

Konfidenciintervallumok, hipotézisvizsgálat

1. Konfidenciintervallum

Általában kérdés, hogy hogyan járhatunk el, ha az ismeretlen paraméterre nem csak egyetlen számot mondhatunk becslésként (pontbecslés), hanem egy intervallumot is megadhatunk, amiről azt állítjuk, hogy a valódi paramétert nagy valószínűséggel tartalmazza. Az így megadott intervallum hossza a becslésünk bizonytalanságát is kifejezi: minél hosszabb az intervallum, annál „bizonytalanabb” a becslésünk.

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

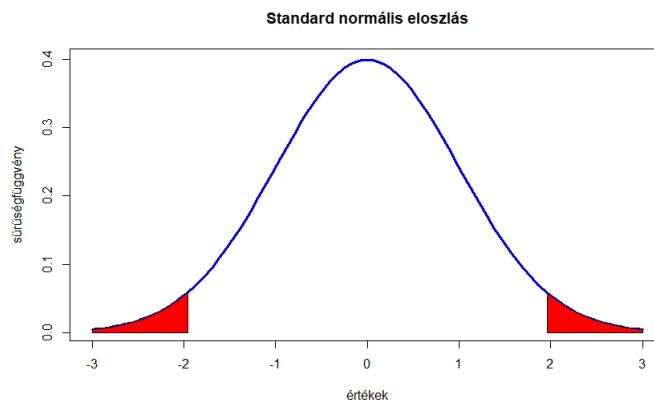
Alapstatisztikák: $n = 60$ (méret), $\bar{X} = 5,99$ (átlag), $s_n^* = 0,18$ (korrigált tapasztalati szórás)

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 konfidenciintervallumnak** hívni, azonban vegyük észre, hogy most az „igazi” várható érték rögzített, ismeretlen (visszatérünk a frekventista hozzáálláshoz), és az intervallum végpontjai lesznek véletlenek, ezek függenek a mintától

1.1. Konfidenciintervallum a várható értékre ismert szórás esetén

Tegyük fel, hogy X_1, \dots, X_n független normális eloszlású valószínűségi változók, a várható értékük, m ismeretlen paraméter, míg a σ szórás ismert (például mert a mért értékek szóródása csak a mérési eljárás, mérőeszköz hibájából adódik, és azt már korábbról ismerjük).



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

Legyen

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

ahol $\Phi(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$ a standard normális eloszlás eloszlásfüggvénye, Φ^{-1} pedig ennek az inverze. Vagyis például ha $\alpha = 0,05$,

$$\alpha = 0,05 \Rightarrow \Phi^{-1}(1 - \alpha/2) = 1,96 \Rightarrow \mathbb{P}(-1,96 \leq Z \leq 1,96) = 95\%$$

az 1. ábrának megfelelően.

Ekkor

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

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

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

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

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

1.1. Állítás (Konfidenciaintervallum a várható értékre, ismert szórás). *Tegyük fel, hogy X_1, \dots, X_n független azonos eloszlású normális eloszlású valószínűségi változók, melyek szórása, σ ismert.*

Ekkor a

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

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

Példa. Megmértük hatvan ember vércukorszintjét, tegyük fel, hogy ez normális eloszlású, valódi szórása 0,2.

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

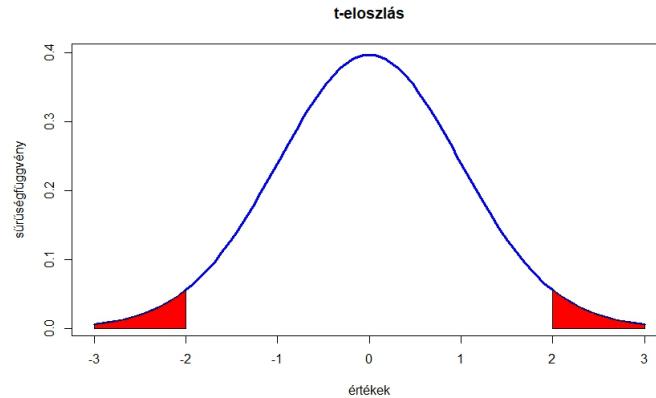
Adjunk meg olyan intervallumot, ami legalább 95% valószínűséggel tartalmazza a vércukorszint valódi várható értékét.

