Kurdics János, Toledo Rodolfo

Statisztika feladatgyűjtemény Excel™ támogatással[1]

A projektek az Európai Unió támogatásával,
az Európai Szociális Alap társfinanszírozásával valósulnak meg.

Nyíregyházi Főiskola

Matematika és Informatika Intézet

2010

[1] Microsoft Inc.

Tartalomjegyzék

Bevezető

1. Nevezetes eloszlások a statisztikában

1. Készítsen táblázatot a standard normális eloszlás- és sűrűségfüggvényhez, és szemléltesse a normális eloszlás- és sűrűségfüggvényeket grafikonon $m = 0$ és $\sigma = 1, 2, 3, 4, 5$ esetén! Használja a mellékelt megoldást és a sablont!
   Tanulmányozza a leíró statisztika elméletét!
   Javaslat: Alkalmazza a `STNORMELOSZL`, `NORM.ELOSZL` beépített függvényeket!

2. Készítsen táblázatot a χ^2 eloszlásfüggvényhez $n = 1 - 30$ szabadságfokkal, illetve készítse el a χ^2-próba táblázatát $n = 1 - 30$ szabadságfokra $90 - 99,9$%-os szinteken, és szemléltesse a χ^2 eloszlás- és sűrűségfüggvényeket grafikonon $n = 5, 10, 15, 20, 25$ szabadságfokra! Használja a mellékelt megoldást és sablont!
   Tanulmányozza a leíró statisztika elméletét!
   Javaslat: Alkalmazza a `KHI.ELOSZLÁS` (ami $1 - F_n(x)$ értéket ad vissza), `INVERZ.KHI` (ami $F_n^{-1}(1 - x)$ értéket ad vissza) beépített függvényeket! A sűrűségfüggvényhez nincsen beépített függvény, meg kell adni azt explicit képletével, amihez alkalmazza a `KITEVŐ` és `GAMMALN` (ami a gammafüggvény természetes alapú logaritmusát adja vissza) beépített függvényeket!

3. Készítsen táblázatot a Student eloszlásfüggvényhez $n = 1 - 30$ szabadságfokkal, illetve készítse el a t-próba táblázatát $n = 1 - 30$ szabadságfokra $90 - 99,9$%-os (kétoldali) szinteken, és szemléltesse a Student eloszlás- és sűrűségfüggvényeket grafikonon $n = 10, 30, 50, 70, 90$ szabadságfokra! Használja a mellékelt megoldást és sablont!
   Tanulmányozza a leíró statisztika elméletét!
   Javaslat: Alkalmazza a
   `T.ELOSZLÁS` (ami $1 - F_n(x)$ értéket ad vissza), `INVERZ.T` (ami $F_n^{-1}(1 - x)$ értéket ad vissza) beépített függvényeket! A sűrűségfüggvényhez nincsen beépített függvény, meg kell adni azt explicit képletével, amihez alkalmazza a `KITEVŐ` és `GAMMALN` (ami a gammafüggvény természetes alapú logaritmusát adja vissza) beépített függvényeket!

4. Készítsen táblázatot a standard gamma eloszlás-és sűrűségfüggvényhez $n = 1 - 30$ szabadságfokkal, és szemléltesse a standard gamma eloszlás- és sűrűségfüggvényeket grafikonon $n = 5, 10, 15, 20, 25$ szabadságfokra! Használja a mellékelt megoldást és sablont!
   Tanulmányozza a leíró statisztika elméletét!
   Javaslat: Alkalmazza a
   `GAMMA.ELOSZLÁS` beépített függvényt!

5. Készítsen táblázatot a béta eloszlás- és sűrűségfüggvényhez $p = 1,5$, $q = 0,2 - 6$ paraméterrel, és szemléltesse a béta eloszlás- és sűrűségfüggvényeket grafikonon $p = 1,5$, $q = 0,4; 0,8; 1,2; 1,6; 4$ paraméterekkel! Használja a mellékelt megoldást és sablont!
   Tanulmányozza a leíró statisztika elméletét!
   Javaslat: Alkalmazza a
   `BÉTA.ELOSZLÁS` beépített függvényt! A sűrűségfüggvényhez nincsen beépített függvény, meg kell adni azt explicit képletével, amihez alkalmazza a `KITEVŐ` és `GAMMALN` (ami a gammafüggvény természetes alapú logaritmusát adja vissza) beépített függvényeket!

6. Készítsen táblázatot a Kolmogorov függvényhez $x = 0,5$ és $x = 2,02$ közötti értékekre és szemléltesse grafikonon! Használja a mellékelt megoldást és sablont! Részletes megoldási útmutató:
   Tanulmányozza a leíró statisztika elméletét!

