Matematikai statisztika előadás, 6. hét, március 25.

1. 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 $\Theta \subseteq \mathbb{R}$.

1.1. 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: $\inf_{\vartheta\in\Theta}\{\mathbb{P}_\vartheta(\vartheta \in (T_1, T_2))\}$.

1.1. Egyoldali konfidenciaintervallum a várható értékre

1.1. 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*

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

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

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

1.2. Konfidenciaintervallum a szórásra

Valójában az eloszlás valódi szórása a legtöbb esetben nem ismert. Kérdés, hogy hogyan változik így a várható értékre adott konfidenciaintervallumban számolt mennyiség eloszlása, vagyis honnan tudjuk, hogy a normális eloszlás helyett a t-eloszlás lesz használható. Másrészt, ha a korrigált tapasztalati szórás eloszlásáról is tudunk valamit, akkor a szórásra is tudunk konfidenciaintervallumot adni – egy olyan intervallumot, ami nagy valószínűséggel tartalmazza a valódi szórást. Ezeket az összefüggéseket foglalja össze az alábbi tétel.

1.1. 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) $\overline{X} \sim N\big(m, \frac{\sigma^2}{n}\big)$;

(2) $\overline{X}$ és s_n^* függetlenek;

(3) $(n-1)s_n^{*2}/\sigma^2$ eloszlása $n-1$ szabadsági fokú χ^2-eloszlás;

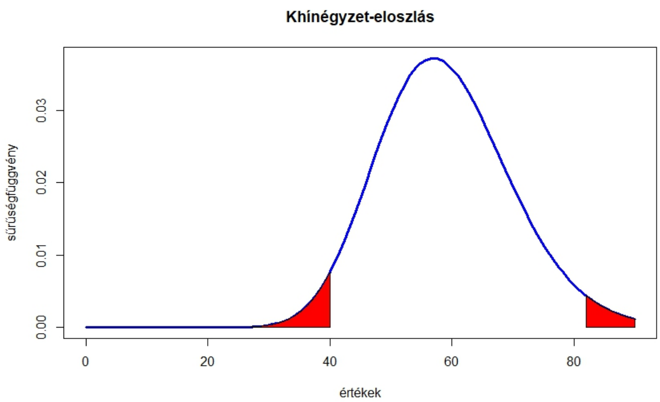
(4) $\frac{\overline{X}-m}{s_n^*} \cdot \sqrt{n}$ eloszlása $n-1$ szabadsági fokú t-eloszlás.

Itt a korrigált tapasztalati szórás:

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

A Fisher–Bartlett-tételből a szórásra vonatkozó részt kiemelve:

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



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

azaz a hányados eloszlása $n-1$ szabadsági fokú χ^2-eloszlás. Ezért a konfidenciaintervallum készítésekor a χ^2-eloszlás kvantiliseire lesz szükség.

Emlékeztető valószínűségszámításból: az $n-1$ szabadsági fokú χ^2-eloszlás a $Z_1^2+Z_2^2+\ldots+Z_{n-1}^2$ eloszlása, ahol Z_j-k független standard normális eloszlásúak.

1.2. 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 az f szabadsági fokú χ^2-eloszlás q-kvantilise.

Ezzel a választással nem a legrövidebb intervallumot kapjuk.

1.3. 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.

Itt $I_1(\vartheta)$ az egyelemű mintából számolt Fisher-információt jelöli.

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. Másképpen felírva:

$$\lim_{n\to\infty} \mathbb{P}\left(\hat{\vartheta}_n - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \frac{1}{\sqrt{\hat{I}_n(\vartheta)}} \le \vartheta \le \hat{\vartheta}_n + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \frac{1}{\sqrt{\hat{I}_n(\vartheta)}} \right) = 1 - \alpha.$$

2. Hipotézisvizsgálat

A hipotézisvizsgálat célja, hogy egy, az ismeretlen paraméterre vonatkozó állításról (nullhipotézis) eldöntsük, hogy elfogadható-e az adatok alapján, vagy az adatok szignifikánsan eltérnek attól, amit ennek az állításnak a teljesülése esetén várnánk, vagyis az adatok alapján az állítás statisztikai értelemben megcáfolható.

