8. feladatsor, statisztika, 2017. ápr. 10.

(1) Az egy adott email címre egy nap alatt érkező emailek számának eloszlását vizsgálták. Azt feltételezték, hogy ez Poisson-eloszlású 3 paraméterrel, azaz annak valószínűsége, hogy egy nap alatt k email érkezik, $3^k/k! \cdot e^{-3}$ (ahol $k = 0, 1, 2, \ldots$). Ennek a feltételezésnek az ellenőrzésére negyven napon át megfigyelték, hogy hány email érkezik. Az adatokat az alábbi táblázat tartalmazza.

emailek száma	0	1	2	3	4	5	6
napok száma	2	5	7	5	7	3	1

95%-os szignifikanciaszinten elfogadható-e az a hipotézis, hogy az egy nap alatt érkező emailek száma 3 paraméterű Poisson-eloszlású?

(2) Betegek egy hatvanfős csoportját véletlenszerűen két részre osztva egy gyógyszerrel, illetve placebo tablettával kezelték őket, majd feljegyeztés, hogy melyik csoportban hányan gyógyultak meg adott időn belül.

	gyógyszeres kezelés	placebo
meggyógyult	25	13
nem gyógyult meg	14	8

(*a*) Állíthatjuk-e 95%-os szignifikanciaszinten, hogy a valódi gyógyszeres kezelés és a gyógyulás között szignifikáns összefüggés van?

(*b*) Állíthatjuk-e 95%-os szignifikanciaszinten, hogy a valódi gyógyszeres kezelés és a gyógyulás között között szignifikáns pozitív korreláció van?

(3) Egy egyetemi évfolyamon 300-an tanulnak. Megszámolták, hogy a legutóbbi vizsgaidőszakban hányszor buktak az egyes hallgatók. Az eredményeket tartalmazza az alábbi táblázat.

bukások száma	0	1	2	3	4
hallgatók száma	80	113	77	26	4

95%-os szignifikanciaszinten elfogadható-e az a hipotézis, hogy az egy hallgatóhoz tartozó bukások száma binomiális eloszlás 4 renddel és $0,25$ paraméterrel?

(4) Negyven embert kérdeztek arról, hogy milyen italokat szoktak inni. Huszonnyolcan fogyasztanak rendszeresen sört, közülük tizennyolcan rendszeresen bort is. Azok között, akik nem fogyasztanak rendszeresen sört, heten voltak, akik bort viszont rendszeresen szoktak fogyasztani.

(*a*) Állíthatjuk-e 95%-os szignifikanciaszinten, hogy a sör– és borfogyasztás között szignifikáns összefüggés van?

(*b*) Állíthatjuk-e $\alpha = 0,05$ terjedelem mellett, hogy a sör– és borfogyasztás között szignifikáns pozitív korreláció van?

---

7. feladatsor, statisztika, 2017. ápr. 3.

(1) Megmérték nyolc beagle fajtájú kutya testtömegét, az alábbi eredmények adódtak (kg):

16.9, 11.5, 9.7, 14.2, 12.0, 10.8, 13.9, 15.6

Tegyük fel, hogy a beagle esetén a testtömeg normális eloszlású, és 2 a szórása.

(a) Elfogadható-e az a hipotézis 95% szinten, hogy a beagle testtömegének várható értéke 12?
(b) Elfogadható-e 98%-os szinten ugyanez a hipotézis?
(c) Elfogadható-e 95%-os szinten az a hipotézis, hogy a beagle testtömege nem több 13-nál?
(d) Néhány terrier testtömegét is megmérték:

12.0, 11.9, 9.6, 10.6, 14.5

Itt a mérések szórását 1.2-nek feltételezzük. Állíthatjuk-e 95%-os szinten, hogy a beagle fajtájú kutyák nehezebbek a terriereknél?

(2) Egy napon tíz budapesti helyszínen megmérték a NO_2-koncentrációt. Az átlag $352\,\mu\text{g}/\text{m}^3$, a korrigált tapasztalati szórás 8 lett.

