

Statisztika

Tóbiás András

2024. január 3-i állapot

Ez a dokumentum a Mészáros Szabolcs-féle Valószínűségszámítás előadásjegyzet kiegészítése a 2023/24-es tanév őszi félévére. A tételek, definíciók stb. számozása ezt az előadásjegyzetet követi. Aki a dokumentumban hibát talál, kérem, jelezze nekem.

Tartalomjegyzék

13.A matematikai statisztika alapfogalmai	1
13.1. Minta és realizáció	3
13.2. Alapstatisztikák	4
13.3. Empirikus eloszlásfüggvény	7
14.Becslésméleti módszerek	8
14.1. Becslések tulajdonságai, átlag és szórás becslései	9
14.2. Maximum likelihood-becslés	11
14.3. Momentumbecslés	16
15.Konfidenciaintervallumok	18
15.1. Alapfogalmak: konfidenciaintervallum, folytonos eloszlás kvantilisei	18
15.2. Konfidenciaintervallum normális eloszlás várható értékére ismert szórás esetén	20
15.3. χ^2 -eloszlás, Student-eloszlás és tulajdonságaik	21
15.4. Konfidenciaintervallum normális eloszlás szórásnégyzetére ismert várható érték esetén (a 2023/24-es tanévben kiegészítő anyag)	23
15.5. Konfidenciaintervallum normális eloszlás várható értékére ismeretlen szórás esetén	24
16.Hipotézisvizsgálati módszerek	24
16.1. Bevezetés	24
16.2. u -próba	26
16.3. t -próba	32
16.4. Megjegyzések a hipotézisvizsgálati képletgyűjtemény használatához	35

13. A matematikai statisztika alapfogalmai

Az előadás eddigi részében a valószínűségi változók eloszlása ismert volt. A matematikai statisztikára viszont az a jellemző, hogy a valószínűségi változók valamilyen mérési eredményeknek felelnek meg, és ezért az eloszlásukat nem ismerjük pontosan. Gyakran előfordul, hogy elméleti megfontolások alapján tudjuk, hogy jó közelítéssel milyen típusú eloszlásúak a mérési eredményeink, de az eloszlás valamilyen ismeretlen ϑ paramétertől függ, amelynek lehetséges értékei a $\theta \subseteq \mathbb{R}^d$ paramétertartomány elemei lehetnek valamely $d \geq 1$ -re.¹ Például:

¹A ϑ és a θ a görög „teta” betű kétféle kisbetűs írásmódja.

- Ha az utcán találunk egy pénzérmét, amiről nem tudjuk, hogy szabályos-e, az egy ismeretlen $\vartheta \in \theta = [0, 1] \subset \mathbb{R}^1$ valószínűséggel mutat fejet; a pénzérmét feldobva próbálhatjuk ϑ értékét megbecsülni, illetve igazolni vagy cáfolni, hogy a pénzérme szabályos. Az

$$\mathbb{1}_{\{\text{a pénzérme egy adott feldobásának eredménye fej}\}}$$

indikátorváltozó paramétere tehát az ismeretlen ϑ .

- A Magyarországon egy adott hónapban vasúti útátjárókban bekövetkező balesetek számáról feltehetjük, hogy jó közelítéssel Poisson-eloszlású, mert a közúton közlekedők száma nagy, egy-egy közlekedő kis valószínűséggel szenved ilyen balesetet, és a többi közlekedőtől többé-kevésbé függetlenül. A Poisson-eloszlás paramétere egy *ismeretlen* $\vartheta \in \theta = (0, \infty) \subset \mathbb{R}^1$ paraméter.

- A BME egy véletlenszerűen kiválasztott női hallgatójának testmagasságát modellezhetjük normális eloszlással². Itt alapesetben a $\mu \in \mathbb{R}$ várható érték és a $\sigma^2 > 0$ szórásnégyzet is ismeretlen, így a paramétertartomány

$$\theta = \{(\mu, \sigma^2) \in \mathbb{R}^2 : \sigma^2 > 0\} \subset \mathbb{R}^2.$$

(Persze nyilvánvaló, hogy μ -nek a gyakorlatban pozitívnak kell lennie, de ettől most az általánosság kedvéért eltekintünk.)

A matematikai statisztikában jellemző, hogy az ismeretlen paraméter tanulmányozása céljából *mintát veszünk*, azaz „generálunk” olyan X_1, \dots, X_n független, azonos eloszlású valószínűségi változókat, amelyek az ismeretlen paraméterű eloszlást követik. Például az előbbi példákban:

- A pénzérmét n -szer feldobjuk, ekkor $i = 1, \dots, n$ esetén X_i értéke legyen 1, ha az i -edik dobás eredménye fej, és 0, ha írás. Ekkor az X_1, \dots, X_n *mintaelemek* független és azonos eloszlású indikátorváltozók, amelyek paramétere az ismeretlen ϑ .
- A Magyarországon vasúti útátjárókban bekövetkező balesetek havi számát n hónapon át megfigyeljük, legyen X_i az i -edik hónapban bekövetkező balesetek száma. Ekkor X_1, \dots, X_n (többé-kevésbé) függetlenek és (nagyjából) Poisson(ϑ) eloszlásúak.
- A BME női hallgatóinak névsorából egyenletesen és egymástól függetlenül választva kisorsolunk n hallgatót, legyen X_i az i -edik hallgató testmagassága. Ekkor X_1, \dots, X_n (hozzávetőleg) függetlenek és (nagyjából) $N(\mu; \sigma^2)$ eloszlásúak, ahol mindkét paraméter ismeretlen.

A matematikai statisztika két legfontosabb ága a *becslélmélet* és a *hipotézisvizsgálat*. A becslélmélet célkitűzése az ismeretlen paraméter értékének a minta alapján történő, minél pontosabban meghatározása. (A „minél pontosabban” persze matematikailag egyáltalán nem precíz kifejezés, később legalábbis konkrét becslélméleti módszereknél tisztázni fogjuk, hogy mit is értünk ez alatt.) A hipotézisvizsgálat célja valamilyen előzetesen rögzített hipotézis (feltevés) igazolása vagy megcáfolása a minta segítségével. Ilyen hipotézis lehet például:

- A pénzérme szabályos.
- Az egy hónapban bekövetkező balesetek átlagos száma 2.
- A BME női hallgatóinak átlagos testmagassága legfeljebb 166 cm.
- A BME női hallgatói testmagasságának szórása 3 cm.

²Ugyanígy egy véletlenszerűen kiválasztott férfi hallgatóét is (más várható értékkel), azonban a BME egy tetszőleges hallgatóját nem, mert a nők és a férfiak átlagos testmagassága különbözik, ami 2 db lokális maximummal rendelkező – tehát közelítőleg sem normális – sűrűségfüggvényhez vezet.

13.1. Minta és realizáció

A minta fogalmát úgy definiáljuk, hogy abban nem utalunk a ϑ ismeretlen paraméterre. Ez azért lesz hasznos, mert a későbbiekben a matematikai statisztika olyan módszereivel is fogunk találkozni, amelyek nemcsak egy ismeretlen paramétertől függő eloszlásra, hanem teljesen ismeretlen eloszlásokra is alkalmazhatóak. (Például a 13.2 és a 13.3 fejezetekben nem fog szerepet játszani ismeretlen paraméter.)

13.1.1. Definíció. \blacktriangle Legyenek X_1, \dots, X_n független, azonos eloszlású változók, amelyek peremeloszlása nem feltétlenül ismert. Ekkor az

$$\mathbf{X} = (X_1, \dots, X_n)$$

valószínűségi vektorváltozót n elemű független, azonos eloszlású mintának (röviden n elemű fae. mintának) nevezzük.

Most pedig bevezetünk néhány hasznos jelölést a paraméteres esetre:

13.1.2. Jelölés. Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta, ahol a mintaelemek eloszlása valamilyen ϑ paramétertől függ (azaz az eloszlásuk a ϑ ismeretében meghatározható), ahol a ϑ lehetséges értékeinek halmaza $\theta \subseteq \mathbb{R}^d$ valamely $d \geq 1$ -re. Ekkor egy $\vartheta \in \theta$ paraméter esetén:

- (1) az X_1 valószínűségi változó ϑ paraméterhez tartozó eloszlásfüggvényét F_ϑ -val jelöljük,
- (2) ha X_1 -nek ezen paraméter esetén létezik sűrűségfüggvénye, akkor ezt f_ϑ -val jelöljük. Vagyis:

$$f_\vartheta(x) = \begin{cases} F'_\vartheta(x), & \text{ha } F_\vartheta \text{ differenciálható } x\text{-ben,} \\ 0, & \text{különben,} \end{cases}$$

- (3) illetve ha X_1 -nek ezen paraméter esetén diszkrét az eloszlása, akkor p_ϑ -val jelöljük az ehhez az eloszláshoz tartozó súlyfüggvényt. Vagyis $x \in \mathbb{R}$ esetén $p_\vartheta(x)$ jelöli annak a valószínűségét, hogy $X_1 = x$, feltéve, hogy ϑ a valódi paraméter.

13.1.3. *Példa.* • Az utcán talált pénzérme esetén $p_\vartheta(0) = 1 - \vartheta$ és $p_\vartheta(1) = \vartheta$, $\vartheta \in [0, 1]$.

- Az útátjárókban bekövetkező balesetek száma esetén

$$p_\vartheta(k) = \frac{\vartheta^k}{k!} e^{-\vartheta}, \quad \vartheta > 0, k = 0, 1, \dots$$

- A BME női hallgatóinak magassága esetén az ismeretlen paraméterünk (μ, σ^2) , így tehát

$$f_{(\mu, \sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}, \quad \mu \in \mathbb{R}, \sigma^2 > 0.$$

Az eloszlásfüggvények felírását az olvasóra bizzuk.

Sokszor a mintaelemek értékkészlete is függ a ϑ paramétertől, amint ezt a következő példa is mutatja (ellentétben az eddigi példákkal).

13.1.4. *Példa.* Egy barátunk számítógéppel 5 db $U(0; 1)$ eloszlású véletlen számot generált, majd a kapott számokat mind megszorozta ugyanazzal a szigorúan 1 és 2 közé eső ϑ számmal, amit ő választott és nekünk nem árult el. Az így kapott $\mathbf{x} = (x_1, \dots, x_5) \in \mathbb{R}^5$ vektort (vagyis az eredetileg generált véletlen számok ϑ -szorosainak vektorát) átadta nekünk. Ez a vektor tehát egy 5 elemű $\mathbf{X} = (X_1, \dots, X_5)$ fae. mintából származik, annak egy *realizációját* adja. A minta elemei függetlenek és $U(0; \vartheta)$ eloszlásúak (világos? Ha nem, határozzuk meg az egyes mintaelemek sűrűségfüggvényét!).

A ϑ paraméter ismeretében az X_1, \dots, X_n valószínűségi változók értékkészlete (ahol a sűrűségfüggvényük pozitív) a $(0, \vartheta)$ intervallum, ami tehát függ ϑ -tól.

Ahhoz, hogy a *realizáció* szó precíz definícióját megadhassuk, először bevezetjük a valószínűségi változó (ϑ paramétertől függő) *lényeges értékei* halmazának fogalmát. Diszkrét esetben ezzel a definícióval a regresszió témakörénél már találkoztunk (lásd a 11.1.5. Definíciót Mészáros Szabolcs jegyzetében), csak akkor nem volt szó ismeretlen paraméterről.

