

# Leíró és matematikai statisztika gyakorlat

Matematikai elemző szakirány

2016/2017 tavaszi félév

## Játékszabályok

- A gyakorlatokról maximum 3-szor lehet hiányozni. Aki többször hiányzik, nem kap gyakjegyet.
- $100 + x$  pontot lehet szerezni a félév során:
  - 30 pont: 1. ZH: III.21.
  - 30 pont: 2. ZH: V.16.
  - 40 pont: két, 20 pontos beadandó feladat
  - $x$  pont: szorgalmi feladatok
- Mindkét ZH-n minimálisan teljesíteni kell 30 %-ot.
- Ha egy ZH sikertelen, nem írod meg, vagy javítani szeretnél, akkor vizsgaidőszak első hetén lesz lehetőség a pótZH megírására vagy a javításra. Csak az egyik ZH anyagából javíthatsz, és a jobbik eredményt veszem figyelembe, azaz nem lehet rontani. Két sikertelen vagy meg nem írt ZH esetén gyakUV-t kell írnod a teljes féléves anyagból (60 pontért, a beadandókért kapott pontok megmaradnak).
- A ZH-kon használható: **számológép** és egy legfeljebb A4-es méretű lapra KÉZZEL írott "puska".

elégtelen (1)	0	–	34,99
elégséges (2)	35	–	49,99
közepes (3)	50	–	64,99
jó (4)	65	–	79,99
jeles (5)	80	–	1000
- Osztályozás:

## Infók a gyakvezetőről

Név Varga László  
Tanszék Valószínűségelméleti és Statisztika Tanszék (ELTE TTK)  
Szoba D 3-309  
E-mail vargal4@cs.elte.hu  
Honlap vargal4.elte.hu

## Ajánlott irodalom

- Molnárné-Tóthné: Általános statisztika példatár I.
- Móri-Szeidl-Zempléni: Matematikai statisztikai feladatok

1.) Legyenek  $X_i \sim N(10, 5^2)$  ( $i = 1, \dots, 9$ ) függetlenek. Számítsuk ki az alábbi mennyiségeket, majd közelítsük őket számítógépes szimuláció segítségével:

- $P(X_8 > 3)$ ,
- $P(\bar{X} < 9)$ ,
- $E\bar{X}$ .

2.) Egy termelő vállalatnál a fizikai munkát végzők összesen 18000 db alkatrészt állítottak elő, amiből a nők teljesítménye 8500 db volt. A vállalatnak 950 férfi fizikai dolgozója van. A nőknél a termelékenység, azaz az egy főre jutó termelt mennyiség 17 db/fő.

- Milyen viszonyszám található a feladat szövegében? Add meg precíz kiszámítási módját (mit osztunk mivel, mértékegységek)!
- Szerkessz statisztikai táblát a megadott adatokból és töltsd ki a hiányzó adatokat! Milyen típusú a tábla és milyen típusú sorokat tartalmaz?

3.) Néhány információ az ELTE matematika alapszakjára 2016-ban jelentkezőkről: az állami finanszírozásos képzésre 348-an jelentkeztek, 36,494%-uk első helyen jelentkezett, végül 110-et vettek fel. A költségtérítéses képzési formára jelentkezők 10,227%-át, 9 főt vették fel. Összesen 141 ember jelölt be az ELTE matematika szakát első helyen.

- Milyen viszonyszám(ok) található(k) a feladat szövegében? Add meg precíz kiszámítási módjukat (mit osztunk mivel, mértékegységek)!
- Szerkessz statisztikai táblát a megadott adatokból és töltsd ki a hiányzó adatokat! Milyen típusú a tábla és milyen típusú sorokat tartalmaz?

4.) Egy vállalat négy részleggel rendelkezik, az ott dolgozók bruttó fizetéséről az alábbi adatok állnak rendelkezésünkre:

Részleg	Átlagkereset (e Ft/fő)	Dolgozók létszáma (fő)
Raktár	200	10
Összeszerelő	250	16
Tampóműhely	250	8
Irodaház	300	10
Összesen	...	...

- Milyen viszonyszám található a táblázatban? Add meg precíz kiszámítási módját (mit osztunk mivel, mértékegységek)!
- Milyen típusú a tábla és milyen típusú sorokat tartalmaz?
- Számítsd ki a hiányzó pontozott értékeket!

5.) Magyarország népességéről az alábbiakat ismerjük:

Település jellege	Népesség megoszlása 2012-ben (%)	Népesség változása 1990-ről 2012-re (%)
Budapest	17,4	-14,4
Többi város	51,9	-2,4
Községek	30,7	-0,8
Összesen	100,0	...