7. A (λ, α)-paraméterű Weibull eloszlású valószínűségi változó eloszlásfüggvénye $F(x) = 1 - e^{-\lambda x^\alpha}$ $(\lambda > 0,\ \alpha > 0,\ x \geq 0)$, sűrűségfüggvénye $f(x) = \lambda\alpha x^{\alpha-1}e^{-\lambda x^\alpha}$ $(x \geq 0)$. Határozza meg a várható értéket és a szórást! Készítsen táblázatot a Weibull eloszlás- és sűrűségfüggvényhez $\lambda = 1$, $\lambda = 1 - 30$ paraméterrel, és szemléltesse a Weibull eloszlás- és sűrűségfüggvényeket grafikonon $\lambda = 1$, $\alpha = 1, 2, 3, 4, 5$ paraméterekkel!
   Tanulmányozza a leíró statisztika elméletét!
   Segítség: a várható értékre $\lambda^{\frac{1}{\alpha}}\Gamma(1+\frac{1}{\alpha})$, a szórásnégyzetre pedig $\lambda^{\frac{-2}{\alpha}}\left(\Gamma(1+\frac{2}{\alpha}) - \Gamma^2(1+\frac{1}{\alpha})\right)$ eredményt kell kapnia.

8. Az (m, σ)-paraméterű lognormális eloszlású valószínűségi változó olyan, hogy természetes alapú logaritmusa normál eloszlású. Határozza meg az eloszlás- és sűrűségfüggvényét, várható értékét és a szórását! Készítsen táblázatot a standard lognormális eloszlás- és sűrűségfüggvényhez , és szemléltesse a lognormális eloszlás- és sűrűségfüggvényeket grafikonon $m = 0$,

$\sigma = 1, 2, 3, 4, 5$ paraméterekkel!
Tanulmányozza a leíró statisztika elméletét!
Segítség: a sűrűségfüggvényre $f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-(\ln x - m)^2/2\sigma^2}$ $(x > 0)$ eredményt kell kapnia, illetve a várható értékre $e^{m+\sigma^2/2}$, a szórásnégyzetre pedig $e^{2m+\sigma^2}(e^{\sigma^2} - 1)$ eredményt.

9. Határozza meg a fenti eloszlásokra az $\alpha_3 = \frac{E(\xi - E\xi)^3}{D^3\xi}$ ferdeségi mutatót és az $\alpha_4 = \frac{E(\xi - E\xi)^4}{D^4\xi}$ kurtózist!
   Tanulmányozza a leíró statisztika elméletét!
   Segítség: a normál eloszlásra $\alpha_3 = 0$, $\alpha_4 = 3$.

2. Leíró statisztika

1. Az analizis-vizsga.xls fájl tartalmazza a Nyíregyházi Főiskola PTI szakosainak Analízis II vizsgái végeredményeit a 2009/2010-es tanévben. Ezekkel az adatokkal:

   (a) Készítsen gyakorisági sort, ha a a csoportképző feltétel az egyes ismérvek!

   (b) Oszlopdiagrammal illusztrálja a kapott gyakorisági sort!

   (c) Számítsa ki az adatsor számtani, mértani, harmonikus és négyzetes közepét!

   (d) Állapítsa meg az adatsor további helyzetmutatóit és szóródási mutatóit!

   Tanulmányozza a leíró statisztika elméletét, amely tartalmazza a feladat részletes megoldását!

2. A Nyíregyházi Főiskola egyik testnevelés szakos csoportján testmagasságot mértek. A mérési eredmények a testmagassag.xls fájlban található.

   (a) Számítsa ki az adatsor számtani, mértani, harmonikus és négyzetes közepét!

   (b) A hallgatók hány százaléka magasabb az átlagnál?

   (c) Állapítsa meg az adatsor további helyzetmutatóit és szóródási mutatóit!

(d) Kördiagrammal illusztrálja a hallgatók nemek szerinti megoszlását!

Tanulmányozza a leíró statisztika elméletét! Használja az Excelt a feladat megoldására!

3. A következő táblázat tartalmazza a felvételi ponthatárokat a Nyíregyházi Főiskolán a 2010-es tanévben.

354	240	401	320	357	374	356	371	236	244
200	200	241	400	318	400	400	323	326	366
382	401	420	205	196	288	250	205	202	398
398	398	402	406	211	252	251	447	278	294
224	202	204	248	288	288	358	358	358	318
280	260	209	206	196	400	408	400	342	318
211	211	326	208	211	206	367	366	211	226
248	248	408	248	244	248	306	407	407	

(a) Készítsen osztályközös gyakorisági sort, ha a a csoportképző feltétel a következő: a 201 pont alatti szakok; a 401 pont feletti szakok; 200 és 400 pont között 4 darab 50 pontos hosszúságú osztály!

(b) Készítsen hisztogramot és gyakorisági poligont!

(c) Állapítsa meg az adatsor helyzetmutatóit és szóródási mutatóit!

Tanulmányozza a leíró statisztika elméletét, amely tartalmazza a feladat részletes megoldását! Használja az Excelt a feladat megoldására!

4. Oldja meg az előző feladatot abban az esetben, ha 7 egyforma hosszúságú osztályt készít! Tanulmányozza a leíró statisztika elméletét! Használja az Excelt a feladat megoldására! Hasonlítsa össze az előző feladatban kapott eredményekkel!

5. Egy Budapesten végzett felmérésben azt kérdezték meg 1000 embertől, hogy hány négyzetméteres lakásban laknak. A felmérés eredményét a következő táblázat mutatja.

Terület	Lakások száma
— 30	102
31 — 50	254
51 — 70	249
71 — 90	199
91 — 110	130
111— 150	55
151—	11
Összesen	1000

(a) Állapítsa meg a gyakorisági sor helyzetmutatóit és szóródási mutatóit!

