Konfidenciaintervallum (7. előadás)

Példa: megmérték hatvan ember vércukorszintjét, egymástól függetlenül.

A minta egy részlete (mmol/l mértékegységben):

5,98	6,1	5,99	6,21	5,97	6,23	...	5,85

Alapstatisztikák: $n = 60$ (méret), $\overline{X} = 5,99$ (átlag), $s_n^* = 0,18$ (korrigált tapasztalati szórás)

Konfidenciaintervallum (7. előadás)

Példa: megmérték hatvan ember vércukorszintjét, egymástól függetlenül.

A minta egy részlete (mmol/l mértékegységben):

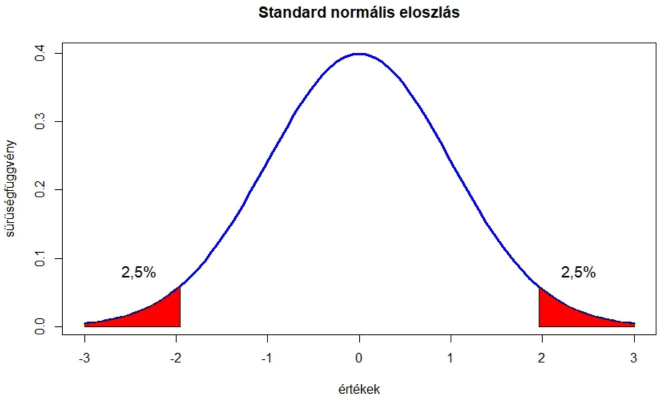
5,98 6,1 5,99 6,21 5,97 6,23 ... 5,85

Alapstatisztikák: $n = 60$ (méret), $\overline{X} = 5,99$ (átlag), $s_n^* = 0,18$ (korrigált tapasztalati szórás)

Cél: a várható érték becslése az adatok alapján

pontosabban: adjunk meg egy olyan intervallumot, ami legalább 95% valószínűséggel tartalmazza az "igazi" várható értéket – ezt fogjuk **95 % megbízhatósági szintű konfidenciaintervallumnak** hívni

A kétoldali z-próba kritikus értéke



Az $\alpha = 0,05$ terjedelmű kétoldali z-próba (u-próba) kritikus értéke: $\Phi^{-1}(1 - \alpha/2) = \Phi^{-1}(0,975) = 1,96$. Vagyis ha $Z \sim N(0,1)$, akkor $\mathbb{P}(-1,96 \leq Z \leq 1,96) = 95\%$.

Konfidenciaintervallum a várható értékre

$$(T_1, T_2) = \left(\overline{X} - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}},\ \overline{X} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}\right)$$

Ekkor

$$\begin{aligned}\mathbb{P}_m(T_1 \leq m \leq T_2) &= \mathbb{P}_m\left(m - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}} \leq \overline{X} \leq m + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\frac{\sigma}{\sqrt{n}}\right) \\ &= \mathbb{P}_m\left(-\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \leq \frac{\overline{X} - m}{\sigma/\sqrt{n}} \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)\right) = \\ &= 2\left(1 - \frac{\alpha}{2}\right) - 1 = 1 - \alpha,\end{aligned}$$

ahol felhasználtuk, hogy mivel $X_1, \ldots, X_n$ független $N(m, \sigma^2)$ eloszlásúak, így $\overline{X} \sim N(m, \sigma^2/n)$ eloszlású. Ezért $Z = (\overline{X} - m)/(\sigma/\sqrt{n})$ standard normális eloszlású, vagyis

$$\mathbb{P}(-a \leq Z \leq a) = \Phi(a) - \Phi(-a) = \Phi(a) - (1 - \Phi(a) = 2\Phi(a) - 1.$$

Példa konfidenciaintervallumra

Minta: megmértük hatvan ember vércukorszintjét, tegyük fel, hogy ez **normális eloszlású, valódi szórása** $0,2$.

$n = 60$ (méret), $\overline{X} = 5,99$ (átlag), $s_n^* = 0,18$ (korrigált tapasztalati szórás)

Adjunk meg olyan intervallumot, ami legalább 95% valószínűséggel tartalmazza a vércukorszint valódi várható értékét.

megbízhatósági szint: $1 - \alpha = 95\%$, azaz $\alpha = 0,05$. Ebből $\Phi^{-1}(1 - \alpha/2) = \Phi^{-1}(0,975) = 1,96$.

95 %-os megbízhatósági szintű konfidenciaintervallum a várható értékre:

$$\left(5,99 - 1,96 \cdot \frac{0,2}{\sqrt{60}},\quad 5,99 + 1,96 \cdot \frac{0,2}{\sqrt{60}}\right) = (5,94; 6,04).$$

Kisebb mintaelemszámhoz hosszabb, nagyobb megbízhatósághoz szintén hosszabb konfidenciaintervallum tartozik.

Konfidenciaintervallumok

Legyen $(\Omega, \mathcal{A}, \mathcal{P})$ statisztikai mező, $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$ és $\underline{X} = (X_1, \ldots, X_n)$ független azonos eloszlású minta. Tegyük fel, hogy ϑ valós paraméter, vagyis $\Theta \subseteq \mathbb{R}$.

Definíció

Azt mondjuk, hogy a $(T_1(\underline{X}), T_2(\underline{X}))$ intervallum legalább $1 - \alpha$ megbízhatósági szintű konfidenciaintervallum ϑ-ra, ha minden $\vartheta \in \mathbb{R}$ esetén teljesül, hogy

$$\mathbb{P}_\vartheta(T_1(\underline{X}) < \vartheta < T_2(\underline{X})) \geq 1 - \alpha.$$

A konfidenciaintervallum megbízhatósági szintje: $\inf_{\vartheta \in \Theta}\{\mathbb{P}_\vartheta(\vartheta \in (T_1, T_2))\}$.

Konfidenciaintervallum a várható értékre

A Φ a standard normális eloszlás eloszlásfüggvénye, azaz ha $Z \sim N(0,1)$:

$$\Phi(t) = \mathbb{P}(Z \leq t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-s^2/2} ds.$$

Állítás (Konfidenciaintervallum a várható értékre, ismert szórás)

Tegyük fel, hogy $X_1, \ldots, X_n$ független azonos eloszlású **normális eloszlású** *való-színűségi változók, melyek szórása, σ ismert.*
Ekkor a

$$(T_1, T_2) = \left(\overline{X} - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}}, \ \overline{X} + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \frac{\sigma}{\sqrt{n}} \right)$$

intervallum $1 - \alpha$ megbízhatósági szintű **kétoldali konfidenciaintervallum** *az el-oszlás várható értékére.*

Fisher–Bartlett-tétel

Valójában az eloszlás valódi szórása a legtöbb esetben nem ismert. A σ szórást az s_n^* korrigált tapasztalati szórással helyettesítjük. Kérdés, hogyan változik így az eloszlás.

Tétel (Fisher–Bartlett)

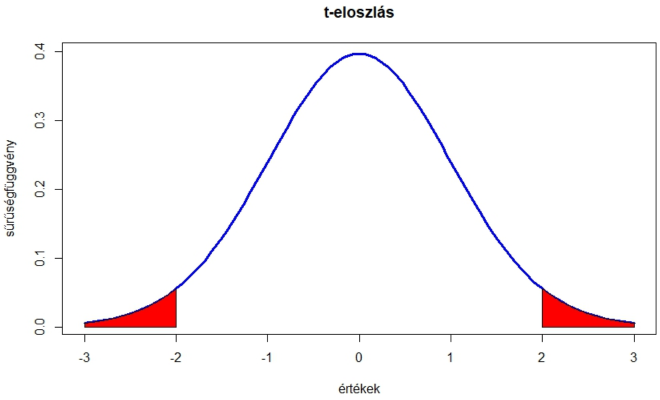
Tegyük fel, hogy $X_1, X_2, \ldots, X_n$ független m várható értékű, σ szórású, **normális eloszlású** *valószínűségi változók. Ekkor*

1. $\overline{X} \sim N\left(m, \frac{\sigma^2}{n}\right)$;
2. $\overline{X}$ *és* s_n^* *függetlenek;*
3. $(n-1)s_n^{*2}/\sigma^2$ *eloszlása* $n-1$ *szabadsági fokú* χ^2*-eloszlás;*
4. $\frac{\overline{X}-m}{s_n^*} \cdot \sqrt{n}$ *eloszlása* $n-1$ *szabadsági fokú* t*-eloszlás.*

Itt

$$s_n^* = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(X_j - \overline{X})^2} = \sqrt{\frac{n}{n-1}\left(\left(\frac{1}{n}\sum_{j=1}^{n}X_j^2\right) - \overline{X}^2\right)}.$$

t-eloszlás kritikus értékei



Az $f = 59$ szabadsági fokú $\alpha = 0,05$ terjedelmű kétoldali t-próba kritikus értéke: $t_{59,0,05} = 2,001$.

Konfidenciaintervallum a várható értékre

Legyenek $Z_0, Z_1, \ldots, Z_n$ független $N(0,1)$ eloszlásúak, és $t_{f,\alpha}$ az f szabadsági fokú α terjedelmű kétoldali t-próba kritikus értéke, azaz az f szabadsági fokú t-eloszlás $1-\alpha/2$-kvantilise:

$$1-\alpha/2 = \mathbb{P}(Y \leq t_{f,\alpha}) = \mathbb{P}\left(\frac{Z_0}{\sqrt{Z_1^2+\ldots+Z_f^2}} \leq t_{f,\alpha}\right).$$

Az $Y = \frac{Z_0}{\sqrt{Z_1^2+\ldots+Z_f^2}}$ valószínűségi változó eloszlása f szabadsági fokú ***t*-eloszlás**.

Konfidenciaintervallum a várható értékre

Legyenek $Z_0, Z_1, \ldots, Z_n$ független $N(0,1)$ eloszlásúak, és $t_{f,\alpha}$ az f szabadsági fokú α terjedelmű kétoldali t-próba kritikus értéke, azaz az f szabadsági fokú t-eloszlás $1-\alpha/2$-kvantilise:

$$1-\alpha/2 = \mathbb{P}(Y \le t_{f,\alpha}) = \mathbb{P}\left(\frac{Z_0}{\sqrt{Z_1^2+\ldots+Z_f^2}} \le t_{f,\alpha}\right).$$

Az $Y = \frac{Z_0}{\sqrt{Z_1^2+\ldots+Z_f^2}}$ valószínűségi változó eloszlása f szabadsági fokú **t-eloszlás**.

Állítás (Konfidenciaintervallum a várható értékre, ismeretlen szórás)

Tegyük fel, hogy $X_1, \ldots, X_n$ független $N(m, \sigma^2)$ normális eloszlású valószínűségi változók (m, σ ismeretlenek). Ekkor a

$$(T_1, T_2) = \left(\overline{X} - t_{n-1,\alpha} \cdot \frac{s_n^*}{\sqrt{n}},\ \overline{X} + t_{n-1,\alpha} \cdot \frac{s_n^*}{\sqrt{n}}\right)$$

intervallum $1-\alpha$ megbízhatósági szintű **kétoldali konfidenciaintervallum** *az eloszlás várható értékére.*

Példa konfidenciaintervallumra

Minta: vércukorszint-értékek, tegyük fel, hogy ez **normális eloszlású**, szórása nem ismert.