(a) 95%-os szinten elfogadható-e az a hipotézis, hogy a koncentráció a $350\,\mu\text{g}/\text{m}^3$ tájékoztatási küszöbérték alatt van?
(b) Elfogadható-e ugyanez a hipotézis 99%-os szinten?
(c) Londonban 20 mérésből az átlagos koncentráció 376, a korrigált tapasztalati szórás 16 lett. $\alpha = 0,05$ terjedelem mellett állíthatjuk-e, hogy Londonban szignifikánsan nagyobb a NO_2 koncentrációja?

(3) Két internet-csatlakozási pont késleltetési idejét mérték meg (másodpercben). Az A ponton $n_1 = 15$ mintából számolva a mérések átlaga $\overline{X}_1 = 0,32$ lett, a korrigált tapasztalati szórás pedig $s^*_{n,1} = 0,0016$. A B ponton $n_2 = 30$ mintából számolva az átlag $\overline{X}_2 = 0,46$, a korrigált tapasztalati szórásnégyzet pedig $s^*_{n,2} = 0,0009$. A késleltetési idő eloszlását normális eloszlásnak feltételezzük, és azt is feltesszük, hogy a két ponton az eloszlások szórása megegyezik.

(a) Elfogadható-e 95%-os szinten, hogy az első ponton $0,3$ másodperc a késleltetési idő?
(b) Elfogadható-e 99%-os szinten, hogy a második ponton a késleltetési idő legfeljebb $0,3$?
(c) $0,05$ terjedelem mellett igaz-e, hogy a két ponton szignifikánsan eltérő a késleltetési idő?

6. feladatsor, 2017. márc. 27.

(1) Egy csoportban a diákok magassága (cm):

180	163	150	157	165	165
191	172	165	168	186	174

(a) Elemezzük a diákok testmagasságát az átlag, a medián, a korrigált tapasztalati szórás, és boxplot ábra (kvartilisek) segítségével.
(b) Mennyi $\hat{F}_n(180)$? Mennyi $\hat{F}_n(190)$?
(c) Mennyi a minta 90%-os kvantilise?
(d) Adjunk 95%-os megbízhatósági szintű konfidenciaintervallumot a testmagasság várható értékére.

(2) Két internet-csatlakozási pont késleltetését mérték meg. Az első ponton $n_1 = 15$ csomagból álló mintából számolva a késleltetési idők átlaga (másodpercben számolva) $\overline{X}_1 = 0,32$ lett, a korrigált tapasztalati szórás pedig $s^*_{n,1} = 0,04$. A második ponton $n_2 = 30$ mintából számolva az átlag $\overline{X}_2 = 0,46$, a korrigált tapasztalati szórás pedig $s^*_{n,2} = 0,03$. A késleltetési idő eloszlását normális eloszlásnak feltételezzük.
(a) Adjunk 95%-os megbízhatósági szintű kétoldali konfidenciaintervallumot az első ponthoz tartozó késleltetési idő várható értékére.
(b) Adjunk 98%-os megbízhatósági szintű kétoldali konfidenciaintervallumot az első ponthoz tartozó késleltetési idő várható értékére.
(c) Adjunk 95%-os megbízhatósági szintű kétoldali konfidenciaintervallumot a második ponthoz tartozó késleltetési idő várható értékére.
(d) Tegyük fel, hogy az első ponthoz tartozó késleltetési idő szórása $0,042$. Ezt felhasználva adjunk újra 95%-os megbízhatósági szintű kétoldali konfidenciaintervallumot az első pont késleltetési idejének várható értékére.
(e) A $(b), (c), (d)$ kérdésekre kapott intervallumok közül melyik (és miért) rövidebb az (a) feladatban kapott intervallumnál?

(3) Egy algoritmust többször futtatva az alábbi futási idők adódtak (másodpercben).

52,8	64,2	63,3	66,4	70,6	69,6	58,2	58,3

(a) Adjunk 95%-os megbízhatósági szintű kétoldali konfidenciaintervallumot a futási idő várható értékére.
(b) Mennyivel lesz hosszabb a konfidenciaintervallum, ha 98%-os megbízhatósági szintet választunk?
(c) Tegyük fel, hogy a futási idő szórása 5. Adjunk most is 95%-os, illetve 98%-os megbízhatósági szintű konfidenciaintervallumot a futási idő várható értékére.
(d) Készítsünk boxplotot a mintából.
(e) Mennyi a tapasztalati eloszlásfüggvény értéke a 65 helyen?