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.

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

2.1. 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.

Egy szokásos választás $\alpha = 0,05$. Ez tehát azt jelenti, hogy ha a nullhipotézis igaz, akkor legfeljebb 5% a téves elutasítás valószínűsége a próba során.

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.

2.1. Próbák tulajdonságai

- 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.

2.2. Neyman–Pearson-lemma

Tegyük fel, hogy a paramétertér összesen két elemből áll, vagyis az ismeretlen paraméternek két lehetséges értéke van: $\Theta = \{\vartheta_0, \vartheta_1\}$. 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.

Például: két telefontöltőnk van, az egyik minden kipróbálásnál a többitől függetlenül $0,2$, a másik $0,9$ valószínűséggel működik. A töltők ránézésre megkülönböztethetlenek. Kiválasztjuk az egyik töltőt, és tízszer kipróbáljuk. Kérdés, hogy hány sikeres kipróbálás esetén döntünk úgy, hogy ez a második töltő lehetett, ha a nullhipotézis az, hogy az első töltő van a kezünkben.

Számítsuk ki, hogy az adott sikeres töltések számának valószínűsége hányszor nagyobb a második, mint az első töltő esetében. Ez éppen a likelihood-függvények hányadosa lesz. 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. Az alábbi állítás arról szól, hogy ez a fajta eljárás (legalább k_0 sikeres töltés esetén utasítjuk el H_0-t) a legerősebb próba, és általában is, két lehetőség esetén a legerősebb próbákat akkor kapjuk, ha a likelihood-hányados alapján döntünk.

2.1. 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.*

2.3. Véletlenített próba

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.

H_0 : hibás töltő, $p = 0,2$

H_1: jó töltő, $p = 0,9$.

Olyan eljárást szeretnénk, amivel a nullhipotézis téves elutasításának valószínűsége **pontosan** $\alpha = 0,05$. A legfeljebb $0,05$ szignifikanciaszintűek közül ez lesz ugyanis a legerősebb, vagyis az ellenhipotézis teljesülése esetén ekkor a legnagyobb a jó döntés valószínűsége.

Ha $k_0 = 5$: az elsőfajú hiba valószínűsége még megfelelő, de nem pontosan $0,05$:

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

Ha $k_0 = 4$: az elsőfajú hiba valószínűsége túl nagy, több, mint a megengedett szignifikanciaszint:

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

Ezért olyan eljárást választunk (ez lesz a véletlenített próba), hogy

- legalább 5 sikeres kísérlet esetén elutasítjuk H_0-t;
- pontosan 4 sikeres kísérlet esetén egy új sorsolást végzünk, és p valószínűséggel utasítjuk el H_0-t – a p valószínűséget úgy választjuk, hogy a fenti feltétel teljesüljön

- legfeljebb 3 sikeres kísérlet esetén elfogadjuk H_0-t (ez tehát nem lehet bizonyíték H_1-re, vagyis arra, hogy a jó töltőt választottuk ki).

Ezzel az eljárással a nullhipotézis téves elutasításának valószínűsége (minden valószínűség a hibás töltőre vonatkozik, hiszen ez a nullhipotézis):

$$\mathbb{P}(X \geq 5) + p \cdot \mathbb{P}(X = 4) = 0,033 + p \cdot \binom{10}{4} 0,2^4 \cdot 0,8^6 = 0,033 + p \cdot 0,088.$$

A feltételünk az volt, hogy ez éppen $0,05$-tel legyen egyenlő, ez $p = 0,19$ esetén teljesül:

$$\mathbb{P}(X \geq 5) + p \cdot \mathbb{P}(X = 4) = 0,033 + 0,19 \cdot 0,088 = 0,05.$$

Tehát legalább 5 sikeres kísérlet esetén biztosan, pontosan 4 sikeres kísérlet esetén $0,19$ valószínűséggel utasítjuk el a nullhipotézist, legfeljebb 3 sikeres kísérlet esetén pedig elfogadjuk.