13.1.5. Definíció (és jelölés). Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta, ahol a mintaelemek eloszlása egy $\vartheta \in \theta \subseteq \mathbb{R}^d$ paramétertől függ. Legyen $\vartheta \in \theta$ és $i = 1, \dots, n$. Ekkor ha létezik az f_ϑ sűrűségfüggvény, akkor definiáljuk az $S_{X_i}^{(\vartheta)}$ halmazt a következőképpen:

$$S_{X_i}^{(\vartheta)} = \{x \in \mathbb{R} \mid f_\vartheta(x) > 0\} \subseteq \mathbb{R}.$$

Ha pedig létezik a p_ϑ súlyfüggvény, akkor definiáljuk az $S_{X_i}^{(\vartheta)}$ halmazt az alábbi képlettel:

$$S_{X_i}^{(\vartheta)} = \{x \in \mathbb{R} \mid p_\vartheta(x) > 0\} \subseteq \mathbb{R}.$$

$S_{X_i}^{(\vartheta)}$ -t mindkét esetben az X_i valószínűségi változó ϑ paraméterhez tartozó **lényeges értékei halmazának** nevezzük.³

13.1.6. *Példa.* Az ismeretlen ϑ valószínűséggel fejlet mutató pénzérme n -szeri feldobásából kapott (X_1, \dots, X_n) fae. mintánál, ahol X_i indikátorváltozó ϑ paraméterrel, $S_{X_i}^{(\vartheta)} = \{0, 1\}$ minden $\vartheta \in \theta = (0, 1)$ -re, $S_{X_i}^{(0)} = \{0\}$ és $S_{X_i}^{(1)} = \{1\}$. Továbbá $p_\vartheta(1) = \vartheta$, $p_\vartheta(0) = 1 - \vartheta$ és $p_\vartheta(x) = 0$ minden $x \in \mathbb{R} \setminus \{0, 1\}$ esetén. A 13.1.4 példában szereplő (X_1, \dots, X_n) fae. mintánál ($n = 5$), ahol X_i egyenletes eloszlású a $(0, \vartheta)$ intervallumon, $S_{X_i}^{(\vartheta)} = (0, \vartheta)$ és

$$f_\vartheta(x) = \begin{cases} \frac{1}{\vartheta}, & \text{ha } x \in (0, \vartheta), \\ 0, & \text{különben} \end{cases}$$

teljesül minden $\vartheta \in \theta = (1, 2)$ paraméterre.

13.1.7. Definíció. Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta, ahol a mintaelemek eloszlása egy $\vartheta \in \theta \subseteq \mathbb{R}^d$ paramétertől függ. Az $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ az $\mathbf{X} = (X_1, \dots, X_n)$ minta egy (lehetséges) **realizációja** a $\vartheta \in \theta$ paraméter esetén, ha $x_i \in S_{X_i}^{(\vartheta)}$ teljesül minden $i \in \{1, \dots, n\}$ -re.

13.1.8. *Példa.* Demonstráljuk a realizáció fogalmát a 13.1.6 példával!

Az utcán talált pénzérme esetén az $(x_1, \dots, x_7) = (1, 0, 0, 0, 1, 1, 0)$ az $n = 7$ elemű $\mathbf{X} = (X_1, \dots, X_7)$ fae. minta egy realizációja minden $0 < \vartheta < 1$ paraméter esetén (de $\vartheta = 0$ és $\vartheta = 1$ esetén nem).

A $(0, \vartheta)$ intervallumon egyenletes eloszlású $\mathbf{X} = (X_1, \dots, X_5)$ fae. mintának

$$(x_1, \dots, x_5) = (0.14, 0.79, 1.13, 1, 1.2),$$

egy lehetséges realizációja, ha $1.2 < \vartheta < 2$, de ha $1 < \vartheta \leq 1.2$, akkor nem.

13.2. Alapstatisztikák

Egy n elemű fae. minta egy statisztikájának a mintaelemek egy olyan függvényét nevezzük, amely szimmetrikus, azaz „minden mintaelemről ugyanúgy függ”. Az utóbbi tulajdonságot formalizálja a következő definíció.

³ X_i lényeges értékeinek halmaza (adott ϑ esetén) csaknem ugyanaz, mint X_i (mint $\Omega \rightarrow \mathbb{R}$ függvény) értékkészlete. Az értékkészlet lehet bővebb, mint a lényeges értékek halmaza, azonban X_i a lényeges értékek halmazán kívüli összes értéket együttvéve 0 valószínűséggel veszi fel.

13.2.1. Definíció. Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta. Ha $T: \mathbb{R}^n \rightarrow \mathbb{R}$ szimmetrikus függvény, azaz

$$T(x_1, \dots, x_n) = T(\pi(x_1), \dots, \pi(x_n))$$

minden $x_1, \dots, x_n \in \mathbb{R}$ és $\pi: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ permutáció⁴ esetén teljesül, akkor a $T(\mathbf{X}) = T(X_1, \dots, X_n)$ valószínűségi vektorváltozót az X_1, \dots, X_n egy **statisztikájának** nevezzük.

Lássunk most néhány klasszikus, részben már az előadás korábbi részeiből is ismert statisztikát! Az első a mintaátlag, amit már a középiskolából is ismerünk.

13.2.2. *Példa.* Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta. Ekkor a 9.2 fejezetben bevezetett

$$\overline{X}_n = \frac{X_1 + \dots + X_n}{n}$$

mintaátlag az \mathbf{X} egy statisztikája. Ha $\mathbf{x} = (x_1, \dots, x_n)$ az $\mathbf{X} = (X_1, \dots, X_n)$ egy realizációja, akkor realizáció átlagát \overline{x}_n -nel fogjuk jelölni: $\overline{x}_n = \frac{x_1 + \dots + x_n}{n}$.

Az előadás eddigi anyaga alapján nem meglepő, hogy a mintaátlag a várható érték egyfajta becslése a minta alapján. Az világos, hogy ha $\mathbb{E}(X_i)$ létezik (vagyis ha $\mathbb{E}[|X_i|] < \infty$), akkor a mintaátlag várható értéke megegyezik az egyes mintaelemek várható értékével:

$$\mathbb{E}[\overline{X}_n] = \frac{1}{n}(\mathbb{E}(X_1) + \dots + \mathbb{E}(X_n)) = \mathbb{E}(X_1).$$

Ezen becslés további tulajdonságait a becslélmélet nyelvezetével a 14 fejezetben fogjuk elemezni.

A következő klasszikus statisztika a korrigált empirikus szórásnégyzet, amely a szórásnégyzet egyfajta közelítése a mintaelemek segítségével.

13.2.3. Definíció. \blacktriangle Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta. Ekkor az

$$(S_n^*)^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X}_n)^2 \quad (13.1)$$

mennyiséget az \mathbf{X} **korrigált empirikus szórásnégyzetének**⁵ nevezzük, $S_n^* = \sqrt{(S_n^*)^2}$ -et pedig az \mathbf{X} **korrigált empirikus szórásának**.

Világos, hogy $(S_n^*)^2$ az \mathbf{X} minta egy statisztikája. Az olvasóban felmerülhet a kérdés, hogy miért éppen $n-1$ -gyel (és nem például n -nel) osztunk a 13.1 formulában. Ennek oka, hogy így kapunk olyan statisztikát, amelynek várható értéke megegyezik a mintaelemek szórásnégyzetével:

13.2.4. Állítás. Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta, amelyre teljesül, hogy $\mathbb{E}(X_i^2) < \infty$. Ekkor

$$\mathbb{E}((S_n^*)^2) = \mathbb{D}^2(X_1).$$

Bizonyítás. A bizonyításban szereplő számítások kissé hosszadalmasak, de a várható érték linearitásán kívül kizárólag azt a tulajdonságot használják, hogy független valószínűségi változók szorzatának várható értéke megegyezik a várható értékek szorzatával (6.1.4. Állítás). Először is,

$$\begin{aligned} \mathbb{E}[(S_n^*)^2] &= \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}(X_i - \overline{X}_n)^2 = \frac{1}{n-1} \sum_{i=1}^n \left(\mathbb{E}(X_i^2) - 2\mathbb{E}(X_i \overline{X}_n) + \mathbb{E}(\overline{X}_n^2) \right) \\ &= \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}(X_i^2) - \frac{2}{n-1} \sum_{i=1}^n \mathbb{E}(X_i \overline{X}_n) + \frac{n}{n-1} \mathbb{E}(\overline{X}_n^2). \end{aligned}$$

⁴Emlékeztetünk arra, hogy permutációnak az $\{1, \dots, n\}$ halmaz önmagára való kölcsönösen egyértelmű leképezéseit (vagyis bijekcióit) nevezzük, lásd például a Szeszlér Dávid-féle BSZ1 jegyzetben (http://cs.bme.hu/bsz1/jegyzet/bsz1_jegyzet.pdf) a determináns témakörénél, illetve az idei Valszám diasorban a kombinatorika témakörénél a permutáció ekvivalens definíciójaként.

⁵Az „empirikus” szó jelentése: „tapasztalati”. Néha használatos a „korrigált tapasztalati szórásnégyzet” kifejezés is.

Számoljuk ki külön-külön a jobb oldalon szereplő három tag várható értékét! Az első tag:

$$\frac{1}{n-1} \sum_{i=1}^n \mathbb{E}(X_i^2) = \frac{n}{n-1} \mathbb{E}(X_1^2),$$

mivel X_1, \dots, X_n azonos eloszlásúak. A második tag -1 -szerese:

$$\begin{aligned} \frac{2}{n-1} \sum_{i=1}^n \mathbb{E}(X_i \bar{X}_n) &= \frac{2}{n-1} \left(\sum_{i=1}^n \frac{1}{n} \mathbb{E}(X_i^2) + \frac{1}{n} \sum_{j \neq i} \mathbb{E}(X_i X_j) \right) = \frac{2}{n-1} \left(\mathbb{E}(X_1^2) + (n-1) \mathbb{E}(X_1)^2 \right) \\ &= \frac{2}{n-1} \mathbb{E}(X_1^2) + 2 \mathbb{E}(X_1)^2, \end{aligned}$$

mivel X_1, \dots, X_n azonos eloszlásúak és függetlenek. A harmadik tag:

$$\begin{aligned} \frac{n}{n-1} \mathbb{E}(\bar{X}_n^2) &= \frac{n}{(n-1)n^2} \mathbb{E} \left[\left(\sum_{i=1}^n X_i \right)^2 \right] = \frac{1}{n(n-1)} \mathbb{E} \left(\sum_{i=1}^n X_i^2 + 2 \sum_{1 \leq i < j \leq n} X_i X_j \right) \\ &= \frac{1}{n(n-1)} \left(\sum_{i=1}^n (\mathbb{E}(X_1^2) + (n-1) \mathbb{E}(X_1)^2) \right) \\ &= \frac{1}{n-1} \mathbb{E}(X_1^2) + \mathbb{E}(X_1)^2. \end{aligned}$$

mivel X_1, \dots, X_n függetlenek és azonos eloszlásúak. Összességében tehát

$$\mathbb{E}[(S_n^*)^2] = \mathbb{E}(X_1^2) \left(\frac{n}{n-1} - \frac{2}{n-1} + \frac{1}{n-1} \right) + \mathbb{E}(X_1)^2 (-2 + 1) = \mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2 = \mathbb{D}^2(X_1).$$

□

13.2.5. *Megjegyzés.* Láthatjuk, hogy ha az (13.1) formulában $\frac{1}{n-1}$ helyett $\frac{1}{n}$ -nel osztanánk, akkor az így kapott statisztika szórásnégyzete (a várható érték linearitása miatt) nem $\mathbb{D}^2(X_1)$, hanem annak $\frac{n-1}{n}$ -szerese lenne. Az $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ statisztikát *empirikus szórásnégyzetnek* (a négyzetgyökét pedig *empirikus szórásnak*) nevezik, $(S_n^*)^2$ -et pedig azért hívják *korrigáltnak*, mert szórásnégyzete megegyezik $\mathbb{D}^2(X_1)$ -gyel. A magyar nyelvű szakirodalomban ez a bevett szóhasználat. Idegen nyelvű művekben előfordulhat az is, hogy $(S_n^*)^2$ -et hívják empirikus szórásnégyzetnek, de ettől mi eltekintünk.