(b) Készítsen hisztogramot és gyakorisági poligont!

Tanulmányozza a leíró statisztika elméletét, amely tartalmazza a feladat részletes megoldását! Használja az Excelt a feladat megoldására!

3. Paraméteres intervallumbecslés

1. Gumiabroncs szállítmányból visszatevéssel 100 elemű mintát veszünk, 8 selejtest találtunk. Adjunk konfindencia intervallumot a p selejtvalószínűségre 95%-os biztonsági szinten. A selejtvalószínűség szabványosan maximum $0,1$ lehet. 90%-os biztonsági szint mellett most 200 elemű mintát veszünk, és 13 selejtest találunk. Szabványos-e a szállítmány? Mit mondhatunk 95%-os biztonsági szinten?
   Tanulmányozza az itt megadott eljárást!
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza a `INVERZ.STNORM` beépített függvényt!

2. Laboratóriumi mérleget vizsgálnak be. A mérőeszköz szórása a gyártó adata alapján ismert, $\sigma = 0,2$. A ξ valószínűségi változó értéke a súlyeltérés mg-ban, normálisnak eloszlásúnak tekinthető. Keressünk konfidencia intervallumot a várható értékre! A szabvány szerint a félhossz nem lehet $0,08$ mg-nál nagyobb. 99%-os biztonsági szinten hány mérést kell elvégezni legalább? Mit mondhatunk 90%-os szinten? Adjunk 99%-os szintű bootstrap konfidencia intervallumot m-re, a szimulációk száma legyen $N = 100$.
   Tanulmányozza az itt megadott eljárást a konfidencia intervallumra illetve

az itteni ismertetést a bootstrap módszerről! Használja a megoldást és a sablont! Részletes megoldási útmutató:

3. Egyfajta chip élettartamát vizsgálják szélsőséges körülmények között. 40 vizsgálatot végeztek, a ξ valószínűségi változó értéke az élettartam órában. Az eloszlás normálisnak tekinthető, a szórása nem ismert előzetesen. Adjunk konfindencia intervallumot a chip várható élettartamára 95%-os biztonsági szinten. Hány órás működés után érdemes a chipet lecserélni? Mit mondhatunk 90%-os szinten? Adjunk 95%-os szintű bootstrap konfidencia intervallumot a várható értékre, a szimulációk száma legyen $N = 100$.
   Tanulmányozza az itt megadott eljárást!
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.T`, `INVERZ.STNORM` és `VÉL` beépített függvényeket!

4. Egyfajta chip élettartamát vizsgálják szélsőséges körülmények között. 40 vizsgálatot végeztek, a ξ valószínűségi változó értéke az élettartam órában. Az eloszlás normálisnak tekinthető, a szórása nem ismert előzetesen. Adjunk konfindencia intervallumot a chip szórására 95%-os biztonsági szinten. Mit mondhatunk 90%-os szinten? Adjunk 95%-os szintű bootstrap konfidencia intervallumot a szórásra, a szimulációk száma legyen $N = 100$.
   Tanulmányozza az itt megadott eljárást!
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.KHI`, `INVERZ.STNORM` és `VÉL` beépített függvényeket!

5. A ξ valószínűségi változó értéke legyen egy adalékanyag viszkozitása. Új gyártási eljárás folyamán más katalizátort alkalmaznak, ez a szórást nem befolyásolja, legyen az η valószínűségi változó értéke az új eljárással készített adalékanyag viszkozitása. 40-40 mérést végeztek, az eloszlások normálisnak tekinthetők. Adjunk a várható értékek különbségére konfindencia intervallumot 95%-os biztonsági szinten. Mit mondhatunk 90%-os szinten? Szerkesszünk 95%-os szintű bootstrap konfidencia intervallumot is, a szimulációk száma legyen $N = 50$.
   Tanulmányozza az itt megadott eljárást!
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.T`, `INVERZ.STNORM` és `VÉL` beépített függvényeket!

6. A ξ valószínűségi változó értéke az az időtartam ms-ban, amíg az igénynek egy szerveren várakoznia kell a kiszolgálásra, az eloszlás jó közelítéssel exponenciálisnak tekinthető. 100 vizsgálatot végeztek. Adjunk a várakozási idő várható értékére konfindencia intervallumot 95%-os biztonsági szinten. Mit mondhatunk 90%-os szinten? Szerkesszünk 95%-os szintű bootstrap konfidencia intervallumot is, a szimulációk száma legyen $N = 50$.
   Tanulmányozza az itt megadott eljárást!
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.GAMMA`, `INDEX` és `VÉL` beépített függvényeket! A véletlen indexeket ne a `RAND.BETWEEN` beépített függvénnyel, hanem inkább az `INT(N*VÉL())+1` képlettel számolja!

4. Paraméteres próbák

1. Azt gyanítjuk, hogy egy kristálycukrot csomagoló gép több cukrot tölt be az 1kg-os csomagokba. A gép 3%-os szórással dolgozik. A gyanúnk igazolására vettünk egy 500 elemű mintát, amit itt találunk. Vizsgáljuk meg a kérdést 95%-os biztonsági szinten!
   Tanulmányozza a paraméteres próbák elméletét! Használja az Excelt a feladat megoldására!
   Segítség: alkalmazzunk z-próbát, az ellenhipotézis egyoldali lesz: $m > 1$.

