

S1: Alapfogalmak és elméleti tudnivalók

- Standardizáljuk a következő kis mintát: 5 9 4 17 12 8 11.
- Egy normális eloszlásból vett minta átlaga 20, modusa 15, mediánja 18, terjedelme 45, variációs együtthatója 50%, első néhány eleme:
 $x_1 = 16$ $x_2 = 28$ $x_3 = 29$ $x_4 = 11$ $x_5 = 22$.
Ezt a mintát kell *standardizálni*.
 - Számítsuk ki a standardizált értékeket (z_1 – z_5) a megadott 5 elemre!
 - Ha a minta összes elemét standardizálnánk, mi lenne az új minta *átlaga* és *szórása*?
 - Próbáljuk meg a többi felsorolt statisztikai jellemző standardizált értékét is kiszámítani! (Vigyázat: nem mindegyiket lehet!)
- Egy $\mu = 26$, $\sigma = 13$ paraméterű normális eloszlásból származnak a következő adatok: 10 23 19 40 27 25 44 33 24 30 56 6 22 25 13 21. Standardizáljuk ezeket (az eloszlás paramétereivel), majd emeljük őket négyzetre. Ha elég jól jellemzi a minta az eloszlást, a kapott számok átlaga 1 körül lesz. Miért?
- Reakcióidő-mérés eredményei a következő, ezredmásodperceket jelentő számok:
311 372 294 305 305 360 297 324
 - Adjuk meg ezen adatok átlagát, mediánját, modulusát, terjedelmét, szórását!
 - Ha azt szeretnénk, hogy mintánk (közelítőleg) normális eloszlású legyen, mit kell tennünk?
 - Ha tudjuk a választ, járjunk el annak megfelelően, és számítsuk ki az új minta átlagát!
 - Ezt a számot az eredeti adatokból is megkaphattuk volna. Hogyan?
(Ellenőrzésképpen számítsuk is ki!)
- Valamely jól meghatározott állatfaj egyedeinek hosszúsága normális eloszlást követ. A hosszúságok várható értéke 27,20 cm, az elméleti szórás 1,48 cm. A normális eloszlás táblázatának segítségével mondjuk meg, hogy milyen határok közé esik az állatok hosszúsága 50, 90, ill. 99,9 százalékos valószínűséggel, és azt is, hogy az állatok hányadrésze esik a 26–30 cm-es határok közé!
- A statisztikai jellemzők eloszlásáról találunk alább néhány igaz és hamis állítást. Válasszuk ki az igaz állításokat!
 - Az eloszlás elméleti.
 - Az eloszlás normális.
 - Az eloszlásgörbe alatti terület a minta elemszámával (a gyakoriságok összegével) egyenlő.
 - Az eloszlásgörbe alatti terület 1-gyel egyenlő.
 - Az eloszlás várható értéke csak akkor egyezik meg a mediánnal, ha az eloszlás szimmetrikus.
 - Az eloszlás általában végtelen terjedelmű.
 - Az eloszlás szórása csökkenthető a mérések pontosabbá tételével.
 - Az eloszlás szórása növelhető a mérések számának csökkentésével.

7. Függetlenség esetén a minta, ill. a változó több jellemzőjének értéke nulla. Az alábbiak közül melyek ezek?
- a) hibakorlát
 - b) korrelációs együttható
 - c) variációs együttható
 - d) meghatározottsági együttható
 - e) kovariancia
 - f) variancia
 - g) regressziós együttható
 - h) Q_h
 - i) Q_r
 - j) szabadságfok
8. Bevezetünk egy kezelést és kísérlettel vizsgáljuk annak hatásosságát. A kísérlet alapján hatásosnak tartjuk a kezelést – de tévedünk. Hogyan nevezzük az így elkövetett hibát, és mit tudunk mondani annak nagyságáról?
9. Bevezetünk egy kezelést és kísérlettel vizsgáljuk annak hatásosságát. A kísérlet alapján hatástalannak tartjuk a kezelést – de tévedünk. Hogyan nevezzük az így elkövetett hibát, és mi módon lehet az ilyen hibát csökkenteni?
10. Ha egyoldali próba helyett kétoldali próbát végzünk ugyanabban a feladatban, a végeredményül kapott valószínűség a) növekszik, b) csökken, c) nem változik, d) nem lehet megmondani. Mi igaz?
11. Az első fajta hiba növelésével a második fajta hiba
- a) nő
 - b) nem változik
 - c) csökken
 - d) nőhet is, csökkenhet is: nem lehet megmondani.
- Válasszuk ki a helyes választ!
12. Egy hatásos gyógyszert kipróbálva, azt hatástalannak találtuk, mert túlságosan kevés kísérletet végeztünk. Ezzel
- a) első fajta
 - b) második fajta
 - c) standard
 - d) metodikai
 - e) valószínűségi
 - f) véletlen
- hibát követtünk el. Melyek a helyes állítások?
13. Két gyógyszerkészítmény (Q és W) hatását akarjuk összehasonlítani, és e célból ugyanavval a kórokozóval beoltott patkányok két csoportja közül az egyiket Q-val, a másikat W-vel kezeljük. A gyógyító hatást az egyes csoportokban életben maradt patkányok számával mérjük. Fogalmazzuk meg, hogy mi ebben a feladatban a nullhipotézis, és adjuk meg – névvel, képlettel – az ennek vizsgálatára alkalmas statisztikai próbát!
14. Egy gyógyszer mellékhatását kutatva, a leölt kísérleti állatok máját szignifikánsan nagyobbak találtuk a kontroll állatokénál.
- a) Melyik az a statisztikai próba (a tanult eljárások közül), amelynek a segítségével ezt a megállapítást tehetjük?
 - b) Mint minden statisztikai próba, ez is bizonyos feltételekhez van kötve. Említsünk néhányat a legfontosabb feltételek közül!