$n = 60$ (méret), $\overline{X} = 5,99$ (átlag), $s_n^* = 0,18$ (korrigált tapasztalati szórás)

Adjunk meg olyan intervallumot, ami legalább 95% valószínűséggel tartalmazza a vércukorszint valódi várható értékét.

megbízhatósági szint: $1 - \alpha = 95\%$, azaz $\alpha = 0,05$. Az $f = 59$ szabadsági fokú $\alpha = 0,05$ terjedelmű kétoldali t-próba kritikus értéke: $t_{59,0,975} = 2$.

95 %-os megbízhatósági szintű konfidenciaintervallum a várható értékre:

$$\left(5,99 - 2 \cdot \frac{0,18}{\sqrt{60}},\quad 5,99 + 2 \cdot \frac{0,18}{\sqrt{60}}\right) = (5,94; 6,04).$$

Hosszabb intervallumot kaptunk, mint ismert szórás esetén.

Egyoldali konfidenciaintervallum a várható értékre

Állítás (Egyoldali konfidenciaintervallum)

Tegyük fel, hogy $X_1, \ldots, X_n$ független $N(m, \sigma^2)$ normális eloszlású valószínűségi változók (m, σ ismeretlenek). Ekkor a

$$\left(-\infty, \quad \overline{X} + \overline{t}_{n-1,1-\alpha} \cdot \frac{s_n^*}{\sqrt{n}} \right)$$

intervallum $1-\alpha$ megbízhatósági szintű **egyoldali konfidenciaintervallum** *az eloszlás várható értékére.*

Itt $\overline{t}_{n-1,1-\alpha/2}$ az $f = n-1$ szabadsági fokú α terjedelmű **egyoldali t-próba** kritikus értéke.

Konfidenciaintervallum a szórásra

Fisher–Bartlett-tétel:

$$\frac{(n-1)s_n^{*2}}{\sigma^2} = \frac{\sum_{j=1}^{n}(X_j - \overline{X})^2}{\sigma^2} \sim \chi^2_{n-1},$$

azaz a hányados eloszlása $n-1$ szabadsági fokú χ^2-eloszlás (ami megegyezik $Z_1^2 + Z_2^2 + \ldots + Z_{n-1}^2$ eloszlásával, ahol Z_j-k független standard normális eloszlásúak).

Állítás

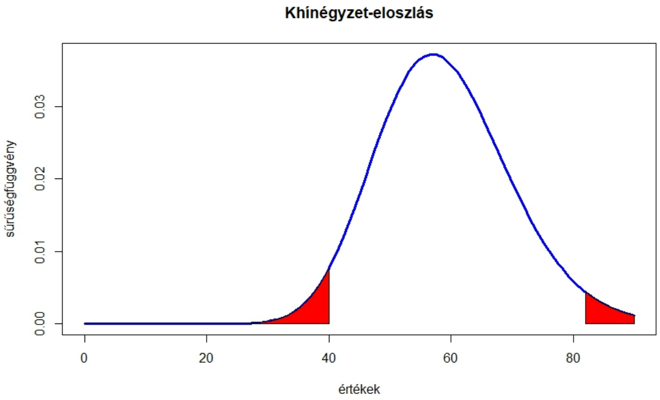
Legyen $X_1, X_2, \ldots, X_n$ független normális eloszlású minta. Ekkor az eloszlás σ^2 szórásnégyzetére $1-\alpha$ megbízhatósági szintű konfidenciaintervallum az alábbi:

$$(T_1, T_2) = \left(\frac{\sum_{j=1}^{n}(X_j - \overline{X})^2}{c_{n-1,1-\alpha/2}}; \frac{\sum_{j=1}^{n}(X_j - \overline{X})^2}{c_{n-1,\alpha/2}} \right),$$

ahol $c_{f,q}$ az f-szabadsági fokú q terjedelmű χ^2-próba kritikus értéke, azaz az f szabadsági fokú χ^2-eloszlás q-kvantilise.

Ezzel a választással nem a legrövidebb intervallumot kapjuk.

A χ^2-eloszlás kvantilisei



Az $f = 59$ szabadsági fokú χ^2-próba kvantilisei $q = \alpha/2 = 0,025$-tel és $q = 1 - \alpha/2 = 0,975$-tel.

Konfidenciaintervallum a valószínűségre

Az A esemény valószínűsége $p \in [0,1]$ ismeretlen paraméter. Ezt szeretnénk megbecsülni. Ha n kísérletből az A esemény X-szer következett be:

$$\mathbb{E}_p(X) = np; \quad D_p(X) = \sqrt{np(1-p)}; \quad \mathbb{E}_p\left(\frac{X}{n}\right) = p; \quad D_p(X) = \frac{\sqrt{p(1-p)}}{\sqrt{n}}.$$

$X \sim \mathrm{Bin}(n,p)$ binomiális eloszlású, viszont a relatív gyakoriságot, X/n-t nem tudjuk egyetlen eloszlással közelíteni.

Ezért X/n eloszlását $\hat{p}$ várható értékű, $\frac{\hat{p}(1-\hat{p})}{\sqrt{n}}$ szórású normális eloszlással közelítjük, ahol $\hat{p}$ az esemény relatív gyakorisága a mintában.

$1-\alpha$ megbízhatósági szintű konfidenciaintervallum p-re:

$$\left(\hat{p} - \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \cdot \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}; \ \hat{p} + \Phi^{-1}\left(1-\frac{\alpha}{2}\right) \cdot \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}\right).$$

Konfidenciaintervallum az ML-becslés alapján

Ha likelihoodfüggvény teljesít bizonyos regularitási feltételeket, akkor a ϑ paraméternek az $X_1, X_2, \ldots, X_n$ mintából számolt $\hat{\vartheta}_n$ maximumlikelihood-becslése

- létezik;
- aszimptotikusan torzítatlan: $\lim_{n\to\infty} \mathbb{E}_\vartheta(\hat{\vartheta}_n) = \vartheta$ minden $\vartheta \in \Theta$-ra;
- aszimptotikusan hatásos: $\lim_{n\to\infty} \sqrt{nI_1(\vartheta)}D_\vartheta(\hat{\vartheta}_n) = 1$ minden $\vartheta \in \Theta$-ra;
- aszimptotikusan normális eloszlású: $\sqrt{nI_1(\vartheta)}(\hat{\vartheta}_n - \vartheta)$ eloszlásban tart a standard normális eloszláshoz minden $\vartheta \in \Theta$-ra $n \to \infty$ esetén.

Konfidenciaintervallum az ML-becslés alapján

Ha likelihoodfüggvény teljesít bizonyos regularitási feltételeket, akkor a ϑ paraméternek az $X_1, X_2, \ldots, X_n$ mintából számolt $\hat{\vartheta}_n$ maximumlikelihood-becslése

- létezik;
- aszimptotikusan torzítatlan: $\lim_{n\to\infty} \mathbb{E}_\vartheta(\hat{\vartheta}_n) = \vartheta$ minden $\vartheta \in \Theta$-ra;
- aszimptotikusan hatásos: $\lim_{n\to\infty} \sqrt{nI_1(\vartheta)}D_\vartheta(\hat{\vartheta}_n) = 1$ minden $\vartheta \in \Theta$-ra;
- aszimptotikusan normális eloszlású: $\sqrt{nI_1(\vartheta)}(\hat{\vartheta}_n - \vartheta)$ eloszlásban tart a standard normális eloszláshoz minden $\vartheta \in \Theta$-ra $n \to \infty$ esetén.

Ez alapján **aszimptotikus** konfidenciaintervallum, ami $n \to \infty$ esetén $1-\alpha$-hoz tartó valószínűséggel tartalmazza ϑ-t:

$$\left(\hat{\vartheta}_n - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{1}{\sqrt{\hat{I}_n(\vartheta)}};\ \hat{\vartheta}_n + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{1}{\sqrt{\hat{I}_n(\vartheta)}}\right),$$

ahol $\hat{I}_n(\vartheta) = n\cdot I_1(\hat{\vartheta}_n)$, vagyis a Fisher-információ kifejezésébe a maximumlikelihood-becslést írjuk be.

Konfidenciaintervallum az ML-becslés alapján

Tekintsünk egy n elemű független mintát az alábbi sűrűségfüggvénnyel, ahol $\lambda > 0$ ismeretlen paraméter:

$$f_\lambda(x) = \frac{\lambda^4}{6} x^3 e^{-\lambda x} \mathbb{I}(x > 0).$$

Konfidenciaintervallum az ML-becslés alapján

Tekintsünk egy n elemű független mintát az alábbi sűrűségfüggvénnyel, ahol $\lambda > 0$ ismeretlen paraméter:

$$f_\lambda(x) = \frac{\lambda^4}{6} x^3 e^{-\lambda x} \mathbb{I}(x > 0).$$

A likelihood-függvény, felhasználva, hogy az X_j-k 1 valószínűséggel pozitívak:

$$f(\lambda) = \prod_{j=1}^{n} f_\lambda(X_j) = \prod_{j=1}^{n} \left(\frac{\lambda^4}{6} X_j^3 e^{-\lambda X_j} \right) = \frac{\prod_{j=1}^{n} X_j^3}{6^n} \cdot \lambda^{4n} \cdot e^{-\lambda \sum_{j=1}^{n} X_j}.$$

Konfidenciaintervallum az ML-becslés alapján

Tekintsünk egy n elemű független mintát az alábbi sűrűségfüggvénnyel, ahol $\lambda > 0$ ismeretlen paraméter:

$$f_\lambda(x) = \frac{\lambda^4}{6} x^3 e^{-\lambda x} \mathbb{I}(x > 0).$$

A likelihood-függvény, felhasználva, hogy az X_j-k 1 valószínűséggel pozitívak:

$$f(\lambda) = \prod_{j=1}^{n} f_\lambda(X_j) = \prod_{j=1}^{n} \left(\frac{\lambda^4}{6} X_j^3 e^{-\lambda X_j} \right) = \frac{\prod_{j=1}^{n} X_j^3}{6^n} \cdot \lambda^{4n} \cdot e^{-\lambda \sum_{j=1}^{n} X_j}.$$

$$\log f(\lambda) = \log \frac{\prod_{j=1}^{n} X_j^3}{6^n} + 4n \log \lambda - \lambda \cdot \sum_{j=1}^{n} X_j.$$

$$\frac{\partial}{\partial \lambda} \log f(\lambda) = \frac{4n}{\lambda} - \sum_{j=1}^{n} X_j.$$

Konfidenciaintervallum az ML-becslés alapján

Tekintsünk egy n elemű független mintát az alábbi sűrűségfüggvénnyel, ahol $\lambda > 0$ ismeretlen paraméter:

$$f_\lambda(x) = \frac{\lambda^4}{6} x^3 e^{-\lambda x} \mathbb{I}(x > 0).$$

A likelihood-függvény, felhasználva, hogy az X_j-k 1 valószínűséggel pozitívak:

$$f(\lambda) = \prod_{j=1}^{n} f_\lambda(X_j) = \prod_{j=1}^{n} \left(\frac{\lambda^4}{6} X_j^3 e^{-\lambda X_j} \right) = \frac{\prod_{j=1}^{n} X_j^3}{6^n} \cdot \lambda^{4n} \cdot e^{-\lambda \sum_{j=1}^{n} X_j}.$$

$$\log f(\lambda) = \log \frac{\prod_{j=1}^{n} X_j^3}{6^n} + 4n \log \lambda - \lambda \cdot \sum_{j=1}^{n} X_j.$$

$$\frac{\partial}{\partial \lambda} \log f(\lambda) = \frac{4n}{\lambda} - \sum_{j=1}^{n} X_j.$$

Ez pontosan akkor pozitív, ha $\lambda < \frac{4}{\overline{X}}$, és negatív különben. Ezért a $\log f(\lambda)$ függvény globális maximumhelye $\lambda = \frac{4}{\overline{X}}$, és a logaritmusfüggvény monotonitása miatt ez lesz a maximumlikelihood-becslés is.