2. Egy alkatrészeket gyártó gyár egyik raktárában találtak olyan menetes orsókat, melyeket régebben gyártották. Az címkéjük szerint az orsók átmérője 2 cm, de ellenőrizni kivánják ezt. Erre egy 40 elemű mintát vettek, amit itt találunk. Vizsgáljuk meg a kérdést 90%-os biztonsági szinten!
   Tanulmányozza a paraméteres próbák elméletét! Használja az Excelt a feladat megoldására!
   Segítség: alkalmazzunk t-próbát, az ellenhipotézis kétoldali lesz: $m \neq 2$.

3. Az előző feladatban említett gyárban szeretnének egy új gépet üzembe helyezni. A közölt paraméterei szerint a gép olyan menetes orsókat tud gyártani, melyeknek átmérője szórása kisebb, mint 0,03 cm. Ellenőrzéskor vettünk egy 100 elemű mintát, amit itt találunk. Vizsgáljuk meg a kérdést 98%-os biztonsági szinten!
   Tanulmányozza a paraméteres próbák elméletét! Használja az Excelt a feladat megoldására!

Segítség: χ^2-próbát alkalmazzunk, az ellenhipotézis egyoldali lesz: $\sigma \neq 0,03$.

4. Minden évben kompetenciamérést végeznek a 6., 8. és 10. osztályokban. Matematikából 600 képességpontot lehet szerezni egy ilyen felmérőlapon. 2009-ben a 6. évfolyamra járó 100 654 tanuló átlag eredménye 489 képességpont volt 99 szórással. Egy kisvárosban 517 hatodikos tanuló írta meg a felmérést. Az egyes eredményeket itt találjuk. Mondhatjuk-e 99 %-os biztonsággal, hogy a kisváros tanulói az országos eredményeknek megfelelően teljesítettek?
   Tanulmányozza a paraméteres próbák elméletét! Használja az Excelt a feladat megoldására!
   Segítség: alkalmazzunk kétmintás t-próbát, az ellenhipotézis kétoldali lesz. Előtte F-próbát alkalmazunk a szórás egyezésének eldöntésére.

5. Az előző feladatban szereplő kompetenciamérésben a tanulók 10 %-a 361 képességpont alatt teljesít. Mondhatjuk-e ugyanezt 95%-os biztonsági szinten a kisvárosban hatodikos tanulói esetén?
   Tanulmányozza a paraméteres próbák elméletét! Használja az Excelt a feladat megoldására!
   Segítség: alkalmazzunk a sokasági arányszámmal kapcsolatos próbát, az ellenhipotézis kétoldali lesz.

6. Egy tanulmányt készítenek a férfiak és a nők különböző szokásairól, többek között azt vizsgálták, hogy hány darab ünneplő ruha van egy embernek. Erről a kérdésről egy véleménykutató cég felmérést végzett. A lekérdezés eredményét itt találjuk (x: férfi, x: nő). Mondhatjuk-e 99 %-os biztonsággal, hogy egyforma szokás tapasztalható a férfiak és a nők között, vagy a nőknek több ünneplő ruhájuk van?
   Tanulmányozza a paraméteres próbák elméletét! Használja az Excelt a feladat megoldására!
   Segítség: Ha az F-próbából nem tudunk szórás egyezést feltételezni, akkor alkalmazzunk a Welch-próbát, az ellenhipotézis egyoldali lesz.
   Megjegyzés: Kétmintás z-próba is alkalmazható, mert a minták elemszáma nagyon nagy.

5. Nemparaméteres próbák

5.1. Illeszkedésvizsgálat

1. Szabályosnak tekinthető-e az a kocka, amelyet $n = 500$-szor feldobva az alábbi gyakoriságokat kaptuk? Legyen az A_i esemény az, hogy i-t dobtunk. A megfelelő relatív gyakoriság legyen k_i (i=1,2,3,4,5,6=r). Nullhipotézis: $p_i = P(A_i) = 1/6$. A módszer a tiszta illeszkedésvizsgálat χ^2 próbával. Döntsünk 95%-os szinten!
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.KHI` beépített függvényt, amely az $F_n^{-1}(1-x)$ értéket adja!

2. Plazmatévék gyártása során egy négyzetcentiméterre jutó hibás pixelek számát vizsgálták a minőségellenőrzés során. 100 mérést végeztek. Döntsünk 95%-os biztonsági szinten, hogy a minta Poisson eloszlásúnak tekinthető-e? A módszer a becsléses illeszkedésvizsgálat χ^2 próbával. Osztályba sorolással A_i teljes eseményrendszert alakítunk ki, $p_i = P(A_i)$, ügyelve arra, hogy az $np_i \geq 10$ feltétel teljesüljön.
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.KHI` és `GYAKORISÁG` beépített függvényeket! Az `INVERZ.KHI` függvény az $F_n^{-1}(1-x)$ értéket adja meg. A `GYAKORISÁG` függvényt tömbképletként kell beírnunk. Az első cellába írjuk a képletet, majd jelöljük ki a megfelelő tartományt, a képletet tartalmazó cellával kezdve. Nyomjuk meg az F2 billentyűt, majd a CTRL+SHIFT+ENTER billentyűket.