- 1990 és 2012 között évente átlagosan mennyivel változott a budapesti lakosság?
- Hány százalékkal változott a népesség száma 1990-ről 2012-re?
- Melyik településen élők részaránya csökkent?

6.) Egy vállalatnál az alkalmazottak számára vonatkozóan tartalmaz január 1-jei adatokat a következő táblázat:

Év	Alkalmazottak száma	Változás		
	január 1-jén (fő)	2012-höz képest (%)	az előző évhez képest (%)	az előző évhez képest (fő)
2010	...	...	...	...
2011	95	...	...	+5
2012	100	...	...	...
2013	...	...	+20	...
2014	...	+50	...	...

- a.) Milyen típusú a tábla és milyen típusú sorokat tartalmaz?
- b.) Töltsd ki a táblázatot! Értelmezzünk a táblázatban néhány értéket!
- c.) Határozd meg az alkalmazottak átlagos számát a 2011-es évben, valamint 2010. január 1. és 2013. január 1. között!
- d.) Jellemezd az alkalmazottak számának évi átlagos változását 2010. január 1. és 2014. január 1. között!
- e.) Ábrázold az alkalmazottak számának alakulását megfelelő diagrammal!
- 7.) Az egyetem büféjében egy adott napon az összes vendég fogyasztását megvizsgálták, és ez alapján az elköltött összegekről az alábbi táblázatot készítették el:

Fogyasztás összege (Ft)	Vendégek száma (fő)
0 – 200	40
201 – 500	42
501 – 800	80
801 – 1200	22
1201 –	16
Összesen	200

- a.) Átlagosan mennyit költöttek a büfében? Készítsünk hisztogramot!
- b.) Vizsgáld az elköltött összegek koncentrációját Lorenz-görbével, koncentrációs együtthatóval és Herfindahl-indexszel!
- 8.) Van két piac, az elsőn 10 azonos méretű vállalat tevékenykedik, a másodikon pedig 5 azonos méretű vállalat van. Hasonlítsuk össze a két piac koncentrációját!
- SZ1.)** Tulajdonosa vagy egy vállalkozásnak, év végén a könyvelőd jelentést készít, amiben a következőket írja: "Remek évet zártál, a tervezetthez képest magasabb lett az árbevétel, bár a költségek csak stagnáltak. A árbevétel tervezett 10%-os növelését 10%-kal túlteljesítetted, így 20%-os növekedést hoztál össze. Az előző évhez képest 10%-os költségcsökkenést tervezted, azonban a tervezetthez képest 10%-kal magasabbak lettek a költségek, így összességében a költségek összege nem változott."
- Értékelj a könyvelőd érvelését! Van-e benne valami, ami sántít; és ha igen, miért? (1p)
- SZ2.)** Egy vállalat alkalmazottainak száma 2010-ről 2014-re évente átlagosan 5,7371 %-kal, azaz 3 fővel nőtt. A kifizetett összes éves bér 2014-ben 216 millió Ft volt, míg a 2010-es havi átlagbér 250 ezer Ft volt. A rendelkezésre álló adatokból készíts táblát és töltsd ki az üres rubrikákat! Átlagosan mennyivel változott az átlagbér? (2p)

- 9.) Legyenek  $X_1, \dots, X_{100}$  független, standard normális eloszlású valószínűségi változók. Szimulációval becsüljük meg a maximum sűrűségfüggvényét (készítsünk hisztogramot)!
- 10.) Legyen  $X \sim \text{Exp}(\lambda)$ . Határozd meg  $X$  móduszát és tetszőleges kvantilisét! Hasonlítsd össze a mediánt és a várható értéket!
- 11.) Legyen  $X \sim \text{Ind}(p)$ . Határozd meg  $X$  móduszát, kvantilisfüggvényét, ferdeségét és csúcosságát!
- 12.) Legyen  $X \sim N(m, \sigma^2)$ . Határozd meg  $X$  móduszát, mediánját, ferdeségét és csúcosságát!
- 13.) Egy csoportban a hallgatók magassága (cm):  
180 163 1500 157 165 165 174 191 172 165 1-68 186
- a.) Nézzük át nagy vonalakban az adatokat, reálisak-e! Próbáljuk javítani az esetleges adathibákat!
- b.) Rajzold fel a tapasztalati eloszlásfüggvényt! Mennyi a tapasztalati eloszlásfüggvény értéke a 180 helyen? Értelmezd szövegesen!
- c.) Elemezd a hallgatók testmagasságát alapstatisztikák: átlag, korrigált tapasztalati szórás, szórási együttható, kvartilisek, terjedelem, interkvartilis terjedelem, tapasztalati ferdeség, tapasztalati csúcosság segítségével! Értelmezd szövegesen az eredményeket!
- d.) Készíts boxplot ábrát!
- e.) Készíts alkalmas osztályközös gyakorisági sort, majd abból hisztogramot!
- 14.) A honlapomon található `Nyarhom.Rdata` nevű fájl a 2014. nyári napi maximum-hőmérsékleteket tartalmazza egy településen ( $^{\circ}\text{C}$ ). Elemezd az adatokat az előző feladat elemzési szempontjai alapján!
- 15.) 2014-ben egy bútorboltban eladott konyhabútorok értékéről az alábbi adatok ismerjünk:

