-edik méreggel) és a $j$-edik oszlopban ($j$-edik gyógyszerrel) kapott $k$-adik eredményt, és vezessük be a következő átlagokat:

$$y_{ij} = \frac{\sum_{k=1}^{p} y_{ijk}}{p}$$

az $i$-edik sorbeli és $j$-edik oszlopbeli eredmények átlaga az ismétlések ($k$ index) szerint,

$$y_{i.} = \frac{\sum_{j=1}^{q} y_{ij}}{q}$$

az $i$-edik sorbeli ($i$-edik méreggel végzett) kísérletek eredményeinek átlaga (vagyis az $y_{ij}$ átlagokata gyógyszerek ($j$ index) szerint átlagoljuk,
Két faktor szerinti keresztosztályozás

\[ y_{ij} = \frac{\sum_{j=1}^{r} y_{ij}}{r} \]

a \( j \)-edik oszlopbeli (\( j \)-edik gyógyszerrel végzett) kísérletek eredményeinek átlaga (vagyis az \( y_{ij} \) átlagokat mérvek (\( i \) index) szerint átlagoljuk,

\[ y_\cdot = \frac{\sum_{i=1}^{r} y_{i\cdot}}{r} = \frac{\sum_{j=1}^{q} y_{j\cdot}}{q} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{q} y_{ij}}{rq} = \frac{\sum_{i=1}^{r} \sum_{j=1}^{q} \sum_{k=1}^{p} y_{ijk}}{rqp} \]

az összes kísérletek eredményeinek átlaga.

A jelölés azt jelenti, hogy amelyik index helyett pont van, a szerint átlagolunk.

A nullhipotéziseket most a következőképpen fogalmazhatjuk meg:

- Az \( A \) faktornak nincs hatása, vagyis az eredmény az \( A \) faktor (a méreg) különböző szintjein az ingadozástól eltekintve ugyanaz:

  \[ H_0^A : \alpha_i = 0, i = 1, \ldots, r \]

- A \( B \) faktornak nincs hatása:

  \[ H_0^B : \beta_j = 0, j = 1, \ldots, q \]

- A kölcsönhatás nem létezik:

  \[ H_0^{AB} : \alpha_j \beta_i = 0, i = 1, \ldots, r, j = 1, \ldots, q \]

A modell paramétereinek becslése:

\[ \hat{\mu} = y_\cdot \]
\[ \hat{\beta}_j = y_{j\cdot} - y_\cdot \]
\[ \hat{\alpha}_i = y_{i\cdot} - y_\cdot \]

Ha pl. a mérgek hatására vonatkozó egyenletet figyelmesen megnézzük, láthatjuk, hogy az így kifejezett hatások közül csak \( r-1 \)-et számíthatunk ki függetlenül, az \( r \)-edik abból adódik, hogy az átlagtól (itt \( y_\cdot \) től) való eltérések összege zérus.

A becsült függvényérték a modell-egyenletbe helyettesítve adódik: \( y_{ij} = ? + ? + ? + \ldots + ?_{ij} = y_{ij} \)

\[ \hat{y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\alpha}_{ij} = y_{ij} \]

Láthatjuk, hogy az egyes cellákra (\( i-j \) kombinációkra) a függvény becsült értéke a cella átlaga.

Az ANOVA-tábla először általánosan, a jelölésekkkel (13.2. táblázat - Az ANOVA-tábla általánosan 2. táblázat):

13.2. táblázat - Az ANOVA-tábla általánosan
A 13.3. táblázat - 3. táblázatban S jelöli a négyzetösszegeket, az s² szórásnégyzeteket úgy kapjuk, hogy a négyzetösszegeket elosztjuk a megfelelő szabadsági fokszámmal. A nullhipotézisek igazságát F-próbával vizsgáljuk, az F oszlopban látható az ehhez szükséges próbastatisztika, ami mindig két szórásnégyzet hányadosa. Ha a nullhipotézis igaz (pl. a mérgek között nincs különbség), a próbastatisztika F-eloszlást követ, melynél a számláló szabadsági foka r-1, a nevezőé \(rq(p-1)\).

A 13.3. táblázat -

<table>
<thead>
<tr>
<th>négyzetösszeg</th>
<th>Szabadsági fok</th>
<th>szórásnégyzet</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>gyógyszer</td>
<td>0,921</td>
<td>3</td>
<td>0,307</td>
<td>13,81</td>
</tr>
<tr>
<td>mérég*gyógyszer</td>
<td>0,250</td>
<td>6</td>
<td>0,042</td>
<td>1,87</td>
</tr>
<tr>
<td>maradék</td>
<td>0,801</td>
<td>36</td>
<td>0,022</td>
<td></td>
</tr>
<tr>
<td>mérég (sor)</td>
<td>1,033</td>
<td>2</td>
<td>0,517</td>
<td>23,22</td>
</tr>
</tbody>
</table>

A p oszlopban lévő értékek annak valószínűségét mutatják, hogy ha az illető nullhipotézis igaz, a véletlen műveként ekkora vagy még nagyobb F érték adódjék. Például ha a mérég-faktornak a valóságban nincs hatása (a mérég-faktor különböző szintjein, a háromféle mérégnél) egyforma a túlélési idő, csak a kísérleti ingadozás miatt kapunk különbségeket, akkor annak a valószínűsége, hogy ezek a különbségek 23,22 vagy nagyobb F értéket eredményezzenek, 3,3310⁻⁷, nagyon kicsiny, tehát a nullhipotézist elutasítjuk. Ugyancsak elutasítjuk a gyógyszerek hatásának azonosságára vonatkozó nullhipotézist. A kölcsönhatásra a p érték 0,11, a szokásos 0,05 (5%)-os küszöbértékénél nagyobb, tehát a nullhipotézist elfogadjuk.

Tehát az elemzés szerint a mérgek között illetve a gyógyszerek között szignifikáns a különbség, a kölcsönhatás nem szignifikáns. Következhet az egyes mérések illetve gyógyszerek túlélési időre gyakorolt hatásának becslese, de előtte ellenőrizzük a varianciaanalízis alkalmazásának feltételeit. Erre a reziduumok vizsgálata ad lehetőséget. A reziduum a mért és számított y értékek közötti különbség:

\[ \Delta y_{ij} = y_{ij} - \hat{y}_{ij} \]
Emlékeztetőül, ha a modellben mindegyik tagot (a kölcsönhatást is) szerepeltetjük, a becsült érték a cella-átlag:

\[ \hat{Y}_{ij} = \mu + \alpha_i + \beta_j + e_{ij}. \]

Ábrázoljuk először ezeket a reziduumokat a becsült értékek, tehát a cella-átlagok függvényében (3. ábra).

A feltételek teljesülése esetén azt várnánk, hogy ahogy a kisérleti hibák (ingadozások) is, a reziduumok véletlenszerűen ingadoznák zérus körül, azonos szórással. Látszik az ábrán, hogy ez nem teljesül, a túlélési idő növekedésével annak ingadozása is nő, tehát a varianciaanalízis eredményei nem érvényesek.

Ugyanerre a következtetésre jutunk, ha a reziduumok eloszlásának normalitúsát ellenőrizzük (3. ábra).
4. ábra. Ha az ingadozás normális eloszlást követ, a következő ábrán a pontoknak véletlenszerű ingadozást kell mutatniuk a vonal körül.

Esetünkben ez nem teljesül, a pontok rendszeres viselkedést mutatnak.

Próbáljuk fölrendezni a szórás $y$-függésének matematikai formáját! Ehhez tételezzük föl azt, hogy a szórás a túlélési idő valamilyen hatványával arányos: $?, ~ y''$ (vagy másképpen $?, = ky''$)

Ha ez igaz, a szórás logaritma az $y$ logaritmus függvényében meredekségű egyenest ad (
Igen, az kitevő kb. 2. A következő táblázat mutatja, hogy az ún. Box-Cox-transzformáció, amely szerint \( y \) helyett annak \( y^\alpha = y^\alpha \) transzformáltját kell vennünk, az kitevő egyes értékeire milyen transzformációt jelent (13.4. táblázat - 4. táblázat 4. táblázat)

**13.4. táblázat - 4. táblázat**

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( \gamma = 1 )</th>
<th>transzformáció ( y^\alpha = y^\gamma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,5</td>
<td>0,5</td>
<td>( 1/y )</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>ln ( y )</td>
</tr>
<tr>
<td>0,5</td>
<td>0,5</td>
<td>( y )</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>( y ) (nincs transzformáció)</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>( 1/y )</td>
</tr>
</tbody>
</table>

Esetünkben \( \alpha = 2 \), tehát a túlélési idő reciprokát kell vennünk, hogy a transzformált szórása konstans legyen. Egyes statisztikai programok gombnyomásra is szolgáltatják a kitevő alkalmas értékét, példa a

![Diagram](image-url)
6. ábra.

Itt -1 és -0,5 közötti értéket kell választanunk a kitevőre, a -1 jobban interpretálható.
14. fejezet - Két faktor szerinti keresztosztályozás

Kemény Sándor dr.

Végezzük el az elemzést a túlélési idő reciprokára (14.1. táblázat - A reciprok értékek táblázata, kiegészítve a sor- és oszlop-átlagokkal 1. táblázat)!

14.1. táblázat - A reciprok értékek táblázata, kiegészítve a sor- és oszlop-átlagokkal

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>Soros átlag</th>
<th>Oszlopos átlag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gyógyszer</td>
<td>3,226</td>
<td>2,222</td>
<td>2,226</td>
<td>2,222</td>
</tr>
<tr>
<td>Méreg</td>
<td>2,222</td>
<td>2,174</td>
<td>2,326</td>
<td>2,326</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I</th>
<th>3,226</th>
<th>1,220</th>
<th>2,326</th>
<th>2,222</th>
<th>1,8007</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>2,222</td>
<td>0,909</td>
<td>2,222</td>
<td>1,408</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,174</td>
<td>1,136</td>
<td>1,587</td>
<td>1,515</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2,326</td>
<td>1,389</td>
<td>1,316</td>
<td>1,613</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Méreg</th>
<th>II</th>
<th>2,778</th>
<th>1,087</th>
<th>2,273</th>
<th>1,786</th>
<th>2,2693</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,448</td>
<td>1,639</td>
<td>2,857</td>
<td>0,980</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,500</td>
<td>2,041</td>
<td>3,226</td>
<td>1,408</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4,348</td>
<td>0,806</td>
<td>2,500</td>
<td>2,632</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| III  | 4,545 | 3,333 | 4,348 | 3,333 | 3,7971 |
|      | 4,762 | 2,703 | 4,000 | 2,778 |       |
|      | 5,556 | 2,632 | 4,167 | 3,226 |       |
|      | 4,348 | 3,448 | 4,545 | 3,030 |       |

\[ y = 3,5193, 1,8619, 2,9472, \frac{2,1610}{2,6224} \]

A méreg-faktor hatására (pontosabban annak hiányára) vonatkozó nullhipotézist elutasítjuk, a másik, a gyógyszerek hatásának azonosságára vonatkozó nullhipotézissel együtt, de itt a \( p \) érték mindkét esetben még

14.2. táblázat - Az ANOVA-táblázat

<table>
<thead>
<tr>
<th></th>
<th>négyzetösszeg</th>
<th>Szabadsági fok</th>
<th>szórásnögyzet</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gyógyszer (oszlop)</td>
<td>20,414</td>
<td>3</td>
<td>6,805</td>
<td>28,33</td>
<td>1,38E-9</td>
</tr>
<tr>
<td>Méreg*gyógyszer</td>
<td>1,571</td>
<td>6</td>
<td>0,262</td>
<td>1,09</td>
<td>0,39</td>
</tr>
<tr>
<td>Maradék (ismétlés)</td>
<td>8,643</td>
<td>36</td>
<td>0,240</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Méreg (sor)</td>
<td>34,877</td>
<td>2</td>
<td>17,439</td>
<td>72,64</td>
<td>2,31E-13</td>
</tr>
</tbody>
</table>

A méreg-faktor hatására (pontosabban annak hiányára) vonatkozó nullhipotézist elutasítjuk, a másik, a gyógyszerek hatásának azonosságára vonatkozó nullhipotézissel együtt, de itt a \( p \) érték mindkét esetben még
Két faktor szerinti keresztosztályozás

Kisebb, tehát még biztosabban lehetünk a döntésünkben (14.2. táblázat - Az ANOVA-táblázat 2. táblázat). A kölcsönhatásra a p érték 0,39, a korábbi 0,11-nál is nagyobb, tehát a nullhipotézist elfogadjuk.