3. A ξ valószínűségi változó értéke az az időtartam ms-ban, amíg az igénynek egy szerveren várakoznia kell a kiszolgálásra. 100 vizsgálatot végeztek. Kérdés, hogy az eloszlás jó közelítéssel exponenciálisnak tekinthető-e? Adjunk választ 95%-os biztonsági szinten. A módszer a becsléses illeszkedésvizsgálat χ^2 próbával.
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.KHI` és `GYAKORISÁG` beépített függvényeket! Az `INVERZ.KHI` függvény az $F_n^{-1}(1-x)$ értéket adja meg. A `GYAKORISÁG` függvényt tömbképletként kell beírnunk. Az első cellába írjuk a képletet, majd jelöljük ki a megfelelő tartományt, a képletet tartalmazó cellával kezdve. Nyomjuk meg az F2 billentyűt, majd a CTRL+SHIFT+ENTER billentyűket.

4. Egy folyó vízhozamát mérik egy adott időszakban köbméter/min-ben. 30 évre visszamenőleg van adat, a nullhipotézis az, hogy a minta gamma-eloszlású. Ellenhipotézis: nem gamma eloszlású. Döntsünk 95%-os szinten! Szemléltesse a tapasztalati illetve az elméleti eloszlásfüggvényt grafikonon! Az eljárás az egymintás (becsléses) Kolmogorov-próba.
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `HOL.VAN`, `GAMMA.ELOSZLÁS` és `LOG` beépített függvényeket!

5.2. Homogenitásvizsgálat

1. Egy éven keresztül vizsgálják a hétfői és keddi közlekedési balesetek számát. Nullhipotézis: A két eloszlás megegyező. Ellenhipotézis: Nem egyeznek meg. Döntsünk χ^2 próbával 95%-os szinten! Az eljárás a homogenitásvizsgálat χ^2 próbával.
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.KHI` és `GYAKORISÁG` beépített függvényeket! Az `INVERZ.KHI` függvény az $F_n^{-1}(1 - x)$ értéket adja meg. A `GYAKORISÁG` függvényt tömbképletként kell beírnunk. Az első cellába írjuk a képletet, majd jelöljük ki a megfelelő tartományt, a képletet tartalmazó cellával kezdve. Nyomjuk meg az F2 billentyűt, majd a CTRL+SHIFT+ENTER billentyűket.

2. Véletlenszerűen kihelyezett kvadrátokban kétféle hangyafaj egyedszámát vizsgálják. Nullhipotézis: A két faj egyedszámának eloszlása megegyező. Ellenhipotézis: Nem egyeznek meg. Döntsünk 95%-os szinten! Az eljárás a homogenitás-vizsgálat χ^2 próbával.
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.KHI` és `GYAKORISÁG` beépített függvényeket! Az `INVERZ.KHI` függvény az $F_n^{-1}(1 - x)$ értéket adja meg. A `GYAKORISÁG` függvényt tömbképletként kell beírnunk. Az első cellába írjuk a képletet, majd jelöljük ki a megfelelő tartományt, a képletet tartalmazó cellával kezdve. Nyomjuk meg az F2 billentyűt, majd a CTRL+SHIFT+ENTER billentyűket.

5.3. Függetlenségvizsgálat

1. Piackutatást végezve 200 fős mintán azt vizsgálják, hogy kétféle márkájú termék előnyben részesítése függ-e a vásárlók nemétől. Nullhipotézis: a valószínűségi változók függetlenek. Ellenhipotézis: a valószínűségi változók nem függetlenek. Döntsünk 95%-os szinten χ^2-próbával! Az eljárás a függetlenségvizsgálat. (A próba semmit sem mond arról, hogy melyik a független és melyik a függő változó.)
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.KHI` beépített függvényt, amely az $F_n^{-1}(1-x)$ értéket adja meg.

2. Új hatóanyagot vizsgálnak. A kórokozóval megfertőzött 100 egér felét a hatóanyaggal kezelik, a másik fele a kontrollcsoport, és egy hét után megvizsgálják az egereket. Nullhipotézis: a valószínűségi változók függetlenek, azaz a gyógyszer hatástalan. Ellenhipotézis: a valószínűségi változók nem függetlenek. Döntsünk 95%-os szinten χ^2-próbával! Az eljárás a függetlenségvizsgálat.
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.KHI` beépített függvényt, amely az $F_n^{-1}(1-x)$ értéket adja meg.

3. Mobiltelefonhoz egyfajta chipet három beszállítótól vásárolja a gyártó. A minőségellenőrzés jónak, megfelelőnek vagy selejetesnek minősíti az alkatrészt. 150 darabos mintát vizsgáltak be. Nullhipotézis: a chipek minősége nem függ attól, hogy honnan származik. Ellenhipotézis: a valószínűségi változók nem függetlenek. Döntsünk 95%-os szinten χ^2-próbával! Az eljárás a függetlenségvizsgálat.
   Használja a megoldást és a sablont!
   Javaslat: Alkalmazza az `INVERZ.KHI` beépített függvényt, amely az $F_n^{-1}(1-x)$ értéket adja meg.