---

5. feladatsor, 2017. márc. 13.

(1) Egy osztályba 16 fiú és 20 lány jár. Tegyük fel, hogy minden tanítási napon egymástól függetlenül a fiúk $0,04$, a lányok $0,05$ valószínűséggel hiányoznak. Legyen X a jövő hétfőn hiányzó fiúk, Y pedig a jövő hétfőn hiányzó lányok száma.
(a) Számítsuk ki az összes jövő hétfői hiányzó, vagyis $X+Y$ várható értékét.
(b) Számítsuk ki X, Y és $X+Y$ szórását.
(c) Mennyi X és Y kovarianciája?
(d) Mennyi X és $X+Y$ kovarianciája?
(e) Mennyi X és $X+Y$ korrelációs együtthatója?

(2) Legyenek X és Y független Poisson-eloszlású valószínűségi változók, X várható értéke 4, Y várható értéke pedig 10.
(a) Mennyi $X+Y$ várható értéke?
(b) Mennyi $X+Y$ szórása?
(c) Mennyi X és $X+Y$ korrelációs együtthatója?
(d) Mennyi $2X+3Y$ és $X-Y$ kovarianciája?
(e) Mennyi $2X-Y$ szórása?

(3) Tegyük fel, hogy egy ember (szisztolés) vérnyomása minden mérésnél 120 Hgmm várható értékű, 10 szórású normális eloszlású valószínűségi változó. Legyen X és Y két vérnyomásmérés eredménye, és tegyük fel, hogy elég sok idő eltelt a két mérés között ahhoz, hogy feltehessük, hogy a mérési eredmények egymástól függetlenek.
(a) Mennyi $X+Y$ várható értéke és szórása?
(b) Mennyi a két mérés átlagának várható értéke és szórása?
(c) Mennyi lenne a mérések átlagának várható értéke és szórása $n = 10$, illetve $n = 100$ mérés esetén?

---

4. feladatsor, 2017. márc. 6.

(1) Tegyük fel, hogy az egyetemisták IQ-teszten elért eredménye normális eloszlású 110 várható értékkel és 10 szórással. Mennyi a valószínűsége, hogy egy véletlenszerűen kiválasztott ember
(a) 140-nél kevesebb pontot ér el?
(b) 120-nál több pontot ér el?
(c) 80-nál kevesebb vagy 140-nél több pontot ér el?

(2) Tegyük fel, hogy egy szerver válaszideje 50 ms várható értékű exponenciális eloszlású valószínűségi változó.
(a) Mennyi a válaszidő szórása?
(b) Mennyi a valószínűsége, hogy a válaszidő legfeljebb 30 ms?
(c) Mennyi a valószínűsége, hogy a válaszidő több 100 ms-nál?
(d) Mennyi a valószínűsége, hogy a válaszidő 40 és 60 ms közé esik?

(3) A postás egy, a $[2, 4]$ intervallumból véletlenszerűen választott időpontban hoz levelet.
(a) Mennyi a postás érkezési időpontjának várható értéke?
(b) Mennyi az érkezési időpontjának szórása?
(c) Mennyi a valószínűsége, hogy 3 óra 45 perc előtt megérkezik?
(d) Feltéve, hogy 2 óra 30 percig nem érkezett meg, mennyi a valószínűsége, hogy 3 óra 45 perc előtt megérkezik?

(4) Mennyi garanciát adjunk, ha azt szeretnénk, hogy eladott termékeink legfeljebb 10%-át kelljen garanciaidőn belül javítani, ha azt feltételezzük, hogy a készülék élettartama
(a) egyenletes eloszlású a $[6, 14]$ intervallumon;
(b) 10 év várható értékű és 2 szórású normális eloszlású;
(c) 10 év várható értékű exponenciális eloszlású?

(5) Tegyük fel, hogy egy véletlenszerűen választott férfi testmagassága 176 cm várható értékű és 8 szórású normális eloszlású valószínűségi változó.
(a) Mennyi a valószínűsége, hogy egy véletlenszerűen választott férfi magassága 168 és 192 cm közé esik?
(b) Mennyi a valószínűsége, hogy két véletlenszerűen választott férfi magasságának az átlaga 168 és 192 cm közé esik?