Szintén alapstatisztika a minta **módusza**, azaz a leggyakoribb érték a mintában. Ha több ilyen érték van, akkor mindegyiket módusznak tekintjük.

További alapstatisztikák definiálása érdekében először bevezetjük a rendezett minta fogalmát.

13.2.6. Definíció. Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta. Legyen $(X_1^*, X_2^*, \dots, X_n^*)$ az X_1, \dots, X_n mintaelemek egy olyan felsorolása, hogy

$$X_1^* \leq X_2^* \leq \dots \leq X_n^*.$$

Ekkor $(X_1^*, X_2^*, \dots, X_n^*)$ -t **rendezett mintának** nevezzük.

13.2.7. *Példa.* Legyen $\mathbf{x} = (x_1, \dots, x_8) = (1, 2, 1, 3, 4, 6, 5, 2)$ egy szabályos dobókocka nyolcszori feldobásából származó $\mathbf{X} = (X_1, \dots, X_8)$ minta egy realizációja. Ekkor ehhez a realizációhoz az (X_1^*, \dots, X_8^*) rendezett minta $(x_1^*, \dots, x_8^*) = (1, 1, 2, 2, 3, 4, 5, 6)$ realizációja tartozik. Tehát a rendezett minta realizációja az eredeti (nem rendezett) minta realizációjából egyértelműen megkapható, akkor is, ha vannak ismétlődő elemek.

Fontos kiemelni, hogy X_1^*, \dots, X_n^* nem függetlenek (ezt formálisan nem bizonyítjuk, de érezhető, hogy pl. a legkisebb mintaelem eloszlása befolyásolja az összes többi rendezett mintaelem eloszlását, hiszen azoknak a legkisebhnél nagyobbknak kell lenniük) és jellemzően nem is azonos eloszlásúak. A rendezett mintaelemek segítségével definiálható az empirikus medián, ami páratlan elemszámú mintánál a rendezett minta középső elemét, páros elemszámúnál pedig a két középső számtani közepét jelenti:

13.2.8. Definíció. Az $\mathbf{X} = (X_1, \dots, X_n)$ n elemű fae. minta **empirikus mediánján** az $m_{n,\mathbf{X}} = X_{k+1}^*$ rendezett mintaelemet értjük, ha $n = 2k + 1$, és az $m_{n,\mathbf{X}} = \frac{X_k^* + X_{k+1}^*}{2}$ kifejezést, ha $n = 2k$.

13.2.9. *Példa.* Ha egy n elemű minta realizációja $(x_1, \dots, x_4) = (\sqrt{2}, \pi, -1, e)$, akkor a rendezett minta realizációja $(x_1^*, x_2^*, x_3^*, x_4^*) = (-1, \sqrt{2}, e, \pi)$ és empirikus mediánja $\frac{\sqrt{2}+e}{2}$. Ha az n elemű minta realizációja $(x_1, x_2, x_3) = (-1, 3, 2)$, akkor a rendezett minta realizációja $(x_1^*, x_2^*, x_3^*) = (-1, 2, 3)$ és az empirikus medián 2 .

13.3. Empirikus eloszlásfüggvény

Ebben az alfejezetben azt tesszük fel, hogy az X_1, \dots, X_n valószínűségi változók függetlenek és azonos eloszlásúak valamilyen ismeretlen eloszlással, ennek eloszlásfüggvénye legyen $x \mapsto F(x) = \mathbb{P}(X_i < x)$. (A ϑ paraméter itt sem fog szerepet játszani.) Az X_i -k eloszlásáról semmilyen ismeretünk nincs, például azt sem tudjuk, hogy az eloszlásuk diszkrét-e, folytonos vagy egyik sem⁶. Hogyan tudjuk az $\mathbf{X} = (X_1, \dots, X_n)$ egy $\mathbf{x} = (x_1, \dots, x_n)$ realizációja alapján az F eloszlásfüggvényt „megbecsülni”, úgy, hogy $n \rightarrow \infty$ esetén a valódi eloszlásfüggvényhez tartó eredményt kapjunk? Ezt a célt szolgálja az empirikus eloszlásfüggvény.

13.3.1. Definíció. \blacktriangle Legyen $n \in \mathbb{N}$ és legyenek X_1, \dots, X_n fae. valószínűségi változók. Az $\mathbb{R} \rightarrow \mathbb{R}$,

$$x \mapsto F_n^*(x) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i < x\}} = \frac{1}{n} |\{i \in \{1, \dots, n\} \mid X_i < x\}| \quad (13.2)$$

függvényt az $\underline{X} = (X_1, \dots, X_n)$ n elemű fae. mintához tartozó **empirikus eloszlásfüggvénynek** nevezzük.

Fontos az elején rögzíteni, hogy F_n^* egy *véletlen* függvény: a függvény értékei az X_1, \dots, X_n valószínűségi változók értékeitől függenek.

A (13.2) egyenletben szereplő definíciót az (X_1^*, \dots, X_n^*) rendezett minta (lásd 13.2.6. Definíció) segítségével a következő egyszerű, explicit formulával is megfogalmazhatjuk:

$$F_n^*(x) = \begin{cases} 0, & \text{ha } x \leq X_1^*, \\ \frac{k}{n}, & \text{ha } X_k^* < x \leq X_{k+1}^*, \quad k = 1, \dots, n-1, \\ 1, & \text{ha } x > X_n^*. \end{cases} \quad (13.3)$$

Ez az átírás azt is mutatja, hogy az $\mathbf{X} = (X_1, \dots, X_n)$ bármely fix $\mathbf{x} = (x_1, \dots, x_n)$ realizációja esetén valóban eloszlásfüggvényt kapunk: a $-\infty$ -ben vett határérték valóban 0 , az ∞ -ben vett határérték 1 , és a függvény monoton növekvő és (a (13.2) definícióban szereplő szigorú egyenlőségeknek köszönhetően) balról folytonos.

A következő állítás megmutatja, hogy az empirikus eloszlásfüggvény „jó becslése” a valódi eloszlásfüggvénynek: várható értékben minden pontban megegyezik vele, a szórása pedig n függvényében csökken. Továbbá az empirikus eloszlásfüggvény minden pontban (1 valószínűséggel) a valódi eloszlásfüggvényhez konvergál $n \rightarrow \infty$ esetén, vagyis ha kellően nagy elemszámú mintát veszünk, akkor egy tetszőleges $x \in \mathbb{R}$ pontban (1 valószínűséggel) tetszőlegesen megközelíthetjük az eloszlásfüggvény $F(x)$ értékét az empirikus eloszlásfüggvény $F_n^*(x)$ értékével.

13.3.2. Állítás. Legyen $n \in \mathbb{N}$ és legyenek X_1, \dots, X_n független, azonos eloszlású valószínűségi változók F eloszlásfüggvényrel. Ekkor minden $x \in \mathbb{R}$ esetén teljesül, hogy

$$(1) \quad \mathbb{E}(F_n^*(x)) = F(x),$$

⁶Utóbbira példa a következőképpen definiált X valószínűségi változó: dobjunk fel egy pénzérmét, és ha fejet kapunk, dobjunk fel egy kockát és legyen X a dobás eredménye, ha pedig írást, akkor legyen X egy $U(0;1)$ eloszlású, a kockadobástól független véletlen szám.

$$(2) \mathbb{D}^2(F_n^*(x)) = \frac{F(x)(1-F(x))}{n}, \text{ és}$$

$$(3) \mathbb{P}(\lim_{n \rightarrow \infty} F_n^*(x) = F(x)) = 1.$$

Bizonyítás. Legyen $x \in \mathbb{R}$. Az (13.2) definíció alapján észrevehetjük, hogy $n \cdot F_n^*(x)$ a független, azonos eloszlású $\mathbb{1}_{\{X_i < x\}}$ indikátorváltozók összege, amelyek paramétere (és egyben várható értéke) $\mathbb{P}(X_i < x) = F(x)$. Emiatt $nF_n^*(x) \sim B(n; F(x))$. A binomiális eloszlás, illetve a várható érték és a szórásnégyzet ismert tulajdonságai miatt tehát

$$\mathbb{E}(F_n^*(x)) = \frac{1}{n} \mathbb{E}(nF_n^*(x)) = \frac{1}{n} nF(x) = F(x)$$

és

$$\mathbb{D}^2(F_n^*(x)) = \frac{1}{n^2} \mathbb{D}^2(nF_n^*(x)) = \frac{1}{n^2} nF(x)(1-F(x)) = \frac{F(x)(1-F(x))}{n},$$

amint állítottuk. A (3) állítás igazolása érdekében alkalmazzuk a független, azonos eloszlású $\mathbb{1}_{\{X_i < x\}}$ indikátorváltozókra a nagy számok erős törvényét (azaz a 9.2.1. Tételt). Ezek várható értéke

$$\mathbb{E}(\mathbb{1}_{\{X_1 < x\}}) = \mathbb{P}(X_1 < x) = F(x)$$

(és a szórásnégyzetük véges, pontosabban $F(x)(1-F(x))$ -szel egyenlő), ezért a 9.2.1. Tétel szerint az átlaguk 1 valószínűséggel a várható értékükhöz tart:

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} F_n^*(x) = F(x)\right) = \mathbb{P}\left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{X_i < x\}} = \mathbb{E}(\mathbb{1}_{\{X_1 < x\}})\right) = 1.$$

Ezzel a (3) állítást is beláttuk. □

A 13.3.2 Állítás (3) részénél több is igaz: az empirikus eloszlásfüggvény nemcsak hogy minden pontban (1 valószínűséggel) konvergál a valódihoz, hanem (1 valószínűséggel) az egész számegyenesen egyenletesen konvergál hozzá, azaz a két függvény közötti távolság nullához tart. Ezt fogalmazza meg a következő tétel:

13.3.3. Tétel (Glivenko–Cantelli-tétel). *Legyen $n \in \mathbb{N}$ és legyenek X_1, \dots, X_n független, azonos eloszlású valószínűségi változók F eloszlásfüggvénnyel. Ekkor az $x \mapsto F_n^*(x)$ függvény 1 valószínűséggel egyenletesen konvergál $x \mapsto F(x)$ -hez, azaz*

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_n^*(x) - F(x)| = 0\right) = 1.$$

Miért erősebb ez a tétel a 13.3.2 Állítás (3) részénél? A 13.3.2. Állításban az a 0 valószínűségi esemény, ahol a (3)-beli konvergencia nem teljesül, elvileg x -től is függ. A Glivenko–Cantelli-tétel alapján azonban ha vesszük ezen események (megszámlálhatatlanul végtelenül sok nulla valószínűségű esemény!) unióját, akkor még mindig 0 valószínűségű eseményt kapunk. A Glivenko–Cantelli tételt ezen az előadáson nem bizonyítjuk, a bizonyítás megtalálható szinte bármelyik matematikai statisztika tankönyben.