Konfidenciaintervallum az ML-becslés alapján

Az előző alapján a Fisher-információt is kiszámíthatjuk, ha tudjuk, hogy

$$\mathbb{E}_\lambda(X_j) = \frac{4}{\lambda}, \text{ és } D_\lambda(X_j) = \frac{4}{\lambda^2}.$$

Ugyanis

$$I_1(\vartheta) = \mathbb{E}_\lambda\left(\left(\frac{\partial}{\partial\lambda}\log f_\lambda(X_1)\right)^2\right) = \mathbb{E}_\lambda\left(\left(\frac{4n}{\lambda} - \sum_{j=1}^{n} X_j\right)^2\right) =$$

$$= D_\lambda^2\left(\sum_{j=1}^{n} X_j\right) = n \cdot D_\lambda^2(X_1) = n \cdot \frac{4}{\lambda^2}.$$

Konfidenciaintervallum az ML-becslés alapján

Az előző alapján a Fisher-információt is kiszámíthatjuk, ha tudjuk, hogy

$$\mathbb{E}_\lambda(X_j) = \frac{4}{\lambda}, \text{ és } D_\lambda(X_j) = \frac{4}{\lambda^2}.$$

Ugyanis

$$I_1(\vartheta) = \mathbb{E}_\lambda\left(\left(\frac{\partial}{\partial\lambda}\log f_\lambda(X_1)\right)^2\right) = \mathbb{E}_\lambda\left(\left(\frac{4n}{\lambda} - \sum_{j=1}^n X_j\right)^2\right) =$$

$$= D_\lambda^2\left(\sum_{j=1}^n X_j\right) = n \cdot D_\lambda^2(X_1) = n \cdot \frac{4}{\lambda^2}.$$

A fentiek alapján 95%-os megbízhatósági szintű aszimptotikus konfidenciainterval-lum λ-ra:

$$\left(\frac{4}{\overline{\overline{X}}} - 1,96 \cdot \frac{4\overline{X}^2}{\sqrt{n}}, \frac{4}{\overline{\overline{X}}} + 1,96 \cdot \frac{4\overline{X}^2}{\sqrt{n}}\right).$$

Hipotézisvizsgálat

Legyen $(\Omega, \mathcal{A}, \mathcal{P})$ paraméteres statisztikai mező, azaz $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$ valamilyen Θ paramétertérrel. A paraméterteret bontsuk fel két diszjunkt halmaz uniójára: $\Theta = \Theta_0 \cup \Theta_1$, ahol tehát $\Theta_0 \cap \Theta_1 = \emptyset$.

Nullhipotézis. $H_0 : \vartheta \in \Theta_0$.

Ellenhipotézis. $H_1 : \vartheta \in \Theta_1$.

Hipotézisvizsgálat

Legyen $(\Omega, \mathcal{A}, \mathcal{P})$ paraméteres statisztikai mező, azaz $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$ valamilyen Θ paramétertérrel. A paraméterteret bontsuk fel két diszjunkt halmaz uniójára: $\Theta = \Theta_0 \cup \Theta_1$, ahol tehát $\Theta_0 \cap \Theta_1 = \emptyset$.

Nullhipotézis. $H_0 : \vartheta \in \Theta_0$.

Ellenhipotézis. $H_1 : \vartheta \in \Theta_1$.

A minta $\underline{X} = (X_1, \ldots, X_n)$, a mintatér legyen B (vagyis $(X_1, \ldots, X_n)$ a $B \subseteq \mathbb{R}^n$ halmaz egy véletlen eleme). A mintateret is felbontjuk két diszjunkt halmaz uniójára: $B = B_0 \cup B_1$, ahol $B_0 \cap B_1 = \emptyset$.

Elfogadási tartomány: B_0. Ha $(X_1, \ldots, X_n) \in B_0$, akkor H_0-t elfogadjuk.

Elutasítási (kritikus) tartomány: B_1. Ha $(X_1, \ldots, X_n) \in B_1$, akkor H_0-t elutasítjuk.

Tehát a nullhipotézist akkor fogadjuk el, ha a minta az elfogadási tartományba esik, különben elutasítjuk.

Hipotézisvizsgálat

Nullhipotézis (null hypothesis). $H_0 : \vartheta \in \Theta_0$.

Ellenhipotézis (alternative hypothesis). $H_1 : \vartheta \in \Theta_1$.

- **Elsőfajú hibát** vétünk, ha H_0 igaz, és elutasítjuk.
- A próba **szignifikanciaszintje vagy terjedelme** (level of significance):

$$\alpha = \sup_{\vartheta \in \Theta_0} \mathbb{P}_\vartheta(\underline{X} \in B_1).$$

- **Másodfajú hibát** vétünk, ha H_0 nem igaz, és elfogadjuk.
- A próba **erőfüggvénye** az alábbi $\beta : \Theta_1 \to [0,1]$ függvény:

$$\beta(\vartheta) = \mathbb{P}_\vartheta(\underline{X} \in B_1) \qquad (\vartheta \in \Theta_1).$$

Hipotézisvizsgálat: p-érték

Definíció

Egy hipotézisvizsgálati feladatban a p-érték (p-value) a legnagyobb olyan szignifikanciaszint, ami mellett H_0-t elfogadjuk.

Vagyis ha α a szignifikanciaszint, akkor

$p < \alpha$ esetén elutasítjuk H_0-t, szignifikáns eltérés H_0-tól.

$p \geq \alpha$ esetén elfogadjuk H_0-t, nincs szignifikáns eltérés H_0-tól, nem volt elég bizonyíték H_1-re.

A szokásos $\alpha = 0,05$ értékkel: $p < 0,05$ esetén **elutasítjuk a nullhipotézist, szignifikáns eltérés van**, különben elfogadjuk a nullhipotézist, nincs szignifikáns eltérés.

Nagy mintaelemszám esetén kis eltérés is szignifikáns. A próba ereje használható annak ellenőrzésére, hogy nem volt-e túl érzékeny az eljárás.

Próbák tulajdonságai

- A próba **szignifikanciaszintje vagy terjedelme** (level of significance) az elsőfajú hiba valószínűségének szuprémuma:

$$\alpha = \sup_{\vartheta \in \Theta_0} \mathbb{P}_\vartheta(\underline{X} \in B_1).$$

- A próba **erőfüggvénye** az alábbi $\beta : \Theta_1 \to [0, 1]$ függvény:

$$\beta(\vartheta) = \mathbb{P}_\vartheta(\underline{X} \in B_1) \qquad (\vartheta \in \Theta_1).$$

- Azonos terjedelmű próbák közül az az **erősebb**, amelynek az erőfüggvénye minden $\vartheta \in \Theta_1$-re nagyobb vagy egyenlő, mint a másiké:

$$\beta^*(\vartheta) = \mathbb{P}_\vartheta(\underline{X} \in B_1^*) \geq \beta(\vartheta) = \mathbb{P}_\vartheta(\underline{X} \in B_1) \qquad (\vartheta \in \Theta_1).$$

- Egy próba **konzisztens**, ha az erőfüggvénye 1-hez tart minden $\vartheta \in \Theta_1$-re.
- Egy próba **torzítatlan**, ha

$$\mathbb{P}_{\vartheta_0}(\underline{X} \in B_1) \leq \mathbb{P}_{\vartheta_1}(\underline{X} \in B_1)$$

minden $\vartheta_0 \in \Theta_0$-ra és $\vartheta_1 \in \Theta_1$-re.

Neyman–Pearson-lemma

Tegyük fel, hogy a paramétertér két elemből áll: $\Theta = \{\vartheta_0, \vartheta_1\}$. A likelihood-függvények: $L_{n,0}$ illetve $L_{n,1}$.

Likelihood-hányados próba: válasszunk egy c számot. Ha

$$\frac{L_{n,1}(X_1, \ldots, X_n)}{L_{n,0}(X_1, \ldots, X_n)} \geq c$$

akkor utasítsuk el a nullhipotézist, különben fogadjuk el.

Lemma (Neyman–Pearson-lemma, 1. rész)

Tegyük fel, hogy a likelihood-hányados próba szignifikanciaszintje α. Ekkor a likelihood-hányados próba a legerősebb próba a legfeljebb α szignifikanciaszintű próbák között.

Például: ϑ_0 esetén indikátor eloszlás $0,2$ paraméterrel, ϑ_1 esetén indikátor eloszlás $0,9$ paraméterrel. Ha 10 megfigyelésből k-nál következik be az esemény:

$$\frac{L_{n,1}(X_1, \ldots, X_n)}{L_{n,0}(X_1, \ldots, X_n)} = \frac{\binom{10}{k} 0,9^k 0,1^{n-k}}{\binom{10}{k} 0,2^k 0,8^{n-k}} = \left(\frac{0,9}{0,2}\right)^k \left(\frac{0,1}{0,8}\right)^{n-k}.$$

Ez annál nagyobb, minél nagyobb k. Elutasítjuk H_0-t, ha $k \geq k_0$ megfelelő k_0-lal.

Neyman–Pearson-lemma

Például: ϑ_0 esetén indikátor eloszlás $0,2$ paraméterrel, ϑ_1 esetén indikátor eloszlás $0,9$ paraméterrel. Legyen X az, hogy 10 megfigyelésből hányszor következik be az esemény. A Neyman–Pearson-lemma szerint akkor utasítjuk el H_0-t, ha $X \geq k_0$ megfelelő k_0-lal.