Konyhabútor ára (e Ft)	Eladott bútorok száma (db)
– 100	10
101 – 200	18
201 – 300	25
301 – 500	30
501 –	17
Összesen	100

Jellemezd (szövegesen is) az értékesített konyhabútorok árának eloszlását alapstatisztikák (módusz, kvartilisek, átlag, szórás) segítségével! Milyen az eloszlás ferdesége? Készíts hisztogramot!

- SZ3.)** Vezesd le az  $n$  és  $p$  paraméterű binomiális eloszlás móduszát! (1p)
- SZ4.)** Legyen  $X_1, \dots, X_{100}$  i.i.d. minta egy  $\text{NegBin}(3; \frac{1}{4})$  eloszlásból. Becsüld meg  $\mathbf{R}$ -es szimuláció segítségével a szórási együttható várható értékét és szórását! Küldd el E-mail-ben a forráskódot! (1p)
- SZ5.)** Vezesd le az  $\text{Exp}(\lambda)$  eloszlás ferdeségét és csúcosságát! (2p)

16.) Tegyük fel, hogy a minta kétparaméteres eloszláscsaládból származik, a paraméterek  $a$  és  $b$ .

Ekkor mutassuk meg, hogy az  $\begin{cases} E_{a,b}X = m_1 \\ E_{a,b}X^2 = m_2 \end{cases}$  egyenletrendszer megoldása

megegyezik az  $\begin{cases} E_{a,b}X = m_1 \\ D_{a,b}^2X = s_n^2 \end{cases}$  egyenletrendszer megoldásával.

17.) Határozzuk meg az ismeretlen paraméter(ek) maximum likelihood és momentum becslését, ha a minta i.i.d.

- Poi( $\lambda$ ) eloszlású;
- Exp( $\lambda$ ) eloszlású;
- $E(a; b)$  eloszlású, ahol  $a < b$ , mindkettő paraméter;
- $E(-a; a)$  eloszlású;

e.)  $f_\vartheta(x) = \frac{2x}{\vartheta^2} I(0 \leq x \leq \vartheta)$  sűrűségfüggvénnyel,  $\vartheta$  paraméter;

f.)  $f_\vartheta(x) = \vartheta x^{\vartheta-1} I(0 \leq x \leq 1)$  sűrűségfüggvénnyel,  $\vartheta$  paraméter.

18.) Legyenek  $X_1, \dots, X_n$  és  $Y_1, \dots, Y_m$  egymástól független  $\lambda$  illetve  $\frac{1}{\lambda}$  paraméterű exponenciális eloszlású minták. Határozzuk meg az ismeretlen paraméter (együttes) ML becslését!

19.) Legyen  $X_1, \dots, X_n$  i.i.d. minta Poi( $\lambda = 2$ ) eloszlásból. Vizsgáljuk meg szimulációval  $n = 10, 40, 100$  és  $500$  nagyságú minták esetén a  $\lambda$  paraméter ML-becslésének eloszlását (készítsünk hisztogramokat)! Mire következtethetünk az ábrák alapján? Aszimptotikusan torzítatlan a becslés? Szimulációk segítségével adjunk 95%-os megbízhatóságú konfidenciaintervallumot az ML-becslésre!

20.) Legyen a  $Z_1, \dots, Z_5$  minta **I.)**  $N(m, 2^2)$  **II.)**  $N(2m + 5, 2^2)$  eloszlású. A megfigyelt értékek a következők: 6; 4,5; 2,5; 2; 1.