14. Becslésméleti módszerek

Ebben a fejezetben ismét azt a helyzetet tekintjük, amikor az X_1, \dots, X_n mintaelemek eloszlása egy $\vartheta \in \theta$ paramétertől függ, és az ismeretlen ϑ paraméterre vagy annak valamilyen $\psi(\vartheta)$ függvényére⁷ kívánunk egy n elemű fae. minta valamilyen statisztikája alapján minél jobb becslést adni. Az itt részletezett becslési módszerek idő hiányában nem lesznek kimerítőek; az érdeklődőknek a mérnökinformatikus és gazdaságinformatikus MSc képzés Matematikai statisztika tárgyát ajánljuk.

⁷A ψ is görög betű, pszí-nek ejtjük.

14.1. Becslések tulajdonságai, átlag és szórás becslései

14.1.1. Definíció. \blacktriangle Legyen adott egy $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta, ahol a mintaelemek eloszlása egy ismeretlen $\vartheta \in \theta \subseteq \mathbb{R}^d$ paramétertől függ, és legyen $\psi: \mathbb{R}^d \rightarrow \mathbb{R}$ egy függvény. Legyen továbbá $T(\mathbf{X}) = T(X_1, \dots, X_n)$ a minta egy statisztikája.

Azt mondjuk, hogy a $T(\mathbf{X})$ statisztika

- (1) a $\psi(\vartheta)$ paraméterfüggvény egy **torzítatlan** becslése, ha minden $\vartheta \in \theta$ esetén

$$\mathbb{E}_\vartheta(T(X_1, \dots, X_n)) = \psi(\vartheta),$$

ahol \mathbb{E}_ϑ jelöli a \mathbb{P}_ϑ valószínűség szerinti várható értéket,

- (2) a $\psi(\vartheta)$ paraméterfüggvény egy **aszimptotikusan torzítatlan** becslése, ha minden $\vartheta \in \theta$ esetén

$$\lim_{n \rightarrow \infty} \mathbb{E}_\vartheta(T(X_1, \dots, X_n)) = \psi(\vartheta),$$

- (3) a $\psi(\vartheta)$ paraméterfüggvény egy **erősen konzisztens** becslése⁸, ha minden $\vartheta \in \theta$ esetén

$$\mathbb{P}_\vartheta\left(\lim_{n \rightarrow \infty} T(X_1, \dots, X_n) = \psi(\vartheta)\right) = 1. \quad (14.1)$$

- (4) Ha $T(\mathbf{X})$ és a minta egy másik $T'(\mathbf{X}) = T'(X_1, \dots, X_n)$ statisztikája is torzítatlan becslése $\psi(\vartheta)$ -nak, akkor azt mondjuk, hogy $T(\mathbf{X})$ **legalább annyira hatásos, mint $T'(\mathbf{X})$** , ha

$$\mathbb{D}_\vartheta^2(T(\mathbf{X})) \leq \mathbb{D}_\vartheta^2(T'(\mathbf{X})),$$

ahol \mathbb{D}_ϑ^2 jelöli a \mathbb{P}_ϑ szerinti szórásnégyzetet.⁹ Ha $T(\mathbf{X})$ legalább annyira hatásos, mint $\psi(\vartheta)$ bármely torzítatlan becslése, akkor azt mondjuk, hogy $T(\mathbf{X})$ a $\psi(\vartheta)$ paraméterfüggvény egy **hatásos** becslése.

Nyilvánvaló, hogy ha egy statisztika torzítatlan becslést ad egy adott paraméterfüggvényre, akkor aszimptotikusan torzítatlan becslést is ad. Lássunk példát torzítatlan, aszimptotikusan torzítatlan, illetve erősen konzisztens becslésekre!

14.1.2. *Példa.* A 14.1.1. Definícióban szereplő szituációban egy lehetséges ψ függvény a

$$\theta \ni \vartheta \mapsto \psi(\vartheta) = \mathbb{E}_\vartheta(X_1),$$

ha az $\mathbb{E}_\vartheta(X_1)$ minden $\vartheta \in \theta$ esetén létezik (azaz ha $\mathbb{E}_\vartheta(|X_1|) < \infty$ minden $\vartheta \in \theta$ esetén teljesül). Ebben az esetben a

$$T(X_1, \dots, X_n) = \bar{X}_n = \frac{X_1 + \dots + X_n}{n}$$

mintaátlag a $\psi(\vartheta) = \mathbb{E}_\vartheta(X_1)$ paraméterfüggvény egy torzítatlan becslése, mivel minden $\vartheta \in \theta$ esetén

$$\mathbb{E}_\vartheta(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}_\vartheta(X_i) = \mathbb{E}_\vartheta(X_1)$$

(lásd a 13.2.2. Példát). Továbbá a nagy számok erős törvénye miatt ha feltesszük, hogy ϑ a valódi paraméter, akkor a mintaátlag 1 valószínűséggel $\psi(\vartheta) = \mathbb{E}_\vartheta(X_1)$ -hez konvergál. Formálisan: minden $\vartheta \in \theta$ esetén

$$\mathbb{P}_\vartheta\left(\lim_{n \rightarrow \infty} \bar{X}_n = \mathbb{E}_\vartheta(X_1)\right) = 1.$$

Emiatt $T(X_1, \dots, X_n) = \bar{X}_n$ erősen konzisztens becslése is a $\psi(\vartheta) = \mathbb{E}_\vartheta(X_1)$ paraméternek.

⁸Léteznek más fogalmak is becslések konzisztenciájára, például a *gyenge konzisztencia*, ami ugyanazt fejezi ki, mint az erős konzisztencia, csak az (14.1) egyenletben szereplő 1 valószínűségű konvergencia helyett valószínűségben való konvergenciával. Ha azt mondjuk, hogy egy becslés *konzisztens*, az alatt azt értjük, hogy gyengén konzisztens. Amint a nagy számok törvényeinek témájánál már láttuk, az 1 valószínűségű konvergenciából következik a valószínűségben való konvergencia, így minden erősen konzisztens becslés konzisztens is. Vannak más konzisztenciafogalmak is, amiket itt nem részletezünk, például a *négyzetes középben való konzisztencia*, lásd a mérnök- és gazdaságinformatikus MSc képzés Matematikai statisztika tárgyát.

⁹Hasonlóan és értelemszerűen vezethetők be a „(szigorúan) hatásosabb”, a „legfeljebb annyira hatásos” és a „(szigorúan) kevésbé hatásos” relációk két torzítatlan becslés között.

14.1.3. *Példa.* Egy másik lehetséges ψ függvény a

$$\theta \ni \vartheta \mapsto \psi(\vartheta) = \mathbb{D}_{\vartheta}^2(X_1),$$

ha $\mathbb{E}_{\vartheta}(X_1^2)$ minden $\vartheta \in \theta$ esetén véges. Ebben az esetben a

$$T(X_1, \dots, X_n) = (S_n^*)^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \overline{X_n})^2$$

korrigált empirikus szórásnégyzet torzítatlan becslése a

$$\psi(\vartheta) = \mathbb{D}_{\vartheta}^2(X_1)$$

paraméterfüggvénynek. Valóban, a 13.2.4. Állítás szerint minden $\vartheta \in \theta$ esetén teljesül, hogy

$$\mathbb{E}_{\vartheta}((S_n^*)^2) = \mathbb{D}_{\vartheta}^2(X_1).$$

Így azt is láthatjuk, hogy a

$$T'(X_1, \dots, X_n) = S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X_n})^2 \quad (14.2)$$

statisztika nem torzítatlan becslése $\mathbb{D}_{\vartheta}^2(X_1)$ -nek, hiszen (a várható érték linearitása miatt, lásd a 13.2.5. Megjegyzést is)

$$\mathbb{E}_{\vartheta}[S_n^2] = \frac{n-1}{n} \mathbb{E}_{\vartheta}[(S_n^*)^2] = \frac{n-1}{n} \mathbb{D}_{\vartheta}^2(X_1).$$

Ugyanakkor $T'(X_1, \dots, X_n) = S_n^2$ is aszimptotikusan torzítatlan becslése $\mathbb{D}_{\vartheta}^2(X_1)$ -nek, mert

$$\lim_{n \rightarrow \infty} \mathbb{E}_{\vartheta}[S_n^2] = \lim_{n \rightarrow \infty} \frac{n-1}{n} \mathbb{D}_{\vartheta}^2(X_1) = \mathbb{D}_{\vartheta}^2(X_1).$$

14.1.4. *Példa.* Egy becslés torzítatlansága önmagában nem garantálja a becslés jóságát, például a hatásosságát; előfordulhat, hogy egy másik torzítatlan becslés szigorúan hatásosabb nála. Legyen pl. (X_1, X_2, X_3) fae. minta $U(\vartheta; 1 + \vartheta)$ eloszlással, ahol $\vartheta \geq 0$ ismeretlen paraméter, és tekintsük a $T(X_1, X_2, X_3) = X_2^*$ rendezett mintaelemet mint (X_1, X_2, X_3) egy statisztikáját. Gyakorlaton korábban beláttuk, hogy $\vartheta = 0$ esetén X_2^* sűrűségfüggvénye

$$f_{X_2^*}(x) = \begin{cases} 6x - 6x^2, & \text{ha } 0 < x < 1, \\ 0, & \text{különben.} \end{cases}$$

Azt is kiszámoltuk már, hogy $\mathbb{E}(X_2^*) = 1/2$, illetve a sűrűségfüggvény alapján könnyen kiszámolható, hogy $\mathbb{D}^2(X_2^*) = 0,05$. Általános $\vartheta > 0$ esetén X_1, X_2, X_3 -mat úgy is megkaphatjuk, hogy a $(0, 1)$ intervallumon egyenletes fae. Y_1, Y_2, Y_3 val. változók mindegyikéhez ϑ -t adunk (közülük a 2. rendezett mintaelemet jelölje Y_2^*). Ezért általános ϑ esetén $\mathbb{E}(X_2^*) = \mathbb{E}(Y_2^* + \vartheta) = 1/2 + \vartheta = \mathbb{E}_{\vartheta}(X_1)$ és $\mathbb{D}^2(X_2^*) = \mathbb{D}^2(Y_2^* + \vartheta) = \mathbb{D}^2(Y_2^*) = 0,05$. Ezért tehát a $T(X_1, X_2, X_3) = X_2^*$ statisztika torzítatlan becslése a $\psi(\vartheta) = \mathbb{E}_{\vartheta}(X_1) = \vartheta + 1/2$ paraméterfüggvénynek. Ugyanakkor ez a statisztika nem hatásos: pl. $S(X_1, X_2, X_3) = \overline{X_3}$ is torzítatlan becslése $\psi(\vartheta)$ -nak, a szórásnégyzete viszont

$$\mathbb{D}^2(\overline{X_3}) = \mathbb{D}^2\left(\frac{X_1 + X_2 + X_3}{3}\right) = \frac{1}{3} \mathbb{D}^2(X_1) = \frac{1}{12 \cdot 3} = \frac{1}{36} < 0,05,$$

vagyis kisebb, mint $T(X_1, X_2, X_3)$ -é. Tehát $S(X_1, X_2, X_3)$ hatásosabb, mint $T(X_1, X_2, X_3)$.