Az alábbi állítás szerint ez legerősebb próba a legfeljebb $\alpha = 0,05$ szignfikanciaszintű próbák között, és erősebb, mint a $k_0 = 5$-tel kapott determinisztikus próbáé.

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

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

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.

2.2. 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.*

2.4. Szekvenciális próbák

A valószínűséghányados n elemű mintából:

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

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.

3. A 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 $n \geq 100$; ez a **centrális határeloszlástételen** múlik: tetszőleges véges szórású, független azonos eloszlású valószínűségi változók átlagának eloszlása normális eloszláshoz hasonló, ha nagy a mintaelemszám

- **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.

3.1. z-próba

Nézzük meg, hogy mit ad a Neyman–Pearson-lemma abban az esetben, ha a nullhipotézis és az ellenhipotézis szerint is normális eloszlású az eloszlás, a szórások is megegyeznek, de a várható értékek nem. Ez fog segíteni abban, hogy a normális eloszlás várható értékére vonatkozó általános legerősebb próbát készítsünk, ismert szórás mellett.

Legyen $\Theta = \{m_0, m_1\}$, és $\mathcal{P}$ álljon az $N(m_0, \sigma)$ és $N(m_1, \sigma)$ eloszlásokból. Vagyis a nullhipotézis az, hogy az eloszlás $N(m_0, \sigma)$, az ellenhipotézis pedig az, hogy $N(m_1, \sigma)$. A likelihood-hányados:

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

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

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

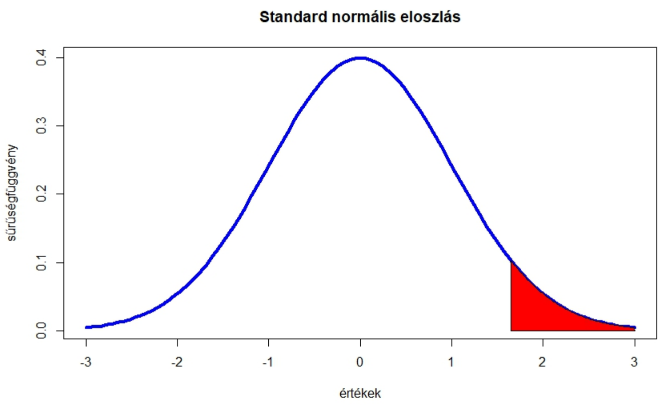
$$= \exp\left(\frac{2(m_1 - m_0)\sum_{j=1}^n X_j + (m_0^2 - m_1^2)n}{2\sigma^2}\right).$$

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, hogy akkor utasítjuk el a nullhipotézist (egyoldali esetben), ha az összeg nagyobb egy bizonyos értéknél. Hiszen ez ugyanaz, mint hogy az átlag nagyobb egy bizonyos értéknél. Azt pedig tudjuk korábbról, hogy ha $X_1, \ldots, X_n$ független azonos eloszlásúak és normális eloszlásúak m_0 várható értékkel és σ szórással, akkor az átlaguk is normális eloszlású, sőt, a normális eloszlások átlagára vonatkozó összefüggés (a Fisher–Bartlett-tétel, az 1.1. tétel (1) része szerint)

$$\overline{X} \sim N\left(m, \frac{\sigma^2}{n}\right) \qquad \Rightarrow \qquad \frac{\overline{X} - m}{\sigma}\sqrt{n} \sim N(0,1).$$

Ezért a kritikus értéket a standard normális eloszlás kvantiliséből számolhatjuk ki. Pontosabban, úgy kell ezt választanunk, hogy H_0 teljesülése esetén a téves elutasítás valószínűsége α legyen (2. ábra):

$$\mathbb{P}\left(\frac{\overline{X} - m}{\sigma}\sqrt{n} > c\right) = 1 - \alpha \qquad \Rightarrow \qquad c = \Phi^{-1}(1-\alpha).$$