4. Vizsgáljuk meg, hogy független-e egymástól az országok GDP-je és a születéskor várható élettartam?[2] Nullhipotézis: a GDP és a várható élettartam függetlenek. Ellenhipotézis: a valószínűségi változók nem függetlenek. Döntsünk 95%-os szinten χ^2-próbával! Az eljárás a függetlenségvizsgálat.
   Használja a megoldást és a sablont! Részletes megoldási útmutató:

[2] Adatforrás: Human Development Reports, ENSZ.

6. Regressziószámítás

1. Szeretnénk megállapítani, hogy van-e összefüggés a csecsemők születési súlya és testhossza között. Ehhez megfigyeltünk 300 csecsemőt, a megfigyeléseket itt találjuk.

   (a) Számítsa ki a Pearson-féle korrelációs együtthatót!

   (b) A determinációs együttható segítségével fejezze ki a kapcsolat szorosságát!

   (c) Állapítsa meg t-próbával 95 %-os biztonsággal, hogy a csecsemők születési súlya és testhossza közötti korreláció szignifikáns-e vagy sem!

   (d) Adja meg és ábrázolja a lineáris regresszió egyenesét!

   Tanulmányozza a regressziószámítás elméletét! Használja az Excelt a feladat megoldására!
   Javaslat: a lineáris regresszió egyenes együtthatóit a `LIN.ILL` beépített függvénnyel lehet megadni, amit tömbképletként kell bevinnünk.

2. Egy digitális pH-merővel végeztünk néhány mérést. Utána a mintákat laboratóriumi körülmények között igényesebb pH-merővel újra mértük. Az eredményeket itt találjuk.

   (a) Adja meg és ábrázolja a lineáris regresszió egyenesét!

   (b) A determinációs együttható segítségével fejezze ki a kapcsolat szorosságát!

   (c) Becsülje meg a regresszió segítségével egy 8-as ph értékű minta laboratóriumi értékét!

   (d) Értelmezze a regresszió egyenes meredekségét!

   Tanulmányozza a regressziószámítás elméletét! Használja az Excelt a feladat megoldására!

3. A szén-dioxid-kibocsátás (1000 tonna) alakulását Magyarországon itt tekintheti meg.

   (a) Készítse el és ábrázolja a lineáris trend egyenesét!

   (b) Állapítsa meg, hogy a lineáris trenddel becsült értékek hány százaléka tér el a valódi értékektől!

   (c) Készítse el és ábrázolja a harmadfokú polinomiális trend görbéjét!

   (d) Állapítsa meg, hogy a harmadfokú polinomiális trenddel becsült értékek hány százaléka tér el a valódi értékektől!

   (e) Becsülje meg a harmadfokú polinomiális trenddel a szén-dioxid-kibocsátás a 2009. évre!

   Tanulmányozza a regressziószámítás elméletét! Használja az Excelt a feladat megoldására!
   Segítség: a trend olyan regressziós görbe, amelynek x komponense időszakok monoton növekvő sorozata 1-től n-ig jelölve. A trenddel becsült értékek a valódi értékektől való eltérés százaléka nem más, mint az eltérés négyzetes közepének és a valódi értékek átlagának aránya százalékban kifejezve.

4. A Magyarországon az 1 éves életkorban várható élettartam százalékos eredményei itt találhatók.

   (a) Készítse el és ábrázolja a másodfokú polinomiális trend görbéjét!

   (b) Állapítsa meg, hogy a trenddel becsült értékek hány százaléka tér el a valódi értékektől!

   (c) Keresse meg azt az évet, amitől kezdve a trend növekedése tapasztalható!

   (d) Becsülje meg a trenddel az 1 éves életkorban várható élettartamot a 2009. évre!

   Tanulmányozza a regressziószámítás elméletét! Használja az Excelt a feladat megoldására!
   Segítség: egy regressziós görbe csúcsait, mint általában más differenciálható görbék esetén, differenciálszámítással oldjuk meg. Másodfokú polinom esetén ez egyszerű feladat.

7. Vegyes feladatok

1. Szennyvíztisztítás során háromféle biológiai szűrőt próbálnak ki. Harminc-harminc mintavételt végeznek el, a szűrt vízben található kólibaktériumok köbmilliméterenkénti számát mérik, az adatokat itt találja.

   (a) A méréssorozattal tűréshatárokat kívánnak adni szűrők hatékonyságára és szórására 95%-os szinten. Keressen konfidencia intervallumot az egyesített minta várható értékére és szórására!

   (b) Nem tételezzük fel a normális eloszlást, mégis tudni szeretnénk, milyen határok közötti az A szűrő hatékonysága. Keressen 95%-os bootstrap konfidencia intervallumot az A minta várható értékére!

   (c) Mekkora lehet a különbség az A és B szűrő szűrőképessége között? Keressen 95%-os konfidencia intervallumot az A és B minta várható értékei különbségére!

   (d) Igaz-e 95%-os szinten, hogy az A szűrő szűrőképessége 8 $1/mm^3$? (Szórását nem ismerjük.)

   (e) Az C szűrő specifikációja szerint a szűrőképessége 10 $1/mm^3$, szórása 3 $1/mm^3$. Megfelel-e a specifikációnak a szűrőképesség 95%-os szinten?

   (f) 95%-os szinten modhatjuk-e azt, hogy az A és C szűrők hatékonysága megegyezik?