Felhasználhatjuk az alábbiakat: $\Phi(1) = 0,84$; $\Phi(1,28) = 0,9$; $\Phi(1,64) = 0,95$; $\Phi(2) = 0,98$; $\Phi(2,33) = 0,99$; $\Phi(3) = 0,997$.

3. feladatsor, 2017. febr. 27.

(1) Egy rendszerben 10 szerver működik. Egy adott napon mindegyik $0,01$ valószínűséggel romlik el, egymástól függetlenül. Mennyi
(a) annak valószínűsége, hogy holnap pontosan 2 szerver romlik el?
(b) a holnap elromló szerverek számának várható értéke?
(c) a holnap elromló szerverek számának szórása?

(2) Sokéves megfigyelések szerint egy évben átlagosan 3,42 alkalommal van jégeső. Feltételezzük, hogy a jégesők éves száma Poisson-eloszlású, és hogy a jégesők várható száma megegyezik a megfigyelt átlaggal.
(a) Mennyi az egy év alatt bekövetkezett jégesők számának szórása?
(b) Mennyi annak valószínűsége, hogy egy évben pontosan háromszor van jégeső?
(c) Mennyi annak valószínűsége, hogy egy évben legalább négyszer van jégeső?

(3) Gábor holnaptól motorral jár dolgozni. Minden nap a többitől függetlenül $0,002$ valószínűséggel éri baleset útközben.
(a) Mennyi a valószínűsége, hogy Gábort (holnaptól számítva) a 10. munkanapján éri először motorbaleset?
(b) Jelölje X, hogy holnaptól hányadik munkanapon éri először baleset Gábort. Milyen eloszlású X?
(c) Mennyi X várható értéke?
(d) Mennyi X szórása?

(4) Öt dobókockával dobunk egyszerre. Jelölje X azt, hogy hány hatost dobtunk.
(a) Mennyi $\mathbb{P}(X=5), \mathbb{P}(X=4), \mathbb{P}(X=3)$?
(b) Milyen eloszlású X?
(c) Mennyi a hatosok számának várható értéke?
(d) Mennyi a hatosok számának szórása?

(5) Bálint minden nap a többitől függetlenül 0,01 valószínűséggel késik el az egyetemről.
(a) Mennyi annak valószínűsége, hogy egy hét (öt hétköznap) alatt pontosan kétszer késik?
(b) Milyen eloszlású a novemberi késéseinek száma, ha novemberben 21 tanítási nap van?
(c) Mennyi a novemberi késéseinek számának várható értéke?
(d) Mennyi a novemberi késéseinek számának szórása?

(6) Egy boltban az egy óra alatt bejövő vevők száma 10 paraméterű Poisson-eloszlású. Mennyi a valószínűsége, hogy reggel 8 és 9 között legfeljebb ketten jönnek? Várhatóan hányan jönnek be reggel 8 és 9 között? Mennyi az ezalatt érkezők számának szórása?

---

2. feladatsor, 2017. febr. 20.

(1) Két szabályos dobókockával dobunk, egy pirossal és egy kékkel. Tekintsük a következő eseményeket. A: az első dobás ötös. B: a második dobás hármas. C: a dobott számok összege 6. D: a dobott számok összege 7. Független-e A és B? Független-e A és C? Független-e A és D? Független-e A, B és C?

(2) Egy betegségben a lakosság 2%-a szenved. A betegség kimutatására szolgáló teszt beteg embereknél 95% valószínűséggel mutatja ki a betegséget, ugyanakkor az egészséges embereknél 1% valószínűséggel tévesen betegséget jelez.

(a) Egy véletlenszerűen választott embernél elvégezve a vizsgálatot, mennyi a valószínűsége, hogy a teszt betegséget jelez? *(b)* Tamásnál elvégezték a tesztet, az eredmény szerint beteg. Mennyi a feltételes valószínűsége, hogy valóban beteg? *(c)* Megismételték a vizsgálatot, az előzőtől függetlenül. Az újabb tesztnél ismét betegséget jelzett a teszt. A két eredmény alapján mennyi a feltételes valószínűsége, hogy Tamás beteg?