Ha $k_0 = 5$: az elsőfajú hiba valószínűsége még megfelelő:

$$\mathbb{P}_0(X \geq 5) = \sum_{j=5}^{10} \binom{10}{j} 0,2^j 0,8^{10-j} = 0,033 \leq 0,05.$$

Ha $k_0 = 4$: az elsőfajú hiba valószínűsége túl nagy:

$$\mathbb{P}_0(X \geq 4) = \sum_{j=4}^{10} \binom{10}{j} 0,2^j 0,8^{10-j} = 0,12 > 0,05.$$

Neyman–Pearson-lemma

Például: ϑ_0 esetén indikátor eloszlás $0,2$ paraméterrel, ϑ_1 esetén indikátor eloszlás $0,9$ paraméterrel. Legyen X az, hogy 10 megfigyelésből hányszor következik be az esemény. A Neyman–Pearson-lemma szerint akkor utasítjuk el H_0-t, ha $X \geq k_0$ megfelelő k_0-lal.

Ha $k_0 = 5$: az elsőfajú hiba valószínűsége még megfelelő:

$$\mathbb{P}_0(X \geq 5) = \sum_{j=5}^{10} \binom{10}{j} 0,2^j 0,8^{10-j} = 0,033 \leq 0,05.$$

Ha $k_0 = 4$: az elsőfajú hiba valószínűsége túl nagy:

$$\mathbb{P}_0(X \geq 4) = \sum_{j=4}^{10} \binom{10}{j} 0,2^j 0,8^{10-j} = 0,12 > 0,05.$$

Pontosan $\alpha = 0,05$ terjedelem: $X \geq 5$ esetén biztosan, $X = 4$ esetén 19% valószínűséggel utasítjuk el a nullhipotézist ($0,19 \cdot \mathbb{P}_0(X = 4) = 0,05 - 0,03$). Ez legerősebb próba a legfeljebb $\alpha = 0,05$ szignfikanciaszintű próbák között, és erősebb, mint a $k_0 = 5$-tel kapott determinisztikus próbáé.

Véletlenített próba

Tegyük fel, hogy a paramétertér két elemből áll: $\Theta = \{\vartheta_0, \vartheta_1\}$. A likelihood-függvények: $L_{n,0}$ illetve $L_{n,1}$.

Likelihood-hányados próba: válasszunk egy c számot. Ha

$$\frac{L_{n,1}(X_1, \dots, X_n)}{L_{n,0}(X_1, \dots, X_n)} > c$$

akkor utasítsuk el a nullhipotézist. Ha a hányados egyenlő c-vel, akkor p valószínűséggel utasítsuk el, és $1 - p$ valószínűséggel fogadjuk el. Ha kisebb c-nél, fogadjuk el a nullhipotézist.

Lemma (Neyman–Pearson-lemma, 2. rész)

Legyen $\alpha \in (0, 1)$ tetszőleges. Ekkor lehet olyan c és p számokat választani, hogy a likelihood-hányados próba a szignifikanciaszintje éppen α, és így ez a legerősebb próba a legfeljebb α szignifikanciaszintű próbák között.

Szekvenciális próbák

A valószínűséghányados n elemű mintából:

$$V_n = \frac{L_{n,1}(X_1, \dots, X_n)}{L_{n,0}(X_1, \dots, X_n)} = \frac{\prod_{j=1}^n f_1(X_j)}{\prod_{j=1}^n f_0(X_j)},$$

ha az eloszlás abszolút folytonos.

A, B rögzített, a próbára jellemző számok. Addig veszünk mintaelemeket, amíg $V_n \geq B$ vagy $V_n \leq A$ nem teljesül. Vagyis:

- ha $V_n \geq B$, elutasítjuk H_0-t;
- ha $V_n \leq A$, elfogadjuk H_0-t;
- ha $A < V_n < B$: új mintaelemet veszünk.

Kétlépcsős változat: n_1 elemű mintát veszünk. Ha $V_{n_1} \geq B$, elutasítjuk H_0-t, ha $V_{n_2} \leq A$, elfogadjuk H_0-t, különben további n_2 darab mintaelemet veszünk, és akkor utasítjuk el a nullhipotézist, ha $V_{n_1+n_2} > C$ teljesül.

Neyman–Pearson-lemma

Példa: legyen $\Theta = \{m_0, m_1\}$, és $\mathcal{P}$ álljon az $N(m_1, \sigma)$ és $N(m_2, \sigma)$ eloszlásokból. A likelihood-hányados:

$$\frac{L_{n,1}(X_1,\ldots,X_n)}{L_{n,0}(X_1,\ldots,X_n)} = \frac{\left(\frac{1}{\sqrt{2\pi}\cdot\sigma}\right)^n \prod_{j=1}^n \exp(-(X_j - m_1)^2/(2\sigma^2))}{\left(\frac{1}{\sqrt{2\pi}\cdot\sigma}\right)^n \prod_{j=1}^n \exp(-(X_j - m_0)^2/(2\sigma^2))} =$$

$$= \exp\left(-\frac{\sum_{j=1}^n (X_j - m_1)^2}{2\sigma^2} + \frac{\sum_{j=1}^n (X_j - m_0)^2}{2\sigma^2} \right) =$$

$$= \exp\left(-\frac{\sum_{j=1}^n (X_j^2 - 2m_1X_j + m_1^2)}{2\sigma^2} + \frac{\sum_{j=1}^n (X_j^2 - 2m_0X_j + m_0^2)}{2\sigma^2} \right) =$$

$$= \quad \exp\left(\frac{2(m_1 - m_0)\sum_{j=1}^n X_j + (m_0^2 - m_1^2)n}{2\sigma^2} \right).$$

Ha $m_1 > m_0$: akkor utasítjuk el a nullhipotézist, ha $\frac{L_{n,1}}{L_{n,0}}$ nagyobb egy c kritikus értéknél, vagyis ha $\sum_{j=1}^n X_j$ nagyobb egy c' kritikus értéknél. Ezért a lesz a z-próba legerősebb próba: ez is ilyen alakú.

Normális eloszlás paramétereire vonatkozó próbák

Az alábbi próbák akkor használhatók, ha

- a megfigyelések függetlenek, és feltételezhetjük, hogy normális eloszlásúak
- a megfigyelések függetlenek, véges szórású eloszlásból származnak, és a minta mérete, azaz n "elég nagy", például $n \geq 100$

Normális eloszlás paramétereire vonatkozó próbák

Az alábbi próbák akkor használhatók, ha

- a megfigyelések függetlenek, és feltételezhetjük, hogy normális eloszlásúak
- a megfigyelések függetlenek, véges szórású eloszlásból származnak, és a minta mérete, azaz n "elég nagy", például $n \geq 100$

- z-**próba** (vagy u-próba): **várható értékre** vonatkozó hipotézis esetén, ha **a** σ **szórás ismert** – egymintás esetben legerősebb próba

Normális eloszlás paramétereire vonatkozó próbák

Az alábbi próbák akkor használhatók, ha

- a megfigyelések függetlenek, és feltételezhetjük, hogy normális eloszlásúak
- a megfigyelések függetlenek, véges szórású eloszlásból származnak, és a minta mérete, azaz n "elég nagy", például $n \geq 100$

- z-**próba** (vagy u-próba): **várható értékre** vonatkozó hipotézis esetén, ha **a σ szórás ismert** – egymintás esetben legerősebb próba
- t-**próba** (vagy Student-próba): **várható értékre** vonatkozó hipotézis esetén, ha **a σ szórás nem ismert** (csak az s_n^* tapasztalati szórás)
- F-**próba**: **szórásra** vonatkozó hipotézis esetén

Kapcsolat a konfidenciaintervallummal: egymintás próbánál akkor fogadjuk el a nullhipotézist α terjedelem mellett, ha a benne megadott érték (várható érték vagy szórás) az $1 - \alpha$ megbízhatósági szintű konfidenciaintervallumba esik.

Egymintás egyoldali z-próba (one-sample one-sided z test)

A próba a normális eloszlás várható értékére vonatkozik ismert szórás mellett. Torzítatlan, konzisztens, **legerősebb próba** egyoldali esetben (a Neyman–Pearson-lemma alapján bizonyítható).

- $X_1, X_2, \ldots, X_n \sim N(m, \sigma^2)$, ahol m ismeretlen paraméter, $\sigma > 0$ ismert.
- Próbastatisztika (eloszlása standard normális H_0 mellett, ezt beláttuk):

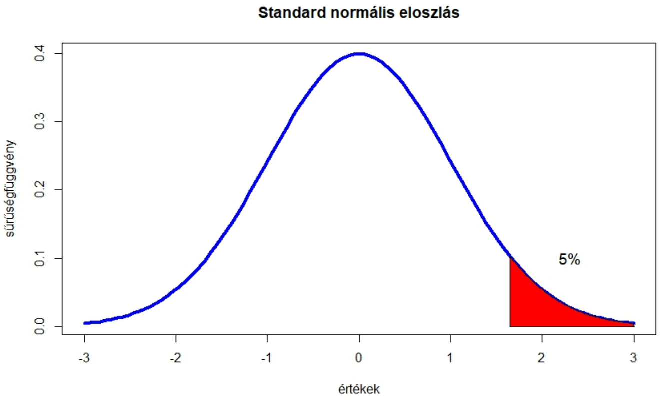
$$z = \frac{\overline{X} - m_0}{\sigma} \cdot \sqrt{n}.$$

- **Egyoldali ellenhipotézis** (one-sided): $H_0 : m \leq m_0$; $H_1 : m > m_0$.
- Ha $z > \Phi^{-1}(1-\alpha)$, akkor elvetjük a nullhipotézist, különben elfogadjuk.
- A p-érték ilyenkor $1 - \Phi(z)$.

$p < 0,05$: a várható érték szignifikánsan több m_0-nál.

$p \geq 0,05$: a várható érték nem több szignifikánsan m_0-nál.

Az egyoldali z-próba kritikus értéke



Az $\alpha = 0,05$ szignifikanciaszintű egyoldali z-próba kritikus értéke:
$\Phi^{-1}(1-\alpha) = \Phi^{-1}(0,95) = 1,645.$

Példa: egymintás egyoldali z-próba

Feltételezés: a testmagasság normális eloszlású.

- Az európai férfiak átlagos testmagassága $177,6$ cm.
- Megmértük 90 holland férfi testmagasságát, a magasságok átlaga 181,7 cm lett. A szórást 8,5 cm-nek feltételezve mondhatjuk-e, hogy a holland férfiak testmagassága szignifikánsan több az európai átlagnál?

Példa: egymintás egyoldali z-próba

Feltételezés: a testmagasság normális eloszlású.

