

Kiváltott agyi jelek informatikai feldolgozása 2018

Statisztika - Gyakorlat

Kiss Gábor

IB.157.

kiss.gabor@tmit.bme.hu

Példa I (Vonat probléma)

- Aladár 25 éves és mindkét nagymamája él még: Borbála és Cecília. Mindkét nagymamájához vonattal kell menni, amik ugyanarról a peronról indulnak de különböző vonatok. Mindkét vonat 10 percenként jár. Aladár ugyanannyira szereti a két nagymamáját, ezért a véletlenre bízta, hogy melyikhez menjen és mindig arra a vonatra száll fel amelyik hamarabb jön. Az elmúlt 30 napon felírta melyik nagymamánál volt: „B, B, B, C, B, B, B, B, B, C, B, B, B, B, B, C, B, B, B, B, B, B, B, B, C, B, B, B, B, B”. Jól működött-e a stratégiája? Ha nem, mi lehetett az oka?

Példa I (Vonat probléma) - Megoldás

- 4-szer volt C-nél, 26 szor B-nél. (A 15-15 helyett) Annak az esélye, hogy 5 nél kevesebbszer jön C feltéve ha $\frac{1}{2}$ valószínűséggel választaná:
 - $27405 + 4060 + 435 + 30 + 1 = 31931$
 - $2^{30} = 1073741824$
 - Vizsgált esetek/Összes esetek $\approx 0,003 \%$ \Rightarrow Nem valószínű.
- Feltehetőleg a C vonat körülbelül 1 perccel a B vonat után indul.

Átlag és szórás

- Legyen egy minta átlag értéke m . Adjunk hozzá a minta minden eleméhez c konstans. Mennyi lesz az új átlag? Szorozzuk meg c konstanssal. Most?
- Legyen egy minta korrigált szórás értéke s . Adjunk hozzá a val. változóhoz c konstans. Mennyi lesz az új szórás? Szorozzuk meg c konstanssal. Most?

Átlag és szórás

- Átlag: $\frac{\sum_{i=0}^n x_i}{n}$ ($= \bar{x}$),
- $\frac{\sum_{i=0}^n (x_i+c)}{n} = c + \frac{\sum_{i=0}^n x_i}{n}$ ($= c + \bar{x}$)
- $\frac{\sum_{i=0}^n (x_i*c)}{n} = c * \frac{\sum_{i=0}^n x_i}{n}$ ($= c * \bar{x}$)
- Korrigált szórás: $\sqrt{\frac{\sum_{i=0}^n (x_i-\bar{x})^2}{n-1}}$ ($= s_n^*$)
- $\sqrt{\frac{\sum_{i=0}^n ((x_i+c)-(\bar{x}+c))^2}{n-1}}$ ($= s_n^*$)
- $\sqrt{\frac{\sum_{i=0}^n ((x_i*c)-(\bar{x}*c))^2}{n-1}} = \sqrt{\frac{\sum_{i=0}^n (x_i-\bar{x})^2 * c^2}{n-1}}$ ($= c * s_n^*$)

Példa II (Fizetés)

- Megmértük 25 ember keresetét (ezer ft) Magyarországon és Ausztriában 1990-ben és 2000-ben és normáltuk az infláció mértékével.

1990, Mo: 123, 133, 134, 105, 200, 113, 144, 132, 127, 130, 131, 140, 124, 128, 141, 127, 134, 144, 127, 135, 113, 132, 132, 125, 100

1990, Au: 133, 145, 154, 143, 110, 133, 200, 142, 132, 140, 151, 152, 134, 123, 121, 147, 127, 154, 132, 132, 139, 128, 123, 101, 128

2000, Mo: 127, 132, 154, 135, 175, 123, 137, 155, 120, 135, 136, 154, 116, 127, 131, 130, 136, 154, 128, 142, 101, 132, 157, 144, 135