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:

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

Kisebb mintaelemszámhoz hosszabb, nagyobb megbízhatósághoz szintén hosszabb konfidenciaintervallum tartozik.



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

1.2. Konfidenciaintervallum és megbízhatósági szint

Legyen $(\Omega, \mathcal{A}, \mathcal{P})$ statisztikai mező, $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$ és $\underline{X} = (X_1, \dots, X_n)$ 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 $\vartheta \in \vartheta$ esetén teljesül, hogy

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

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

Vegyük észre, hogy

- a feltételnek az ismeretlen paraméter minden lehetséges értékére teljesülnie kell (frekvencia hozzáállás);
- ϑ rögzített, az intervallum végpontjai azonban véletlenek, a mintától függőek, erre vonatkozik a valószínűség;
- nagyobb megbízhatósági szinthez hosszabb konfidenciaintervallum tartozik.

1.3. Konfidenciaintervallum a várható értékre ismeretlen szórás esetén

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. Erről szól az alábbi tétel (mely a t -próbának is az alapja).

1.1. Tétel (Fisher–Bartlett). Tegyük fel, hogy X_1, X_2, \dots, X_n független m várható értékű, σ szórású, **normális eloszlású** valószínűségi változók. Ekkor

(1) $\bar{X} \sim N(m, \frac{\sigma^2}{n})$;

(2) \bar{X} és s_n^* függetlenek;

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

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

Itt

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

Illetve, legyenek Z_0, Z_1, \dots, Z_f 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:

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

Az $Y = \frac{Z_0}{\sqrt{Z_1^2 + \dots + Z_f^2}}$ valószínűségi változó eloszlása f szabadsági fokú **t -eloszlás** (emlékeztető valószínűségi számításból: https://backhauszagi.web.elte.hu/gyak/sst_vsz4ea_1/sst_vsz4je_110.pdf).

Ismert szórás esetén $(\bar{X} - m)/\sigma \sim N(0, 1)$ teljesült, és ennek a standard normális eloszlásnak a kvantiliseit használtuk a konfidenciaintervallum hosszához. Most tehát a standard normális eloszlás helyett az $n-1$ szabadsági fokú t -eloszlást használjuk. A kritikus értékeket a 2. ábra mutatja.

1.2. Állítás (Konfidenciaintervallum a várható értékre, ismeretlen szórás). *Tegyük fel, hogy X_1, \dots, X_n független $N(m, \sigma^2)$ normális eloszlású valószínűségi változók (m, σ ismeretlenek). Ekkor a*

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

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

Példa. Tekintjük a korábban látott vércukorszint-értékeket, és tegyük fel, hogy ez **normális eloszlású**, szórása nem ismert.

$n = 60$ (méret), $\bar{X} = 5,99$ (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:

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

Hosszabb intervallumot kaptunk, mint ismert szórás esetén.

1.4. Egyoldali konfidenciaintervallum a várható értékre

1.3. Állítás (Egyoldali konfidenciaintervallum). *Tegyük fel, hogy X_1, \dots, X_n független $N(m, \sigma^2)$ normális eloszlású valószínűségi változók (m, σ ismeretlenek). Ekkor a*

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

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.

Más eloszlások. Bár a fenti állítások a normális eloszlásra vonatkoztak, a centrális határeloszlástétel alapján, ha a mintaelemszám elég nagy (például $n \geq 100$), más eloszlású **véges szórású** mintára is használhatjuk a fenti összefüggéseket, ugyanis az átlag és a korrigált tapasztalati szórás hasonlóan viselkedik, mint normális eloszlás esetén.