- Határozzunk meg 95%-os (99%-os) megbízhatóságú konfidenciaintervallumot  $m$ -re!
- Hány elemű mintára van szükségünk, ha azt szeretnénk, hogy a konfidenciaintervallum legfeljebb 0,01 hosszúságú legyen?
- Mi változik az a.) esetben, ha a szórást nem ismerjük?
- Adjunk a szórásra 98%-os megbízhatóságú konfidenciaintervallumot.

$$\sum_{i=1}^5 z_i = 16 \quad \sum_{i=1}^5 (z_i - \bar{z})^2 = 16,3$$

$$\chi_{4;0,01}^2 = 0,3 \quad \chi_{4;0,99}^2 = 13,28$$

21.) Az előző évben figyelemmel kísértük a sárkányföldi tőzsdeindex, a SÜSÜX változását. Az alapstatisztikák: átlag: 1113,18; szórás: 95,31. A tőzsde 200 napon keresztül volt nyitva. Adjunk ezek alapján 95%-os megbízhatóságú konfidenciaintervallumot az index adott évre vonatkozó várható értékére!

SZ6.) Öt véletlen számot jegyeztünk fel: 100,32,76,52,17. Ha tudjuk, hogy ezek az  $\{1, 2, \dots, N\}$  halmazból vett véletlen minta elemei, akkor határozzuk meg az ismeretlen  $N$  paraméter ML-becslését és momentum becslését! (1p)

SZ7.) Egy CASCO biztosítás kárai 2013-ban 200, 1200, 1800, 125, 485 ezer Ft

voltak. A káreloszlásról feltételezzük, hogy  $(\alpha, \beta)$  paraméterű Pareto-eloszlású, azaz az eloszlásfüggvény

$$F(x) = \begin{cases} 1 - \left(\frac{\beta}{\beta+x}\right)^\alpha & \text{ha } x > 0 \\ 0 & \text{különben} \end{cases}$$

Számítsd ki a Pareto-eloszlás várható értékét és szórásnégyzetét, majd határozd meg az ismeretlen paraméterek momentum módszeres becslését a minta alapján! (2p)

SZ8.) Tegyük fel, hogy az  $n$  elemű mintánk lognormális eloszlású, azaz a mintaelemek logaritmus  $N(m, \sigma^2)$  eloszlású. Határozd meg az ismeretlen paraméterek maximum likelihood és momentum becslését! Segítség: használható a Wikipédiáról a sűrűségfüggvény képletét és a kiszámított várható értéket/szórásnégyzetet. (2p)

22.) Legyen  $X_1, \dots, X_n$  i.i.d. minta  $N(m, 1^2)$  eloszlásból. Célunk az ismeretlen  $m$  paraméter becslése. Tekintsük az alábbi statisztikákat:

$$T_1(\mathbf{X}) = X_8 \quad T_2(\mathbf{X}) = \frac{X_3 + X_7}{2} \quad T_3(\mathbf{X}) = \frac{X_9 + X_{19}}{8} \quad T_4(\mathbf{X}) = \bar{X}$$

- A fenti statisztikák közül melyek torzítatlanok, illetve konzisztensek? Amelyik nem torzítatlan, hogyan tudnánk torzítatlanná tenni?
- Vizsgáljuk meg a fenti statisztikák közül a torzítatlanokat hatásosság szempontjából!

23.)  $X_1, \dots, X_n \sim \text{Exp}(\lambda)$  i.i.d. minta esetén adj torzítatlan becslést  $e^{-3\lambda}$ -ra és  $\frac{1}{\lambda}$ -ra!

24.) Legyen  $X_1, \dots, X_n$  Pascal-eloszlású (geometriai) minta  $p$  paraméterrel.

- Adjunk meg  $X_3$  függvényeként torzítatlan becslést  $p(1-p)^4$ -re!
- Adjunk maximum likelihood becslést  $p(1-p)$ -re!

25.)  $X_1, \dots, X_n \sim \text{Poi}(\lambda)$  i.i.d. minta esetén adj torzítatlan becslést  $e^{-\lambda}$ -ra és  $\lambda^2$ -re!

26.) Legyen  $X_1, \dots, X_n$  független, azonos abszolút folytonos eloszlású valószínűségi változók sorozata. Adjuk meg  $X_1^* = \min(X_1, \dots, X_n)$  és  $X_n^* = \max(X_1, \dots, X_n)$  eloszlás- és sűrűségfüggvényét! A minimumnál külön is vizsgáljuk meg azt az esetet, ha az  $X_i$  változók exponenciális eloszlásúak!

27.) Adj torzítatlan becslést az i.i.d.  $E[0; \vartheta]$  eloszlású minta  $\vartheta$  paraméterére a

- mintaátlag
  - maximum
- segítségével. Melyik a hatásosabb? Melyik konzisztens?

28.) Mutassuk meg, hogy exponenciális eloszlású minta esetén  $T(\mathbf{X}) = n \cdot X_1^*$  statisztika torzítatlan, de nem konzisztens becslése a várható értéknek!