14.2. Maximum likelihood-becslés

A paraméterbecslés egyik módszere a *maximum likelihood módszer*, amelyet ebben az alfejezetben ismertetünk. A módszer lényege, hogy az $\mathbf{X} = (X_1, \dots, X_n)$ n elemű fae. minta $\mathbf{x} = (x_1, \dots, x_n)$ realizációjára alapján meghatározzuk azt a $\vartheta_* = \vartheta_*(x_1, \dots, x_n)$ -gal jelölt paramétert, amely esetén a legvalószínűbb, hogy (közelítőleg) ezt a realizációt kapjuk. Miután meghatároztuk a $T(x_1, \dots, x_n) = \vartheta_*(x_1, \dots, x_n)$ függvényt, a becslés eredménye a

$$T(X_1, \dots, X_n) = \vartheta_*(X_1, \dots, X_n)$$

statisztika lesz. (A becslés eredménye tehát a véletlen minta egy függvénye, nem pedig a realizációé – ezt érdemes mindig észben tartani!)

Annak a matematikai leírására, hogy egy adott paraméter esetén milyen valószínűséggel kapunk (közelítőleg) egy adott x_1, \dots, x_n realizációt, bevezetjük a *likelihood-függvény* fogalmát. A következő definíció használja a 13.1.1. Definícióban bevezetett p_ϑ súlyfüggvényt és f_ϑ sűrűségfüggvényt.

14.2.1. Definíció. \blacktriangle Legyenek X_1, \dots, X_n fae. valószínűségi változók, amelyek eloszlása egy $\vartheta \in \theta \subseteq \mathbb{R}^d$ paramétertől függ.

(1) Ha X_1 eloszlása bármely $\vartheta \in \theta$ esetén diszkrét és p_ϑ jelöli a ϑ paraméterhez tartozó súlyfüggvényt, akkor az $\mathbf{X} = (X_1, \dots, X_n)$ n elemű fae. minta bármely (x_1, \dots, x_n) realizációja esetén legyen

$$L_\vartheta(x_1, \dots, x_n) = \prod_{i=1}^n p_\vartheta(x_i).$$

(2) Ha X_1 eloszlása bármely $\vartheta \in \theta$ esetén folytonos és $f_\vartheta(x)$ jelöli a ϑ paraméterhez tartozó sűrűségfüggvényt, akkor az $\mathbf{X} = (X_1, \dots, X_n)$ n elemű fae. minta bármely (x_1, \dots, x_n) realizációja esetén legyen

$$L_\vartheta(x_1, \dots, x_n) = \prod_{i=1}^n f_\vartheta(x_i).$$

Az $\mathbb{R}^n \rightarrow \mathbb{R}$, $(x_1, \dots, x_n) \mapsto L_\vartheta(x_1, \dots, x_n)$ függvényt mindkét esetben (a ϑ paraméterhez tartozó) **likelihood-függvénynek** nevezzük.

14.2.2. *Megjegyzés.* Diszkrét esetben az L_ϑ likelihood-függvény a mintaelemek együttes eloszlását írja le abban az esetben, ha ϑ a valódi paraméter. Valóban, ha A egy esemény, jelölje $\mathbb{P}_\vartheta(A)$ az A valószínűségét, feltéve, hogy ϑ a valódi paraméter. Ezzel a jelöléssel az X_1, \dots, X_n függetlensége miatt azt kapjuk, hogy

$$L_\vartheta(x_1, \dots, x_n) = \prod_{i=1}^n p_\vartheta(x_i) = \prod_{i=1}^n \mathbb{P}_\vartheta(X_i = x_i) = \mathbb{P}_\vartheta(X_1 = x_1, \dots, X_n = x_n).$$

Folytonos esetben pedig, ha ϑ a valódi paraméter, akkor mivel $f_\vartheta(x_i)$ az X_i valószínűségi változó sűrűségfüggvényének értéke az x_i helyen,

$$L_\vartheta(x_1, \dots, x_n) = \prod_{i=1}^n f_\vartheta(x_i)$$

a függetlenség miatt megegyezik az X_1, \dots, X_n valószínűségi változók együttes sűrűségfüggvényével. A likelihood-függvény szemléletes jelentése tehát ebben az esetben

$$\begin{aligned} L_\vartheta(x_1, \dots, x_n) &= \prod_{i=1}^n f_\vartheta(x_i) = \prod_{i=1}^n \lim_{\Delta x_i \downarrow 0} \frac{1}{\Delta x_i} \mathbb{P}_\vartheta(X_i \in [x_i, x_i + \Delta x_i)) \\ &= \lim_{\Delta x_1 \downarrow 0, \dots, \Delta x_n \downarrow 0} \frac{1}{\Delta x_1 \dots \Delta x_n} \mathbb{P}_\vartheta(X_1 \in [x_1, x_1 + \Delta x_1), \dots, X_n \in [x_n, x_n + \Delta x_n)). \end{aligned}$$

Ezáltal nyer értelmet az a pongyola megfogalmazás, hogy $L_\vartheta(x_1, \dots, x_n)$ azt adja meg, hogy milyen valószínűséggel kapjuk (közelítőleg) (x_1, \dots, x_n) -t az (X_1, \dots, X_n) realizációjaként, ha ϑ a valódi paraméter. Diszkrét esetben közelítésre nincs is szükség: $L_\vartheta(x_1, \dots, x_n)$ pontosan annak a valószínűsége, hogy az (x_1, \dots, x_n) realizációt kapjuk. Folytonos esetben ugyanez a valószínűség azonban 0, ezért $L_\vartheta(x_1, \dots, x_n)$ valóban csak azt adja meg, hogy milyen valószínűséggel kapunk egy (x_1, \dots, x_n) -hez nagyon közeli realizációt.

14.2.3. Definíció. Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta, ahol a mintaelemek egy $\vartheta \in \theta \subset \mathbb{R}^d$ paramétertől függő eloszlással rendelkeznek, és tegyük fel, hogy (minden $\vartheta \in \theta$ esetén) létezik az L_ϑ likelihood-függvény. Legyen $\mathbf{x} = (x_1, \dots, x_n)$ az $\mathbf{X} = (X_1, \dots, X_n)$ egy realizációja.

A **maximum likelihood becslés** ϑ -ra az (x_1, \dots, x_n) **realizáció alapján** az a $\vartheta_*(x_1, \dots, x_n) \in \theta$ paraméter, amely esetén a likelihood-függvény $L_\vartheta(x_1, \dots, x_n)$ értéke maximális. Formálisan kifejezve,

$$\vartheta_*(x_1, \dots, x_n) = \arg \max_{\vartheta \in \theta} L_\vartheta(x_1, \dots, x_n). \quad (14.3)$$

Más szavakkal, $\vartheta_*(x_1, \dots, x_n)$ az a $\vartheta \in \theta$ paraméter, amelyre teljesül, hogy bármely $\vartheta' \in \theta$ esetén

$$L_{\vartheta_*(x_1, \dots, x_n)}(x_1, \dots, x_n) \geq L_{\vartheta'}(x_1, \dots, x_n). \quad (14.4)$$

Elvileg elképzelhető, hogy a (14.4) egyenlet $\vartheta_*(x_1, \dots, x_n)$ több értéke esetén is teljesül, vagyis a maximum nem szigorú és a maximumhely nem egyértelmű, de a gyakorlatban ilyen probléma ritkán fordul elő. Mint ahogy az sem jellemző probléma, hogy a likelihood-függvénynek egyáltalán nincs maximuma.

Amint már utaltunk rá, a maximum likelihood becslés végeredményeként egy statisztikát kapunk, vagyis nem az (x_1, \dots, x_n) realizáció, hanem az $\mathbf{X} = (X_1, \dots, X_n)$ valószínűségi vektorváltozó egy függvényét. Erről szól a következő definíció:

14.2.4. Definíció. A 14.2.3. Definícióban leírt helyzetben vezessük be a

$$T: \mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto \vartheta_*(x_1, \dots, x_n)$$

függvényt. Ekkor a

$$T(X_1, \dots, X_n) = \vartheta_*(X_1, \dots, X_n)$$

statisztika neve: **maximum likelihood becslés** ϑ -ra.

Hogyan lehet a maximum likelihood-becslést a gyakorlatban kiszámítani? A 14.3 egyenletben szereplő, ϑ szerinti maximalizálási probléma egy egydimenziósszélsőérték-számítási feladat, ami elvileg az első félévés analízistudásunkkal megoldható, ha a likelihood-függvény differenciálható. Azonban a likelihood-függvény még viszonylag egyszerű súly- vagy sűrűségfüggvénnyel rendelkező eloszlások esetén is egy n tényező szorzat, amelynek a szorzat- és láncszabály alkalmazásával való deriválása hosszadalmas és kellemetlen művelet. Gyakran segít, ha a likelihood-függvény helyett annak a logaritmusát próbáljuk ϑ szerint maximalizálni, ugyanis a logaritmus a szorzatokat összegekké alakítja, amit lényegesen könnyebb deriválni. Ezért vezetjük be a következő fogalmat.

14.2.5. Definíció. Ha $(x_1, \dots, x_n) \mapsto L_\vartheta(x_1, \dots, x_n)$ egy likelihood-függvény, akkor az

$$\mathbb{R}^n \rightarrow \mathbb{R}, \quad (x_1, \dots, x_n) \mapsto l_\vartheta(x_1, \dots, x_n) = \ln L_\vartheta(x_1, \dots, x_n)$$

függvényt **log-likelihood-függvénynek** nevezzük.

Jegyezzük meg, hogy a log-likelihood-függvény minden likelihood-függvény esetén jóldefiniált, mert a likelihood-függvény definíciójában szereplő súly- és sűrűségfüggvények minden realizációban vett értéke pozitív. (Ezért is volt szükségünk a lényeges értékek halmaza és realizáció – viszonylag bonyolult – definíciójára!) A log-likelihood-függvény pedig ugyanott veszi fel a maximumát, mint a likelihood-függvény:

14.2.6. Állítás. Ha $(x_1, \dots, x_n) \mapsto L_\vartheta(x_1, \dots, x_n)$ egy likelihood-függvény és $(x_1, \dots, x_n) \mapsto l_\vartheta(x_1, \dots, x_n) = \ln L_\vartheta(x_1, \dots, x_n)$ az ehhez tartozó log-likelihood-függvény, akkor $\vartheta \in \theta$ esetén $L_\vartheta(x_1, \dots, x_n)$ pontosan akkor maximális ϑ -ban, ha $l_\vartheta(x_1, \dots, x_n)$ maximális ϑ -ban.

Bizonyítás. Az állítás a (természetes alapú) logaritmusfüggvény szigorú monotonitásából következik. \square

Ha a likelihood-függvényt log-likelihood-függvénné alakítjuk, akkor a maximum likelihood-becslésre az alábbi eljárást adhatjuk (ez az eljárás csak akkor működik, ha a log-likelihood-függvény differenciálható):

Eljárás (Maximum likelihood-becslés log-likelihood-függvénnel).