Számítsuk ki, mikkora az I méreg hatása (a közepes értékhez képest)!

\[ \hat{\alpha}_1 = y_{11} - y_{..} = 1,8007 - 2,6224 = -0,8217 \]

A közepes érték becslése:

\[ \hat{\mu} = y_{..} = 2,6224 \]

Számítsuk ki, mikkora a B gyógyszer hatása (a közepes értékhez képest)!

\[ \hat{\beta}_2 = y_{22} - y_{..} = 1,8619 - 2,6224 = -0,7605 \]

Mikkora reciprok túlélési időre számíthatunk az I méreggel és a B gyógyszerrel (a kölcsönhatás, minthogy nem volt szignifikáns, nem szerepel a modellben)?

\[ \hat{Y}_{12} = \hat{\mu} + \hat{\alpha}_1 + \hat{\beta}_2 = 2,6224 - 0,8217 - 0,7605 = 1,0402 \]

Ha a várható túlélési időt akarjuk megkapni, az érték reciprokát kell venni: \(1/1,0402=0,9614\) (10 órás egységekben, tehát 9,6 óra).

Az alábbi hatás-ábrák:

![Hatás-ábrák](image)
Két faktor szerinti keresztosztályozás

1. Rögzített és véletlen faktor

A rögzített faktorok jellemzője, hogy szintjeiket a kísérletekhez megválaszthatjuk és beállíthatjuk. A véletlen (random) faktor szintjeit nincs módunk tetszőlegesen beállítani, azokat egy elképzelt sokaságból véletlenszerűen választhatjuk ki. Ha a kíséréteket egy következő alkalommal megismételjük, nem tudnánk a sokaságból ismét ugyanazokat a szinteket kiválasztani.

Az adatok földolgozásának rögzített faktorok (például négyfélé technológiai változat) esetén az a célja, hogy megállapitsuk, van-e különbség az egyes faktorok különböző szintjei között, majd rögtön ezután azt is, hogy melyik közülük a legjobb, illetve jobb-e valamelyik szint, mint egy másik, sőt becsüljük az egyes szintek hatását, és arra konfidencia-intervallumot is adhatunk.

Véletlen faktor esetén csak arra vagyunk kíváncsiak, hogy az illető faktornak van-e hatása (ennyiben ugyanaz a kérdés, mint rögzített faktor esetén), de azt nem kérdezzük, hogy melyik szintje jobb, és nem is becsüljük az egyes szintek hatását. Ehelyett inkább arra vagyunk kíváncsiak, hogy a faktornak a jövőben ugyancsak véletlenszerűen kiválasztandó (vagy előállító) más szintjén milyen mértékű eltérésre kell számítani, mekkora a véletlen faktor szintjeinek ingadozása miatt a függő változóban. Több véletlen faktor esetén azt is kérdezhetjük, hogy közülük melyik milyen mértékben járul hozzá az y függő változó ingadozásához.

Az adatok összefoglalásában szerepel a „1. a., b. ábra” két faktor hatását és annak 95%-os konfidencia-intervallumát (még mindig a reciprok túlélési időre).

2. Egy véletlen faktor szerinti varianciaanalízis

Példa (Kemény S., Deák A.: Kísérletek tervezése és értékelése, Műszaki Könyvkiadó, Budapest, 2000, p. 249)

Egy kémiai elemzést három napon naponta két szem végztek el. Az eredményeket (yij) a 14.3. táblázat - Kémiai elemzés 3. táblázat tartalmazza. Itt a „nap” véletlen faktor, mert nem azt kérdezzük, hogy az egyik napi (pl. márc. 21-i) elemzési eredmény jobb-e a másiknál (mert akkor ezentúl mindig márc. 21- edikén akarnánk mérni), hanem azt, hogy nagyobb-e a különböző napokon végzett elemzések ingadozása az egy napon belüli
Két faktor szerinti keresztosztályozás

ismétléseknél, vagyis a „nap” ad-e az ingadozást többletet. Ha a kíséreltsorozatot megismételjük, nem ugyanazon napokat választanánk, hanem a napok sokaságából másikakat. A következtetés eszerint nincs az adott napokra lesz érvényes, hanem az ezekkel azonos sokasághoz tartozó más napokra is. Nyilvánvalóan nem a naptári nap hatásáról van egyébként szó, hanem a különböző napokon különbözőképpen megnyilvánuló egyéb zavaró körülményekről (hőmérséklet, a levegő nedvességtartalma, az elektromos hálózat frekvenciája stb.).

14.3. táblázat - Kémiai elemzés

<table>
<thead>
<tr>
<th></th>
<th>1. nap</th>
<th>2. nap</th>
<th>3. nap</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_{ij} )</td>
<td>96,963</td>
<td>97,567</td>
<td>97,195</td>
</tr>
<tr>
<td>( y_{ij} )</td>
<td>96,897</td>
<td>97,236</td>
<td>97,345</td>
</tr>
<tr>
<td>( y_{ij} )</td>
<td>97,005</td>
<td>97,495</td>
<td></td>
</tr>
</tbody>
</table>

Kérdés: Okoz-e többlet-ingadozást az, hogy különböző napokon végezték a méréseket? Fogalmazzuk meg a modellt és a hipotéziseket véletlen faktor esetére!

A modell:

\[ y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \]

ahol \( i \) most a véletlen jellegű A faktor (pl. nap) i-edik szintjének hatását jelenti. Itt a kíséreltnél beállított szint (pl. „apr. 4. szerda”) és, maga is valószínűségi változó abban az értelemben, hogy amikor valamelyik napot kijelöljük a kíséreltre, akkor a napok (és az ezekkel összekapcsolt eltérő körülmények) sokaságából választunk.

Az A faktornak legyen most is \( r \) szintje, és ahogy az egy rögzített faktor szerinti varianciaanalízisnél, nem szükséges, hogy az egy i szinten végrehajtott ismétlések \( p \) száma azonos legyen.

Változatlanul érvényesek az kísérleti hibára a rögzített faktoroknál ismertetett feltételezések, de további feltételezésekkel is elünk:

- az \( \mu \) hatás várható értéke zérus;
- szóráségyzete (varianciája) \( \sigma^2_{\alpha} \), konstans;
- az \( \alpha \) hatások a különböző i szinteken egymástól és az \( \varepsilon \) kísérleti ingadozástól (hibától) is függetlenek;
- az \( \varepsilon \) hatás normális eloszlású.

Ebben a modellben az ismétlési hiba kiküszöbölésével kapható \( Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} \), is valószínűségi változó (varianciája \( \sigma^2_{\varepsilon} \)), mert az ismétlési hibával is terhelt mérési eredmények más érték körül ingadoznak hőfűn, mint szerdán. Az \( Y_{ij} \) kísérleti adatok varianciájára írhatjuk, hogy

\[
Var(Y_{ij}) = \sigma^2_A + \sigma^2_{\varepsilon},
\]

ezért a modellt a varianciakomponensek modelljének is nevezik.

Az A véletlen faktor hatástalanságára vonatkozó nullhipotézis nem úgy fogalmazzuk meg, hogy \( \alpha = 0 \), hanem úgy, hogy

\[
Var(\alpha) = \sigma^2_A = 0
\]

A nullhipotézis tehát: \( H_0: \alpha = 0 \) (nem okoz többlet-ingadozást az, hogy különböző napokon végezték a méréseket) (4. táblázat).

14.4. táblázat - Az ANOVA-táblázat
Két faktor szerinti keresztosztályozás

<table>
<thead>
<tr>
<th>Az forrása</th>
<th>eltérés-</th>
<th>Szabadsági fok</th>
<th>Szórásnégyzet</th>
<th>A szórásnégyzet várható értéke</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ismétlések (csoportokon belüli)</td>
<td>( S_R ) = ( \sum_{i} \sum_{j} ) r szorzva (p-1)</td>
<td></td>
<td>( \frac{S_R^2}{r(p-1)} )</td>
<td>?^2 _A</td>
<td>?^2 _A</td>
</tr>
<tr>
<td>Teljes</td>
<td>( S_0 ) = ( \sum_{i} \sum_{j} ) r szorzva p-1</td>
<td></td>
<td>( \frac{S_A^2}{r-1} )</td>
<td>?^2 A + p szorzva</td>
<td>( \frac{S_A^2}{S_R^2} )</td>
</tr>
<tr>
<td>A hatása (csoportok közötti)</td>
<td>( S_A ) = ( p \sum_{i} ) r-1</td>
<td></td>
<td>( \frac{S_A^2}{r-1} )</td>
<td>( \frac{S_A^2}{S_R^2} )</td>
<td></td>
</tr>
</tbody>
</table>

Az egy rögzített faktor szerinti osztályozás táblázatához képest új „a szórásnégyzet várható értéke” oszlop, amiből azt látjuk, hogy az \( s_A^2 \) szórásnégyzet várható értéke a kísérleti (ismétlési) ingadozás \( \sigma^2 \) szórásnégyzetén (varianciánál) kívül tartalmazza a napok közötti esetleges ingadozás \( \sigma^2 \) szórásnégyzetét (varianciáját) is. Ha arra a kérdésre akarunk válaszolni, hogy létezik-e a napok közötti ingadozás (vagyis hogy \( \sigma^2 > 0 \)), az \( s_A^2 \) szórásnégyzetet kell összehasonlítanunk az \( s_R^2 \) szórásnégyzettel, ezt mutatja az F oszlop. Ha a nullhipotézis igaz, az \( s_A^2 \) szórásnégyzet várható értéke megegyezik az \( s_R^2 \) szórásnégyzetével.

Ha a nullhipotézis igaz, az A véletlen faktor hatásának vizsgálatára szolgáló F arány F-eloszlást követ, a számítás szabadsági foka \( r-1 \), a nevezőé \( r(p-1) \). Ez az \( F \) próbastatisztika pontosan ugyanaz, mint amit az egy rögzített faktor szerinti varianciánalízisnál használtunk.

3. A példa megoldásának folytatása (5. táblázat)

<table>
<thead>
<tr>
<th>Az forrása</th>
<th>eltérés-</th>
<th>Szabadsági fok</th>
<th>Szórásnégyzet</th>
<th>Szórásnégyzet várható értéke</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ismétlések</td>
<td>3</td>
<td>0.088587</td>
<td>( \sigma^2 )</td>
<td>( \sigma^2 )</td>
<td>( 0.088587 )</td>
<td></td>
</tr>
<tr>
<td>A: nap</td>
<td>2</td>
<td>0.092760</td>
<td>0.092760</td>
<td>0.092760</td>
<td>0.092760 + 0.092760</td>
<td></td>
</tr>
</tbody>
</table>

Látható az eredményekből, hogy nem okoz többletingadozást a napok közötti különbség. Annak \( p \) valószínűsége ugyanis, hogy két szórásnégyzet aránya az egységnyitől ennire (vagy jobban) eltérjen, amikor a nullhipotézis igaz, 0.452, ezért elfogadjuk a nullhipotézist, miszerint az, hogy különböző napokon végzik az elemzést, nem okoz többlet-ingadozást.