29.) Február 17-én Budapesten az elmúlt 10 évben az alábbi középhőmérsékleteket mérték: 2; 2,5; 1,6; -4,5; 5,3; 7,9; 1,5; -1,6; -2,2; 1,6.

- Számítsuk ki és ábrázoljuk a középhőmérséklet sűrűségfüggvényének Parzen-

Rosenblatt becslését, ha  $h = 0,25$  és a magfüggvény  $k(x) = \begin{cases} \frac{1}{2} & \text{ha } |x| < 1 \\ 0 & \text{különben} \end{cases}$

b.) Készítsük el a Parzen-Rosenblatt-féle sűrűségfüggvénybecslést Gauss-magfüggvény esetén különböző sávszélességekre (R segítségével)!

**SZ9.)** Adjunk torzítatlan becslést a  $[0, \vartheta]$  intervallumon egyenletes eloszlás paraméterére  $n$  elemű i.i.d. mintából a minimum segítségével. Hatásosabb a becslés, mint a 27. feladat a.) részében kapott torzítatlan becslés? Konzisztens a becslés? (2p)

**SZ10.)** Legyen  $X_1, \dots, X_n$  i.i.d. minta az  $E(a, b)$  eloszlásból,  $a$  és  $b$  paraméterek. Mutassuk meg, hogy  $b$  ML-becslése nem torzítatlan, de aszimptotikusan torzítatlan becslése  $b$ -nek! (2p)

**SZ11.)** Legyen  $X_1, \dots, X_n$  i.i.d. minta  $\text{Bin}(k, p)$ -ből,  $Y_1, \dots, Y_n$  i.i.d. minta  $\text{Bin}(l, p)$ -ből, és tegyük fel, hogy a két minta egymástól is független. Milyen  $(a, b)$  számpárokra lesz  $a\bar{X} + b\bar{Y}$  a  $p$  paraméter torzítatlan becslése? Ezen számpárok közül melyikre lesz a becslés szórása minimális? (2p)

**30.)** Valaki azt állítja, hogy a klíma változik, és ezt azzal véli bizonyítottattnak, hogy az elmúlt 10 évben 2-szer is volt jégeső, pedig korábban az egyes évekre a jégeső valószínűsége a hivatalos adatok alapján csupán  $p=0.1$  volt. Írjuk fel a hipotéziseket, a próbát és állapítsuk meg az elsőfajú hiba valószínűségét, valamint az erőfüggvényt a  $p=0.2$  pontban!

**31.)** Az  $X$  valószínűségi változó egyenletes eloszlású a  $(-b; 1+2b)$  intervallumon. A  $H_0: b=0$  hipotézist szeretnénk ellenőrizni a  $H_1: b>0$  hipotézis ellenében, e célból a következő próbát alkalmazzuk: egy megfigyelést végzünk és ha ez a  $(0, 1; 0, 85)$  intervallumba esik, elfogadjuk  $H_0$ -t, különben elvetjük.

Írjuk fel a próba erőfüggvényét! Mekkora a próba terjedelme?

**32.)** Legyen  $X_1, \dots, X_n$  minta az  $f(x) = \begin{cases} \frac{2x}{a^2} & \text{ha } 0 \leq x \leq a \\ 0 & \text{különben} \end{cases}$  sűrűségfüggvényű eloszlásból.

a.) Tekintsük a következő hipotéziseket:  $H_0: a=1$   $H_1: 0 < a < 1$   
Adjunk meg  $X_n^*$  függvényében 5 %-os terjedelmű próbát, keressük a kritikus tartományt  $\mathcal{X}_k = \{X : T(X) < c\}$  alakban! Mi lesz az erőfüggvény?

b.) Tekintsük a következő hipotéziseket:  $H_0: a=1$   $H_1: a = \frac{1}{2}$   
Adjunk meg  $\alpha$  terjedelemhez egyenletesen legerősebb próbát!

**33.)** Legyen két megfigyelésünk a  $(3; p)$  paraméterű binomiális eloszlásból. Adjuk meg a legjobb olyan próbát az alábbi hipotézisekre, melynek elsőfajú hiba valószínűsége 0,04:  $H_0: p = \frac{1}{2}$   $H_1: p = \frac{1}{4}$

**34.)** Az alábbi minta 4 év október 18-án Budapesten mért napi középhőmérséklet adatait tartalmazza. Ellenőrizzük a  $H_0: m = 15$  hipotézist  $\alpha = 0.05$  elsőfajú hiba-  
valószínűség mellett *értelmes* alternatív hipotézissel szemben.