1. Rögzített x_1, \dots, x_n realizáció és $\vartheta \in \theta$ paraméter esetén meghatározzuk a likelihood-függvény $L_\vartheta(x_1, \dots, x_n)$ értékét.
2. Vesszük ennek (e alapú) logaritmusát, így megkapjuk $l_\vartheta(x_1, \dots, x_n)$ értékét.
3. Az a $\vartheta \mapsto l_\vartheta(x_1, \dots, x_n)$ függvényt deriváljuk (ϑ szerint).
4. Megvizsgáljuk, hogy a derivált hol egyenlő 0-val.
5. Megvizsgáljuk, hogy a derivált zérushelyei lokális maximumhelyek-e. (Emlékeztetünk arra, hogy lokális maximumhellyel állunk szemben, ha a második derivált az adott helyen negatív, vagy ha az első derivált pozitívól negatívba vált előjelet az adott pontban.) Ha esetleg több lokális maximumhely is van, akkor ezek közül az lesz a globális maximumhely, ahol a likelihood-függvény értéke a legnagyobb.¹⁰
6. Tegyük fel, hogy egy egyértelmű $\vartheta_*(x_1, \dots, x_n)$ globális maximumhelyet kaptunk. Ekkor $\vartheta_*(X_1, \dots, X_n)$ lesz a keresett maximum likelihood becslés.

Az alábbiakban bemutatjuk a módszert egy diszkrét és egy folytonos példán, egyéb példákkal pedig a gyakorlatokon (és a vizsgákon) találkozhatunk.

14.2.7. *Példa* (Maximum likelihood módszer, diszkrét eset: Poisson-eloszlás). A fenti eljárást alkalmazzuk a Poisson-eloszlás esetére: tegyük fel, hogy az X_1, \dots, X_n mintaelemek peremeloszlása Poisson-eloszlás ismeretlen $\vartheta > 0$ paraméterrel. Legyen (x_1, \dots, x_n) az (X_1, \dots, X_n) egy realizációja, vagyis egy n darab nemnegatív egész számból álló vektor. A továbbiakban először tegyük fel, hogy az x_1, \dots, x_n számok közül legalább az egyik nem nulla.

1. Mivel

$$p_\vartheta(x_i) = \mathbb{P}_\vartheta(X_i = x_i) = \frac{\vartheta^{x_i}}{x_i!} e^{-\vartheta}, \quad i = 1, \dots, n,$$

ezért

$$L_\vartheta(x_1, \dots, x_n) = \prod_{i=1}^n p_\vartheta(x_i) = \prod_{i=1}^n \frac{\vartheta^{x_i}}{x_i!} e^{-\vartheta} = \frac{\vartheta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} e^{-n\vartheta}.$$

2. A log-likelihood függvényre tehát a logaritmus azonosságait használva azt kapjuk, hogy

$$l_\vartheta(x_1, \dots, x_n) = \ln \left(\frac{\vartheta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} e^{-n\vartheta} \right) = \ln \vartheta \sum_{i=1}^n x_i - \sum_{i=1}^n \ln(x_i!) - n\vartheta.$$

3. A log-likelihood függvényt ϑ szerint deriváljuk:

$$\frac{\partial}{\partial \vartheta} l_\vartheta(x_1, \dots, x_n) = \frac{1}{\vartheta} \sum_{i=1}^n x_i - n.$$

¹⁰Amennyiben a θ paramétertartomány nem az egész \mathbb{R}^d , akkor vizsgálni kell azt is, hogy a tartomány határán nincs-e maximumhely.

4. Láthatjuk, hogy a derivált alakja meglehetősen egyszerű, vagyis a logaritmus alkalmazása valóban hasznos volt. A derivált zérushelyét is könnyű meghatározni:

$$\frac{\partial}{\partial \vartheta} l_{\vartheta}(x_1, \dots, x_n) = 0 \quad \Leftrightarrow \quad \frac{1}{\vartheta} \sum_{i=1}^n x_i - n = 0 \quad \Leftrightarrow \quad \vartheta = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}_n.$$

Vagyis a log-likelihood-függvény (és a likelihood-függvény) egyetlen zérushelye a mintaátlag realizációja, vagyis a realizáció elemeinek számtani közepe.

5. Ebben a pontban a log-likelihood-függvénynek szigorú lokális maximuma van, mivel a második derivált értéke bármely $\vartheta \in \theta$ pontban negatív:

$$\frac{\partial^2}{\partial \vartheta^2} l_{\vartheta}(x_1, \dots, x_n) = -\frac{1}{\vartheta^2} \sum_{i=1}^n x_i < 0.$$

Mivel ez az egyetlen lokális szélsőérték van és a $\theta = (0, \infty)$ paramétertartomány egy nyílt intervallum, a kapott maximum globális is.

6. A log-likelihood-függvény tehát a

$$\vartheta_*(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}_n$$

pontban lesz maximális. Így a maximum likelihood becslésünk:

$$\vartheta_*(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n. \quad (14.5)$$

Ha $x_1 = \dots = x_n = 0$, akkor a likelihood-függvénynek nincs a $(0, \infty)$ intervallumon lokális maximuma. A globális maximum ekkor a θ határán, a $\vartheta = 0$ pontban vétetik fel. (Interpretáció: minél kisebb $\vartheta > 0$, annál valószínűbb, hogy minden mintaelem értéke 0.) Ebben az esetben tehát $\vartheta_*(x_1, \dots, x_n) = 0$, vagyis ekkor is igaz, hogy $\vartheta_*(x_1, \dots, x_n) = \bar{x}_n$. A 14.5 egyenlet tehát mindig teljesül.

14.2.8. *Példa* (Maximum likelihood módszer, folytonos eset: exponenciális eloszlás). A fenti eljárást alkalmazzuk az exponenciális eloszlás esetére: tegyük fel, hogy az X_1, \dots, X_n mintaelemek peremeloszlása exponenciális eloszlás ismeretlen $\vartheta > 0$ paraméterrel. Legyen (x_1, \dots, x_n) az (X_1, \dots, X_n) egy realizációja, vagyis egy n darab pozitív valós számból álló vektor.

1. A mintaelemek sűrűségfüggvénye

$$f_{\vartheta}(x_i) = \vartheta e^{-\vartheta x_i}, \quad i = 1, \dots, n,$$

így a likelihood-függvényre a következő kifejezés adódik:

$$L_{\vartheta}(x_1, \dots, x_n) = \prod_{i=1}^n f_{\vartheta}(x_i) = \prod_{i=1}^n \vartheta e^{-\vartheta x_i} = \vartheta^n e^{-\vartheta \sum_{i=1}^n x_i}.$$

2. A log-likelihood-függvény tehát

$$l_{\vartheta}(x_1, \dots, x_n) = \ln \left(\vartheta^n e^{-\vartheta \sum_{i=1}^n x_i} \right) = n \ln \vartheta - \vartheta \sum_{i=1}^n x_i.$$

3. A log-likelihood függvényt ϑ szerint deriváljuk:

$$\frac{\partial}{\partial \vartheta} \ln l_{\vartheta}(x_1, \dots, x_n) = \frac{n}{\vartheta} - \sum_{i=1}^n x_i.$$

4. A derivált zérushelyét ismét egyszerű meghatározni:

$$\frac{\partial}{\partial \vartheta} \ln l_{\vartheta}(x_1, \dots, x_n) = 0 \quad \Leftrightarrow \quad \frac{n}{\vartheta} - \sum_{i=1}^n x_i = 0 \quad \Leftrightarrow \quad \vartheta = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}_n}.$$

5. Mivel

$$\frac{\partial^2}{\partial \vartheta^2} \ln l_{\vartheta}(x_1, \dots, x_n) = -\frac{n}{\vartheta^2}$$

minden $\vartheta > 0$ esetén negatív, a kapott zérushely lokális maximumhelye a log-likelihood-függvénynek. Mivel ez az egyetlen lokális maximumhely és $\theta = (0, \infty)$ egy nyílt intervallum, ez egyben a (log-)likelihood-függvény egyetlen globális maximumhelye.

6. A maximum-likelihood becslés az (x_1, \dots, x_n) realizáció alapján tehát

$$\vartheta_*(x_1, \dots, x_n) = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}_n}.$$

Így tehát a keresett maximum likelihood becslés

$$\vartheta_*(X_1, \dots, X_n) = \frac{n}{\sum_{i=1}^n X_i} = \frac{1}{\bar{X}_n}.$$

14.2.9. *Megjegyzés.* Az előző két példa azt is megmutatja, hogy a maximum likelihood becslés lehet torzítatlan becslése ϑ -nak, de nem mindig az. Ugyanis az ismeretlen ϑ paraméterű Poisson-eloszlás esetén ϑ éppen megegyezik $\mathbb{E}_{\vartheta}(X_1)$ -gyel, a mintaelemek várható értékeire pedig az $\vartheta_*(X_1, \dots, X_n) = \bar{X}_n$ mintaátlag mindig torzítatlan becslés (vö. 14.1.2. Példa). Ezért ebben az esetben a maximum likelihood becslés torzítatlan becslése a ϑ paraméternek.

Ugyanakkor az exponenciális eloszlás esetén

$$\mathbb{E}_{\vartheta}(\vartheta_*(X_1, \dots, X_n)) = \mathbb{E}_{\vartheta}\left(\frac{n}{\sum_{i=1}^n X_i}\right) \neq \frac{n}{\sum_{i=1}^n \mathbb{E}_{\vartheta}(X_i)} = \frac{1}{\frac{1}{\vartheta}} = \vartheta.$$

Tehát ha ϑ a valódi paraméter, akkor $\mathbb{E}_{\vartheta}(\vartheta_*(X_1, \dots, X_n))$ nem egyezik meg ϑ -val, tehát a maximum likelihood becslés nem torzítatlan becslése ϑ -nak.¹¹

A likelihood-függvény log-likelihood-függvényé alakítása különösen abban az esetben segíthet, ha a mintát alkotó független, azonos eloszlású X_i valószínűségi változók súly- illetve sűrűségfüggvényében exponenciális tényezők is szerepelnek. Erre példa a Poisson-eloszlás, az exponenciális eloszlás vagy független exponenciálisok összegének eloszlása, illetve a normális eloszlás – amelynek a szituációtól függően lehet egy vagy két ismeretlen paramétere is. Vannak azonban olyan eloszlások is, ahol érdemes közvetlenül a likelihood-függvényt maximalizálni, anélkül, hogy a logaritmusát vennénk, sőt a maximalizálást nem feltétlenül deriválás segítségével érdemes végezni. Lássunk erre is egy példát!

¹¹Az előbbi „ \neq ” állítást itt részletesen nem bizonyítjuk, de általában igaz az, hogy bármilyen szigorúan konvex f függvény és X valószínűségi változó esetén $\mathbb{E}(f(X)) \geq f(\mathbb{E}(X))$, és egyenlőség akkor és csak akkor áll fenn, ha X 1 valószínűséggel konstans. (Ennek egy speciális esete az az általunk már ismert állítás, hogy $\mathbb{E}(X^2) \geq \mathbb{E}(X)^2$, – tehát $\mathbb{D}^2(X) \geq 0$ – és egyenlőség – vagyis $\mathbb{D}^2(X) = 0$ – akkor és csak akkor áll fenn, ha X 1 valószínűséggel konstans.) Itt $f(x) = 1/x$ (ami $(0, \infty)$ -n szigorúan konvex) és $X = \sum_{i=1}^n X_i$ (és \mathbb{P} helyett \mathbb{P}_{ϑ} -ről van szó).