- Az európai férfiak átlagos testmagassága $177,6$ cm.
- Megmértük 90 holland férfi testmagasságát, a magasságok átlaga 181,7 cm lett. A szórást 8,5 cm-nek feltételezve mondhatjuk-e, hogy a holland férfiak testmagassága szignifikánsan több az európai átlagnál?
- $H_0 : m \leq 177,6$; $H_1 : m > 177,6$.
- $$z = \frac{\overline{X} - m_0}{\sigma} \cdot \sqrt{n} = \frac{181,7 - 177,6}{8,5}\sqrt{90} = 4,57.$$
- $\alpha = 0,05$ szignifikanciaszint mellett $\Phi^{-1}(1-\alpha) = 1,645$, így $z > \Phi^{-1}(1-\alpha)$.
  p-érték: $1 - \Phi(4,57) < 0,0001 < 0,05$.
- Elutasítjuk a nullhipotézist. Az adatok alapján a holland férfiak testmagasságának várható értéke szignifikánsan több 177,6 cm-nél, vagyis az európai átlagnál.

Egymintás kétoldali z-próba

A próba a normális eloszlás várható értékére vonatkozik ismert szórás mellett. Nem legerősebb (nincs legerősebb próba ebben a feladatban).

- $X_1, X_2, \ldots, X_n \sim N(m, \sigma^2)$, ahol m ismeretlen paraméter, $\sigma > 0$ ismert.
- Próbastatisztika (eloszlása standard normális H_0 mellett):

$$z = \frac{\overline{X} - m_0}{\sigma} \cdot \sqrt{n}.$$

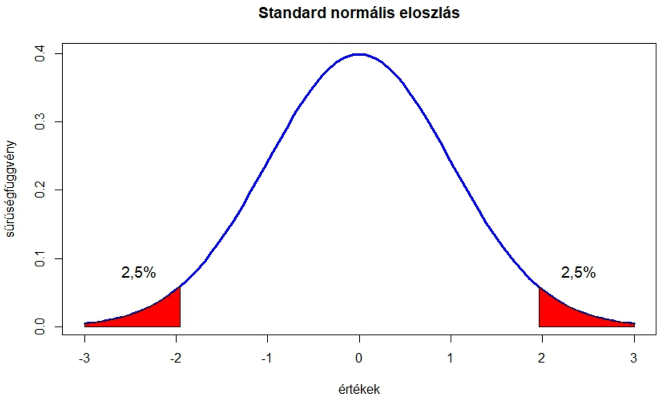
- **Kétoldali ellenhipotézis** (two-sided): $H_0 : m = m_0$; $H_1 : m \neq m_0$.
- Ha $|z| > \Phi^{-1}(1 - \alpha/2)$, akkor elvetjük a nullhipotézist, különben elfogadjuk.
- A p-érték ilyenkor $2 - 2\Phi(|z|)$.

Φ a standard normális eloszlásfüggvény: $\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$.

$p < 0,05$: a várható érték szignifikánsan eltér m_0-tól.

$p \geq 0,05$: nincs szignifikáns eltérés m_0-tól.

A kétoldali z-próba kritikus értéke



Az $\alpha = 0,05$ szignifikanciaszintű kétoldali z-próba kritikus értéke: $\Phi^{-1}(1 - \alpha/2) = \Phi^{-1}(0,975) = 1,96$.

Példa: egymintás z-próba

Egy gyárban a minőségellenőrzésnél olyan mérleget használnak, melynél egy m tömegű tárgyat mérve a mérési eredmények független normális eloszlású valószínűségi változók m várható értékkel és $\sigma = 3$ gramm szórással.

- A termékkatalógus szerint egy adott típusú kalapács fejének 364 g tömegűnek kell lennie.
- A fenti mérlegen megmérték 20 kalapács fejének tömegét. Az átlag 367,2 gramm lett. Ez alapján állítható-e, hogy a kalapácsok fejének tömege szignifikánsan eltér az előírt 364 grammtól?

Példa: egymintás z-próba

Egy gyárban a minőségellenőrzésnél olyan mérleget használnak, melynél egy m tömegű tárgyat mérve a mérési eredmények független normális eloszlású valószínűségi változók m várható értékkel és $\sigma = 3$ gramm szórással.

- A termékkatalógus szerint egy adott típusú kalapács fejének 364 g tömegűnek kell lennie.
- A fenti mérlegen megmérték 20 kalapács fejének tömegét. Az átlag 367,2 gramm lett. Ez alapján állítható-e, hogy a kalapácsok fejének tömege szignifikánsan eltér az előírt 364 grammtól?
- $H_0 : m = 364$; $H_1 : m \neq 364$.
- $$z = \frac{\overline{X} - m_0}{\sigma} \cdot \sqrt{n} = \frac{367,2 - 364}{3}\sqrt{20} = 4,77.$$
- $\alpha = 0,05$ szignifikanciaszint mellett $\Phi^{-1}(1-\alpha/2) = 1,96$. $p = 1,84 \cdot 10^{-6} < 0,05$.
- $|z| < \Phi^{-1}(1-\alpha/2)$, elutasítjuk a nullhipotézist. A kalapácsok fejének tömegének várható értéke a minta alapján szignifikánsan eltér az előírt 364 grammtól.

Fisher–Bartlett-tétel

Valójában az eloszlás valódi szórása a legtöbb esetben nem ismert. A σ szórást az s_n^* korrigált tapasztalati szórással helyettesítjük. Kérdés, hogyan változik így az eloszlás.

Tétel (Fisher–Bartlett)

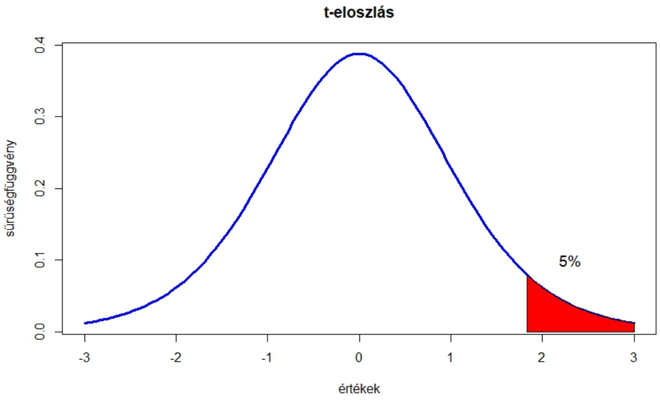
Tegyük fel, hogy $X_1, X_2, \ldots, X_n$ független m várható értékű, σ szórású, **normális eloszlású** *valószínűségi változók. Ekkor*

1. $\overline{X} \sim N\left(m, \frac{\sigma^2}{n}\right)$;
2. $\overline{X}$ *és* s_n^* *függetlenek;*
3. $(n-1)s_n^{*2}/\sigma^2$ *eloszlása* $n-1$ *szabadsági fokú* χ^2*-eloszlás;*
4. $(\overline{X}-m)\sqrt{n}/s_n^*$ *eloszlása* $n-1$ *szabadsági fokú t-eloszlás.*

Itt

$$s_n^* = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(X_j-\overline{X})^2} = \sqrt{\frac{n}{n-1}\left(\left(\frac{1}{n}\sum_{j=1}^{n}X_j^2\right)-\overline{X}^2\right)}.$$

t-eloszlás egyoldali kritikus értékei



Az $f = 9$ szabadsági fokú $\alpha = 0,05$ szignfikianciaszintű egyoldali t-próba kritikus értéke: $\overline{t}_{9,0,05} = 1,83$.

Egymintás egyoldali t-próba (one-sample one-sided t-test)

- **A normális eloszlás várható értékére, ismeretlen szórás esetén – legerősebb próba.**
- $X_1, X_2, \ldots, X_n \sim N(m, \sigma^2)$, ahol m, σ ismeretlen paraméterek.
- Próbastatisztika, aminek eloszlása t-eloszlás H_0 mellett a Fisher–Bartlett-tétel szerint:
$$t = \frac{\overline{X} - m_0}{s_n^*} \cdot \sqrt{n}.$$
- **Egyoldali ellenhipotézis** (one-sided): $H_0 : m \leq m_0$; $\quad H_1 : m > m_0$.
- Ha $t > \bar{t}_{n-1,\alpha}$, azaz $p < \alpha$, elutasítjuk a nullhipotézist; ilyenkor a várható érték szignifikánsan több m_0-nál.
- Ha $t \leq \bar{t}_{n-1,\alpha}$, azaz $p \geq \alpha$, elfogadjuk a nullhipotézist, a várható érték nem több szignifikánsan m_0-nál az adatok alapján.
- A kritikus érték: $\bar{t}_{n-1,\alpha}$ az $f = n-1$ szabadsági fokú (degree of freedom) t-eloszlás $1-\alpha$-kvantilise, vagyis az $f = n-1$ szabadsági fokú egyoldali t-próba kritikus értéke α szignifikanciaszint (level of significance) mellett.

Példa: egymintás egyoldali t-próba

Egy adott helyen vett tíz mintából megmértük az ivóvíz keménységét. Az alábbi eredmények adódtak (mg/l CaO):

351	370	352	340	362	363	366	355	374	347

Állíthatjuk-e az adatok alapján, hogy az ivóvíz keménységének várható értéke szignifikánsan meghaladja a 350 mg/l egészségügyi határértéket?

Példa: egymintás egyoldali t-próba

Egy adott helyen vett tíz mintából megmértük az ivóvíz keménységét. Az alábbi eredmények adódtak (mg/l CaO):

351	370	352	340	362	363	366	355	374	347

Állíthatjuk-e az adatok alapján, hogy az ivóvíz keménységének várható értéke szignifikánsan meghaladja a 350 mg/l egészségügyi határértéket?

$$n = 10; \qquad \overline{X} = 358; \qquad s_n^* = 10,77$$

Feltételezzük, hogy a mérési eredmények normális eloszlásúak, **az egymintás egyoldali t-próbát** alkalmazzuk: $H_0 : m \leq 350$; $H_1 : m > 350$.

$$t = \frac{\overline{X} - m_0}{s_n^*} \cdot \sqrt{n} = \frac{358 - 350}{10,77}\sqrt{10} = 2,35.$$

Az $f = n - 1 = 9$ szabadsági fokú egyoldali t-próba kritikus értéke $\alpha = 0,05$ szignifikanciaszint mellett $\overline{t}_{9;0,05} = 1,833$.

Példa: egymintás egyoldali t-próba

Egy adott helyen vett tíz mintából megmértük az ivóvíz keménységét. Az alábbi eredmények adódtak (mg/l CaO):

351	370	352	340	362	363	366	355	374	347

Állíthatjuk-e az adatok alapján, hogy az ivóvíz keménységének várható értéke szignifikánsan meghaladja a 350 mg/l egészségügyi határértéket?

$$n = 10; \qquad \overline{X} = 358; \qquad s_n^* = 10,77$$

Feltételezzük, hogy a mérési eredmények normális eloszlásúak, **az egymintás egyoldali t-próbát** alkalmazzuk: $H_0 : m \leq 350$; $H_1 : m > 350$.

$$t = \frac{\overline{X} - m_0}{s_n^*} \cdot \sqrt{n} = \frac{358 - 350}{10,77}\sqrt{10} = 2,35.$$

Az $f = n - 1 = 9$ szabadsági fokú egyoldali t-próba kritikus értéke $\alpha = 0,05$ szignifikanciaszint mellett $\bar{t}_{9;0,05} = 1,833$.