Középhőm. (C fok) adatok:

14,8	12,2	16,8	11,1
------	------	------	------

a.) A korábbi tapasztalatok alapján tekintsük az értékek szórását 2-nek. Adjuk meg a  $p$ -értéket is.

b.) Ne használjunk a szórásra vonatkozóan előzetes információt.

**35.)** Tegyük fel, hogy az emberi magasság normális eloszlású.

a.) Végezzünk statisztikai próbát arra vonatkozóan, hogy a gyakorlaton lévő lányok átlagmagassága 170 cm!

b.) Végezzünk statisztikai próbát arra vonatkozóan, hogy a gyakorlaton lévő fiúk átlagmagassága 180 cm!

**36.)** A Természettudományi Kar II. évfolyamán az egyik gyakorlati csoportban 10-en írtak statisztika zárthelyit. Két feladatsor volt, mindkettőben 30 pontot lehetett elérni. Tegyük fel, hogy az elért pontszámok normális eloszlásúak. A pontszámokat tartalmazza az alábbi táblázat:

1. feladatsor	12	11	8	14	10
2. feladatsor	15	14	9	16	11

a.) Vajon az első feladatsor nehezebb volt?

b.) Mennyiben változik a helyzet, ha nem 10 diákról, hanem csak 5-ről van szó, és a 2. feladatsor a pótZH eredménye?

**37.)** Az alábbi két minta 10 egyforma képességűnek feltételezett sportoló súlylökésben elért eredményeit tartalmazza. A sportolók két ötfős csoportban készültek az edzőtáborban. Edzéstervük ugyanaz volt, de az első csoportban készülők minden reggel fejenként 10 tojást és 25 túró rudit ettek meg. A második csoportban készülőknek reggel és este 1-1 kg szalonnát és 1-1 kg madártejet kellett megenni. 2 hét felkészülés után értékelték az eredményeket. Tételezzük fel, hogy normális eloszlásból származnak a minták és a terjedelem 5%.

1. csoport	15,8	15,2	16,3	17,1	16,1
2. csoport	19,0	12,1	17,2	14,7	21,0

a.) Melyik diéta volt jobb, ha a dobások szórását 2-nek tekintjük?

b.) Állíthatjuk-e, hogy a második csoportban nagyobb változékonyságot mutat a sportolók teljesítménye?

c.) Ha nem ismerjük a szórást, akkor tekinthetjük-e valamelyik diétát jobbnak?

$$F_{4,4}^{0,95} = 4,4 \quad F_{5,5}^{0,95} = 5,05 \quad F_{4,4}^{0,975} = 9,6 \quad F_{5,5}^{0,975} = 7,15$$

**38.)** Tegyük fel, hogy az emberi magasság normális eloszlású. Végezzünk alkalmas statisztikai próbát arra vonatkozóan, hogy a gyakorlaton lévő lányok átlagmagassága megegyezik a fiúk átlagmagasságával!

**SZ12.)** Oldd meg a 32.) feladatot abban az esetben, ha baloldali kritikus tartományt választunk:  $\mathcal{X}_k = \{X : T(X) > c\}$ . (1p)

**SZ13.)** Egy érme szabályosságát (a  $H_1: p > 0,5$  ellenhipotézissel szemben;  $p$  a fejdobás valószínűsége) az alábbi módszerrel teszteljük:  $n$ -szer feldobjuk az érmét, és ha legalább 2 írást dobtunk, akkor elfogadjuk  $H_0$ -t.

a.) Mekkora legyen  $n$ , hogy az elsőfajú hiba kisebb legyen, mint 0,05?

b.) Adjuk meg a próba erőfüggvényét! ( $1+1=2$ p)

**SZ14.)** A Politikatudományi Kar HÖK elnöke nagyon fontosnak tartja népszerűségét. Amennyiben a hallgatók legfeljebb 70%-a utálja, az számára elfogadható ( $H_0$  hipotézis). Az ennél nagyobb arány esetén ( $H_1$  hipotézis) lemond. Minden

negyedév végén 10 hallgatót kérdez meg (közvélemény-kutatást tart). Az elnök akkor mond le, ha a tízből legalább 8 diák utálja.