4. Keresztosztályozás két véletlen faktor szerint

Példa (Kemény S., Deák A.: Kísérletek tervezése és értékelése, Műszaki Könyvkiadó, Budapest, 2000, p. 254)

Egy elemzést nemsak különböző napokon végeztek el, hanem különböző személyek is. Úgy képzeljük el a helyzetet, hogy a napok illetve a személyek sokaságából választunk szinteket az adott kísérlet sorozathoz, tehát a nap is, a személy is véletlen faktor. A következtetés eszerint nemsak az adott napokra és adott személyekre lesz érvényes, hanem az ezekkel azonos sokasághoz tartozó más napokra és más személyekre is. A személyek például akkor vélhetők egy sokasághoz tartozók, ha azonos képzettségük és gyakorlati tapasztalataik is hasonló. Minden kombinációban azonos \( p \) számú ismétlést végzünk. Az eredményeket a 6. táblázat mutatja.

<table>
<thead>
<tr>
<th>1. nap</th>
<th>2. nap</th>
<th>3. nap</th>
<th>y_j</th>
</tr>
</thead>
</table>

184

Created by XMLmind XSL-FO Converter.
Két faktor szerinti keresztosztályozás

<table>
<thead>
<tr>
<th>1. nap</th>
<th>2. nap</th>
<th>3. nap</th>
<th>y <em>j</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>96,963</td>
<td>97,567</td>
<td>97,195</td>
<td>97,170</td>
</tr>
<tr>
<td>2. személy</td>
<td>97,232</td>
<td>97,241</td>
<td>97,215</td>
</tr>
<tr>
<td>97,184</td>
<td>97,025</td>
<td>97,581</td>
<td>97,247</td>
</tr>
<tr>
<td>3. személy</td>
<td>96,988</td>
<td>97,202</td>
<td>97,352</td>
</tr>
<tr>
<td>96,797</td>
<td>97,327</td>
<td>97,283</td>
<td>97,158</td>
</tr>
<tr>
<td>4. személy</td>
<td>97,035</td>
<td>97,339</td>
<td>97,388</td>
</tr>
<tr>
<td>97,095</td>
<td>97,318</td>
<td>97,215</td>
<td>97,247</td>
</tr>
<tr>
<td>y <em>i</em></td>
<td>97,024</td>
<td>97,240</td>
<td>97,335</td>
</tr>
<tr>
<td>y <em>i, j</em></td>
<td>97,200</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A kérdés: Az, hogy a mérést különböző napokon és különböző személyek végzik, okoz-e többletingadozást az egy nap egy személy végezte ismétlések szóródásához képest?

Itt két véletlen jellegű faktorunk van, közöttük kölcsönhatás is lehet, pl. mert egy-egy személy teljesítőképessége az egyes napokon nem szükségképpen azonos.

Jelölje a napok hatását, a személyeket, ekkor a modell a következő: y _i, j_ = ? + ? _i_ + ? _j_ + ? _i, j_ Az ismétlési hiba indexében a k(ij) azt jelöli, hogy a k-adik ismétlés egy i-j (nap-személy) kombinációhoz tartozik.

A véletlen faktorokkal kapcsolatos feltételezések itt a következőképpen fogalmazandók meg:

- az ? _i_ , ? _j_ , ? _i, j_ hatások valószínűségi változók, várható értékük zérus;
- ? _A_ , ? _B_ , ? _AB_ szórásnégyzetük (varianciájuk) konstans;
- az ? _i_ , ? _j_ , ? _i, j_ hatások a különböző i és j szinteken egymástól és az ? _ij_ kísérleti ingadozástól (hibától) is függetlenek;

Fontos, hogy a hibák itt is függetlenek, a napnak és személynek tulajdonítható tényezők az illető két véletlen faktorban vannak, az egyéb körülmények véletlenszerűségét pedig a kísérletek randomizálásával biztosítjuk.

Nullhipotézisek:

H _A_ : ? _A_ = 0 (a napoknak nincs hatása, nem okoz többletingadozást);
H _B_ : ? _B_ = 0 (a személyek között nincs különbség, nem okoznak többletingadozást);
H _AB_ : ? _AB_ = 0 (nincs kölcsönhatás a napok és a személyek között).

Legyen általánosan az A faktornak r szintje, a B faktornak q és végezzünk p-szer ismétlést (14. táblázat - 7. táblázat).

14.7. táblázat

<table>
<thead>
<tr>
<th>Az elterés forrása</th>
<th>eltérés- négyzetösszeg</th>
<th>Szabadsági fok</th>
<th>szórásnégyzet</th>
<th>a szórásnégyzet várható értéke</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>S <em>B</em> = rp Σ j y <em>j</em></td>
<td>q-1</td>
<td>S <em>B</em> ^2</td>
<td>rp S <em>B</em> ^2 + p S <em>A</em> ^2</td>
<td>Sn/2 / S <em>AB</em> ^2</td>
</tr>
</tbody>
</table>

185

Created by XMLmind XSL-FO Converter.
Két faktor szerinti keresztszátozás

<table>
<thead>
<tr>
<th>Az forrása</th>
<th>eltérés-négyzetösszeg</th>
<th>Szabadsági fok</th>
<th>a szórásnégyzet várható értéke</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>$S_{AB} = p \sum_{i} \sum_{j} (Y_{ij} - \bar{y})^2$</td>
<td>(r-1), (q-1)</td>
<td>$S_{AB}^2 = \frac{S_{AB}^2}{(r-1)(q-1)}$</td>
<td>$\frac{S_{AB}^2}{S_A^2}$</td>
</tr>
<tr>
<td>Ism.</td>
<td>$S_{R} = \sum_{i} \sum_{j} \sum_{k} (p_{ijk} - \bar{Y}<em>{i.j} - \bar{Y}</em>{.k} + \bar{Y}_{..})^2$</td>
<td>q</td>
<td>$S_{R}^2 = \frac{S_{R}^2}{r-1}$</td>
<td>$\gamma_q^2$</td>
</tr>
<tr>
<td>Teljes</td>
<td>$S_{0} = \sum_{i} \sum_{j} \sum_{k} (Y_{ijk} - \bar{Y}_{..})^2$</td>
<td></td>
<td>$S_{0}^2 = \frac{S_{0}^2}{r-1}$</td>
<td>$\gamma_r^2$</td>
</tr>
<tr>
<td>A</td>
<td>$S_{A}^2 = gq \sum_{i} (\bar{Y}<em>{i..} - \bar{Y}</em>{..})^2$</td>
<td>$\gamma_{r-1}$</td>
<td>$S_{A}^2 = \frac{S_{A}^2}{\gamma_{r-1}}$</td>
<td>$gq \sigma_A^2 + \sigma_e^2$</td>
</tr>
</tbody>
</table>

A két rögzített faktor szerinti osztályozás táblázatához képest új „a szórásnégyzet várható értéke” oszlop.

Ha a $\gamma_{r-1}$ = 1 nullhipotézist (a napoknak nincs hatása) kívánjuk vizsgálni, az $s_A^2$ szórásnégyzetet kell összehasonlítanunk $s_{AB}^2$ szórásnégyzettel, vagyis az $s_{AB}^2 / s_{AB}^2$ arányra végzünk $F$-próbát, mert a $H_0: \gamma_{r-1} = 0, \gamma_q^2 = 0$ fennállása esetén mindkét szórásnégyzet várható értéke $\gamma_{r-1}^2 + \gamma_q^2$ eloszlású. A rögzített faktoroknál megszokott $s_A^2 / s_{AB}^2$ vizsgálat itt nem alkalmazható, mert az $s_A^2 / s_{AB}^2$ arány eloszlása akkor is eltér az $F$-eloszlástól, ha $\gamma_{r-1}^2 = 0, \gamma_q^2 = 0$. Hasonlóan a $\gamma_{r-1}^2 = 0$ nullhipotézist (a személyeknek nincs hatása) az $s_{AB}^2 / s_{AB}^2$ arányval vizsgáljuk. Az $AB$ kölcsönhatás nagyságának vizsgálatára az $s_{AB}^2 / s_{AB}^2$ arány alkalmas.

Ha nincsenek ismételt mérések, nincs $s_{AB}^2$ szórásnégyzet (szabadsági foka zérus), így a $\gamma_{r-1}^2$ hipotézis nem vizsgálható (14.8. táblázat - A példa adatainak kiértékelése 8. táblázat).

14.8. táblázat - A példa adatainak kiértékelése

<table>
<thead>
<tr>
<th>Az forrása</th>
<th>eltérés-négyzetösszeg</th>
<th>Szabadsági fok</th>
<th>a szórásnégyzet várható értéke</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>B: személy</td>
<td>3</td>
<td>$s_A^2 = 0,01078$</td>
<td>$6?<em>{A}^2 + 2?</em>{AB}^2 + ?_{r-1}^2$</td>
<td>$s_A^2 / s_{AB}^2 = 0,4435$</td>
<td>0,0184</td>
</tr>
<tr>
<td>AB: nap*személy</td>
<td>6</td>
<td>$s_{AB}^2 = 0,02431$</td>
<td>$2?<em>{AB}^2 + ?</em>{r-1}^2$</td>
<td>$s_{AB}^2 / s_{AB}^2 = 0,7096$</td>
<td>0,6487</td>
</tr>
<tr>
<td>ismétlések</td>
<td>12</td>
<td>$s_{AB}^2 = 0,03426$</td>
<td>$?_{r-1}^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A: nap</td>
<td>2</td>
<td>$s_A^2 = 0,20317$</td>
<td>$8?<em>{A}^2 + 2?</em>{AB}^2 + ?_{r-1}^2$</td>
<td>$s_A^2 / s_{AB}^2 = 0,8357$</td>
<td>0,0184</td>
</tr>
</tbody>
</table>

Látható, hogy a személyek különbözősége és a kölcsönhatás elhanyagolható mértékben növeli meg az ingadozást (a véletlen 0,73 illetve 0,65 valószínűséggel okozna ekkora vagy nagyobb $F$-arányt), a napok hatása viszont jelentős.

A szórásnégyzetekből becsülhetjük az egyes hatások variánciáját is. Például az $s_A^2$ szórásnégyzet várható értéke $?_{r-1}^2$, ez maga a becslés:

$$\hat{\sigma}_e^2 = \hat{\sigma}_R^2 = 0,03426$$

Az $s_A^2$ szórásnégyzet várható értéke $8?_{A}^2 + 2?_{AB}^2 + ?_{r-1}^2$, tehát a $?_{A}^2$ becsleséhez számolnunk kell:

$$\hat{\sigma}_A^2 = \frac{S_A^2 - S_{AB}^2}{8} = 0,0223$$

Hasonlóan

Created by XMLmind XSL-FO Converter.
Két faktor szerinti keresztosztályozás

\[ \hat{\sigma}_{AB}^2 = \frac{s_{AB}^2 - s_R^2}{2} = -0,004975 \]

A negatív érték meglepő. Gyakran előfordul, hogy a becsült szórásnégyzetekre (varianciákra) negatív érték adódik, ami nyilvánvalóan értelmetlen, ilyenkor zérusra igazítjuk a becsült értéket. A statisztikai programok bizonyos beállításoknál ezt a korrekciónál elvégzik.

Ha figyelembe vesszük, hogy a B faktornak nincs hatása, és az AB kölcsönhatás sem létezik, ezek szórásnégyzetét az ismétlések szórásnégyzetével egyesíthetjük (pooling), amelynek eredménye:

\[ \sigma_e^2 = \frac{0,03426 \cdot 12 + 0,02431 \cdot 6 + 0,01078 \cdot 3}{21} = 0,0281 \]

A \( \sigma_e^2 \) és \( \sigma_A^2 \) hasonló nagysága azt jelenti, hogy a különböző napokon végzett ismételt elemzések varianciája (szórásnégyzete) kb. kétszerese az azonos napokon végzett ismétlésekének, mivel a teljes variancia a két variancia összege:

\[ \sigma_{teljes}^2 = \sigma_e^2 + \sigma_A^2 = 0,0223 + 0,0281 = 0,0504 \]
15. fejezet - Keresztosztályozás egy rögzített és egy véletlen faktor szerint: véletlen blokk

Kemény Sándor dr.