   (g) 95%-os szinten modhatjuk-e azt, hogy a három szűrő hatékonysága megegyezik?

   (h) Csoportosítsuk az adatokat $12 - 171/mm^3$ gyenge, $7 - 121/mm^3$ közepes, $2 - 71/mm^3$ kiváló szűrés kategóriák szerint. Lehet-e 95%-os szinten normáleloszlásúnak tekinteni a B szűrőképességet?

   (i) Azonosnak tekinthető-e az A és B szűrőképesség eloszlása?

2. Szilárdtest félvezető vezetőképességét mérjük $\frac{1}{m\Omega cm}$-ben különböző hőmérsékleteken, az adatokat itt találja.

   (a) Adjon tűréshatárokat a vezetőképesség várható értékére és szórására 95%-os szinten a -20, +20C^o-os hőmérsékletintervallumban!

   (b) Nem tételezzük fel a normális eloszlást, mégis tudni szeretnénk, milyen határok közötti a vezetőképesség. Keressen 95%-os bootstrap konfidencia intervallumot a várható értékre!

   (c) Igaz-e 95%-os szinten, hogy a vezetőképesség 2,10 $\frac{1}{m\Omega cm}$? (Szórását nem ismerjük.)

   (d) Adjunk lineáris modellt a jelenségre! Számítsuk ki a D determinációs együtthatót! A korrelációra végezzen t-próbát! A számított értékekre adjon konfidencia-intervallumot!

3. A vér ionizált kalcium és magnézium szintjét mérik 30 páciensben $mmol/L$-ben, az adatokat itt találja.

   (a) Adjunk a két ásvány szintjének várható értékére és szórására 95%-os szinten konfidencia intervallumot!

   (b) Nem tételezzük fel a normális eloszlást, mégis tudni szeretnénk, milyen határok közötti a két ásvány szintjének várható értéke. Keressen 95%-os bootstrap konfidencia intervallumot a minták várható értékeire!

   (c) Igaz-e 95%-os szinten, hogy a Ca-szint átlagosan 1,24 $mmol/L$? (Szórását nem ismerjük.)

   (d) Igaz-e 95%-os szinten, hogy az Mg-szint átlagosan 0,85 $mmol/L$? (Szórását nem ismerjük.)

   (e) Csoportosítsuk az adatokat $1,00-1,16 mmol/L$ alacsony, $1,16-1,32 mmol/L$ normál, $1,16-1,32/mm^3$ magas Ca-szint kategóriák szerint. Lehet-e 95%-os szinten normáleloszlásúnak tekinteni a vér Ca-szintjét?

   (f) Csoportosítsuk az adatokat $0,25-0,65 mmol/L$ alacsony, $0,65-1,05 mmol/L$ normál, $1,05-1,45/mm^3$ magas Mg-szint kategóriák szerint. Lehet-e 95%-os szinten normáleloszlásúnak tekinteni a vér Mg-szintjét?

   (g) Feltételezések szerint az abnormális Ca-szintet a rendellenes Mg-szint okozza. Függetlennek tekinthető-e a Ca és Mg-szint?

4. Akváriumi kísérletekben átokhínár (*Elodea*) relatív növekedési faktorát (RGR) és 1mg szárazanyag klorofiltartalmát (Chl) vizsgálták különböző nitrogén és fényintenzitás szintek mellett[3], az adatokat itt találja.

   (a) Adjunk az Elodea relatív növekedési faktorának a megadott N szintek melletti várható értékére és szórására 95%-os szinten konfidencia intervallumot.

   (b) Adjunk az Elodea Chl tartalmának a megadott fényszintek melletti várható értékére és szórására 95%-os szinten konfidencia intervallumot.

   (c) Nem tételezzük fel a normális eloszlást, mégis tudni szeretnénk, milyen határok közötti a relatív növekedési faktor és a Chl tartalom várható értéke. Keressen 95%-os bootstrap konfidencia intervallumot a minták várható értékeire!

   (d) Igaz-e 95%-os szinten, hogy a relatív növekedési faktor átlagosan 0,04 $1/nap$? (Szórását nem ismerjük.)

   (e) Igaz-e 95%-os szinten, hogy a Chl tartalom átlagosan 20 $\mu g/mg$? (Szórását nem ismerjük.)

   (f) Csoportosítsuk a relatív növekedési faktor adatokat a $-0,035 - 0,00$, $0,00 - 0,035$, $0,035 - 0,07$, $0,07 - 0,105$, $0,105 - 0,14\ 1/nap$ intervallumok szerint. Lehet-e 95%-os szinten normáleloszlásúnak az RGR-t?

   (g) Csoportosítsuk a Chl adatokat $3,2 - 9,8$, $9,8 - 16,4$, $16,4 - 23,0$, $23,0 - 29,6$, $29,6 - 36,2\ \mu g/mg$ kategóriák szerint. Lehet-e 95%-os szinten normáleloszlásúnak tekinteni az Elodea Chl-tartalmát?

   (h) Függetlennek tekinthető-e az N szint és az RGR?

   (i) Függetlennek tekinthetők-e a fényviszonyok és a Chl-tartalom?

   (j) Illesszen másodfokú görbét az RGR adatokra a N szint függvényében!

   (k) Illesszen másodfokú görbét a Chl adatokra a fényviszonyok függvényében!

