Statisztika kidolgozott példák

1. *Közvéleménykutatást végzünk, megkérdezünk 256 embert. Mennyi a valószínűsége, hogy a megkérdezettek között a férfiak száma 121 és 135 között van?*

   A férfiak száma: K_n binomiális eloszlású.
   $n = 256$, $p = 1 - p = 1/2 \quad \Rightarrow \quad np = 128$, $\sqrt{np(1-p)} = 8$.
   Mivel nagy számokkal kell számolni, közelítjük a valószínűséget.

$$P(121 \leq K_n \leq 135) \overset{(1)}{=} P\left(\frac{121-128}{8} \leq \frac{K_n - 128}{8} \leq \frac{135-128}{8}\right) =$$

$$P\left(\frac{K_n - 128}{8} \leq \frac{7}{8}\right) - P\left(\frac{K_n - 128}{8} < -\frac{7}{8}\right) \overset{(2)}{\approx} \Phi\left(\frac{7}{8}\right) - \Phi\left(-\frac{7}{8}\right) \overset{(3)}{=}$$

$$2 \cdot \Phi\left(\frac{7}{8}\right) - 1 \overset{(4)}{\approx} 2 \cdot 0,81 - 1 = 0,62.$$

   (1) sztenderdizálás
   (2) Moivre-Laplace tétel alkalmazása: $\dfrac{K_n - np}{\sqrt{np(1-p)}}$ közelítőleg sztenderd normális eloszlású
   (3) $\Phi(-z) = 1 - \Phi(z)$ alkalmazása
   (4) $\Phi(7/8)$ kikeresése a sztenderd normális eloszlás táblázatából.

2. *Egy szabályos pénzérmét 400-szor feldobunk. Legalább mennyi fejet dobunk 0,95 valószínűséggel?*

   A dobott fejek száma: K_n binomiális eloszlású.
   $n = 400$, $p = 1 - p = 1/2 \quad \Rightarrow \quad np = 200$, $\sqrt{np(1-p)} = 10$.
   Mivel nagy számokkal kell számolni, közelítjük n-t.

$$0,95 = P(K_n \geq x) = 1 - P(K_n < x) \overset{(1)}{=} 1 - P\left(\frac{K_n - 200}{10} < \frac{x-200}{10}\right) \overset{(2)}{\approx}$$

$$1 - \Phi\left(\frac{x-200}{10}\right) \quad \Rightarrow \quad \Phi\Bigg(\underbrace{\frac{x-200}{10}}_{z}\Bigg) = \Phi(z) = 1 - \Phi(-z) = 0,05 \quad \Rightarrow$$

$$\Phi(-z) = 0,95 \quad \overset{(3)}{\Rightarrow} \quad -z \approx 1,65 \quad \Rightarrow \quad \frac{x-200}{10} \approx -1,65 \quad \Rightarrow \quad x \approx 183,5.$$

   (1) sztenderdizálás
   (2) Moivre-Laplace tétel alkalmazása
   (3) $-z$ kikeresése a sztenderd normális eloszlás táblázatából.

3. *Egy gyermek játékszer fizikai terhelhetőségére elvégzett próbatesztek a következő mérési eredményekre vezettek:* $40, 45, 39, 42, 37$ *(kg).*

   (a) *Adjunk becslést a terhelhetőség várható értékére és szórására!*

   (b) *Tudjuk, hogy a terhelhetőség normális eloszlást mutat. Adjunk meg olyan intervallumot, amely a várható értéket 0,95 valószínűséggel tartalmaza.*

   (c) *Tegyük fel, hogy a szórás a korrigált empirikus szórással egyezik meg. Adjunk a várható értékre konfidencia-intervallumot ebben az esetben is* $\alpha = 0,05$*-hoz!*

   (d) *Teszteljük azt a nullhipotézist, hogy a terhelhetőség várható értéke 40 kg (a szórás ismeretlen)!*

   (e) *Teszteljük azt a nullhipotézist, hogy a terhelhetőség várható értéke 40 kg (a szórás ismert, megegyezik a korrigált empirikus szórással)!*