(3) Egy rendszerbe beérkező igényeket három szerver valamelyike dolgozza fel. Minden új igény 40% valószínűséggel az A, 30% valószínűséggel a B, szintén 30% valószínűséggel a C szerverhez kerül. Az A szerver 4%, a B szerver 1%, a C szerver 5% valószínűséggel ad hibaüzenetet. *(a)* Mennyi annak valószínűsége, hogy egy beküldött igényre hibaüzenetet kapunk válaszul? *(b)* Egy beküldött igényre hibaüzenetet kaptunk vissza. Mennyi a valószínűsége, hogy az igényünket a B szerver dolgozta fel?

(4) Egy dobozban egy jó és egy rossz kábel van. A jó 95% valószínűséggel, a rossz 30% valószínűséggel működik minden kipróbálásnál függetlenül. Találomra kivesszük valamelyik kábelt (mindkettőt azonos valószínűséggel választva). Tízszer kipróbáltuk, ebből nyolcszor működött, kétszer nem. Mennyi a valószínűsége, hogy a jó kábelt vettük ki a dobozból?

(5) Tegyük fel, hogy a férfiak 5%, a nők 1% valószínűséggel szenvednek színvakságban, és hogy a férfiak aránya a teljes népességben 50%.

(a) Mennyi a valószínűsége, hogy egy véletlenszerűen kiválasztott ember színvak?

(b) Gábor elmondja, hogy a testvére színvakságban szenved. Mennyi a feltételes valószínűsége, hogy Gábor testvére nő?

(6) Két szabályos dobókockával dobunk. Tekintsük az alábbi eseményeket. A: az első dobás hatos. B: legalább az egyik dobás hatos. C: mindkét dobás hatos.
(a) Mennyi B valószínűsége?
(b) Mennyi a $\mathbb{P}(A|B)$ feltételes valószínűség?
(c) Mennyi a $\mathbb{P}(C|A)$ feltételes valószínűség?

---

1. feladatsor, 2017. febr. 13.

(1) Vegyünk egy 100 bit hosszú véletlen bitsorozatot (minden lehetséges sorozatot egyformán valószínűnek tekintve). Mennyi a valószínűsége, hogy pontosan 43 darab 1-es van benne? Mennyi a valószínűsége, hogy több 1-es van, mint 0?

(2) 100-szor dobunk egy kockával. Melyik a legvalószínűbb az alábbiak közül: pontosan 10-szer dobunk hatost, pontosan 14-szer dobunk hatost, pontosan 18-szor dobunk hatost.

(3) A lottósorsoláson 1–90-ig számozott golyók közül húznak visszatevés nélkül 5-öt, és egy szelvényen öt számra lehet tippelni. Mennyi a valószínűsége, hogy pontosan 4 találatunk lesz? Mennyi az öttalálatos valószínűsége? Mennyi a valószínűsége, hogy legfeljebb 2 számot sikerül eltalálni?

(4) Egy 32 lapos kártyapakliból, melyben nyolc piros lap van, egy játékhoz kiosztunk Annának öt, véletlenszerűen választott lapot (minden ötös csoport egyformán valószínű). Mennyi annak valószínűsége, hogy Anna kap legalább egy piros lapot? Mennyi annak valószínűsége, hogy pontosan két piros lapot kap?

(5) Péternek nyolc szürke, három fekete és két fehér pólója van a szekrényben. Kivesz találomra három darabot a hétvégi utazáshoz, minden pólót azonos valószínűséggel választva. Mennyi a valószínűsége, hogy pontosan két fekete pólót visz magával? Mennyi a valószínűsége, hogy csak szürke pólót visz magával?

(6) Három szabályos dobókockával dobunk. Mennyi a valószínűsége, hogy a dobott számok összege legfeljebb 5?

(7) Öt szabályos dobókockával dobunk. Mennyi a valószínűsége a "full" dobásnak, azaz hogy lesz három egyforma szám, és a maradék kettő is egyforma, de nem mind az öt dobott szám azonos? Mennyi a valószínűsége, hogy mind az öt szám egyforma?