Mivel $t > \bar{t}_{9;0,05}$, **elutasítjuk a nullhipotézist**, a vízkeménység szignifikánsan meghaladja 350 mg/l határértéket. A p-érték: $p = 0,0217 < 0,05.$

Hipotézisvizsgálat: példa az R szoftverben

Tíz mintából mértük meg a víz keménységét.

Nullhipotézis (null hypothesis, H_0): $m \leq 350$

Ellenhipotézis (alternative hypothesis, H_1): $m > 350$

```
> viz<-c(348, 367, 349, 337, 359, 360, 363, 352, 371, 344)

> t.test(viz, mu=350, alternative="greater")
```

Hipotézisvizsgálat: példa az R szoftverben

Tíz mintából mértük meg a víz keménységét.

Nullhipotézis (null hypothesis, H_0): $m \leq 350$

Ellenhipotézis (alternative hypothesis, H_1): $m > 350$

> viz<-c(348, 367, 349, 337, 359, 360, 363, 352, 371, 344)

> t.test(viz, mu=350, alternative="greater")

```
One Sample t-test

data:  viz

t = 2.3489, df = 9, p-value = 0.02169

alternative hypothesis:  true mean is greater than 350

95 percent confidence interval:  351.7566 Inf

mean of x :  358
```

Most $p = 0.02169 < 0,05 = \alpha$, elutasítjuk a nullhipotézist.

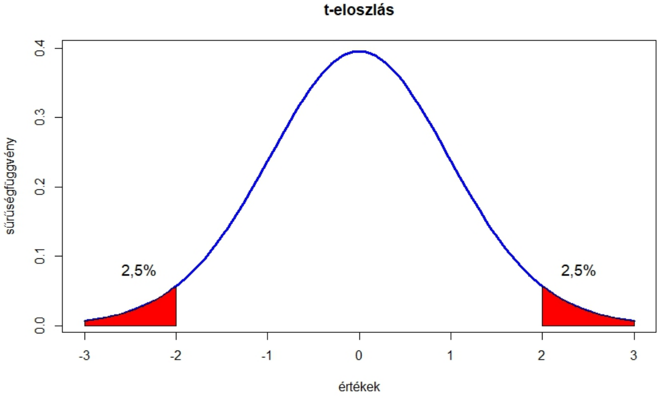
Egymintás kétoldali t-próba (one-sample two-sided t-test)

- **A normális eloszlás várható értékére, ismeretlen szórás esetén.** Nem legerősebb (nincs ilyen).
- $X_1, X_2, \ldots, X_n \sim N(m, \sigma^2)$, ahol m, σ ismeretlen paraméterek.
- Próbastatisztika (eloszlása t-eloszlás/Student-eloszlás H_0 mellett):

$$t = \frac{\overline{X} - m_0}{s_n^*} \cdot \sqrt{n}.$$

- **Kétoldali ellenhipotézis** (two-sided): $H_0 : m = m_0$; $\qquad H_1 : m \neq m_0$.
- Ha $|t| > t_{n-1,\alpha}$, azaz $p < \alpha$, akkor elutasítjuk a nullhipotézist, a várható érték szignifikánsan eltér m_0-tól.
- Ha $|t| \leq t_{n-1,\alpha}$, azaz $p \geq \alpha$, akkor elfogadjuk H_0-t, a várható érték nem tér el szignifikánsan m_0-tól.
- A kritikus érték: $t_{n-1,\alpha}$ az $f = n - 1$ szabadsági fokú (degree of freedom) t-eloszlás $1 - \alpha/2$-kvantilise, vagyis az $f = n - 1$ szabadsági fokú (degree of freedom) kétoldali t-próba kritikus értéke α szignifikanciaszint mellett.

Kétoldali t-próba kritikus értékei



Az $f = 29$ szabadsági fokú $\alpha = 0,05$ szignfikianciaszintű kétoldali t-próba kritikus értéke: $t_{29;0,05} = 2,04$.

Példa: Egymintás kétoldali t-próba

Egy gyógyszer hatóanyagtartalma a csomagolás szerint 10 mg. Harminc tabletta hatóanyag-tartalmát megmérve a mérések átlaga $9,8$, korrigált tapasztalati szórása $0,62$ lett. A szignifikanciaszintet $\alpha = 0,05$-nek választva az adatok alapján szignifikánsan eltér-e a hatóanyag-tartalom várható értéke a 10 mg-tól?

Példa: Egymintás kétoldali t-próba

Egy gyógyszer hatóanyagtartalma a csomagolás szerint 10 mg. Harminc tabletta hatóanyag-tartalmát megmérve a mérések átlaga $9,8$, korrigált tapasztalati szórása $0,62$ lett. A szignifikanciaszintet $\alpha = 0,05$-nek választva az adatok alapján szignifikánsan eltér-e a hatóanyag-tartalom várható értéke a 10 mg-tól?

$$n = 30; \qquad \overline{X} = 9,8; \qquad s_n^* = 0,62$$

Egymintás kétoldali t-próbát végezhetünk, normális eloszlást feltételezve.

$$H_0 : m = 10; \qquad H_1 : m \neq 10; \qquad \alpha = 0,05; \qquad f = n - 1 = 29.$$

$$t = \frac{\overline{X} - m_0}{s_n^*} \cdot \sqrt{n} = \frac{9,8 - 10}{0,62} \cdot \sqrt{30} = -1,77.$$

A kritikus érték: $t_{29,0,975} = 2,045 \Rightarrow |t| = 1,77 \leq 2,045$, nincs szignifikáns eltérés.
p-érték: $p = 0,0867 \geq 0,05$.

Kétmintás, egyoldali, párosítatlan Student-féle t-próba

A **várható érték összehasonlítására** azonos szórás esetén (two-sample one-sided unpaired Student t-test).

$X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ **független normális eloszlású azonos szórású** valószínűségi változók: $X_i \sim N(m_1, \sigma^2)$, $Y_i \sim N(m_2, \sigma^2)$, ahol m_1, m_2, σ ismeretlen paraméterek.

Egyoldali ellenhipotézis: $H_0 : m_1 \leq m_2$; $\quad H_1 : m_1 > m_2$.

Kétmintás, egyoldali, párosítatlan Student-féle t-próba

A **várható érték összehasonlítására** azonos szórás esetén (two-sample one-sided unpaired Student t-test).

$X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ **független normális eloszlású azonos szórású** valószínűségi változók: $X_i \sim N(m_1, \sigma^2)$, $Y_i \sim N(m_2, \sigma^2)$, ahol m_1, m_2, σ ismeretlen paraméterek.

Egyoldali ellenhipotézis: $H_0 : m_1 \leq m_2$; $H_1 : m_1 > m_2$.

Próbastatisztika (eloszlása t-eloszlás $m_1 = m_2$ mellett):

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{(n_1 - 1)s_{n_1}^{*2}(X) + (n_2 - 1)s_{n_2}^{*2}(Y)}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}.$$

Ha $t > \overline{t}_{n_1+n_2-2,\alpha}$, akkor elvetjük a nullhipotézist, különben elfogadjuk. A $\overline{t}_{n_1+n_2-2,\alpha}$ kritikus érték az $f = n_1 + n_2 - 2$ szabadsági fokú **egyoldali** t-próba kritikus értéke α szignifikanciaszint mellett (a megfelelő eloszlás $1 - \alpha$-kvantilise).

Ha $p < \alpha$: elvetjük H_0-t, az első várható érték szignifikánsan nagyobb a másodiknál.

Kétmintás t-próba: példa

Két különböző márkájú vaj tömegét mértük meg, az X típusúét tízszer, az Y típusúét nyolcszor. Az átlagok és korrigált tapasztalati szórások (kg-ban):

$$\overline{X} = 0,217, \qquad s_n^*(X) = 0,027, \qquad \overline{Y} = 0,203, \qquad s_n^*(Y) = 0,03.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy az X típusú vajak tömege szignifikánsan több az Y típusú vajénál? Feltételezzük, hogy a mérések **szórása azonos**.

Kétmintás t-próba: példa

Két különböző márkájú vaj tömegét mértük meg, az X típusúét tízszer, az Y típusúét nyolcszor. Az átlagok és korrigált tapasztalati szórások (kg-ban):

$$\overline{X} = 0,217, \qquad s_n^*(X) = 0,027, \qquad \overline{Y} = 0,203, \qquad s_n^*(Y) = 0,03.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy az X típusú vajak tömege szignifikánsan több az Y típusú vajénál? Feltételezzük, hogy a mérések **szórása azonos**.

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{(n_1 - 1)s_{n_1}^{*2}(X) + (n_2 - 1)s_{n_2}^{*2}(Y)}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}.$$

Behelyettesítve:

$$t = \frac{0,217 - 0,203}{\sqrt{9 \cdot 0,027^2 + 7 \cdot 0,03^2}} \cdot \sqrt{\frac{80 \cdot 16}{18}} = 1,04.$$

Kétmintás t-próba: példa

Két különböző márkájú vaj tömegét mértük meg, az X típusúét tízszer, az Y típusúét nyolcszor. Az Az átlagok és korrigált tapasztalati szórások (kg-ban):

$$\overline{X} = 0,217, \qquad s_n^*(X) = 0,027, \qquad \overline{Y} = 0,203, \qquad s_n^*(Y) = 0,03.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy az X típusú vajak tömege szignifikánsan több az Y típusú vajénál? Feltételezzük, hogy a mérések **szórása azonos**.

$H_0 : m_1 \leq m_2$, $H_1 : m_1 > m_2$

A próbastatisztika értéke: $t = 1,04$.

Az $f = n_1 + n_2 - 2 = 10 + 8 - 2 = 16$ szabadsági fokú egyoldali t-próba kritikus értéke $\alpha = 0,05$ szignifikanciaszint mellett: $\overline{t}_{16,0,05} = 1,746$.

Kétmintás t-próba: példa

Két különböző márkájú vaj tömegét mértük meg, az X típusúét tízszer, az Y típusúét nyolcszor. Az Az átlagok és korrigált tapasztalati szórások (kg-ban):

$$\overline{X} = 0,217, \qquad s_n^*(X) = 0,027, \qquad \overline{Y} = 0,203, \qquad s_n^*(Y) = 0,03.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy az X típusú vajak tömege szignifikánsan több az Y típusú vajénál? Feltételezzük, hogy a mérések **szórása azonos**.

$H_0 : m_1 \leq m_2$, $H_1 : m_1 > m_2$

A próbastatisztika értéke: $t = 1,04$.

Az $f = n_1 + n_2 - 2 = 10 + 8 - 2 = 16$ szabadsági fokú egyoldali t-próba kritikus értéke $\alpha = 0,05$ szignifikanciaszint mellett: $\overline{t}_{16,0,05} = 1,746$.