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

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

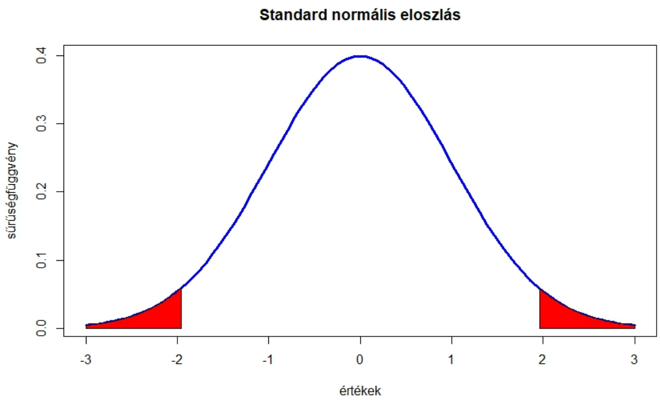
$p < 0,05$: a várható érték szignifikánsan több m_0-nál.

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

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?
- $H_0 : m \leq 177,6$; $\quad H_1 : m > 177,6.$
- $$z = \frac{\overline{X} - m_0}{\sigma} \cdot \sqrt{n} = \frac{181,7 - 177,6}{8,5}\sqrt{90} = 4,57.$$
- $\alpha = 0,05$ 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.



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

3.1.2. Egymintás kétoldali z-próba

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

- $X_1, X_2, \dots, X_n \sim N(m, \sigma^2)$, 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): $H_0 : m = m_0$; $\quad H_1 : m \neq m_0$.
- 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.

Példa. 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?
- $H_0 : m = 364$; $\quad H_1 : m \neq 364$.
- $$z = \frac{\overline{X} - m_0}{\sigma} \cdot \sqrt{n} = \frac{367,2 - 364}{3}\sqrt{20} = 4,77.$$
- $\alpha = 0,05$ 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.

Példa. Mennyi a z-próba erőfüggvényének az értéke a 2 helyen, ha 16 elemű mintánk van, mely 9 szórású normális eloszlásból származik és a hipotézisek: $H_0 : m = 1$, $H_1 : m > 1$. Tegyük fel, hogy a kritikus érték $z_\alpha = 2$.

Az erőfüggvény a helyes döntés valószínűségét adja meg, az ellenhipotézishez tartozó paraméterek esetén. Most

$$\beta(2) = \mathbb{P}_2(\text{elutasítjuk } H_0\text{-t})$$

a kérdés, vagyis hogy mennyi valószínűséggel utasítjuk el a nullhipotézist, ha a minta $N(2, 9^2)$ eloszlású.

Az egymintás z-próba próbastatisztikája $z = \frac{\overline{X}-m_0}{\sigma}\sqrt{n}$. Most $m_0 = 1, \sigma = 9, n = 16$, a kritikus érték pedig 2. Akkor utasítjuk el a nullhipotézist (az $m > 1$ egyoldali ellenhipotézisnek megfelelően), ha $z > 2$ (ez jelenti azt, hogy az átlag nagy, ami a várható érték nagy értékére utal). Tehát a kérdés, amiből átrendezésekkel juthatunk el a megoldáshoz:

$$\beta(2) = \mathbb{P}_2(\text{elutasítjuk } H_0\text{-t}) = \mathbb{P}_2(z > 2) = \mathbb{P}_2\left(\frac{\overline{X}-m_0}{\sigma}\sqrt{n} > 2\right) = \mathbb{P}_2\left(\overline{X} > \frac{2\sigma}{\sqrt{n}} + m_0\right) =$$
$$= \mathbb{P}\left(\frac{\overline{X}-2}{\sigma}\sqrt{n} > \frac{m_0-2}{\sigma}\sqrt{n} + 2\right) = \mathbb{P}\left(\frac{\overline{X}-2}{\sigma}\sqrt{n} > \frac{1-2}{9}\cdot 4 + 2\right) = \mathbb{P}\left(\frac{\overline{X}-2}{\sigma}\sqrt{n} > \frac{14}{9}\right) =$$
$$= 1 - \Phi\left(\frac{14}{9}\right),$$

ugyanis $\overline{X} \sim N(2, \sigma/\sqrt{n})$, ha független, azonos eloszlású normális eloszlású valószínűségi változókat átlagolunk (1.1. tétel).