14.2.10. *Példa* (Maximum likelihood módszer: egyenletes eloszlás ismeretlen alsó végpontú intervallumon). Legyenek X_1, \dots, X_n független és $U(-\vartheta; 0)$ eloszlású, tehát a $(-\vartheta, 0)$ intervallumon egyenletes eloszlású valószínűségi változók, ahol $\vartheta > 0$ egy (ismeretlen) paraméter. Adjuk meg a $\vartheta_*(X_1, \dots, X_n)$ maximum likelihood becslést!

Az egyenletes eloszlás tulajdonságai miatt az f_ϑ sűrűségfüggvény

$$f_\vartheta(x) = \begin{cases} \frac{1}{\vartheta}, & \text{ha } x \in (-\vartheta, 0), \\ 0, & \text{különben.} \end{cases}$$

Így a likelihood-függvény (X_1, \dots, X_n) bármely (x_1, \dots, x_n) realizációja esetén

$$L_\vartheta(x_1, \dots, x_n) = \begin{cases} \frac{1}{\vartheta^n}, & \text{ha } x_1, x_2, \dots, x_n \in (-\vartheta, 0), \\ 0, & \text{különben.} \end{cases}$$

Vegyük észre, hogy ebben az esetben X_1 lényeges értékeinek $S_{X_1}^{(\vartheta)} = (-\vartheta, 0)$ halmaza függ a ϑ paramétertől, az intervallum hosszának csökkentésével pedig a likelihood-függvény értéke az (x_1, \dots, x_n) pontban szigorúan monoton csökken. Az intervallum hosszát mindaddig csökkenthetjük, amíg az (x_1, \dots, x_n) az (X_1, \dots, X_n) egy realizációja marad, azaz amíg

$$\min\{x_1, \dots, x_n\} > -\vartheta,$$

azaz $\vartheta > -\min\{x_1, \dots, x_n\}$. Így a likelihood-függvény éppen abban a határesetben lesz maximális, amikor $-\min\{x_1, \dots, x_n\} = \vartheta$. (Ebben az esetben (x_1, \dots, x_n) már éppen nem realizáció, de bármely $\vartheta > -\min\{x_1, \dots, x_n\}$ esetén igen.) Ezért az (x_1, \dots, x_n) realizációhoz tartozó maximum likelihood becslésünk

$$\vartheta_*(x_1, \dots, x_n) = -\min\{x_1, \dots, x_n\},$$

a maximum likelihood becslésünk pedig

$$\vartheta_*(X_1, \dots, X_n) = -\min\{X_1, \dots, X_n\} = -X_1^*.$$

14.3. Momentumbecslés

Egy másik fontos paraméterbecslési módszer a *momentumok módszere*. Ehhez először vezessük be a momentumok és az empirikus momentumok fogalmát:

14.3.1. Definíció. \blacktriangle Legyen k egy pozitív egész szám és X egy valószínűségi változó. Ha $\mathbb{E}[|X|^k] < \infty$, akkor az X k . **momentumán** az $\mathbb{E}[X^k] \in \mathbb{R}$ mennyiséget értjük.

14.3.2. Definíció. Legyen k egy pozitív egész szám és legyenek X_1, \dots, X_n független, azonos eloszlású valószínűségi változók. Az (X_1, \dots, X_n) n -elemű fae. minta **k -adik empirikus momentumán** az

$$\widehat{m}_k = \frac{1}{n} \sum_{i=1}^n X_i^k$$

statisztikát értjük.

A momentumok módszere különösen akkor lesz hasznos, ha egyenél több paraméterünk van. Ilyen például a normális eloszlás ismeretlen μ várható értékkel és ismeretlen σ^2 szórásnégyzettel, ahol az ismeretlen paraméterek vektora $\underline{\vartheta} = (\mu, \sigma^2) \in \mathbb{R}^2$, ahol $\mu \in \mathbb{R}$ és $\sigma^2 > 0$. Vagy ilyen például az egyenletes eloszlás egy ismeretlen kezdő- és végpontú (a, b) intervallumon, ahol a paramétervektor $\underline{\vartheta} = (a, b) \in \mathbb{R}^2$ valamely $-\infty < a < b < \infty$ valós számokra.

Általában tegyük fel, hogy az $\mathbf{X} = (X_1, \dots, X_n)$ fae. minta elemeinek eloszlását meghatározó ismeretlen paraméterek vektora $\underline{\vartheta} = (\vartheta_1, \dots, \vartheta_k)$, amely egy $\theta \subseteq \mathbb{R}^k$ halmaz egy eleme. Tegyük fel azt is, hogy az X_1, \dots, X_n mintaelemeknek a paraméter bármely választása esetén létezik az első k momentuma. Ezek a momentumok természetesen a $\vartheta_1, \dots, \vartheta_k$ paraméterek függvényei:

$$\begin{aligned} m_1 &= g_1(\vartheta_1, \dots, \vartheta_k), \\ m_2 &= g_2(\vartheta_1, \dots, \vartheta_k), \\ &\vdots \\ m_k &= g_k(\vartheta_1, \dots, \vartheta_k) \end{aligned}$$

valamely $g_1, g_2, \dots, g_k: \mathbb{R}^k \rightarrow \mathbb{R}$ függvényre. A

$$(\vartheta_1, \dots, \vartheta_k) \mapsto \underline{g}(\vartheta_1, \dots, \vartheta_k) := (g_1(\vartheta_1, \dots, \vartheta_k), \dots, g_k(\vartheta_1, \dots, \vartheta_k))$$

leképezés tehát egy $\underline{g}: \mathbb{R}^k \rightarrow \mathbb{R}^k$ függvény, vagyis egy vektormező. Ha ennek a vektormezőnek létezik inverze, akkor jelöljük ezt $\underline{h} = (h_1, \dots, h_k)$ -val. Ezen inverz segítségével a momentumok ismeretében megkaphatjuk a $\underline{\vartheta} = (\vartheta_1, \dots, \vartheta_k)$ paramétert:

$$\begin{aligned} \vartheta_1 &= h_1(m_1, \dots, m_k), \\ \vartheta_2 &= h_2(m_1, \dots, m_k), \\ &\vdots \\ \vartheta_k &= h_k(m_1, \dots, m_k) \end{aligned}$$

14.3.3. Definíció. A fenti szituációban az $\underline{\vartheta} = (\vartheta_1, \dots, \vartheta_n)$ paramétervektor **momentumbecslésén** az $\widehat{\underline{\vartheta}} = (\widehat{\vartheta}_1, \dots, \widehat{\vartheta}_k)$ vektort értjük, ahol

$$\widehat{\vartheta}_i = h_i(\widehat{m}_1, \dots, \widehat{m}_k), \quad i = 1, \dots, k,$$

és \widehat{m}_i az i -edik empirikus momentumot jelöli.

Differenciálható \underline{g} leképezés esetén az $\underline{h} = (h_1, \dots, h_k)$ inverz létezésének szükséges feltétele, hogy \underline{g} Jacobi-determinánsa a θ halmazon sehol se legyen 0. Lássunk példákat a momentumbecslésre:

14.3.4. *Példa* (Momentumbecslés: normális eloszlás). Legyenek X_1, \dots, X_n független és egyenként $N(\mu; \sigma^2)$ eloszlású valószínűségi változók, ahol $\mu \in \mathbb{R}$ és $\sigma^2 > 0$. Az első $k = 2$ momentumot $\underline{\vartheta} = (\mu, \sigma^2)$ esetében ekkor a következőképpen kaphatjuk meg:

$$\begin{aligned} m_1 &= \mathbb{E}_{\underline{\vartheta}}(X_1) = \mu, \\ m_2 &= \mathbb{E}_{\underline{\vartheta}}(X_1^2) = \mathbb{D}_{\underline{\vartheta}}^2(X_1) + \mathbb{E}_{\underline{\vartheta}}(X_1)^2 = \sigma^2 + \mu^2. \end{aligned}$$

Ebben az esetben a $\underline{g}: (\mu, \sigma^2) \mapsto (\mu, \sigma^2 + \mu^2)$ leképezés természetesen invertálható¹²:

$$\begin{aligned} \mu &= m_1, \\ \sigma^2 &= m_2 - m_1^2, \end{aligned}$$

azaz az inverz függvény képlete $\underline{h}: (m_1, m_2) \mapsto (m_1, m_2 - m_1^2)$. Így a momentumbecslésünk

$$\begin{aligned} \widehat{\mu} &= \widehat{m}_1 = \overline{X_n}, \\ \widehat{\sigma^2} &= \widehat{m}_2 - \widehat{m}_1^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \overline{X_n}^2. \end{aligned}$$

¹²Ebben az esetben a Jacobi-mátrixot úgy kell kiszámítani, hogy változóként úgy σ^2 -et (és nem σ -t) tekintjük, így a Jacobi-mátrix:

$$\begin{bmatrix} \frac{\partial g_1}{\partial \mu}(\mu, \sigma^2) & \frac{\partial g_1}{\partial(\sigma^2)}(\mu, \sigma^2) \\ \frac{\partial g_2}{\partial \mu}(\mu, \sigma^2) & \frac{\partial g_2}{\partial(\sigma^2)}(\mu, \sigma^2) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2\mu & \sigma^2 \end{bmatrix}.$$

Ennek determinánsa σ^2 , ami a θ paramétertartományon sehol sem 0.

Az $\hat{m}_2 - \hat{m}_1^2$ pedig éppen a (14.2) egyenletben definiált S_n^2 statisztikával (azaz az empirikus szórásnégyzettel, vagyis a korrigált empirikus szórásnégyzet $(n-1)/n$ -szeresével) egyezik meg, ugyanis

$$S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \frac{2}{n} \sum_{i=1}^n X_i \bar{X}_n + \frac{1}{n} \sum_{i=1}^n \bar{X}_n^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - 2\bar{X}_n + \bar{X}_n^2 = \hat{m}_2 - \hat{m}_1^2. \quad (14.6)$$

Ez az egyenlőség pedig bármilyen eloszlású minta esetén igaz.