Itt $t = 1,04 < 1,746 = \overline{t}_{16,0,05}$, ezért **elfogadjuk H_0-t**. Az X típusú vaj tömege **nem haladja meg szignifikánsan** az Y típusúét.

p-érték: 0,149>0,05

Kétmintás t-próba: példa

Kétféle joghurt cukortartalmát szeretnénk összehasonlítani. Az elsőből $n_1 = 20$, a másodikból $n_2 = 12$ dobozban mértük meg a cukortartalmat (grammban).

Az átlagok és korrigált tapasztalati szórások grammban számolva ($X_1, \ldots, X_{20}$ az első minta, $Y_1, \ldots, Y_{12}$ a második):

$$\overline{X} = 18,4, \qquad s_n^*(X) = 1,2, \qquad \overline{Y} = 19,9, \qquad s_n^*(Y) = 1,3.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy a kétféle joghurtban szignifikánsan eltérő a cukortartalom?

Kétmintás, kétoldali, párosítatlan Student-féle t-próba

A **várható érték összehasonlítására** azonos szórás esetén (two-sample two-sided unpaired Student t-test).

$X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ **független normális eloszlású azonos szórású** valószínűségi változók: $X_i \sim N(m_1, \sigma^2)$, $Y_i \sim N(m_2, \sigma^2)$, ahol m_1, m_2, σ ismeretlen paraméterek.

Kétoldali ellenhipotézis: $H_0 : m_1 = m_2$; $\quad H_1 : m_1 \neq m_2$.

Kétmintás, kétoldali, párosítatlan Student-féle t-próba

A **várható érték összehasonlítására** azonos szórás esetén (two-sample two-sided unpaired Student t-test).

$X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ **független normális eloszlású azonos szórású** valószínűségi változók: $X_i \sim N(m_1, \sigma^2)$, $Y_i \sim N(m_2, \sigma^2)$, ahol m_1, m_2, σ ismeretlen paraméterek.

Kétoldali ellenhipotézis: $H_0 : m_1 = m_2$; $\quad H_1 : m_1 \neq m_2$.

Próbastatisztika (eloszlása t-eloszlás H_0 mellett):

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{(n_1 - 1)s_{n_1}^{*2}(X) + (n_2 - 1)s_{n_2}^{*2}(Y)}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}.$$

Ha $|t| > t_{n_1+n_2-2,1-\alpha}$, akkor elvetjük a nullhipotézist, különben elfogadjuk. A $t_{n_1+n_2-2,1-\alpha}$ kritikus érték az $f = n_1 + n_2 - 2$ szabadsági fokú **kétoldali** t-próba kritikus értéke α szignifikanciaszint mellett (a megfelelő eloszlás $1-\alpha/2$-kvantilise).

Ha $p < \alpha$: elvetjük H_0-t, az várható értékek szignifikánsan eltérnek egymástól.

Kétmintás t-próba: példa

Kétféle joghurt cukortartalmát szeretnénk összehasonlítani. Az elsőből $n_1 = 20$, a másodikból $n_2 = 12$ dobozban mértük meg a cukortartalmat (grammban).

Az átlagok és korrigált tapasztalati szórások grammban számolva ($X_1, \ldots, X_{20}$ az első minta, $Y_1, \ldots, Y_{12}$ a második):

$$\overline{X} = 18,4, \qquad s_n^*(X) = 1,2, \qquad \overline{Y} = 19,9, \qquad s_n^*(Y) = 1,3.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy a kétféle joghurtban szignifikánsan eltérő a cukortartalom? Feltételezzük, hogy a minták **függetlenek, normális eloszlásúak**, **azonos szórásúak**.

Kétmintás t-próba: példa

Kétféle joghurt cukortartalmát szeretnénk összehasonlítani. Az elsőből $n_1 = 20$, a másodikból $n_2 = 12$ dobozban mértük meg a cukortartalmat (grammban).

Az átlagok és korrigált tapasztalati szórások grammban számolva ($X_1, \ldots, X_{20}$ az első minta, $Y_1, \ldots, Y_{12}$ a második):

$$\overline{X} = 18,4, \qquad s_n^*(X) = 1,2, \qquad \overline{Y} = 19,9, \qquad s_n^*(Y) = 1,3.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy a kétféle joghurtban szignifikánsan eltérő a cukortartalom? Feltételezzük, hogy a minták **függetlenek, normális eloszlásúak**, **azonos szórásúak**.

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{(n_1 - 1)s_{n_1}^{*2}(X) + (n_2 - 1)s_{n_2}^{*2}(Y)}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}.$$

Behelyettesítve:

$$t = \frac{18,4 - 19,9}{\sqrt{19 \cdot 1,2^2 + 11 \cdot 1,3^2}} \cdot \sqrt{\frac{20 \cdot 12 \cdot 30}{32}} = -3,3.$$

Kétmintás t-próba: példa

Kétféle joghurt cukortartalmát szeretnénk összehasonlítani. Az elsőből $n_1 = 20$, a másodikból $n_2 = 12$ dobozban mértük meg a cukortartalmat (grammban).

Az átlagok és korrigált tapasztalati szórások grammban számolva ($X_1, \ldots, X_{20}$ az első minta, $Y_1, \ldots, Y_{12}$ a második):

$$\overline{X} = 18,4, \qquad s_n^*(X) = 1,2, \qquad \overline{Y} = 19,9, \qquad s_n^*(Y) = 1,3.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy a kétféle joghurtban szignifikánsan eltérő a cukortartalom? Feltételezzük, hogy a minták **függetlenek, normális eloszlásúak, azonos szórásúak**.

$H_0 : m_1 = m_2$, $H_1 : m_1 \neq m_2$

A próbastatisztika értéke: $t = -3,3$.

Az $f = n_1 + n_2 - 2 = 20 + 12 - 2 = 30$ szabadsági fokú **kétoldali** t-próba kritikus értéke $\alpha = 0,05$ szignifikanciaszint mellett: $t_{16,0,05} = 2,042$.

Kétmintás t-próba: példa

Kétféle joghurt cukortartalmát szeretnénk összehasonlítani. Az elsőből $n_1 = 20$, a másodikból $n_2 = 12$ dobozban mértük meg a cukortartalmat (grammban).

Az átlagok és korrigált tapasztalati szórások grammban számolva ($X_1, \ldots, X_{20}$ az első minta, $Y_1, \ldots, Y_{12}$ a második):

$$\overline{X} = 18,4, \qquad s_n^*(X) = 1,2, \qquad \overline{Y} = 19,9, \qquad s_n^*(Y) = 1,3.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy a kétféle joghurtban szignifikánsan eltérő a cukortartalom? Feltételezzük, hogy a minták **függetlenek, normális eloszlásúak, azonos szórásúak**.

$H_0 : m_1 = m_2$, $H_1 : m_1 \neq m_2$

A próbastatisztika értéke: $t = -3,3$.

Az $f = n_1 + n_2 - 2 = 20 + 12 - 2 = 30$ szabadsági fokú **kétoldali** t-próba kritikus értéke $\alpha = 0,05$ szignifikanciaszint mellett: $t_{16,0,05} = 2,042$.

Itt $|t| = 3,3 > 2,042 = t_{30,0,05}$, ezért **elutasítjuk** H_0**-t**. A kétféle joghurt cukortartalma **szignifikánsan különböző**

Kétmintás t-próba: példa

Kétféle joghurt cukortartalmát szeretnénk összehasonlítani. Az elsőből $n_1 = 20$, a másodikból $n_2 = 12$ dobozban mértük meg a cukortartalmat (grammban).

Az átlagok és korrigált tapasztalati szórások grammban számolva ($X_1, \ldots, X_{20}$ az első minta, $Y_1, \ldots, Y_{12}$ a második):

$$\overline{X} = 18,4, \qquad s_n^*(X) = 1,2, \qquad \overline{Y} = 19,9, \qquad s_n^*(Y) = 1,3.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy a kétféle joghurtban szignifikánsan eltérő a cukortartalom? Feltételezzük, hogy a minták **függetlenek, normális eloszlásúak, azonos szórásúak**.

$H_0 : m_1 = m_2$, $H_1 : m_1 \neq m_2$

A próbastatisztika értéke: $t = -3,3$.

Az $f = n_1 + n_2 - 2 = 20 + 12 - 2 = 30$ szabadsági fokú **kétoldali** t-próba kritikus értéke $\alpha = 0,05$ szignifikanciaszint mellett: $t_{16,0,05} = 2,042$.

Itt $|t| = 3,3 > 2,042 = t_{30,0,05}$, ezért **elutasítjuk H_0-t**. A kétféle joghurt cukortartalma **szignifikánsan különböző** – ha a szórások azonosak, és a próba alkalmazható (ezt eddig feltettük).

F-próba

Független normális eloszlású minták **szórásának** összehasonlítására.

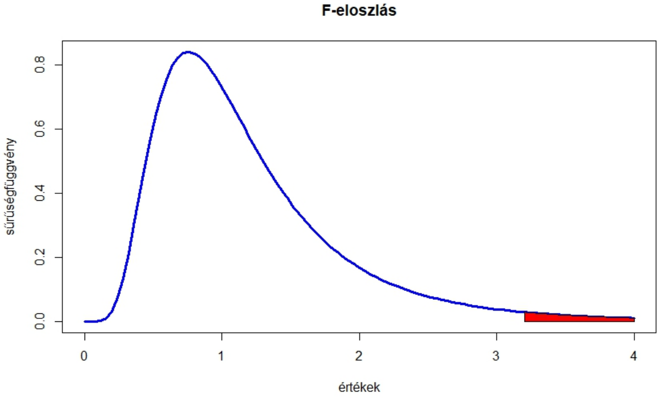
- Legyenek most $X_1, X_2, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ független normális eloszlású valószínűségi változók, ahol $X_i \sim N(m_1, \sigma_1^2)$, $Y_i \sim N(m_2, \sigma_2^2)$. Itt $m_1, m_2, \sigma_1, \sigma_2$ ismeretlen paraméterek.
- Kétoldali ellenhipotézis: $H_0 : \sigma_1 = \sigma_2$; $H_1 : \sigma_1 \neq \sigma_2$.
- Próbastatisztika (eloszlása F-eloszlás H_0 mellett):

$$F = \frac{s_{n_1}^{*2}}{s_{n_2}^{*2}}.$$

- Ha $F > F_{n_1-1, n_2-1}$ vagy $1/F > F_{n_2-1, n_1-1}$, akkor elvetjük a nullhipotézist, különben elfogadjuk, ahol F_{f_1, f_2} az f_1, f_2 szabadsági fokú az F-eloszlás $1-\alpha/2$-kvantilise.

$p < 0,05$: a szórások szignifikánsan eltérnek.

Az F-próba kritikus értéke



Az F-próba kritikus értéke: $F_{19,11} = 3,24$, ez az eloszlás $1-\alpha/2 = 0,975$-kvantilise

Kétmintás F-próba: példa

Kétféle joghurt cukortartalmát szeretnénk összehasonlítani. Az elsőből $n_1 = 20$, a másodikból $n_2 = 12$ dobozban mértük meg a cukortartalmat (grammban). Az átlagok és korrigált tapasztalati szórások ($X_1, \dots, X_{20}$ az első minta, $Y_1, \dots, Y_{12}$ a második):

$$\overline{X} = 18,4, \qquad s_n^*(X) = 1,2, \qquad \overline{Y} = 19,9, \qquad s_n^*(Y) = 1,3.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy a kétféle joghurtban szignifikánsan eltérő a cukortartalom szórása? Feltételezzük, hogy a minták **függetlenek, normális eloszlásúak**.