[3] Szabó Sándor adatai, NYF Biológia Intézet

5. Akváriumi kísérletekben átokhínár (*Elodea*) különböző szinteken való jelenléte mellett békalencse (*Lemna*) száraztömegét (DW) és relatív növekedési faktorát (RGR) vizsgálták [4], az adatokat itt találja.

   (a) Adjunk a Lemna száraztömegének a megadott Elodea szintek melletti várható értékére és szórására 95%-os szinten konfidencia intervallumot.

   (b) Adjunk a Lemna relatív növekedési faktorának a megadott Lemna szintek melletti várható értékére és szórására 95%-os szinten konfidencia intervallumot.

   (c) Nem tételezzük fel a normális eloszlást, mégis tudni szeretnénk, milyen határok közötti a száraztömeg és a relatív növekedési faktor várható értéke. Keressen 95%-os bootstrap konfidencia intervallumot a minták várható értékeire!

   (d) Igaz-e 95%-os szinten, hogy a Lemna száraztömeg átlagosan 0,04 g? (Szórását nem ismerjük.)

   (e) Igaz-e 95%-os szinten, hogy a Lemna növekedési faktor átlagosan 0,1 $1/nap$? (Szórását nem ismerjük.)

   (f) Csoportosítsuk a DW adatokat $0,02-0,05$, $0,05-0,1$, $0,1-0,14$, $0,14-0,18\ g$ kategóriák szerint. Lehet-e 95%-os szinten bétaeloszlásúnak tekinteni a Lemna száraztömegét?

   (g) Csoportosítsuk a relatív növekedési faktor adatokat a $0,05-0,11$, $0,11-0,17$, $0,17-0,23$, $0,23-0,29\ 1/nap$ intervallumok szerint. Lehet-e 95%-os szinten bétaeloszlásúnak az RGR-t?

   (h) Függetlennek tekinthető-e az Elodea szint és a Lemna száraztömeg?

   (i) Függetlennek tekinthető-e az Elodea szint és a Lemna RGR?

   (j) Illesszen másodfokú görbét a Lemna száraztömeg adatokra az Elodea szint függvényében!

   (k) Illesszen másodfokú görbét a Lemna RGR adatokra az Elodea szint függvényében!

---

[4] Szabó Sándor adatai, NYF Biológia Intézet, in: Szabó, S. Scheffer, M., Roijackers, R. M. M., Valuto, B., Braun M., Nagy, P. T., Borics, G, Zambrano, L. (2010): Strong growth limitation of floating plant (Lemna gibba) by a submerged macrophyte (Elodea nuttallii) under laboratory conditions. Freshwater Biology, 55, 681-690.

6. Oldat aranyion tartalmának meghatározásához atom-abszorpciós spektrometria módszerrel történő új mérési eljárást validálnak[5]. Különböző interferáló ionok jelenléte mellett (40-2000-szeres koncentrációban) vizsgálták a mérési módszer szelektivitását, a valós Au koncentráció és a mért érték közötti eltérésre vonatkozó százalékos adatokat itt találja.

   (a) Adjunk az eltérés várható értékére és szórására 95%-os szinten konfidencia intervallumot.

   (b) Nem tételezzük fel a normális eloszlást, mégis tudni szeretnénk, milyen határok közötti az átlagos eltérés. Keressen 95%-os bootstrap konfidencia intervallumot a minta várható értékére!

   (c) Igaz-e 95%-os szinten, hogy átlagosan 1% a hiba? (Szórását nem ismerjük.)

   (d) Csoportosítsuk az adatokat $-8,0 - (-3,5)$, $-3,5 - 1,0$, $1,0 - 5,5$, $5,5 - 10,0$ osztályokba. Lehet-e 95%-os szinten normáleloszlásúnak tekinteni az eltérést?

[5]Balogh József adatai, NYF Kémia Tanszék, in: L. Kocúrová, J.S. Balogh, J. Skrliková, J. Posta, V. Andruch: A novel approach in dispersive liquid-liquid microextraction based on the use of an auxiliary solvent for adjustment of density UV-VIS spectrophotometric and graphite furnace atomic absorption spectrometric determination of gold based on ion pair formation. Talanta 82 (2010), 1958-1964

Hivatkozások

[1] A.C. Aitken. *Statistical mathematics*. Oliver and Boyd, London, Edinburgh, 1947.

[2] P.G. Hoel. *Introduction to mathematical statistics*. John Wiley & Sons, New York, London, Sydney, 1962.

[3] O. Lukács. *Matematikai statisztika*. Műszaki, Budapest, 1987.

[4] G.C. Montgomery, D. C.; Runger. *Applied Statistics and Probability for Engineers*. John Wiley & Sons, Inc., New York,, 2002.

[5] F.A.; Boes D.C. Mood, A.M.; Graybill. *Introduction to the theory of statistics*. MacGraw-Hill, New York, 1974.

[6] A. Rényi. *Valószínűségszámítás*. Tankönyvkiadó, Budapest, 1984.

[7] I. Vincze. *Matematikai statisztika*. Tankönyvkiadó, Budapest, 1986.

[8] C.E. Weatherburn. *A first course in mathematical statistics*. Cambridge University Press, Cambridge, 1961.