Házi feladat április 1., 8:15-ig Az ismerősöktől gyűjtött adatok alapján állíthatjuk-e, hogy az ismerősök utazási idejének várható értéke szignifikánsan több 60 percnél? Ehhez feltételezzük, hogy az utazási idő szórása 30 perc, és az utazási idő normális eloszlású (még ha a kerekítések miatt nem is az).

Az előző házi feladat A házi feladathoz begyűjtött adatok alapján

1. adjunk konfidenciaintervallumot az utazási idők várható értékére külön a férfiak, külön a nők esetében, illetve összesen;
2. adjunk konfidenciaintervallumot annak valószínűségére, hogy egy véletlenszerűen kiválasztott ismerős legalább 5 sorozatot nézett az utóbbi egy hónapban.

Az utazási időkről tételezzük fel, hogy normális eloszlásúak (noha valójában a kerekítések miatt ez nem igazán teljesül).

Megoldás.

Tegyük fel, hogy $X_1, \dots, X_n$ 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.

Az R-ben a t-próbára vonatkozó parancs megadja a konfidenciaintervallumot is, a megbízhatósági szint is állítható:

```
> noiutazas<-c(0, 45, 120, 120, 60, 60, 30, 60, 100, 70, 180, 40, 60, 100, 80, 90, 120)

> t.test(noiutazas, conf.level=0.95)

95 percent confidence interval:  56.60688 100.45194

ferfiutazas<-c(60, 30, 70, 60, 20, 60, 10, 120, 60, 120, 130)

t.test(ferfiutazas, conf.level=0.95)

95 percent confidence interval:  39.88616 94.65930

> utazas<-c(noiutazas, ferfiutazas)

> t.test(ferfiutazas, conf.level=0.95)

95 percent confidence interval:  58.00614 90.20815
```

Tehát a 95%-os megbízhatósági szintű konfidenciaintervallum a nők esetében percben számolva $(56,6; 100,5)$, a férfiak esetében $(39,9; 94,66)$, összességében pedig $(58; 90)$. Összehasonlítva: a nők esetében rövidebb konfidenciaintervallumot tudtunk adni, azaz kisebb a bizonytalanság, a várható érték viszont az ő esetükben tűnik nagyobbnak. Az adatokat összesítve nagyobb lett a mintaelemszám, mint bármelyik esetben külön-külön, így természetes, hogy rövidebb konfidenciaintervallum adódott.

A sorozatokra vonatkozó adatok:

```
sorozat=c(6, 4, 5, 7, 2, 12, 3, 4, 12, 4, 9, 13, 5, 7, 8, 13, 16, 5, 6, 9, 1, 1, 1, 3, 5,
1, 0, 3, 2, 1, 0, 0, 10, 0, 2, 1, 0, 0, 5, 1, 2, 2, 1, 2, 2, 0, 2, 2, 0, 1, 0, 0, 1, 0,
1, 8, 4, 3, 5, 5, 1, 0, 3, 8, 10, 1, 5, 5, 1, 0, 3, 1, 0)
```

A legalább öt sorozatot nézők relatív gyakorisága:

```
> p=length(sorozat[sorozat>=5])/length(sorozat)
> p
[1] 0.3424658
```

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

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

Ez alapján az alsó végpont, ha $\alpha = 0,05$, azaz 95%-os megbízhatósági szintű konfidenciaintervallumot készítünk:

```
> p-qnorm(0.975)*sqrt(p*(1-p))/sqrt(length(sorozat))
[1] 0.2336092
> p+qnorm(0.975)*sqrt(p*(1-p))/sqrt(length(sorozat))
[1] 0.4513223
```

Tehát a $(0,23; 0,54)$ intervallum 95% megbízhatósági szintű konfidenciaintervallum a legalább 5 sorozat nézésének valószínűségére.