Gyakran előfordul, hogy egy rögzített faktor hatásának vizsgálatára végzett kísérlétsorozatban nem biztosítható a körülmények azonosága. Ilyenkor használható a véletlen blokk elrendezés, amelyével egy (vagy több) új, véletlen faktort vezetünk be.

1. Példa


Ha úgy járnának el, hogy valamelyik technológiával az első, más technológiával a második stb. kukoricalekvár-adagból végezzenek a kísérleteket, nem lehetne látni, hogy a tapasztalt különbség a technológiák közötti különbség-e, vagy pedig a főhasznált kukoricalekvár-adagok közötti különbségnél köszönhető. Tehát a kukoricalekvár-adagokat a technológiák között el kell osztani.

A teljes randomizálás azt jelentené, hogy véletlenszerűen választunk egy technológiát, majd kisorsoljuk, hogy milyen kukoricalekvárral végezzünk a kísérletet, majd megint választunk egy technológiát, ismét sorsolunk hozzá kukoricalekvárt, s.t.t. A teljes randomizálás egyetlen faktor (a technológia) szerinti vizsgálatot jelentene, ötszöri ismétlésnél, ez összesen 20 kísérlet. Az ismétlést sok esetben különböző táptalaj-adagokkal végeznének, így a kukoricalekvárok közötti különbség az ismétlések szórásában jelentkezné. Emiatt megnövekedne a szórás, ezzel a másodfajú hiba valószínűsége, tehát egy létező hatást nehezebben vennénk észre.

A példában a kísérleteket 5 blokkban végzik, ezeken belül a kukoricalekvár azonos. Mind az öt kukoricalekvárt a négy technológiával kombinálni kell, císként azonos számú ismétlésnél, hogy kiegyensúlyozott legyen a terv. Randomizálásra itt is szükség lehet, például nem mindegy, hogy egy hordó főlbonása után mennyi idővel kerül sorra a fermentáció, tehát a kísérletek sorrendje esetleg nem közömbös.

A terv teljesen analóg a két faktor szerinti kereszt-osztályozás kísérleti tervével. A különbség a terv kialakításának céljában, a megválaszolandó kérdéseken van. A mérgezés-például azt kérdeztük, hogy van-e a mérgek hatása között különbség, van-e a gyógyszerek hatása között különbség, és van-e kölcsönhatás, tehát a kérdések szempontjából a faktorok egyenrangúak voltak. Itt is azt kérdezhettük, hogy van-e a technológiák között különbség, de azt nem hogy van-e a kukoricalekvár-szempontjából a faktorok egyenrangúak voltak. Itt is azt kérdezhettük, hogy van-e a technológiák között különbség, de azt nem hogy van-e a kukoricalekvár-szempontjából a faktorok egyenrangúak voltak.

A tervben a technológia rögzített faktor, mert ha meg akarnánk ismételni az egész kísérletsorozatot (nem akarjuk, ez csak gondolatkiérlet), választhatnánk ugyanazokat a technológiákat (a technológia faktor ugyanazon szintjeit), mint az eredeti kísérleteknél. A kukoricalekválokra nem választhatnánk ugyanazokat a szinteket (ugyanazokat az adagokat), mert azok már elfogytak, illetve még ha nem fogytak is el, a hosszabb állás miatt megváltozhatnak. Tehát a kukoricalekvár-adag véletlen faktor.

A kísérleti eredmények az 15.1. táblázat - 1. táblázatban láthatók.

15.1. táblázat -

<table>
<thead>
<tr>
<th>Kukoricalekvár</th>
<th>Technológia</th>
<th>y_i</th>
</tr>
</thead>
</table>
Keresztoztszályozás egy rögzített és egy véletlen faktor szerint: véletlen blokk

<table>
<thead>
<tr>
<th>Kukoricalekvár</th>
<th>Technológia</th>
<th>( y_j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>89</td>
<td>88</td>
</tr>
<tr>
<td>2</td>
<td>84</td>
<td>77</td>
</tr>
<tr>
<td>3</td>
<td>81</td>
<td>87</td>
</tr>
<tr>
<td>4</td>
<td>87</td>
<td>92</td>
</tr>
<tr>
<td>5</td>
<td>79</td>
<td>81</td>
</tr>
<tr>
<td>( y_i )</td>
<td>84</td>
<td>85</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Legyen általában \( A \) a rögzített faktor, \( r \) a szintjeinek száma, \( B \) a véletlen faktor, \( q \) szinttel, \( p \) pedig legyen az ismétlések száma. A modell:

\[
y_{ijk} = \mu + \alpha_i + \beta_j + \alpha \beta_{ij} + \varepsilon_{ik}(y)
\]

A véletlen faktor részvételével fellépő kölcsönhatást (AB) is véletlen jellegűnek tekintjük. Az utolsó tag indexelése azt mutatja, hogy adott technológiával és adott kukoricalekvárral végzik az ismétléseket.

A szokásos nullhipotézisek:

- \( H_0^A : \alpha_i = 0 \) (a technológiák között nincs különbség);
- \( H_0^B : \beta_j = 0 \) (a kukoricalekvárok között nincs különbség);
- \( H_0^{AB} : \alpha \beta_{ij} = 0 \) (nincs kölcsönhatás a technológia és a kukoricalekvár között).

A penicillin-példában \( A \) a technológia \((r = 4)\), \( B \) a kukoricalekvár \((q = 5)\), és nincs ismétlés \((p = 1)\). A varianciaanalízis négyzetösszegeit és szórásnégyzeteit a 15.2. táblázat - 2. táblázat mutatja, ebben láthatjuk a szórásnégyzetek várható értékeit és, azt is, hogy milyen összehasonlítással vizsgálhatók az egyes nullhipotézisek.

### 15.2. táblázat -

<table>
<thead>
<tr>
<th>Az forrás</th>
<th>eltérés-négyzetösszeg</th>
<th>Szabadsági fok</th>
<th>szórásnégyzet a szórásnégyzet várható értéke</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>( S_B = rp \sum_{j=1}^{q-1} ) ( j )</td>
<td>( q-1 )</td>
<td>( S_B^2 = \frac{S_B^2}{q-1} )</td>
<td>( pr \sigma_B^2 + p \sigma_B^2 / S_{AB}^2 )</td>
</tr>
<tr>
<td>AB</td>
<td>( S_{AB} = p \sum_{i=1}^{(r-1) szorozva (q-1)} ) ( i )</td>
<td>( r-1 )</td>
<td>( S_{AB}^2 = \frac{S_{AB}^2 + \sigma_e^2}{(r-1)} )</td>
<td>( S_{AB}^2 / S_R^2 )</td>
</tr>
<tr>
<td>ismétlés</td>
<td>( S_R = \sum_{i=1}^{r} \sum_{j=1}^{q-1} ) ( i )</td>
<td>( rq szorozva (p-1) )</td>
<td>( S_R^2 = \frac{S_R^2}{rq(p-1)} )</td>
<td>( S_R^2 )</td>
</tr>
<tr>
<td>teljes</td>
<td>( S_0 = \sum_{i=1}^{r} \sum_{j=1}^{q-1} ) ( i )</td>
<td>( rqp-1 )</td>
<td>( S_0^2 )</td>
<td>( S_0^2 )</td>
</tr>
<tr>
<td>A</td>
<td>( S_A = qp \sum_{i=1}^{r-1} ) ( i )</td>
<td>( r-1 )</td>
<td>( S_A^2 = \frac{S_A^2}{r-1} )</td>
<td>( qp \Phi(A) + ps_A^2 / S_{AB}^2 )</td>
</tr>
</tbody>
</table>
Keresztosztályozás egy rögzített és egy véletlen faktor szerint: véletlen blokk

\[ \Phi(A) = \frac{\sum_{i=1}^{r} \alpha_i^2}{r-1} \]

Az A sorban
\[ \Phi(A) = \frac{\sum_{i=1}^{r} \alpha_i^2}{r-1} \]

ennek értéke zérus, ha minden \( i = 0 \) (a technológiák között nincs különbség), ekkor az \( s_{\alpha}^2 \) szórásnéményzet várható értéke megegyezik az \( s_{\alpha}^2 \) szórásnéményzet várható értékével. Ha \( i = 0 \), az \( s_{\alpha}^2 \) szórásnéményzet várható értéke megegyezik az \( s_{\alpha}^2 \) szórásnéményzet várható értékével. Eszerint az A és a B faktor hatását az \( AB \) kölcsönhatás szórásnéményzetéhez képest vizsgálhatjuk, az \( AB \) kölcsönhatás pedig az ismétlések szórásnéményzetéhez viszonyítva. Ha nincs ismétlés, a kölcsönhatás nem értékelhető ki a modellből. A penicillin-peldában éppen ez a helyzet.

2. A példa folytatása

Végezzük el a varianciaanalízist a példa adataival!

A 15.3. táblázat - 3. táblázat utolsó oszlopában található \( p \) érték (az elsőfajú hiba elkövetésének valószínűsége) alapján megállapítható, hogy a technológia nem szignifikáns, a kukoricalekvár viszont igen, mivel \( p = 0,041 \) 0,05.

<table>
<thead>
<tr>
<th>Az forrása</th>
<th>Szabadsági fok</th>
<th>Szórásnéményzet a szórásnéményzet várható értéke</th>
<th>( F )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>B: kuk.lekvár</td>
<td>4</td>
<td>66,0</td>
<td>( 4\sigma_B^2 + \sigma_{AB}^2 / \sigma_{AB}^2 = 3,51 )</td>
<td>0,041</td>
</tr>
<tr>
<td>AB: kuk.lekv.*technol.</td>
<td>12</td>
<td>18,8</td>
<td>( \sigma_{AB}^2 + \sigma_{\epsilon}^2 )</td>
<td></td>
</tr>
<tr>
<td>A: technol.</td>
<td>3</td>
<td>23,3</td>
<td>( 5\Phi(A) + \sigma_{A}^2 / \sigma_{AB}^2 = 1,24 )</td>
<td>0,339</td>
</tr>
</tbody>
</table>

Adjunk becsület a szignifikáns \( ? \)-re (a kukoricalekvár okozta ingadozás varianciájára):

\[ \sigma_B^2 = \frac{S_B^2 - S_{AB}^2}{4} = 11,8 \]

A penicillin-peldában a kukoricalekvár nem igazi faktor, csak arra használtuk a két faktor szerinti keresztosztályozás tervét, hogy a technológiák (az A faktor) hatását megszüntssük a kukoricalekvárétól (B faktor) és a kölcsönhatástól (AB).

Ugyanezt az elvet (a kísérletek blokkokban való elrendezését) használtuk fől az önkontrollos diéta-kísérletekről végzett páros t-próbáiból arra, hogy a páciensek közötti különbségeket ne hagyjuk belekeveredni a diéta észlelt hatásába. Ott a rögzített faktor két szintje a diéta előtti és utáni állapot (\( r = 2 \), a véletlen faktor a páciens (\( g = 10 \) szinttel). A blokkfaktor mindig egy szintjét kombináltuk a vizsgálati kivánt faktor mindkét szintjével. Úgy is fogalmazhatunk, hogy a páciensek alkotta blokkokon belül vizsgáltuk a diéta hatását.
Keresztosztályozás egy rögzített és egy véletlen faktor szerint: véletlen blokk

Az ANOVA-táblázatokban a szórásnégyzetek várható értékének összevetésével különbségeket látunk a két rögzített, két véletlen illetve egy rögzített és egy véletlen faktort tartalmazó tervek között. Nevezetesen, ha a tervben véletlen faktor van, nem az ismétlés, hanem a két faktor közötti kölcsönhatás szórásnégyzete az F próbabastatisztika nevevoie.

3. Hierarchikus osztályozás


Ha egy B faktor szintjei különbözőek az A faktor különböző szintjein, hierarchikus tervről beszélünk. Az angol nyelvű szakirodalomban nested design a neve, ez arra utal, hogy a B faktor szintjei az A faktoréiba vannak ágyazva. Az előbbi példában az állat-faktor szintjei vannak a gyógyszer-faktoréiba ágyazva.