15. Egy gyógyszerről azt tartják, hogy májnagyobbodást okoz. Két kísérleti állatcsoport közül (mindkettőben 18 állat volt) az egyiket ezzel a gyógyszerrel kezelték, a másikat nem. Mindkét csoport állatait ugyanakkor leölve megmérték a májtömegüket; a gyógyszerrel kezelt csoportban x_i , a másikban y_i jelöli – grammokban kifejezve – az adatokat. Ugyanakkor tettünk egy megállapítást is: nagyobb-e a máj a normálnál vagy nem. A fentiek alapján a kérdés (okoz-e májnagyobbodást a gyógyszer?) eldöntésére melyik statisztikai eljárást tartjuk a legalkalmasabbnak? Válasszuk ki a kevésbé megfelelő, de azért alkalmazható eljárásokat is!
- χ^2 -próba
 - egymintás t -próba
 - kétmintás t -próba
 - becslés a májak valódi tömegére
 - korrelációs t -próba
 - a két csoport közti korreláció vizsgálata
 - az x_i és y_i adatok közti regresszió meghatározása.
16. Két csoportban 10, illetve 22 személyt vizsgáltunk. Egyrészt a magasságokat hasonlítottuk össze, másrészt azt, hogy melyikben hordanak többen szemüveget. A két vizsgálat eredményeképp $\chi^2 = 18,1$ és $t = 1,81$ adódott. Ennek alapján válaszoljunk a következő kérdésekre:
- Milyen szabadságfoknál kell a χ^2 értéket kikeresni?
 - Eszerint az érték szignifikáns az 5%-os szinten?
 - Milyen szabadságfoknál kell a t értéket kikeresni?
 - Szignifikáns a kapott érték a (kétoldali) 5%-os szinten?
 - Eldöntve, hogy a két próba közül melyik volt alkalmas a magasság és melyik a szemüvegviselés összehasonlítására, válaszoljunk, hogy melyik szempontból tér el szignifikánsan a két csoport.
17. Egyes próbák *robustusak* bizonyos feltételek megsértésével szemben, mások azonban nem. Például a t -próba robustus, az F -próba pedig nem – ugyanarra a feltételre vonatkozólag. Melyik ez a feltétel?
18. Tízfényi csoportban végrehajtott kísérlet után egymintás t -próbát végeztünk. A kapott t -érték pontosan a fele volt annak, ami a kétoldali szignifikanciához – minimálisan – kellett volna. (A szignifikanciaszintet a szokásnak megfelelően választottuk.)
- Mennyi volt az általunk kapott t ?
 - Mekkora csoporton végezzük az új kísérletet, hogy az várhatóan szignifikáns eredményt adjon? Melyek azok a statisztikai jellemzők, amelyek változatlanságát – egyébként jogosan – feltételeztük az új kísérlet megtervezésekor?
 - Ezzel tulajdonképpen „túlbiztosítottuk” magunkat. Miért?
- Még így sem lehetünk azonban biztosak a szignifikancia felől, hiszen a véletlen hatására valamennyi újból számolt statisztikai jellemző eltérő értéket adhat a korábbinál.
19. Két minta összehasonlítására kétoldali t -próbát végeztünk. A minták függetlenek és normális eloszlásúak voltak, mindkettőben 8 elem szerepelt. A kapott eredmény: $t = 2,42$. Válaszoljunk a következő kérdésekre!
- Milyen képlettel számoltuk ki t értékét?
 - A próba előtt meg kellett győződnünk a szórások egyformaságáról. Ez hogyan történt?
 - A kapott t érték elbírálásához táblázatot kell használnunk. Ez a példában milyen szabadságfokon történt?
 - Mekkora az az 5%-os táblázatbeli érték, amelyikhez a kapott t -t, 2,42-t hasonlítjuk?
 - Mindezek alapján mit mondhatunk a minták különbözőségéről, illetve egyformaságáról?