Statisztika zárthelyi dolgozat

1. Egy elmélet szerint egy bizonyos kémiai reakció 180 C°-on következik be. 15 független megfigyelés eredménye

$183,2\ 179,4\ 175,4\ 179,0\ 180,6\ 178,1\ 178,1\ 175,4\ 179,3\ 176,9\ 178,3\ 179,1\ 180,2\ 174,2\ 179,4.$

Milyen következtetést vonhatunk le az elméletre az adatokból?

Megoldás: Ha feltételezzük, hogy a mintánk normális eloszlású, akkor egymintás *t*-próbával dönthetünk arról a nullhipotézisről, hogy a minta várhatóértéke $m_0 = 180.$

Az átlag: $\bar{x}_{15} = \frac{1}{15}\cdot(183.2+179.4+175.4+179.0+180.6+178.1+178.1+175.4+179.3+176.9+178.3+179.1+180.2+174.2+179.4) = 178.44$

$\frac{1}{15}\cdot(183.2^2+179.4^2+175.4^2+179.0^2+180.6^2+178.1^2+178.1^2+175.4^2+179.3^2+176.9^2+178.3^2+179.1^2+180.2^2+174.2^2+179.4^2) = 31846.$

$s_{15}^2 = 31846.-178.44^2 = 5.1664, s_{15}^* = \sqrt{\frac{15}{14}\cdot 5.1664} = 2.3527$

A próbastatisztika számított értéke: $t_{próba} = \frac{178.44-180}{2.3527}\cdot\sqrt{15} = -2.5681$

Az $\varepsilon = 0.05$ höz tartozó kritikus érték az $f = 14$ szabadságfokó Student táblázatból: 2.145.

Mivel a számított érték abszolút értéke a nagyobb, ezért ezen a szinten elvetjük a nullhipotézist, azaz a mérésünkkel nem lehet alátámasztani az elméletet.

2. 17 véletlenszerűen választott fogyni akaró ember közül hatan diétáztak, 11-en pedig tornával próbálkoztak. Egy idő után a következő eredmény jelentkezett:

Fogyás mértéke diétázóknál:

$$7,5\ 6,0\ 2,0\ 3,5\ 6,0\ 5,0.$$

Fogyás mértéke tornázóknál:

$$5,0\ 8,0\ 6,5\ 5,5\ 4,5\ 9,0\ 7,5\ 6,5\ 5,0\ 2,5\ 6,0.$$

Igaz-e, hogy a torna hatásosabb a diétánál?

Megoldás: Két független mintás *t*-próbával lehet a kérdést eldönteni, amennyiben feltételezzük a normális eloszlást a mintára, és az *F*-próba elfogadja a szórások egyezését.

A két mintaelemszám $n_f = 6$ és $n_t = 11.$

A két átlag: $\bar{x}_6 = \frac{1}{6}\cdot(7.5+6.0+2.0+3.5+6.0+5.0) = 5.0$

$\bar{y}_{11} = \frac{1}{11}\cdot(5.0+8.0+6.5+5.5+4.5+9.0+7.5+6.5+5.0+2.5+6.0) = 6.0$

A két szórás számítása:

$\frac{1}{6}\cdot(7.5^2+6.0^2+2.0^2+3.5^2+6.0^2+5.0^2) = 28.25$

$s_{x,6}^2 = 28.25-25 = 3.25, s_{x,6}^* = \sqrt{\frac{6}{5}\cdot 3.25} = 1.9748$

$\frac{1}{11}\cdot(5.0^2+8.0^2+6.5^2+5.5^2+4.5^2+9.0^2+7.5^2+6.5^2+5.0^2+2.5^2+6.0^2) = 38.955$

$s_{y,11}^2 = 38.955-36 = 2.955, s_{y,11}^* = \sqrt{\frac{11}{10}\cdot 2.955} = 1.8029.$

Az *F* próba végrehajtása:

$F_{próba} = \frac{3.25}{2.955} = 1.0998$

A szabadságfokok: $f_1 = 5, f_2 = 10.$

Az $\varepsilon = 0.05$ höz tartozó kritikus érték: 4.24.

Tehát a két minta szórásai egyenlőeknek tekinthető, elvégezhető a *t*-próba.

$t_{próba} = \frac{5-6}{\sqrt{6\cdot 3.25+11\cdot 2.955}}\cdot\sqrt{\frac{6\cdot 11\cdot(6+11-2)}{6+11}} = -1.0582$

Az $\varepsilon = 0.05$ höz és a $f = 15$-höz tartozó kritikus érték: 2.131.

Mivel a próbastatisztika számított értéke abszolút értékben kisebb, ezért elfogadjuk a nullhipotézist, miszerint a diétával és a tornával egyenlő mértékben lehet fogyni.

3. Két dobókockával dobva a következő eredmények adódtak:

dobás	1	2	3	4	5	6
1. kocka	7	11	8	10	8	6
2. kocka	16	11	20	19	18	16

Tekinthető-e a két kocka egyformának? Döntésünket $\varepsilon = 0.05$ szinten hozzuk meg. Melyik kocka szabályos?

Megoldás: Döntésünket homogenitásra vonatkozó χ^2-próbával hozhatjuk meg.