4. Példa

(Kemény S., Deák A.: Kísérletek tervezése és értékelése, Műszaki Könyvkiadó, Budapest, 2000, p. 267)


15.4. táblázat -

<table>
<thead>
<tr>
<th>tétel</th>
<th>minta</th>
<th>mg</th>
<th>tétel</th>
<th>minta</th>
<th>mg</th>
<th>tétel</th>
<th>minta</th>
<th>mg</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>eleje</td>
<td>2,59</td>
<td>2</td>
<td>eleje</td>
<td>2,57</td>
<td>3</td>
<td>eleje</td>
<td>2,56</td>
</tr>
<tr>
<td>1</td>
<td>eleje</td>
<td>2,60</td>
<td>2</td>
<td>eleje</td>
<td>2,57</td>
<td>3</td>
<td>eleje</td>
<td>2,58</td>
</tr>
<tr>
<td>1</td>
<td>közep</td>
<td>2,62</td>
<td>2</td>
<td>közep</td>
<td>2,58</td>
<td>3</td>
<td>közep</td>
<td>2,56</td>
</tr>
<tr>
<td>1</td>
<td>közep</td>
<td>2,60</td>
<td>2</td>
<td>közep</td>
<td>2,58</td>
<td>3</td>
<td>közep</td>
<td>2,60</td>
</tr>
<tr>
<td>1</td>
<td>közep</td>
<td>2,62</td>
<td>2</td>
<td>közep</td>
<td>2,58</td>
<td>3</td>
<td>közep</td>
<td>2,57</td>
</tr>
<tr>
<td>1</td>
<td>vége</td>
<td>2,57</td>
<td>2</td>
<td>vége</td>
<td>2,59</td>
<td>3</td>
<td>vége</td>
<td>2,57</td>
</tr>
<tr>
<td>1</td>
<td>vége</td>
<td>2,57</td>
<td>2</td>
<td>vége</td>
<td>2,59</td>
<td>3</td>
<td>vége</td>
<td>2,56</td>
</tr>
<tr>
<td>1</td>
<td>vége</td>
<td>2,58</td>
<td>2</td>
<td>vége</td>
<td>2,57</td>
<td>3</td>
<td>vége</td>
<td>2,57</td>
</tr>
<tr>
<td>1</td>
<td>eleje</td>
<td>2,60</td>
<td>2</td>
<td>eleje</td>
<td>2,58</td>
<td>3</td>
<td>eleje</td>
<td>2,55</td>
</tr>
</tbody>
</table>

A hierarchikus osztályozást erre a példára az

\[ y_{ijk} = \mu + \alpha_i + \beta_j + \alpha \beta_{ij} + \epsilon_{ij} \]

1. ábra szemlélteti.

Az
Gyártott tételek

<table>
<thead>
<tr>
<th>Minták</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. ábrán a minta szintjeinél (beállításainál) zárójelben jelezzük, hogy a kettő közül az elsőről vagy másodikról van szó, de ez nem a minta-faktor két reprodukálható szintjét jelenti, a második adag első mintája semmilyen értelemben nem tartozik jobban össze az első adag első mintájával, mint az első adag második mintájával. Nincs mód a tétel és mintavétel faktorok közötti kölcsönhatás vizsgálatára, mert ahhoz az kellene, hogy ugyanazt a fajta mintát mindig a tételből venné.

A tétel véletlen faktor, mert ha a kísérlet sorozatot még egyszer el akarnánk végezni, természetesen csak másik tételek választásával tehetnénk. A minta is véletlen faktor, mert ha ugyanazzal a tétellel kísérleteznénk még, nem vehetnénk még egyszer ugyanazt a mintát belőle.

Vizsgáljuk meg, hogy a tételek között van-e különbség, és a minták között van-e különbség, és adjunk becslést a gyártott tételek, a vett minták és az ismételt analízisek eltérésének szórásnégyzetére (varianciájára)!

A modell a következő:

\[ y_{ijk} = \mu + \alpha_i + \beta_j(i) + \epsilon_{k(ij)} \]

ahol, az i-edik tétel hatása \((i = 1, \ldots, r)\); \(\alpha_i\) az i-edik tételel vett j-edik minta hatása \((j = 1, \ldots, q)\); \(\beta_j(i)\) az i-edik tételel j-edik mintájának k-adik analízisénél \((k = 1, \ldots, p)\) elkövetett hiba. Az hibák egymástól függetlenek, és a statisztikai próbához feltételezzük, hogy normális eloszlást követnek \(\epsilon\) varianciával. Ugyanezt tételezzük fő a faktorok hatásáról és kölcsönhatásukról.

A második faktor szintjeinek hatását azért jelölj \(A\) helyett \(\alpha\), mert a \(B\) faktor minden egyes szintje csak az \(A\) faktor egy bizonyos (i-edik) szintjéhez tartozik. Ugyanezt jelképezi a \(B(A)\) jelölés.

A nullhipotézisek a következők:

\(H_{0A} : \alpha^2 = 0\) (a tételek között nincs különbség);

\(H_{0B} : \beta^2 = 0\) (az egy tételel vett minták között nincs különbség).

Az 15.5. táblázat - ANOVA táblázat 5. táblázatban látható a négyzetösszegek, szabadsági fokok és szórásnégyzetek kifejezése.

15.5. táblázat - ANOVA táblázat

<table>
<thead>
<tr>
<th>Az forrása</th>
<th>eltérés- négyzetösszeg</th>
<th>Szabadsági fok</th>
<th>szórásnégyzet</th>
<th>a szórásnégyzet várható értéke</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>B(A)</td>
<td>(S_{B(A)} = p \sum_i^r (q-1))</td>
<td>(S_{B(A)}^2 = \frac{S_{B(A)}^2}{r(q-1)})</td>
<td>(S_{B(A)}^2 = \frac{S_{B(A)}^2}{r(q-1)})</td>
<td>(\frac{S_{B(A)}^2}{S_{R}^2})</td>
<td></td>
</tr>
<tr>
<td>Ism.</td>
<td>(S_{R} = \sum_i \sum_j q(p-1))</td>
<td>(S_{R}^2 = \frac{S_{R}^2}{r(q-1)})</td>
<td>(S_{R}^2 = \frac{S_{R}^2}{r(q-1)})</td>
<td>(\frac{S_{R}^2}{S_{R}^2})</td>
<td></td>
</tr>
</tbody>
</table>
Keresztosztályozás egy rögzített és egy véletlen faktor szerint: véletlen blokk

<table>
<thead>
<tr>
<th>Az forrása</th>
<th>eltérés-négyzetösszeg</th>
<th>Szabadsági fok</th>
<th>szórásnégyzet</th>
<th>a szórásnégyzet várható értéke</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$S_A = q p \sum (\cdot)^{r-1}$</td>
<td>$q p \sigma_A^2 + p \sigma^2$</td>
<td>$S_A^2 = \frac{S_A^2}{r-1}$</td>
<td>$S_A^2$</td>
<td>$\frac{S_A^2}{S_B(A)}$</td>
</tr>
</tbody>
</table>

Ha a tételek között a valóságban nincs különbség (az, hogy különböző tételek vannak, nem növeli az ingadozás szórását), $?_I^{2} = 0$, és az $S_A^{2}$ szórásnégyzet várható értéke megegyezik az $S_B(A)$ szórásnégyzet várható értékével. Ha $?_B^{2} = 0$, az $S_B(A)$ szórásnégyzet várható értéke megegyezik az $S_R$ szórásnégyzet (az ismétlés ismétlése) várható értékével, ezerint az A faktor hatását a B beágyazott faktor (a minta) szórásnégyzetéhez képest vizsgálhatjuk, a B faktor hatását pedig az ismétlés szórásnégyzetéhez viszonyítva.

A varianciaanalízis eredményeit a 15.6. táblázat - A varianciaanalízis eredményei 6. táblázat foglalja össze.

15.6. táblázat - A varianciaanalízis eredményei

<table>
<thead>
<tr>
<th>Az forrása</th>
<th>eltérés-négyzetösszeg</th>
<th>Szabadsági fok</th>
<th>szórásnégyzet</th>
<th>a szórásnégyzet várható értéke</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>B(A) hatása (mintavétel)</td>
<td>$r(q-1)=6$</td>
<td>$S_B(A)^2=0,00005$</td>
<td>$3\sigma_B^2 + \sigma_e^2$</td>
<td>$\frac{S_B(A)^2}{S_R^2}$</td>
<td>$=3,97$</td>
<td>0,011</td>
</tr>
<tr>
<td>Ismétlés (analízis)</td>
<td>$r(q-1)=18$</td>
<td>$S_A^2=0,000126$</td>
<td>$\sigma_e^2$</td>
<td>$=0,0001515$</td>
<td>$=0,011$</td>
<td>0,123</td>
</tr>
<tr>
<td>A hatása (tétel)</td>
<td>$r-1=2$</td>
<td>$S_A^2=0,0001515$</td>
<td>$9\sigma_A^2 + 3\sigma_B^2$</td>
<td>$\frac{S_A^2}{S_B(A)}$</td>
<td>$=3,030$</td>
<td>0,123</td>
</tr>
</tbody>
</table>

A $p$ értékek alapján úgy dönthetük, hogy a tételek közötti különbség nem szignifikáns, a tételen belüli minták közötti különbség viszont igen.

4.1. A varianciakomponensek becslése

Az ismétlés (analízis) varianciája: $?_I^{2} = S_B^{2} = 1,26$ szorozva $10^4$

A mintavétel varianciája: $?_{minta}^{2} = S_B^{2} = 0,000050 – 0,000126 / 3 = 1,25$ szorozva $10^4$

A tételek különbözőségének varianciája: $?_{tétel}^{2} = S_A^{2} – S_B^{2} / qp = 0,0001515 – 0,00050 / 9 = 1,13$ szorozva $10^4$

A példában az elemzés, a mintavétel és a tételek közötti eltérés varianciája azonos nagyságrendű.

5. Megjegyzés a kísérelti tervek és a modellek alkalmazásához

Minthogy az adatok elemzése manapság kizárólag statisztikai programok alkalmazásával történik, nincs szükség (és nagyobb adattömeg esetén lehetőség sem) arra, hogy a felhasználó kézzel számoljon a képletek és táblázatok alapján.

Amire szükség van, az az, hogy tisztában legyen az egyes modellek sajátosságaitól (feltételezéseivel és azok következményeivel). Például lássa világosan, hogy többcentrumos vizsgálatnál a centrumokat veheti rögzített faktornak is, ekkor a következtetései azokra a centrumokra lesznek érvényesek, amelyek a vizsgálatban részt vettek, vagy veheti a centrumokat véletlen faktornak, ekkor a következtetéseik a centrumok sokaságára lesznek érvényesek, tehát azokra is, amelyek nem vettek részt a vizsgálatban. A választás következménye az, hogy az összehasonlításoknál (a hatások szignifikánciájának vizsgálatánál) az $F$ próbabestatisztika nevezője nem ugyanaz, a rögzített hatások esetén az ismétlés szórásnégyzete, a véletlen hatások esetén a megfelelő kölcsönhatás.
Nem szabad megfeledkeznie a feltételezések (a szórásnégyzet állandósága, normális eloszlás stb.) érvényességének vizsgálatáról, mihelyt az adatok a kezében vannak. Ha a feltételezések nem érvényesek, az elemzés sem lesz érvényes.

Ugyancsak világosan kell látnia a felhasználónak, hogy a modellt (és a föltehető kérdéseket) nagyban meghatározza a kísérleti terv, például hogy kereszt-osztályozásról vagy hierarchikus osztályozásról van szó. Másrésztt a kísérletek elképzelésekor és megvalósításakor van módunk gondoskodni arról, hogy bizonyos feltételezések az elemzésnél jogosak legyenek, elsősorban a kísérleti hibák függetlensége illetve a véletlen faktor egyes szintjeinek függetlensége a kísérleti hibától, ezeket randomizálással (véletlenszerűsítéssel) valósítjuk meg.
16. fejezet - Több változós elemzések: korreláció és regresszió

Analízis kiterjesztése

Dinya Elek dr.