2000, Au: 143, 166, 144, 152, 142, 133, 160, 152, 154, 140, 139, 166, 164, 95, 181, 167, 211, 134, 137, 134, 155, 175, 154, 132, 140

- Végezzünk statisztikai vizsgálatot

Példa II (Fizetés) - Megoldás

- Átlagok: 130,96 136,96 136,64 150,8
- Szórások: 18,04273 18,54832 15,48892 21,64294 (korrigált)
- Mi kell?
 - Hipotézisek felállítása
 - 90-00 eltérés (2 darab)
 - Mo-Au eltérés (2 darab)
 - Megfelelő próbák kiválasztása
 - Ha átlagot vizsgálunk akkor t vagy u próba (egymintás vagy kétmintás vagy párosított)
 - Ha szórás vizsgálunk akkor f próba (ez kell a kétmintás t próbához)

Példa II (Fizetés) - Megoldás

- Legyen 4 hipotézis: kétmintás t próba
 - Szabadsági fok: 48
 - t értékek: -1,15936 (90 Mo-Au) -1,194325 (Mo 90-00) -2,42761 (Au 90-00) (00 Mo-Au) -2,66022
 - Kritikus érték 95%: kb 2,01 (táblázatból a 97,5%-os kell nincs pontosan ilyen értékünk tehát becsüljük...)
- De lehet az évesnél párosított is...
- F próba:
 - Szabadsági fokok: 24, 24
 - f értékek: 1,056829 1,356944 1,361518 1,952495
 - Kritikus érték 95%: kb. 2,03
- Következtetés...
 - Szignifikáns eltérést lehet kimutatni az „x” próbával az átlagra (vagy szórásra) nézve „y” szignifikancia szint mellett. Az elsőfajú hiba nagysága 1-y. Tehát feltehetőleg a vizsgált valószínűségi változók átlaga (vagy szórása) nem egyezik meg.
 - Nem lehet szignifikáns eltérést kimutatni az „x” próbával az átlagra (vagy szórásra) nézve „y” szignifikancia szint mellett. Tehát feltehetőleg a vizsgált valószínűségi változók átlaga (vagy szórása) megegyezik (legalábbis ezzel a módszerrel nem lehet különbséget kimutatni).

Példa III – Depresszió

- Tudjuk, hogy nők esetén az átlagos alaphang 200 Hz értékű, 13 Hz szórású normális eloszlást követ. Megmértük 10-10 nő átlagos alaphang értékét, ahol az egyik csoport egészséges volt, a másik pedig depressziós:
 - 200, 190, 175, 201, 213, 202, 220, 198, 205, 210 (Hz) (egészséges)
 - 155, 181, 193, 182, 200, 178, 185, 190, 175, 165 (Hz) (depressziós)
- Vizsgáljuk meg, hogy az egészséges nők átlagos alaphang értéke valóban a teljes sokaságból származik (megegyezik-e az átlag értékük a tudottal)!
- Vizsgáljuk meg, hogy a depressziós nők átlagos alaphang értéke mutat-e szignifikáns különbséget az egészséges nők alaphangjához képest!

Példa III – Depresszió

- Átlagok: 201,4 180,4
- Szórások: 12,5 13,2 (korr szórás)
- Mi kell?
 - Mint az előző feladatban
 - U próba (ha tudott az (átlag érték és) szórás)

Példa III – Depresszió

- U próba kritikus értéke 95%-os szign. szint mellett 1,96 (táblázatból a 97,5%-hoz tartozó eloszlás érték kell)
- u érték: 0,340553
- Így az első esetben nem tudunk szignifikáns különbséget kimutatni.
- Második esetben szignifikáns különbség van (u-val: -4,76774)

Példa III – Depresszió

- Mikor melyik próbát használjuk?
 - Amelyiknek a feltételei teljesülnek és minél specifikusabb az adott problémára. Esetleg vizsgáljuk meg több próbával is és ha mindegyik hasonló eredményt mutat, akkor még biztosabb a döntésünk.