Kétmintás F-próba: példa

Kétféle joghurt cukortartalmát szeretnénk összehasonlítani. Az elsőből $n_1 = 20$, a másodikból $n_2 = 12$ dobozban mértük meg a cukortartalmat (grammban). Az átlagok és korrigált tapasztalati szórások ($X_1, \ldots, X_{20}$ az első minta, $Y_1, \ldots, Y_{12}$ a második):

$$\overline{X} = 18,4, \qquad s_n^*(X) = 1,2, \qquad \overline{Y} = 19,9, \qquad s_n^*(Y) = 1,3.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy a kétféle joghurtban szignifikánsan eltérő a cukortartalom szórása? Feltételezzük, hogy a minták **függetlenek, normális eloszlásúak**.

$H_0 : \sigma_1 = \sigma_2$, $H_1 : \sigma_1 \neq \sigma_2$

A próbastatisztika értéke: $F = \frac{s_{n_1}^{*2}}{s_{n_2}^{*2}} = \frac{1,2^2}{1,3^2} = 0,85$, és $\frac{1}{F} = \frac{1,3^2}{1,2^2} = 1,17$.

Az $(f_1, f_2) = (n_1 - 1, n_2 - 1) = (19, 11)$ szabadsági fokú F-próba kritikus értéke $\alpha = 0,05$ esetén: $3,24$, míg az $(f_2, f_1) = (n_2 - 1, n_1 - 1) = (11, 19)$ szabadsági fok esetén $2,76$.

Kétmintás F-próba: példa

Kétféle joghurt cukortartalmát szeretnénk összehasonlítani. Az elsőből $n_1 = 20$, a másodikból $n_2 = 12$ dobozban mértük meg a cukortartalmat (grammban). Az átlagok és korrigált tapasztalati szórások ($X_1, \ldots, X_{20}$ az első minta, $Y_1, \ldots, Y_{12}$ a második):

$$\overline{X} = 18,4, \qquad s_n^*(X) = 1,2, \qquad \overline{Y} = 19,9, \qquad s_n^*(Y) = 1,3.$$

Állíthatjuk-e $\alpha = 0,05$ szignifikanciaszint mellett, hogy a kétféle joghurtban szignifikánsan eltérő a cukortartalom szórása? Feltételezzük, hogy a minták **függetlenek, normális eloszlásúak**.

$H_0 : \sigma_1 = \sigma_2$, $H_1 : \sigma_1 \neq \sigma_2$

A próbastatisztika értéke: $F = \frac{s_{n_1}^{*2}}{s_{n_2}^{*2}} = \frac{1,2^2}{1,3^2} = 0,85$, és $\frac{1}{F} = \frac{1,3^2}{1,2^2} = 1,17$.

Az $(f_1, f_2) = (n_1 - 1, n_2 - 1) = (19, 11)$ szabadsági fokú F-próba kritikus értéke $\alpha = 0,05$ esetén: $3,24$, míg az $(f_2, f_1) = (n_2 - 1, n_1 - 1) = (11, 19)$ szabadsági fok esetén $2,76$.

Mivel $F < 3,24$ és $1/F < 2,76$, **elfogadjuk a nullhipotézist**, a szórások nem térnek el szignifikánsan, és **a kétmintás *t*-próba valóban alkalmazható**.

Normális eloszlásra vonatkozó kétmintás próbák

Az alábbiakat kell ellenőrizni kétmintás próbáknál:

- A minta **normális eloszlású**, vagy a mintaelemszám elég nagy és a szórás feltehetően véges (a centrális határeloszlástétel alapján az átlag közel normális eloszlású).

Normális eloszlásra vonatkozó kétmintás próbák

Az alábbiakat kell ellenőrizni kétmintás próbáknál:

- A minta **normális eloszlású**, vagy a mintaelemszám elég nagy és a szórás feltehetően véges (a centrális határeloszlástétel alapján az átlag közel normális eloszlású).
- Kétmintás esetben: a **két minta egymástól független** (”unpaired” eset). Ha a két minta természetes módon párosítható, **párosított** (”paired”) próba alkalmazható. Példa: megmérjük húsz ember vérnyomását egy adott napon reggel és este. Igaz-e, hogy a reggeli érték jelentősen eltér az estitől?
- Ha a **szórásokról feltételezhetjük, hogy megegyeznek**: a Student-féle t-próba alkalmazható.
- Ha a **szórásokról nem tételezhetjük fel, hogy megegyeznek**: a Welch-féle t-próba alkalmazható.

Példa: párosított t-próba

1991 és 2010 között feljegyezték az éves csapadékösszeget Budapesten, illetve Szegeden. Az átlag Budapesten 533 mm, a korrigált tapasztalati szórás 139, Szegeden az átlag 540 mm, a korrigált tapasztalati szórás 143 lett (forrás: OMSZ). Állíthatjuk-e, hogy Szegeden szignifikánsan nagyobb a csapadékmennyiség várható értéke?

év	1991	1992	1993	1994	1995	...	átlag	s_n^*
Budapest	594	364	505	481	575	...	**533**	139
Szeged	617	457	408	399	562	...	**540**	143

Példa: párosított t-próba

1991 és 2010 között feljegyezték az éves csapadékösszeget Budapesten, illetve Szegeden. Az átlag Budapesten 533 mm, a korrigált tapasztalati szórás 139, Szegeden az átlag 540 mm, a korrigált tapasztalati szórás 143 lett (forrás: OMSZ). Állíthatjuk-e, hogy Szegeden szignifikánsan nagyobb a csapadékmennyiség várható értéke?

év	1991	1992	1993	1994	1995	...	átlag	s_n^*
Budapest	594	364	505	481	575	...	**533**	139
Szeged	617	457	408	399	562	...	**540**	143

A két adatsor **nem független**, mert egy éven belül a két város időjárása nem független (az egyes minták sem teljesen függetlenek, és nem biztos, hogy normális eloszlásúak). Ezért **párosított** (paired) t-próba alkalmazható, egyoldali nullhipotézissel.

$H_0 : m_1 \geq m_2$, $H_1 : m_1 < m_2$, ahol m_1 a budapesti, m_2 a szegedi csapadékmennyiség várható értéke.

Példa: párosított t-próba

1991 és 2010 között feljegyezték az éves csapadékösszeget Budapesten, illetve Szegeden. Az átlag Budapesten 533 mm, a korrigált tapasztalati szórás 139, Szegeden az átlag 540 mm, a korrigált tapasztalati szórás 143 lett (forrás: OMSZ). Állíthatjuk-e, hogy Szegeden szignifikánsan nagyobb a csapadékmennyiség várható értéke?

év	1991	1992	1993	1994	1995	...	átlag	s_n^*
Budapest	594	364	505	481	575	...	**533**	139
Szeged	617	457	408	399	562	...	**540**	143

A két adatsor **nem független**, mert egy éven belül a két város időjárása nem független (az egyes minták sem teljesen függetlenek, és nem biztos, hogy normális eloszlásúak). Ezért **párosított** (paired) t-próba alkalmazható, egyoldali nullhipotézissel.

$H_0 : m_1 \geq m_2$, $H_1 : m_1 < m_2$, ahol m_1 a budapesti, m_2 a szegedi csapadékmennyiség várható értéke.

A próbát elvégezve a p-értékre $0,366$ adódott.

Példa: párosított t-próba

1991 és 2010 között feljegyezték az éves csapadékösszeget Budapesten, illetve Szegeden. Az átlag Budapesten 533 mm, a korrigált tapasztalati szórás 139, Szegeden az átlag 540 mm, a korrigált tapasztalati szórás 143 lett (forrás: OMSZ). Állíthatjuk-e, hogy Szegeden szignifikánsan nagyobb a csapadékmennyiség várható értéke?

év	1991	1992	1993	1994	1995	...	átlag	s_n^*
Budapest	594	364	505	481	575	...	**533**	139
Szeged	617	457	408	399	562	...	**540**	143

A két adatsor **nem független**, mert egy éven belül a két város időjárása nem független (az egyes minták sem teljesen függetlenek, és nem biztos, hogy normális eloszlásúak). Ezért **párosított** (paired) t-próba alkalmazható, egyoldali nullhipotézissel.

$H_0 : m_1 \geq m_2$, $H_1 : m_1 < m_2$, ahol m_1 a budapesti, m_2 a szegedi csapadékmennyiség várható értéke.

A próbát elvégezve a p-értékre $0,366$ adódott.

Elfogadjuk a nullhipotézist, az adatok alapján Szegeden nem több szignifikánsan a csapadékmennyiség várható értéke, mint Budapesten.

Welch-féle t-próba

A **várható érték összehasonlítására** párosítatlan esetben (two-sample two-sided unpaired Welch t-test). Legyenek $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ **független normális eloszlású** valószínűségi változók: $X_i \sim N(m_1, \sigma_1^2)$, $Y_i \sim N(m_2, \sigma_2^2)$, ahol $m_1, m_2, \sigma_1, \sigma_2$ ismeretlen paraméterek.

Kétoldali ellenhipotézis: $H_0 : m_1 = m_2$; $H_1 : m_1 \neq m_2$.

Próbastatisztika:

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{s_{n_1}^{*2}(X)}{n_1} + \frac{s_{n_2}^{*2}(Y)}{n_2}}}.$$

Ha $|t| > t_{f,1-\alpha}$, akkor elvetjük a nullhipotézist, különben elfogadjuk. A $t_{f,1-\alpha}$ kritikus érték az f szabadsági fokú **kétoldali** t-próba kritikus értéke α szignifikanciaszint mellett (a megfelelő eloszlás $1 - \alpha/2$-kvantilise).

Szabadsági fok:

$$f \approx \frac{\left(\frac{s_{n_1}^{*2}(X)}{n_1} + \frac{s_{n_2}^{*2}(Y)}{n_2}\right)^2}{\frac{s_{n_1}^{*4}(X)}{n_1^2(n_1-1)} + \frac{s_{n_2}^{*4}(Y)}{n_2^2(n_2-1)}}.$$

Ha $p < \alpha$: elvetjük H_0-t, az várható értékek szignifikánsan eltérnek egymástól.

Házi feladat április 18., kedd, 12:00-ig

Készítsünk 95%-os, illetve 98%-os megbízhatósági szintű konfidenciaintervallumot az első házi feladathoz begyűjtött adatok alapján a sportolással töltött idő várható értékére, illetve az utazással töltött idő várható értékére.

Mennyire reális eredményeket kaptunk? Mely feltételek teljesülnek a konfidencia-intervallumra megfogalmazott feltételek közül, melyek nem?