- a.) Mekkora a próba terjedelme?  
 b.) Várhatóan hány negyedévet fog tevékenykedni az elnök, ha stabilan a diákok 65%-a utálja? (2+2= 4p)

**SZ15.)** Legyen  $X_1$  minta az  $f(x)$  sűrűségfüggvényű eloszlásból. Tekintsük a következő hipotéziseket:

$$H_0: f(x) = f_0(x) = 2(1-x) \cdot I(0 < x < 1)$$

$$H_1: f(x) = f_1(x) = 2x \cdot I(0 < x < 1)$$

Adjunk meg  $\alpha$  terjedelemhez egyenletesen legerősebb próbát! (1p)

**39.)** Az alábbi táblázatban adatok találhatók azon személyek számáról, akik lórúgás következtében haltak meg 10 porosz hadtestben 20 év alatt (1875–1894) (összesen  $10 \cdot 20 = 200$  adat):

halálesetek száma	0	1	2	3	4
gyakoriság	109	65	22	3	1

Ellenőrizzük azt a hipotézist, hogy a halálesetek száma egy hadtestben egy év alatt Poisson-eloszlású!

**40.)** Rendelkezésünkre áll a következő minta: 0,55; 0,59; 0,34; 0,69; 0,95; 0,34; 0,53; 0,54; 0,03; 0,11; 0,15; 0,67; 0,48; 0,09; 0,55; 0,02; 0,37; 0,76; 0,83; 0,92.

A megoldás során alkalmazzunk diszkretizálást, azaz képezzünk alkalmas gyakorisági sort az adatokból.

- a.) Elfogadhatjuk-e azt a hipotézist, hogy a minta (0,2) intervallumon egyenletes eloszlású? Vizsgáljuk meg Q-Q plot-tal is!  
 b.) Elfogadhatjuk-e azt a hipotézist, hogy a minta egyenletes eloszlású? Vizsgáljuk meg Q-Q plot-tal is!  
 c.) Elfogadhatjuk-e azt a hipotézist, hogy a minta exponenciális eloszlású? Vizsgáljuk meg Q-Q plot-tal is!

**41.)** Az Informatikai Kar III. évfolyamán 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	27	3

- a.) Elfogadhatjuk-e azt a hipotézist, hogy egy hallgató bukásszáma  $\text{Bin}(4; 0,25)$  eloszlású?  
 b.) és azt, hogy  $\text{Bin}(4;p)$  eloszlású?

**42.)** Az alábbi táblázat CASCO biztosítással rendelkezők éves kárszámát tartalmazza 2012-ben és 2013-ban:

Kárszám	0	1	2	3	4	5	>5
Vezetők száma	3692	232	65	7	3	1	0
Vezetők száma	3542	284	135	24	9	5	1

- a.) Vajon tekinthető-e a 2012-es kárszám Poisson-eloszlásúnak?

b.) Vajon tekinthető-e a kárszám azonos eloszlásúnak a két évben?

**43.)** Az alábbi kontingencia-táblázat mutatja, hogy 100 évben a csapadék mennyisége és az átlaghőmérséklet hogyan alakult.

Hőmérséklet \ Csapadék	Kevés	Átlagos	Sok
Hűvös	15	10	5
Átlagos	10	10	20
Meleg	5	20	5

(A cellákban az egyes esetek gyakoriságai találhatóak.) Tekinthető-e a csapadék-mennyiség és a hőmérséklet függetlennek?

**SZ16.)** 100 napon keresztül feljegyezték egy város energiafogyasztását. Az alábbi táblázat azt tartalmazza, hogy az egyes intervallumokba hány megfigyelés esett, valamint azt is, hogy az adott intervallumba eső értékeknek mennyi az átlaga. Az energiafogyasztást normális eloszlásúnak tekinthetjük?

Intervallumok	< 5000	5000 – 6000	6000 – 7000	> 7000
Gyakoriságok	20	31	28	21
Átlagok	3875	5700	6500	7800

(2p)

**44.)** Legyenek adottak a következő (x,y) párok:

$x_i$	0	1	6	5	3
$y_i$	4	3	0	1	2

- a.) Határozzuk meg és ábrázoljuk is az  $aX + b$  alakú regressziós egyenest.  
 b.) Számoljuk ki a reziduálisokat és becsüljük meg a hiba-szórásnégyzetet.  
 c.) Adjunk előrejelzést  $x=10$ -re a regressziós egyenes alapján.