$n = 5$

(a) Képletgyűjteményből:

$$E_n(\xi) = \frac{40+45+39+42+37}{5} = 40,6$$

$$V_n(\xi) = \frac{1}{5}((40-40,6)^2+(45-40,6)^2+(39-40,6)^2+(42-40,6)^2+(37-40,6)^2) = 7,44$$

$$V_n^*(\xi) = \frac{5}{4} \cdot V_n = 9,3 \quad D_n(\xi) = 2,73 \quad D_n^*(\xi) = 3,05$$

(b) A szórás ismeretlen, ezért az intervallumot $P(-x_\alpha < Z_n^*(\mu) < x_\alpha) = 0,95$ átalakításával kapjuk, ahol – mivel a képletgyűjteményben található $Z_n^*(\mu)$ Student eloszlású – az $x_\alpha = \Phi_{n-1}^{-1}(1-\alpha/2) = \Phi_4^{-1}(0,975) = 2,776$ a Student eloszlás táblázatából. Az átalakítással kapjuk:

$$0,95 = P\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i - x_\alpha\sqrt{\frac{V_n^*}{n}} < \mu < \frac{1}{n}\sum_{i=1}^{n}\xi_i + x_\alpha\sqrt{\frac{V_n^*}{n}}\right) =$$

$$P(40,6 - 2,776 \cdot \sqrt{1,86} < \mu < 40,6 + 2,776 \cdot \sqrt{1,86}) = P(36,81 < \mu < 44,39)$$

Tehát az intervallum: $(36,81;\ 44,39)$.

(c) A szórás ($\sigma = 3,05$) ismert, ezért az intervallumot $P(-x_\alpha < Z_n(\mu) < x_\alpha) = 0,95$ átalakításával kapjuk, ahol – mivel a képletgyűjteményben található $Z_n(\mu)$ sztenderd normális eloszlású – az $x_\alpha = \Phi^{-1}(1-\alpha/2) = \Phi^{-1}(0,975) = 1,96$ a sztenderd normális eloszlás táblázatából. Az átalakítással kapjuk:

$$0,95 = P\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i - x_\alpha\frac{\sigma}{\sqrt{n}} < \mu < \frac{1}{n}\sum_{i=1}^{n}\xi_i + x_\alpha\frac{\sigma}{\sqrt{n}}\right) =$$

$$P(40,6 - 1,96 \cdot \sqrt{1,86} < \mu < 40,6 + 1,96 \cdot \sqrt{1,86}) = P(37,93 < \mu < 43,27)$$

Tehát az intervallum: $(37,93;\ 43,27)$.

(d) Mivel normális eloszlású a minta és a szórás ismeretlen, t-próbát alkalmazunk, $\alpha = 0,05$.
$H_0 : \mu = 40$ $H_1 : \mu \neq 40$
A képletgyűjtemény (és a korábbi számolások) alapján:

$$t = Z_n^*(\mu) = \frac{40,6 - 40}{\sqrt{1,86}} \approx 0,44$$

A Student eloszlás táblázatából:

$$t_\alpha = \Phi_{n-1}^{-1}(1 - \alpha/2) = \Phi_4^{-1}(0,975) = 2,776$$

$|t| < t_\alpha$, ezért H_0-t elfogadjuk. (e) Mivel normális eloszlású a minta és a szórás ismert ($\sigma = 3,05$), mű-próbát alkalmazunk, $\alpha = 0,05$.
$H_0 : \mu = 40$
$H_1 : \mu \neq 40$
A képletgyűjtemény (és a korábbi számolások) alapján:

$$u = Z_n(\mu) = \frac{40,6 - 40}{\sqrt{1,86}} \approx 0,44$$

A sztenderd normális eloszlás táblázatából:

$$u_\alpha = \Phi^{-1}(1 - \alpha/2) = \Phi^{-1}(0,975) = 1,98$$

$|u| < u_\alpha$, ezért a nullhipotézist elfogadjuk.