1.5. 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:

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

$X \sim \text{Bin}(n, p)$ 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. Ahhoz, hogy a közelítés megfelelő legyen, az $np(1-p) \geq 10$ feltételt szokták például megkövetelni [1].

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

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

További olvasnivaló, egy lehetséges alkalmazásról: <https://ematlap.hu/tudomany-tortenet-2020-12/992-mennyit-teszteljunk-2-v3>

1.6. Konfidenciaintervallum a szórásra

Nem csak a várható érték lehet az a mennyiség, amire konfidenciaintervallumot szeretnénk adni, hanem a szórás is.

A Fisher–Bartlett-tételből (1.1. tétel) a szórásra vonatkozó részt kiemelve:

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

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

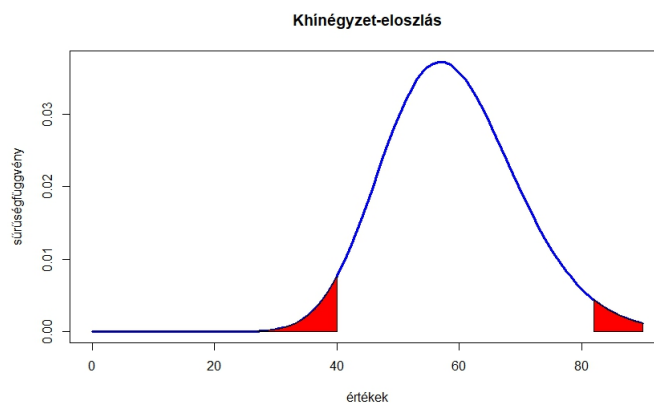
Emlékeztető valószínűségi számításból: az $n - 1$ szabadsági fokú χ^2 -eloszlás a $Z_1^2 + Z_2^2 + \dots + Z_{n-1}^2$ eloszlása, ahol Z_j -k független standard normális eloszlásúak.

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

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

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. Itt is azt látjuk, hogy az átlagtól való eltérések négyzetösszege helyett jelenik meg a független normális eloszlások négyzetösszege.



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

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

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

Elutasítási (kritikus) tartomány: B_1 . Ha $(X_1, \dots, X_n) \in B_1$, 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 $\beta : \Theta_1 \rightarrow [0, 1]$ függvény:

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

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:

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

- 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 paraméterter 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

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

akkor utasítsuk el a nullhipotézist, különben fogadjuk el.

Például: két teleföntöltőnk van, az egyik minden kipróbálásnál a többitől függetlenül 0,2, a másik 0,9 valószínűséggel működik. A töltők ránézésre megkülönböztethetetlenek. 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:

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

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 -al.

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 $\alpha = 0,05$ szignifikanciaszintű 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

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

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.

2.2. lemma (Neyman–Pearson-lemma, 2. rész). *Legyen $\alpha \in (0, 1)$ tetszőleges. Ekkor lehet olyan c és p számokat választani, hogy a véletlenített 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:

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

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 $V_{n_1} \leq A$, 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.

Hivatkozások

[1] Sheldon Ross, A first course in probability. Second edition, Macmillan Co., New York, 1984.

Házi feladat április 7., szerda, 9:00-ig Tekintsük a közösségi médiával töltött időt a félév elején összegyűjtött mintában. Adjunk konfidenciaintervallumot a közösségi médiával töltött idő várható értékére

- (a) külön a 25 évesnél fiatalabbak és külön a 25 évesnél idősebbek esetén 95% megbízhatósággal;
- (b) a teljes mintából 90%-os és 95%-os megbízhatósággal.

Milyen következtetéseket vonhatunk le? Mennyire teljesülnek a fenti módszerek alkalmazásának feltételei? (A házi feladatnál nem baj, ha nem, csak gondoljuk meg ezt is.)