1. Bevezetés

A Bevezető részben áttekintettük a *korreláció és a regresszióanalízis* egyszerű formáit és alkalmazásait. A mindegyik munkában azonban gyakoribbak az olyan problémák, amikor több változó között kell alkalmazni a korreláció és regresszió analízist. Ilyen esetekben már nem érvényesek az eddig tanultak, szükséges az ismeretek kiterjesztése, általánosítása.

A problémák nehezebbek és a megoldásuk is az, matematikai szempontból. A végeredményt nagymértékben befolyásolja az alkalmazó felkészültsége, tapasztalata, hiszen a probléma megoldásához széles tárhából vállogathat, és nem mindegy, hogy milyen eszközökhöz nyúl a cél érdekében. A következőkben a két fogalom általánosításával foglalkozunk, illetve megismerkedünk a nemlineáris regresszióval is.

2. Többszörös korreláció

Ha kettőnél több változó kapcsolatrendszerét vizsgáljuk, akkor lineáris többszörös korrelációról beszélünk és magát az együthatot R-el jelöljük. Az R értéke 0 és 1 között van; 0 érték esetén a változók között nincs lineáris kapcsolat, míg 1 esetén a kapocsolat maximális erejű. Az R² determinációs együthatot a kétváltozós r²-hez hasonlóan értelmezzük:

\[ R^2 = \frac{S_{123}^2}{S_1^2} \]

ahol

\[ S_{123}^2 \]

: az x₁ variációjának az a része amit az x₂ és x₃ változók együttesen magyaráznak,

\[ s_1^2 \]

: az x₁ változó teljes variáciája.

Több változó együttes vizsgálata esetén a változók egymásközi korrelációját a korrelációs mátrix alapján tanulmányozhatjuk (16.1. táblázat - 1. táblázat). Pl. négy változó esetén a korrelációs mátrix alakja

<table>
<thead>
<tr>
<th></th>
<th>X₁</th>
<th>X₂</th>
<th>X₃</th>
<th>X₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>X₁</td>
<td>1</td>
<td>r₁₂</td>
<td>r₁₃</td>
<td>r₁₄</td>
</tr>
<tr>
<td>X₂</td>
<td>r₁₂</td>
<td>1</td>
<td>r₂₃</td>
<td>r₂₄</td>
</tr>
<tr>
<td>X₃</td>
<td>r₁₃</td>
<td>r₂₃</td>
<td>1</td>
<td>r₃₄</td>
</tr>
<tr>
<td>X₄</td>
<td>r₁₄</td>
<td>r₂₄</td>
<td>r₃₄</td>
<td>1</td>
</tr>
</tbody>
</table>

A főátloiban 1 áll, mivel egy változónak önmagával való korrelációja maximális értékű. A mátrix szimmetrikus mivel a főátlo alatt és félelt az elemek azonosok: r₁₂ = r₂₁ vagy r₃₄ = r₄₃, stb. Általában kifejezve az azonosságot: rᵢᵣ = rᵢ. Ezeket a korrelációs együthatókat *inter-korrelációs* (interklassz) együthatóknak is nevezzük. Nem tévesztendő össze az intraklassz-korrelációs együthatóval, amely egy összehasonlító elemzés során nem összfegygett mér, hanem a vizsgált személyek variabilitását hasonlítja pl. a diagnózist végző orvosok variabilitásához.
A korrelációs együttátható a két változó közötti kapcsolat erősségét mutatja, de ebből az értékből nem érzelhető a többi változó befolyásoló hatása. Az is előfordulhat, hogy pl. az $r_{12}$ értékre az $x_3$ vagy $x_4$ esetleg mindkettő változó hatással van. Ilyen esetben ahhoz, hogy az $x_i$ és $x_j$ változók között a kapcsolatot más változó(k) hatásától megtisztítsuk, a zavaró hatást el kell távolítani. Erre szolgál a parciális korreláció, amely két változó kapcsolatát úgy vizsgálja, hogy a többi változó hatását konstansnak tekinti.

Legyen három változónk $x_1$, $x_2$ és $x_3$, a közöttük lévő korrelációk $r_{12}$, $r_{13}$ és $r_{23}$. Az $r_{12}$ hatásából az $x_1$ hatását a következő módon szűrjük ki (elsőrendű parciális korrelációs együttátható):

$$r_{123} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1 - r_{13}^2)(1 - r_{23}^2)}}$$

Az $r_{123}$ együttáthatót parciális együttáthatónak nevezzük: az 12,3 indexben a vessző utáni szám jelenti azt a változót, amelynek hatását kiszűrjük. Az $r_{123}$ a reziduálak közötti korrelációt jelenti, az $x_3$ hatásának kiszűrése után.

Legyen $r_{12} = 0,72$, $r_{13} = 0,66$ és $r_{23} = 0,62$. Az $r_{123}$ parciális együttátható

$$r_{123} = \frac{0,72 - 0,66 \cdot 0,62}{\sqrt{(1 - 0,66^2)(1 - 0,62^2)}} = 0,527$$

A determinációs együttáthatók jelentése alapján az $x_1$ és $x_2$ között a magyarázott variancia $r_{12}^2 = 0,72^2 = 0,518$, az $x_3$ hatásának kiszűrése után megmaradt rész

$$r_{123}^2 = 0,527^2 = 0,278$$

Az $x$, befolyásoló hatása tehát $0,518 - 0,278 = 0,24$ vagy %-osan:

$$\frac{0,24}{0,518} = 0,463$$

= 46,3%. A maradék 53,7%-ért más faktor(ok) a felelős(ek).

A parciális korrelációt több tényező kiszűrésére is ki lehet terjeszteni pl. $x_3$ és $x_4$ kiszűrése

$$r_{1234} = \frac{r_{124} - r_{134}r_{234}}{\sqrt{(1 - r_{134}^2)(1 - r_{234}^2)}}$$

formula alapján lehetséges (másodrendű parciális korrelációs együttátható). A parciális korrelációs együttátható szignifikanciát, a $H_0$: $r_{123} = 0$ hipotézist, a következő statisztikával ellenőrizhetjük

$$t = \frac{r_{123}}{\sqrt{1 - r_{123}^2}} \sqrt{N - 3}$$

amely $df = N - 3$ szabadságfokú t-eloszlást követ.

### 3. Többváltozós regressziószámítás

A kétváltozós regresszió tulajdonságai egyszerűen általánosíthatók három vagy több változó egyidejű vizsgálatára is. Ilyen esetben az egyenlet alakja $y = a_0 + a_1X_1 + a_2X_2 + ... + a_nX_n$, ahol

$X_i$: különböző regressziós változók

$a_i$: regressziós együttáthatók.
Többváltozós elemzések: korreláció és regresszió

A független változók között nemcsak folytonos, hanem nominális (dummy változók) változók is megengedettek. A többváltozós vizsgálatok vagyis több változó bevonása a vizsgálatba értékesebb, komplexebb vizsgálat, mivel a nyerhető információ is sokoldalúbb. Azonban azt is szem előtt kell tartani, hogy több változó esetén az eredmény nehezebben értelmezhető.

A többváltozós vizsgálatnál az egyik legfontosabb szempont, hogy X, változók független legyenek egymástól, vagyis a változók között ne legyen kapcsolat. A problémát multikolinearitásnak nevezzük. Az egymással kapcsolatban lévő változókat ki kell hagyni a vizsgálatból. A multikolinearitás vizsgálatára a változók korrelációs mátrixának determinánsa is felhasználható: $R = 0$ estén a változók között a kapcsolat maximális, $R = 1$–nél a változók függetlenek.

A számítógépes programok kiszámolják az $R^2$-t és az ún. módosított $R^2$-t (adjusted $R^2$). Az $R^2$ jelentése ebben az esetben az, hogy az $Y$ variánciájának a változók hány %-át magyarázzák. A módosított $R^2$ érték kisebb, és megbízhatóbb mértéke a regresszió jóságának, mivel ez az érték már mintafüggetlen.

Az analízis eredményét a kiugróérték erőteljesen befolyásolhatja, hasonlóan a kevés esetszám is. Többváltozós analízisnél az esetszámmal vonatkozó ökológia: a szükséges esetszám legalábbos hatzorosan legyen az $X$ változók számának.

3.1. Regressziós eljárások

A többváltozós lineáris regresszió használatakor a számítógépes programok a független változók szűréséhez – vagyis annak megállapítására, hogy mely független változók gyakorolnak tényleges hatást a függő változóra – automatikus segítséget nyújtanak a stepwise eljárások révén. A módszer lényege abban áll, hogy az algoritmus egy F-próba értéket számol: a maradék négyzetösszeg (elhagyva a kérdéses változót) és a teljes maradék négyzetösszeg (a kérdéses változó is benne van) különbsége osztva a teljes szórásnégyzettel. Ha az F értéke kisebb mint egy előre beállított eltávolítási érték, akkor az éppen vizsgált változó elhagyható, s kimarad a végső modellből.

- **Forward stepwise**
  Az eljárás a megadott változók közül próbál mindig egy újabbat bevonnai a modellbe (esetleg egy már bevontat el is hagyhat) szintén az F-próba segítségével. Ha az F értéke nagyobb mint a beállított érték, akkor megfelelővé válik a változó bevonása, s kezdődik előlről az eljárás.

- **Backwards stepwise**
  Minden egységes változóra meghatározzuk az F értékét, s amelyiknél ez kisebb mint az előre definiált érték, akkor a változó kimegy (esetleg bevonódik egy már kihagyott) a regresszióból, s az eljárás a következő változó vizsgálatával folytatódik.

3.2. Logisztikus regresszió

Kétfajta logisztikus regressziót használhatunk: bináris (a megfigyelt eseménynek csak két állapota van) vagy polychotomus (a megfigyelt esemény több állapotú) regressziót. A továbbiakban csak a bináris regresszióval foglalkozunk.

A vizsgált $Y$ esemény lehet pl. a szívinfarktus (bejövedett vagy nem következett be), transzplantáció eredménye (a beültetett szerv kilődött vagy nem kődött ki) a tüdőrák megfigyelésének az eredménye egy prospectív vizsgálat során (kiáltatva a megfigyelt egyéneknek a tüdőrák vagy sem). Ilyen esetekben – az $x_i$ független változók egyaránt tartalmazhatnak folytonos és nominális adatokat –, az $Y$ esemény bekövetkezési valószínűségét logisztikus regresszióval becsüljük. Az eljárás nagyon hasonlít a korábban megismert lineáris regresszióhoz

$$y = a + \sum_{i=1}^{N} b_i x_i$$

azzal a különbséggel, hogy az egyenlet jobb oldalán álló $x_i$ változók eloszlásáról nem követeljük meg a normális eloszlást. Az $x_i$ változók azok a zártfokfaktorok, amelyek segítségével becsülini akarjuk az $Y$ esemény valószínűségét $P(Y=1)$. 

197
Mivel $Y$ csak két értéket vehet fel, a szokásos lineáris regresszió nem alkalmazható.

Ha vesszük a $p/(1-p)$ kifejezést, ahol a $p$ a vizsgált esemény valószínűsége, akkor ehhez az értékhöz a $(0, +)$ intervallum tartozik, de az $ln[p/(1-p)]$-hez viszont már a $(-, +)$ intervallum. Az utóbbi azonos az

$$y = a + \sum_{i=1}^{N} b_i x_i$$

lineáris regressziós kifejezés intervallumával.