$n = 7+11+8+10+8+6 = 50$

$m = 16+11+20+19+18+16 = 100$

$r = 6$

A próbastatisztika:

$$t_{próba} = 50 \cdot 100 \cdot \left(\frac{\left(\frac{7}{50} - \frac{16}{100}\right)^2}{7+16} + \frac{\left(\frac{11}{50} - \frac{11}{100}\right)^2}{11+11} + \frac{\left(\frac{8}{50} - \frac{20}{100}\right)^2}{8+20} + \frac{\left(\frac{10}{50} - \frac{19}{100}\right)^2}{10+19} + \frac{\left(\frac{8}{50} - \frac{18}{100}\right)^2}{8+18} + \frac{\left(\frac{6}{50} - \frac{16}{100}\right)^2}{6+16} \right) = 3.5805$$

Az $\varepsilon = 0.05$ szignifikancia-szinthez és az $f = 5$-höz tartozó kritikus érték: 11.1.
A nullhipotézist elfogadjuk ezen a szinten, vagyis a két kocka egyforma.
A szabályosságra vonatkozóan tiszta illeszkedésvizsgálatot végezhetünk el szintén χ^2-próbával.
A próbastatisztikák:

$$t_{próba\_1} = \frac{\left(7 - 50 \cdot \frac{1}{6}\right)^2 + \left(11 - 50 \cdot \frac{1}{6}\right)^2 + \left(8 - 50 \cdot \frac{1}{6}\right)^2 + \left(10 - 50 \cdot \frac{1}{6}\right)^2 + \left(8 - 50 \cdot \frac{1}{6}\right)^2 + \left(6 - 50 \cdot \frac{1}{6}\right)^2}{50 \cdot \frac{1}{6}} = 2.08$$

$$t_{próba\_2} = \frac{\left(16 - 100 \cdot \frac{1}{6}\right)^2 + \left(11 - 100 \cdot \frac{1}{6}\right)^2 + \left(20 - 100 \cdot \frac{1}{6}\right)^2 + \left(19 - 100 \cdot \frac{1}{6}\right)^2 + \left(18 - 100 \cdot \frac{1}{6}\right)^2 + \left(16 - 100 \cdot \frac{1}{6}\right)^2}{100 \cdot \frac{1}{6}} = 3.08$$

Ugyan azt a kritikus értéket kell most is használnunk. Azt kaptuk, hogy a két kocka szabályos, de mégsem azonos eloszlásúak!

4. Egy évfolyam statisztika ZH-jának eredményét tartalmazza az alábbi táblázat:

pontszám	0-5	6-10	11-15	16-20
fiúk	12	18	20	10
lányok	5	10	8	7

Függ-e a statisztika tudás a nemektől?
Megoldás: Függetlenség vizsgálatot hajtunk végre χ^2-próbával. $r = 4$ és $s = 2$.
Mintaelemszám: $n = 90$,
Peremösszegek: $n_{.1} = 17, n_{.2} = 28, n_{.3} = 28, n_{.4} = 17$
$n_{1.} = 60, n_{2.} = 30$
A próbastatisztika:

$$t_{próba} = 90 \cdot (\frac{\left(\frac{12}{90} - \frac{17}{90} \cdot \frac{60}{90}\right)^2}{\frac{17}{90} \cdot \frac{60}{90}} + \frac{\left(\frac{18}{90} - \frac{28}{90} \cdot \frac{60}{90}\right)^2}{\frac{28}{90} \cdot \frac{60}{90}} + \frac{\left(\frac{20}{90} - \frac{28}{90} \cdot \frac{60}{90}\right)^2}{\frac{28}{90} \cdot \frac{60}{90}} + \frac{\left(\frac{10}{90} - \frac{17}{90} \cdot \frac{60}{90}\right)^2}{\frac{17}{90} \cdot \frac{60}{90}} + \frac{\left(\frac{5}{90} - \frac{17}{90} \cdot \frac{30}{90}\right)^2}{\cdot \frac{30}{90}}$$
$$+ \frac{\left(\frac{10}{90} - \frac{28}{90} \cdot \frac{30}{90}\right)^2}{\frac{28}{90} \cdot \frac{30}{90}} + \frac{\left(\frac{8}{90} - \frac{28}{90} \cdot \frac{30}{90}\right)^2}{\frac{28}{90} \cdot \frac{30}{90}} + \frac{\left(\frac{7}{90} - \frac{17}{90} \cdot \frac{30}{90}\right)^2}{\frac{17}{90} \cdot \frac{30}{90}}) = 0.94538$$

A szabadságifok: $f = (4-1) \cdot (2-1) = 3$, az $\varepsilon = 0.05$ szignifikancia-szinthez tartozó kritikus érték: 7.81.
A nullhipotézist, azaz azt, hogy a tárgyi tudás független a hallgató nemétől el kell fogadni.

5. Legyen $X_1, X_2, \ldots, X_n$ egy statisztikai minta a következő folytonos eloszlásból:

$$f_\vartheta(x) = \begin{cases} \vartheta^2 \cdot x \cdot e^{-\vartheta x} & \text{ha } x > 0 \\ 0 & \text{különben} \end{cases}$$

Adja meg a $\vartheta > 0$ paraméter maximum-likelihood becslését.
Megoldás: $L(x_1, x_2, \ldots, x_n; \vartheta) = \prod_{i=1}^{n} (\vartheta^2 \cdot x_i \cdot e^{-\vartheta x_i})$

$$l(x_1, x_2, \ldots, x_n; \vartheta) = \sum_{i=1}^{n} (2 \ln \vartheta + \ln x_i - \vartheta x_i) = 2n \ln \vartheta + \sum_{i=1}^{n} \ln x_i - \vartheta \sum_{i=1}^{n} x_i.$$

$$\frac{dl}{d\vartheta} = \frac{2n}{\vartheta} - \sum_{i=1}^{n} x_i = 0 \Rightarrow \vartheta = \frac{2n}{\sum_{i=1}^{n} x_i} = \frac{2}{\bar{x}_n}.$$

$\frac{d^2 l}{d\vartheta^2} = -\frac{2n}{\vartheta^2} < 0 \Rightarrow$maximum!