**45.)** A január 8-án tartott Statisztika II. vizsgát 8 hallgató írta meg, akiktől megkérdeztük, mennyi órát készültek a vizsgára, hány pontot szereztek a tantárgy előfeltételének számító Statisztika I. tantárgyból a vizsgán és milyen magasok:

Statisztika II. pontszám	49	55	56	62	65	70	78	92
Hány órát készült (ó)	15	16	14	13	12	19	21	24
Statisztika I. pontszám	60	50	66	53	67	76	88	87
Testmagasság (cm)	160	174	178	182	173	168	191	167

- a.) Vizsgáljuk meg lineáris regresszióval a tanulási idő hatását a Statisztika II. pontszámra! Ábrázoljuk a regressziós egyenest!  
 b.) Illesszünk négyzetes regressziós függvényt a Statisztika II. pontszámra, ha a magyarázó változó a tanulási idő! Ábrázoljuk a regressziós egyenest!  
 c.) Illesszünk lineáris regressziót a Statisztika II. pontszámra, ha a magyarázó változók a tanulási idő és a Statisztika I. pontszám!  
 d.) Illesszünk lineáris regressziót a Statisztika II. pontszámra, ha a magyarázó változók a tanulási idő, a Statisztika I. pontszám és a testmagasság!  
 e.) Vessük össze a modelleket!  
 f.) A Statisztika II. vizsga sikeres, ha a hallgató legalább 50 pontot elér. Júli 175 cm magas, a Statisztika I.-ből 60 pontot szerzett és a Statisztika II.-re 18 órát tervez tanulni. Várhatóan át fog menni a Statisztika II. vizsgán?

46.) Egy kereskedelmi egység három fajta paprikás chips-et árul, a következő táblázat a 2014/2015-ös értékesítésről tartalmaz adatokat:

Márka	2014.		2015.	
	Értékesített mennyiség (db)	Egységár (Ft/db)	Értékesített mennyiség (db)	Egységár (Ft/db)
Chio	25	300	30	310
Lays	20	250	30	240
Cheetos	10	200	10	220
Összesen	...	...	...	...

a.) Jellemezd az értékesítésben bekövetkezett mennyiségi, ár- és értékváltozásokat egyedi és összetett indexekkel! Értelmezd szövegesen az egyes indexeket!

b.) Számítsd ki az árváltozás miatti többletbevételt!

47.) Egy vállalat termelési értékének (árbevételének) a 35,4%-át 2014-ben az I. számú üzem, a többit pedig a II. számú üzem adta. Az I. számú üzem termékeinek egységára 2014-ről 2015-re átlagosan 5%-kal, a II. számú üzemé pedig átlagosan 3%-kal csökkent.

Számítsuk ki a vállalati termelés volumenének változását, ha ismert, hogy a vállalati termelési érték 3%-kal emelkedett! Értelmezd szövegesen a kapott volumenváltozást!

48.) Egy kisvállalkozás kétféle terméket gyárt. Termelési adatai:

Termék fajtája	Termelési érték 2015-ben (M Ft)		Volumenváltozás (2014=100%)
	folyó áron	2014-es áron	
A	50	60	110
B	60	80	120
Összesen	...	...	...

a.) Határozd meg a termelés értékindexét!

b.) Határozd meg mindkét súlyozással az ár- és volumenindexeket!

c.) Számítsd ki az volumenváltozás miatti többletbevételt 2014-es árakon!

49.) Mari néni kávézójában 3 féle kávé szolgál fel, a családi könyvelésből az alábbi adatok ismertek:

Kávéfajta	A forgalom értéke 2015-ben (e Ft)	Az árak A forgalom értékének alakulása, 2015/2010 (%)	
Cappuccino	2000	130	200
Cafe Latte	1500	120	180
Espresso	1000	120	150

a.) Számítsd ki az ár-, és volumenindexet a kávézó forgalmára vonatkozóan!

b.) A forgalom értékének növekedéséből hány forint volt az ár- és a volumenváltozás hatása?

SZ17.) János gazda kétféle almát termel: jonatánt és starking-ot. A két almafajtából származó bevétel egyaránt 15-15%-kal növekedett 2015-ről 2016-ra. 2016-ban az összes árbevétel 60%-a a jonatán almák eladásából származott. A jonatán al-

mák ára 2015-ről 2016-ra 5%-kal, a starkingoké pedig 8%-kal nőtt. Számítsd ki a Fisher-féle volumenindexet! (1p)

SZ18.) Egy boltban háromféle cigarettát árúsítanak. A cigaretták összes forgalma (árbevétele) 2013-ról 2015-re 20%-kal emelkedett. A cigarettákra vonatkozóan az alábbi adatokat ismerjük:

Márka	Árbevétel megoszlása (%)		Árak alakulása 2014/2013 (%)	Árak alakulása 2015/2014 (%)	Árbevétel alakulása 2015/2014 (%)
	2013.	2014.			
Marlboro	60	50	115	110	105
Helikon	20	30	107	102	110
Sopianae	20	20	125	95	120

Számítsd ki a Fisher-féle volumenindexet, ha a bázisidőszak 2013, a tárgyidőszak pedig 2015! (2p)