Legyen $u=[x_1, x_2, ..., x_N]$ az a vektor, amely a prediktor $x_i$ változókat (rizikófaktorokat) tartalmazza. Vizsgáljuk az $Y=1$ esemény bekövetkezését logisztikus regresszióval. A regressziós modell alakja

$$\ln\left[\frac{P(Y=1|u)}{1-P(Y=1|u)}\right] = \ln\frac{P(Y=1|u)}{P(Y=0|u)} = a + \sum_{i=1}^{N} b_i x_i$$

Az ezzel ekvivalens modell

$$P(Y=1|u) = \frac{\exp\left(a + \sum_{i=1}^{N} b_i x_i\right)}{1 + \exp\left(a + \sum_{i=1}^{N} b_i x_i\right)}$$

Tehát a modell a lineáris regressziót használja az $Y$ esemény valószínűségének a becsleséhez.

Ha egy prediktor változóra igaz, hogy $b_i=0$, akkor az a faktor nincs hatással a vizsgált eseményre. Az eljárás során azt vizsgáljuk, hogy ez a feltevés igaz-e, vagyis teszteljük a $H_0: b_i=0$ hipotézist a

$$z = \frac{b_i}{\text{standard error of } b_i}$$

formulával, ahol $b_i$ a becsült regressziós együttható.

Gyakoriak az olyan vizsgálatok, amikor a prediktor változó hatását csak más zavaró (confounding) változó (pl. az életkor) hatásán keresztül értékelhetjük. A zavaró változóról tudjuk, hogy befolyással van a vizsgált eseményre, ezért figyelembe kell venni az analízis során. Ilyen esetekben a ténylegesen vizsgált rizikófaktorokat korrigáljuk (adjusted) a zavaró változó hatásával, mert csak így kapunk valós eredményt. A logisztikus regresszió alkalmas az ilyen korrekciók elvégzésére.

A módszer további előnye, hogy a független változók eloszlására nincs semmi feltétel. A másik előny, hogy a regressziós koefficienseket ($b_i$) mint relatív kockázati értéket (relatív risk, RR) lehet felhasználni kohort, vagy odds ratio-ként (esély hányadosként, OR) case-control vizsgálatokban. Értelmezésük és számításuk azonos, pl. az odds ratio= $\exp(b_i)$ kifejezéssel határozható meg.

A logisztikus regresszió használata előtt az $y$ függő változót binárisára kell kódolni ($0$ = az esemény nem következett be, $1$ = az esemény bekövetkezett). A számítógépes programok a regressziós koefficiensek mellett ($a$ $b_i$) az OR értékeket és azok 95%-os konfidenциaintervallumát is meghatározzák.

A számítási eljárás bonyolultabb mint a lineáris regresszióról. Általában az iteratív maximum likelihood módszert használnják a számítógépes programok.

A logisztikus regresszió alkalmazásánál vegyük figyelembe a következőket:

az egyének egymástól függetlenül, random módon válasszuk a mintába legalább 5 - 10 esemény jusson mindegyik vizsgált prediktor változóra.
3.3. Probit regresszió

Maga a kifejezés az angol „probability unit” szavak összevonásából keletkezett és a dózis-hatás (tehát kvantálás hatásviszgálat) összefüggés egyik gyakran alkalmazott statisztikai eszköze.

A dózis-hatás összefüggés azt vizsgálja, hogy az egyes dózisok milyen mértékben (%) mértékben változnak a kísérlet során. Abból a feltéveből indulunk ki, hogy az egyedek érzékenysége log-normális eloszlású, tehát véve a dózisok logaritmusát normális eloszlást kapunk. Maga a normális eloszlás arra ad felvilágosítást, hogy egy adott log-dózis értéknél mekkora rész a reagálók relatív gyakorisága. Első lépésként a %-os értékeket probit-transzformáljuk, megpedig úgy, hogy a standard normális eloszlást úgy bontjuk részekre, hogy mindegyikbe egyforma sok eset jusson. A negatív számok elkerülése végett 5öt adunk a normál deviatumokhoz és probitnak nevezzük az így kapott értékeket. Így, a probit 5 érték megfelel az 50%-os cumulative gyakoriságnak, a probit 6 érték a 84,3%-os cumulative gyakoriságnak, a probit 3 érték meg a 3,12%-os cumulative gyakoriságnak. A probitok képzésére táblázatok állnak rendelkezésre.

A probit-transzformáció eredményeképpen a probit (y) és a log-dózis (x) között már lineáris a regressziós kapcsolat:

\[ y = \beta_0 + \beta_1 x \]

A paraméterek becslése (maximum-likelihood módszerrel) iteratív módon történik. A számolás igen hosszadalmas, ezért gépi programokat használunk a feladatra.

A vizsgálatban általában arra vagyunk kíváncsiak, hogy az egyenes kitüntetett pontjaihoz milyen dózis értékek tartoznak pl. az ED\(_50\) esetén. Ez az érték a regressziós egyenes a és b paramétereinek ismeretében – majd az eredményeinek véve az anti logaritmusát – meghatározható.

A probit és a logit-függvény között formailag alig van eltérés, nehéz eldönteni, hogy az adott probléma megoldására melyik módszer az alkalmassabb. Sokoldalúsága miatt ma már inkább a logit metódust alkalmazzák a vizsgálatokban.

3.4. Poisson regresszió

Abban az esetben, ha az adatok nem normális eloszlásúak, de Poisson eloszlást követnek pl. kohort adatok, prospektív vagy retrospektív adatok, standardizált halálozási arány illetve minden, időben ritkán bekövetkező (pl. ritka betegség vizsgálata. különböző földrajzi helyeken) eseménynél, ott az elemzésre ez a regresszió fajta használatos.

Legyen \( r \) az i-edik megfigyelés értéke az \( n \) megfigyelésből, és legyen \( x=(x_1, x_2, ..., x_p)^T \) a \( p \) magyarázó változók vektora. A regressziós modell feltételezi, hogy a \( r \) = \( g(x) \) alakban felírható, ahol a \( g \) egy függvény (multiplikatív modellnél feltesszük, hogy \( g(x')=\exp(x') \), a \( = (r_1, r_2, ..., r_p) \) a regressziós koeficientek vektorá és \( x' = x_1 + x_2 + ...+x_p \) pedig a lineáris prediktor.

Az i-edik megfigyelésre vonatkozó sűrűségfüggvény ekkor a következő alakú

\[ f(r, x) = \frac{\exp(-\exp(x'))}{r!} \]

\( r! \)

A feladat itt is a magyarázó változók együthatóinak a meghatározása, de a számolási eljárás valamelyest bonyolultabb mint a lineáris regresszióval. Az együthatók ismeretében a regresszió egyenlete felfírható.

4. Nem-lineáris regresszió
A regresszió analízisben akkor beszélünk nem-lineáris regressziószámításról, ha a függő és független változók közötti kapcsolatot nem írható le a korábban megismert lineáris függvényel. Ilyen esetben a pontokra egy görbe vonal illeszkedik a legjobban (lásd 1/f. ábra). Azt az eljárást, amelyel a ponthalmazra legjobban illeszkedő görbe egyenletét keressük görbeillesztésnek nevezzük.

Speciális esete a problémának az az eset, amikor a regressziós egyenlet adott \( g_1(x), \ldots, g_m(x) \) nem-lineáris függvényekkel lineáris alakban felírható.

\[
f(x) = y = a_0 + a_1 g_1(x) + \ldots + a_m g_m(x)
\]

ahol \( a \) együttható paraméter értékek ismeretlenek. A paraméterek becsülésére a maximum-likelihood illetve a legkisebb négyzetek módszere használatos.

Általánosan használt eljárás, hogy a regressziós egyenletet a változók transzformációjával linearizáljuk pl. logaritmus alkalmazásával. Sokszor azonban az \( f(x) \) alakja nem ismert, becsülése problémás. Ilyen esetekben az adatok ábrázolása, a scatter-plot diagramm segíthet az \( f(x) \) meghatározásában. Általában a polinomiális megközelítés (másod-, harmad-, esetleg negyedfokú polinom) szokott segíteni a probléma megoldásában. A közéltö polinomok fokszámát variancia-analízissel szekvenciálisan határozhatjuk meg.

Az Excel program például öt fajta nem-lineáris görbeillesztési lehetőséget kínál fel az alábbi lépések megtétele után (természetesen a lineáris kapcsolatot is támogatja a program):

- rajzoljuk meg a két változó \( x \) a független, \( y \) a függő változó) grafikonját
- térjünk át grafikonszerkesztő módba (kattintsunk kétszer a grafikonra)
- kattintsunk magára a pontokat összekötő vonalra és az Insert/Trendline sor kiválasztásával az alábbi görbeillesztési lehetőségek közül választhatunk:

  - logaritmus: \( y = a + b \ln(x) \)
  - polinom: \( y = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \)
  - hatványfüggvény: \( y = a x^b \) lineáris alakja: \( \ln(y) = \ln(a) + b \ln(x) \)
  - exponenciális: \( y = a e^{b x} \) lineáris alakja: \( \ln(y) = \ln(a) + b \ln(x) \)
  - mozgó átlagolás (görbe simítás)

Ezek közül a pontfelhőnek legjobban megfelelő illesztési módszert válasszuk ki. Az Options pontban kérhető, hogy a program a teljes egyenletet, az \( R^2 \) értéket (az illeszkedés jóságát) és az illesztett görbét jelenítse meg.
III. rész - Harmadik rész
Tartalom

17. Szoftverhasznlati útmutató ......................................................... 203
   1. A statisztikai programokról .................................................. 203
18. Az interneten elérhető szabad szoftverek, shareware szoftverek, kereskedelmi programsomagok 204
   1. Ingyenes programok ............................................................. 204
      1.1. Rice Virtual Lab in Statistics: Analysis Lab .......................... 204
      1.2. JavaStat ................................................................. 204
      1.3. StatCrunch (WebStat) .................................................... 204
      1.4. VassarStats (Richard Lowry) ......................................... 204
      1.5. The R Project for Statistical Computing ............................ 204
   2. A statisztikai próbák ereje .................................................... 205
19. Milyen feladathoz milyen szoftvert használjunk? .......................... 206
   1. Adat-tárak az interneten .................................................. 206
20. Az egyszerű statisztikai számításokhoz optimális szoftverek köré ........... 207
   1. Általános célú programok ................................................... 207
   2. Mintaszám meghatározás ..................................................... 207
   3. Adat amputáció ............................................................... 207
21. Feladatmegoldás táblázatkezelővel ........................................... 208
   1. Az első fajta hiba megadása ............................................... 208
22. Biostatisztika, szimulációk az interneten (angol nyelven) ............... 209
   1. Simulációk, kalkulátorok .................................................... 209
23. Biostatisztikai szimulációk statisztikai szoftvercsomagokkal ............... 210
   1. Oktató programok az interneten ........................................ 210
   2. Introductory statistics: concepts, models, and applications (D.W. Stockburger) 210
17. fejezet - Szoftverhasznlati útmutató

1. A statisztikai programokról

Dinya Elek dr.

A gyakorlatban igen sokféle statisztikai programmal találkozhatunk, amelyek vagy valamilyen speciális igényt elégtének ki vagy pedig általános célú, kommersziális szoftverek. Az előbbire példa az adatbányászati programok, az utóbbira pedig az SPSS. A programok színvonalában jól nyomon követhető a számítástechnikai haladás: a programok igényessége, felhasználóbaráti szemlélete, szolgáltatása együtt alakul(t) a számítástechnikai eszközök és az operációs rendszerek fejlődésével.

A szoftverek először nagyszámítógépes környezetben működtek (ilyen volt pl. a BMDP program, amely R-22 számítógépen futott), OS környezetben. Később, ahogy a PC-k kezdtek teret nyerni az alkalmazásokban, úgy jelentek meg az első komolyabb statisztikai szoftverek is ezeken a gépeken (Apple eszközökön a 80-as évek közepén). Ezzel párhuzamosan, a korábbi nagygépek egyre jobban elveszítették jelentőségüket, hiszen teljesítményüket a PC-k túlhaladták. Teljesen kiszorítani azonban nem tudják a mainframe gépeket a PC-k, hiszen egy igazi szupergép többet tud mint egy super PC, és ez a tendencia meg is marad, hiszen erőforrásaik nagyobb teljesítményre képesek mint az asztali gépeké. Ezért van meg az a kettőség, hogy a jelentősebb szoftverek mindkét fajta gépen futtathatók.

