

Hipotézisvizsgálat R-ben

1-mintás u -próba

- ▶ Az elmúlt évben egy, az Antarktison talált királypingvin-kolónia esetén a pingvinek átlagos testtömege 15.4 kg volt. Idén ugyanebből a kolóniából megmérték 35 pingvin tömegét, és átlagos értéknek 14.6 kg-ot kaptak. Tegyük fel, hogy a populációban a pingvinek tömege normális eloszlású, 2.5 kg szórással. Döntsünk 95%-os megbízhatósággal arról, hogy idén ugyanannyi-e a kolóniában az átlagos pingvin tömeg, mint tavaly!
- ▶ Egy teát 200 grammos dobozokban árulnak, a csomagológép szórása 4 gramm. A Fogyasztóvédelmi Felügyelőség lemérte öt véletlenszerűen kiválasztott teásdoboz tömegét, melyekre az alábbi, grammban kifejezett értékek adódtak:

196, 202, 198, 197, 190.

Hipotéziseinket pontosan megfogalmazva és feltételezve, hogy a teásdobozok tömege normális eloszlást követ, döntsünk 98%-os szinten, hogy az átlagos töltőtömeg tényleg 200 gramm, avagy kevesebb annál.

1-mintás t -próba

- ▶ Az atlétikai világbajnokságon résztvevő kokszföldi csapat néhány versenyzője arra panaszkodott, hogy a leadott doppingtesztjeiket nem megfelelően analizálták és az egyik szernek túlságosan magas koncentrációját mutatták ki, minek következtében a versenybíróság törölte az eredményeiket. A Kokszföldi Atlétikai Szövetség a laboratóriumot tesztelendő nyolc mintát küldött, melyek mindegyikében a kérdéses anyag koncentrációja pontosan 0.500 g/l volt. A laboratórium az alábbi eredményeket szolgáltatta:

0.485, 0.518, 0.460, 0.530, 0.560, 0.550, 0.490, 0.575.

A laboratórium mérési hibáját normálisnak tételezve fel döntsön 95%-os szinten, igazuk van-e az atlétáknak.

t-próba beépített függvénnyel

- ▶ Töltsük be az R beépített `immer` adatállományát!
- ▶ `library(MASS)`
- ▶ `immer`: ez az adattömb egy árpamező hozamait tartalmazza 1931-ben és '32-ben. Tegyük fel, hogy a hozamok egy adott évben normális eloszlásúak.
- ▶ Használjuk a `t.test` parancsot annak eldöntésére, hogy 1931-ben 100 egység volt-e az átlagos hozam. Használjunk 5%-os szignifikanciaszintet.
- ▶ Döntsük el – az előbbi szintet használva –, hogy megegyezik-e a hozam a 2 évben.

Paraméteres próbák – folytatás

Kétfajta instant kávé oldódási idejét tesztelték, melyekből minden alkalommal azonos mennyiséget tettek 1 dl forrásban lévő vízbe. A kísérletek eredményeit – az oldódási időket másodpercben – az alábbi táblázat tartalmazza:

Mokka Makka	8.2	5.0	6.8	6.7	5.8	7.3	6.4	7.8
Koffe In	5.1	4.3	3.4	3.7	6.1	4.7		

- ▶ Az oldódási időket normálisnak tételezve fel 5%-os szinten igazoljuk, hogy nincs különbség az oldódási idők szórása között.
- ▶ Az előző pontbeli szinten vizsgáljuk meg azt az állítást, hogy a Mokka Makka kávé lassabban oldódik, mint a Koffe In.

Paraméteres próbák – folytatás

A Mindent Tudás Egyeteme harmadéves matematikus hallgatói két zárthelyi dolgozatot írtak statisztikából. Az alábbi táblázat tíz véletlenszerűen kiválasztott hallgató eredményeit tartalmazza:

Hallgató	A	B	C	D	E	F	G	H	I	J
1. dolgozat	57	63	67	82	45	65	53	32	51	27
2. dolgozat	53	62	63	80	46	64	44	28	50	29

A dolgozateredmények eltérését normális eloszlásúnak tételezve fel, döntsünk 95%-os szinten, van-e különbség a két dolgozat nehézségi foka között.

Paraméteres próbák – folytatás

A Felsőkutyaalvi Kerékpárüzem kerékrészlegének vezetője arra gyanakszik, hogy az egyik beszállító által készített küllők hosszúsága igencsak változékony. Gyanújának ellenőrzése céljából az adott beszállító termékeiből véletlenszerűen kiválasztott 20 darabot és megmérte azok hosszát. A hossz szórásnégyzetének a minta alapján számolt torzítatlan becslése 1.0369mm^2 . A beszállító állítása szerint a küllők hosszának szórása 0.75mm . A küllők hosszát normális eloszlásúnak tételezve fel ellenőrizzük, megalapozott-e a részlegvezető gyanúja. Döntsünk 95%-os szinten.

Nemparaméteres χ^2 -próbák

- ▶ Nagy mintaelemszám szükséges (osztályonként 6-10 már jó).
- ▶ Mindig jobboldali kritikus tartományt használunk.

Illeszkedésvizsgálat: tényleg az adott eloszlásból származik a minta?

```
chisq.test(megfigyeltgyak,p=varhatogyak)
```

Függetlenségvizsgálat: az adatokat 2 ismérv szerint csoportosítva vizsgáljuk, hogy az ismérvek függetlenek-e egymástól.

```
chisq.test(kontingenciatablazat)
```

Előjeldpróba

Feladat: Egy üdítőgyártó cég megváltoztatta egyik italának az összetételét és kíváncsiak, vajon jobban szeretik-e majd az új ízt a vásárlók. Így megkértek 18 főt, hogy teszteljék az új italt. A kísérlet során kiderült, hogy 5 résztvevőnek ízlett jobban az új ital, a többiek a régi ízt preferálják.

Döntsünk 5%-os szignifikanciaszinten arról, hogy egyformán népszerű-e a két íz!

```
binom.test(pozitiv,osszes)
```

Feladat

Az alábbi adatok 12 *Turbo tudás* módszerrel felkészített hallgató vizsgapontszámait tartalmazzák (a maximális pontszám 50):

36 26 30 34 42 24 30 45 32 19 35 38.

Közismert, hogy a hagyományos módszerrel tanulók körében a pontok mediánja 30.

- (a) Az előjelpróba segítségével döntsünk 95%-os szinten, hogy az új módszerrel megszerzett pontok magasabb mediánértékkel bírnak-e.
- (b) Döntsünk a Wilcoxon-féle előjeles rangösszeg próbával is, a szint legyen a korábbi.

```
wilcox.test(...)
```

Szórásanalízis

Egy franchise gyorsétterem 3 új menüt tesztel. Ahhoz, hogy lássák, egyenlően népszerűek-e, kiválasztanak 18 éttermet véletlenszerűen, és 6-ban az első menüt kezdik el árulni, 6-ban a másodikat, 6-ban pedig a harmadikat. Egy hét után a következő eladási mennyiségeket könyvelték:

Menü 1	22	42	44	52	45	37
Menü 2	52	33	8	47	43	32
Menü 3	16	24	19	18	34	39

Döntsünk 95%-os megbízhatósággal arról, hogy egyenlően népszerűek-e a menük, feltéve, hogy az eladások – menüként – normális eloszlást követnek.

```
aov(eredmeny ~ faktor)
```