A statisztikai programok sajátos filozófiával bírnak, amit a program kezelési technikája csak alátámaszt, kiszolgál. Érdemes először mindig a filozófiát megérteni, mert könnyebbé válik a használat elsajátítása.

Az első PC-s statisztikai alkalmazások DOS környezetben főleg BASIC, FORTRAN nyelveken megírt programok formájában kerültek a felhasználókhoz. Alkalmazásukhoz nem ártott ismerni magát a DOS rendszert is. Adatkezelési, futtatási technikájuk nehézkes volt, bár az akkori matematikai statisztika minden ágát lefedték.

A Windows operációs rendszer általánossá válásával a '90-es évek elejétől a programok is nagy változásban mentek át: felhasználóbarátábbá, egyszerűbbé váltak. Az Excel program megjelenése pedig áttörést hozott az adatkezelésben, sőt ki is jelölte az utat az adatbeviteli képernyők formájának számára. Ma már szinte kötelező az Excel formátum használata, a sor-oszlop szerinti mátrix kitöltése. A programok egyre több mindent tudnak, a sok lehetőség között néha az eligazolás sem könnyű. Az átlagos felhasználó számára gyakran szükségtelenek az egyes területek. Ma a programok használatához alapfokú Windows operációs rendszer és Excel (esetleg más táblázatkezelő program pl. Lotus) ismeretek szükségesek.
18. fejezet - Az interneten elérhetőségek szabad szoftverek, shareware szoftverek, kereskedelmi programcsomagok

Dinya Elek dr., Makara Gábor dr.

1. Ingyenes programok

• Dataplot:
  tudományos vizualizáció, nem lineáris modellezés.

• EasyStat:
  t és F teszt, ANOVA és kontingencia táblák kezelése.

• Irristat:
  alapstatisztikák, kísérletek kiértékelése.

• R:
  nyílt forráskódú szoftver (http://www.r-project.org)

• SISA:
  interaktív elemzések sok egyedi modullal.

• Scilab:
  numerikus számításokhoz, 3D grafikák.

1.1. Rice Virtual Lab in Statistics: Analysis Lab

Az adatfile-t vagy kíválaszthatjuk a felkínált változatok közül, vagy magunk viszünk be adatokat. Az adatok bevitelének függvényében válnak aktívvá a különböző statisztikai eljárásokat működtető programrészek.

1.2. JavaStat

Ez az internethely (http://members.aol.com/johnp71/javastat.html) rengeteg internetes információt összgyűjtött a különböző statisztikával kapcsolatos anyagokról. Jól használható az interaktív, internetes statisztikai számításokra utaló rész.

1.3. StatCrunch (WebStat)

Bejelentkezést és regisztrációt igénylő Honlap, melyen egyszerűbb eljárásokat használhatunk.

1.4. VassarStats (Richard Lowry)

Sokféle statisztikai eljárást lehet a Vassar College honlapján lévő programokkal elvégezni. Nagyon hasznos.

1.5. The R Project for Statistical Computing
Ez a szabad szoftver kategóriába tartozó GNU projekt, amely egy számítógépes nyelvi környezetet ad haladó statisztikai számítások és grafikák készítéséhez. Hasonló a széles körben elterjedt S rendszerhez, és a legtöbb esetben futtatja az S-ben írt programokat is. Használata komolyabb tanulást és erőfeszítést igényel.

2. A statisztikai próbák ereje

Power analysis for numerous designs and statistics John Pezzullo

Ezen a lapon összeágyújtották az interneten a statisztikai próbák erejével kapcsolatos számítási lehetőségeket.

Power analysis for ANOVA designs by Michael Friendly

Statistical considerations for clinical trials and scientific experiments

Power analysis for ANOVA designs by Michael Friendly

Statistical considerations for clinical trials and scientific experiments by David Schoenfeld

Power and sample size page by Russ Lenth. Kitűnő honlap, ahol igen hasznos programok segítenek a minta méretének megtervezését.
19. fejezet - Milyen feladathoz milyen szofvert használjunk?

Makara Gábor dr.

Kutatók számára fontos fogalmak, a sporttudományból vett példákon bemutatva.

Journal of Statistics Education

Ez az internetes tudományos folyóirat a statisztikai oktatás kérdéseivel foglalkozik, de cikkei és példái hasznosak lehetnek azok számára is, akik tanulják és használják a biostatisztikát. (http://www.amstat.org/publications/jse/)

1. Adat-tárak az interneten

Számos helyen lehet találni olyan adatokat, amelyek a számítások kipróbálásához tanuláshoz igen alkalmazsak.

20. fejezet - Az egyszerű statisztikai számításokhoz optimális szoftverek köre

Dinya Elek dr.

1. Általános célú programok

- **BMDP (BioMedical Data Processing):**


- **SAS (Statistical Analysis System):**


- **S-Plus:**

  professzionális programrendszer. Nem annyira népszerű mint a SAS rendszer.

- **SPSS (Statistical Programs for Social Sciences):**

  kiváló Windows alapú program. Mély elemzésre is lehetőséget ad, de eszköztára nincs olyan mély mint a SAS-é. A grafikai reprezentációja elsőrangú.

- **STATISTICA:**

  ugyanazt lehet elmondani mint a d) pontban. Saját formátum nyelve van és SAS (szűkített) modulokat is lehet benne készíteni.

2. Mintaszám meghatározás

- **nQuery Advisor:**

  igazi professzionális program az adott feladatra.

- **Power a Precision:**

  jó színvonalú program.

3. Adat amputáció

- **Solas:**

  kiváló program a hiányzó értékek kezelésére.
21. fejezet - Feladatmegoldás táblázatkezelővel

1. Az első fajta hiba megadása

A khi-négyzet teszt elvégzése Excel táblázatkezelővel

A táblázatkezelő programok sikeréhez az is hozzájárult, hogy viszonylag könnyű velük számos statisztikai vizsgálatot elvégezni. A most tárgyalás alatt levő khi-négyzet teszt elvégzésére az Excel táblázatkezelő programban egy egyszerű beépített függvény áll rendelkezésre [CHITEST(megfigyelt gyakoriságok helye, várható gyakoriságok helye)]. A bemutatja az általunk már kiszámolt példára, hogy hogyan szükséges az adatainkat megadni a táblázatkezelőben, hogy a kiértékelés megtörténhessen.

A felső és alsó része ugyanaz az Excel táblázatnál két különböző verziójú megjelenítése. A B3 és C4 közötti területre íródak be a megfigyelt gyakorisági adatok. Tehát ez a kontingencia táblázat helye. Egyaránt látjuk a felső és az alsó táblázatoknál is, a számok cellákban belüli balra illetve jobbra jobb címke alapján elválasztva vele az értékeket. %.

A D3-as cellába az a képlet került, hogy "+SUM(B3:C3)", azaz számmázzuk (adjuk össze) a B3 és C3 közötti cellatartalmakat. Úgy, hogy az értékeket számoljuk ki, emellett a B9 és C10 közötti cellákba kerültek a várható gyakorisági értékek (segéd kontingencia táblázat). Elegendő kézzel beírni a B9 cellába a "$D3*B5/$D5" formulát és lemásolni a tőle jobbra levő, majd az alattuk levő két cellába. A dollár jel az utána álló karakter (oszlopbetű vagy sorszám) abszolút voltára utal. Ha egy hivatkozás abszolút, akkor a másolókor is megmarad az értéke. Ennek a fajta felírásnak a hasznossága talán nem látszik eléggé e 2-szer 2-es táblázatnál, de nagyon megbízható, ha a táblázat pl. 5-ször 10-es. A – mindössze ellenőrzésképpen – kiszámoljuk a segéd kontingencia táblázat peremgyakoriságait.

A D13 cellába van beirva a CHITEST(megfigyelt gyakoriságok helye, várható gyakoriságok helye) függvény, konkrétan: "+CHITEST(B3:C4,B9:C10)". A beépített Excel függvény ilyenkor kiszámítja azt a szignifikanciaszintet, amely mellett a nullhypotézis esetlegesen el lehet vetni. Ha ez az érték nagyobb, mint 0,05 = 5%, akkor mégsem szokás a nullhypotézist elvetni. Ha pedig kisebb, mint 5%, akkor elvetjük, és az 5% helyett használható az a hibavalószínűség, amelyet az Excel kiszámított. A példánkban ez a szignifikanciaszint 0,0384 = 3,84%. Míg a táblázatok alapján csak annyit állíthatunk, hogy a tényleges első fajta hiba kisebb, mint 5% (továbbá esetleg azt, hogy 1% és 5% közé esik).
22. fejezet - Biostatisztika, szimulációk az interneten (angol nyelven)

*Makara Gábor dr.*

**1. Szimulációk, kalkulátorok**

A szimuláció haszna, hogy nehezen érthető, vizualizálható fogalmakat számítással, grafikus illusztrációval segít elsaajátítani. A szimuláció segítségével gyorsan lehet különböző esetszámú kísérletek eredményeit modellezni, tetszőleges számban ismételni.

A szimulációk általában Java programnyelven készülnek, és sok esetben csak egyes böngészők használata, és optimális beállítása esetén futnak jól. Legyünk figyelemmel a technikai követelményekre, a böngésző típusára, a Sun Java környezet meglétére.

A szimulációk, kalkulátorok ma már elégé elterjedtek, és előfordul, hogy néhány szimuláció vagy valószínűségi kalkulátor az oktatóprogramok részeként kerülnek bemutatásra.


Számos statisztikai fogalom és eljárás bemutatása Java appletek segítségével.

Sokféle kalkulátor a GraphPad cég honlapjáról

Elloszlások értékeinek számolására, többszörös összehasonlítás esetén a valószínűség számítására, kísérleti egyedek random csoportosításához, még kémiai számításokhoz is kínál gyors segítséget. ([http://www.graphpad.com/quickcalcs/index.cfm](http://www.graphpad.com/quickcalcs/index.cfm))
23. fejezet - Biostatisztikai szimulációk statisztikai szoftvercsomagokkal

Makara Gábor dr.

1. Oktató programok az interneten


Egyszerű, angol nyelvű bevezető színtű internetes oktatóanyag. Van mellette egy angol nyelvű, önálló matematikai alapok átismerésére szolgáló rész is.

Szimulációs része Java applet-ek (Hotlist for Java applets) segítségével mutatja be a medián és az átlag viszonyát, a hisztogram készítésében az osztályok és a mintaszám szerepét, a lineáris regresszió tulajdonságait, és a diszkrét eloszlások tulajdonságait.

2. Introductory statistics: concepts, models, and applications (D.W. Stockburger)

Jól követhető, részletes internetes tanfolyam angolul. (http://www.psychstat.smsu.edu/SBK00.htm)

A benne lévő mintaeloszlási szimulációt az Internet Explorer újabb változatával lehet jól megtekinteni, használni. (Használatához néhány Active-X bővítményt kell telepíteni.)

Valószínűség kalkulátor: http://www.psychstat.smsu.edu/pdf/pdfj.htm

Segít kiszámítani egy adott eloszlás esetén a paraméterekhez tartozó görbe alatti területet, a normális, a t-, F- és a Khi-négyzet eloszlások esetében. Igen jó demonstrációs eszköz.

A binomiális kalkulátor a binomiális eloszlás egyes értékeinek kiszámításában nyújt segítséget. (http://www.psychstat.smsu.edu/pdf/BinomialCalculator.htm)

Számos példán mutat be statisztikai fogalmakat, eljárásokat. (http://www.psychstat.smsu.edu/scripts/dws148f/exercises.asp)


Irodalomjegyzék