14.3.5. *Példa* (Momentumbecslés: egyenletes eloszlás). Tegyük fel, hogy az X_1, \dots, X_n független, azonos eloszlású valószínűségi változók az (a, b) intervallumon egyenletes eloszlásúak, ahol most az ismeretlen paraméterek vektora $\vartheta = (a, b)$, $-\infty < a < b < \infty$. Ekkor az egyenletes eloszlás tulajdonságai alapján az első $k = 2$ momentum a következő:

$$m_1 = \mathbb{E}_{\vartheta}(X_1) = \frac{a+b}{2},$$

$$m_2 = \mathbb{E}_{\vartheta}(X_1^2) = \mathbb{D}_{\vartheta}^2(X_1) + \mathbb{E}_{\vartheta}(X_1)^2 = \frac{(b-a)^2}{12} + \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2 + 3(a+b)^2}{12}.$$

Tehát azt kapjuk, hogy

$$\frac{(b-a)^2}{12} = m_2 - m_1^2 \quad \Rightarrow \quad b-a = 2\sqrt{3(m_2 - m_1^2)}$$

és

$$a+b = 2m_1.$$

(Vagyis itt is invertálható leképezéssel állunk szemben; ellenőrizhetjük, hogy az a, b szerinti deriválással kapott Jacobi-mátrix determinánsa $-\infty < a < b < \infty$ esetén sohasem 0.) Ezzel tehát egy lineáris egyenletrendszert kaptunk a -ra és b -re, ami $-\infty < a < b < \infty$ esetén mindig egyértelműen megoldható. A megoldás

$$a = m_1 - \sqrt{3(m_2 - m_1^2)},$$

$$b = m_1 + \sqrt{3(m_2 - m_1^2)}.$$

Felhasználva az $\hat{m}_2 - \hat{m}_1^2 = S_n^2$ egyenletet (vö. (14.6)), azt kapjuk, hogy a momentumbecslés

$$\hat{a} = \hat{m}_1 - \sqrt{3(\hat{m}_2 - \hat{m}_1^2)} = \bar{X}_n - \sqrt{3S_n^2},$$

$$\hat{b} = \hat{m}_1 + \sqrt{3(\hat{m}_2 - \hat{m}_1^2)} = \bar{X}_n + \sqrt{3S_n^2}.$$

15. Konfidenciaintervallumok

15.1. Alapfogalmak: konfidenciaintervallum, folytonos eloszlás kvantilisei

Az eddig ismerttetett maximum likelihood becslés és momentumbecslés a *pontbecslések* kategóriájába tartoznak, mivel az ismeretlen ϑ paraméterre (vagy annak valamilyen függvényére) vonatkozó becslésünk eredménye egy, a mintától függő „pont”, vagyis egy valós szám vagy vektor. Bár a pontbecslés bizonyos értelemben a véletlen mintaeredmény ismeretében a „legjobb tippünket” adja meg a paraméterre, a gyakorlatban ritkán fordul elő, hogy a valódi paraméter valóban közel van a becslésünk eredményéhez. Ezért ebben a fejezetben bevezetjük a konfidencia-intervallumokat, amelyek az *intervallumbecslések*hez tartoznak. Itt a becslés eredménye egy olyan, a realizációtól függő intervallum, amelybe a valódi ϑ paraméter egy előre megadott $1 - \varepsilon$ valószínűséggel beletartozik. Ezt a valószínűséget az intervallum hosszának növelésével 1-hez tetszőlegesen közelívé tehetjük, például gyakran választjuk 0.99-nek (itt a hibavalószínűség $\varepsilon = 0.01$) vagy 0.95-nek ($\varepsilon = 0.05$). Ezt írja le formálisan a következő definíció.

15.1.1. Definíció. Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy n elemű fae. minta, ahol a mintaelemek eloszlása egy $\vartheta \in \theta \subseteq \mathbb{R}^d$ paramétertől függ, legyen $\psi: \mathbb{R}^d \rightarrow \mathbb{R}$ egy függvény, és legyen $0 < \varepsilon < 1$. Azt mondjuk, hogy a $[T_1(\mathbf{X}), T_2(\mathbf{X})]$ intervallum (pontosan) $1 - \varepsilon$ **szintű konfidenciaintervallum** a $\psi(\vartheta)$ paraméterfüggvényre, ha

$$\mathbb{P}_\vartheta(T_1(\mathbf{X}) \leq \psi(\vartheta) \leq T_2(\mathbf{X})) = 1 - \varepsilon,$$

minden $\vartheta \in \theta$ esetén.

15.1.2. *Megjegyzés.* 1. A definícióban szereplő $T_1(\mathbf{X})$ és $T_2(\mathbf{X})$ intervallum-végpontok az \mathbf{X} minta statisztikái, tehát a véletlen mintától függenek, de a realizáció ismeretében már nem véletlenek.

2. Pontosán $1 - \varepsilon$ szintű konfidenciaintervallum általában csak akkor adható meg, ha X_1, \dots, X_n peremeloszlása folytonos. Diszkrét esetben általában *legalább* $1 - \varepsilon$ szintű konfidenciaintervallumokról szokás beszélni, ahol a definícióban „ $= 1 - \varepsilon$ ” helyett „ $\geq 1 - \varepsilon$ ” szerepel. Ennek ebben a kurzusban nem lesz nagy jelentősége.

Mielőtt példákat mutathatnánk a konfidenciaintervallumra, szükségünk van még a *kvantilis* fogalmára is. Ezt is csak folytonos valószínűségi változók esetén vezetjük be.

Legyen X folytonos valószínűségi változó. A 4.2.1. Definíció értelmében ez azt jelenti, hogy X -nek létezik az f_X sűrűségfüggvénye. Tudjuk, hogy ebben az esetben az $x \mapsto F_X(x)$ eloszlásfüggvény folytonos. Az egyszerűség kedvéért tegyük fel, hogy a sűrűségfüggvény pontosan egy I nyílt intervallumon nem 0 (azaz I minden pontjában pozitív, $\mathbb{R} \setminus I$ minden pontjában pedig 0). Erre példák:

1. az egyenletes eloszlás, ahol $I = (a, b)$, $-\infty < a < b < \infty$,
2. az exponenciális eloszlás, ahol $I = (0, \infty)$,
3. és a normális eloszlás, ahol $I = (-\infty, \infty) = \mathbb{R}$.

Ekkor az I intervallumon az F_X eloszlásfüggvény szigorúan monoton nő. Továbbá ha az I intervallum (a, b) vagy (a, ∞) alakú, akkor $F_X(x) = 0$ minden $x \leq a$ -ra teljesül, és ha az I intervallum (a, b) vagy $(-\infty, b)$ alakú, akkor $F_X(x) = 1$ minden $x \geq b$ -re teljesül. Ebből az következik, hogy létezik az

$$F_X^{-1}: (0, 1) \rightarrow I, \quad y \mapsto F_X^{-1}(y)$$

inverzfüggvény (ahol $x \in I$ és $0 < y < 1$ esetén $x = F_X^{-1}(y)$ pontosan akkor teljesül, ha $y = F_X(x)$). Ez az inverzfüggvény is folytonos és szigorúan monoton növekvő. Fontos azonban hangsúlyozni, hogy F_X^{-1} értékkészlete általában nem az egész \mathbb{R} , hanem csak I . Valóban, az előbbi példák azt mutatják, hogy a 0 értéket az F_X vagy végtelen sok helyen felveszi (pl. egyenletes és exponenciális eloszlás) vagy sehol nem veszi fel (pl. normális eloszlás), hasonlóan az 1 értékhez.

15.1.3. Definíció. Legyen X folytonos valószínűségi változó F_X eloszlásfüggvénnyel és f_X sűrűségfüggvénnyel, és tegyük fel, hogy létezik egy olyan I nyílt intervallum, hogy minden $x \in \mathbb{R}$ esetén $f_X(x) \neq 0$ pontosan akkor teljesül, ha $x \in I$.

Legyen $0 < y < 1$. Ekkor az $F_X^{-1}(y) \in I$ pontot az X **y -kvantilisének** nevezzük. Az $\frac{1}{2}$ -kvantilist ($y = 1/2$) **mediánnak** nevezzük.

Ha x az X y -kvantilise, akkor teljesül, hogy

$$F_X(x) = \mathbb{P}(X < x) = y,$$

vagyis X pontosan y valószínűséggel vesz fel x -nél kisebb értékeket. A medián esetén $y = 1/2$, vagyis annak a valószínűsége, hogy $X < x$, megegyezik annak a valószínűségével, hogy $X \geq x$ (sőt a folytonosság miatt az $X > x$ valószínűséggel is). Ilyen értelemben a medián valóban az empirikus medián, vagyis a „rendezett minta középső eleme” (páros elemszámnál a két középső átlaga) egyfajta analógiája.

Például diszkrét valószínűségi változó esetén elképzelhető, hogy a $0 < y < 1$ értéket az eloszlásfüggvény átugorja. Például a szabályos kockadobás eloszlásfüggvénye az $1/12$ értéket sehol nem veszi fel, mert 0-ról egyből $1/6$ -ra ugrik (az $x = 1$ pontban). Emiatt az y -kvantilis fenti definíciója csak akkor helyes, ha a valószínűségi változó eloszlásfüggvénye folytonos, de a definíció – némi technikai nehézség árán – általánosítható diszkrét valószínűségi változókra is.

15.2. Konfidenciaintervallum normális eloszlás várható értékére ismert szórás esetén

Következzék most első példánk a konfidenciaintervallumra, amely valószínűleg a legismertebb példa is a témakörben:

15.2.1. *Példa.* Legyenek $X_1, \dots, X_n \sim N(\mu; \sigma^2)$ fae. valószínűségi változók, ahol a normális eloszlás $\mu \in \mathbb{R}$ várható értéke ismeretlen, de $\sigma^2 > 0$ szórásnégyzete adott. $\varepsilon \in (0, 1)$ esetén szerkesszünk $1 - \varepsilon$ szintű konfidenciaintervallumot a normális eloszlás várható értékére úgy, hogy annak hossza a lehető legkisebb legyen!

Mivel az

$$f_\mu(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

sűrűségfüggvény az x pont μ -tól való távolsága függvényében szigorúan monoton csökken (bármilyen is legyen a μ paraméter értéke), ezért adott ε esetén a legrövidebb $1 - \varepsilon$ szintű konfidenciaintervallumot akkor fogjuk kapni, ha azt az \bar{X}_n mintaátlag körül szimmetrikusan választjuk, vagyis

$$[\bar{X}_n - r_\varepsilon, \bar{X}_n + r_\varepsilon]$$

alakban, úgy, hogy

$$\mathbb{P}_\mu(\mu \in [\bar{X}_n - r_\varepsilon, \bar{X}_n + r_\varepsilon]) = 1 - \varepsilon \quad (15.1)$$

minden $\mu \in \mathbb{R}$ esetén teljesüljön. A feladatunk tehát az r_ε (ε -tól függő) értékének meghatározása.

A mintaátlag eloszlása a normális eloszlás ismert tulajdonságai alapján:

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n \underbrace{X_i}_{\sim N(\mu; \sigma^2) \text{ fae.}} \sim N(\mu; \sigma^2/n).$$

$\underbrace{\hspace{10em}}_{\sim N(n\mu; n\sigma^2)}$

Ebből következik, hogy

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0; 1). \quad (15.2)$$

Fejezzük ki most a (15.1) egyenletet az $\frac{\bar{X}_n - \mu}{\sigma} \sqrt{n}$ valószínűségi változó segítségével:

$$\begin{aligned} \mathbb{P}_\mu(\mu \in [\bar{X}_n - r_\varepsilon, \bar{X}_n + r_\varepsilon]) = 1 - \varepsilon &\Leftrightarrow \mathbb{P}_\mu(-r_\varepsilon \leq \bar{X}_n - \mu \leq r_\varepsilon) = 1 - \varepsilon \\ \Leftrightarrow \mathbb{P}_\mu\left(-\frac{r_\varepsilon}{\sigma} \sqrt{n} \leq \frac{\bar{X}_n - \mu}{\sigma} \sqrt{n} \leq \frac{r_\varepsilon}{\sigma} \sqrt{n}\right) = 1 - \varepsilon. \end{aligned}$$

Az utóbbi pedig (szokás szerint Φ -vel jelölve a standard normális eloszlás eloszlásfüggvényét) pontosan akkor teljesül, ha

$$\Phi\left(\frac{r_\varepsilon}{\sigma} \sqrt{n}\right) - \Phi\left(-\frac{r_\varepsilon}{\sigma} \sqrt{n}\right) = 2\Phi\left(\frac{r_\varepsilon}{\sigma} \sqrt{n}\right) - 1 = 1 - \varepsilon,$$

vagyis átrendezve

$$\Phi\left(\frac{r_\varepsilon}{\sigma} \sqrt{n}\right) = 1 - \frac{\varepsilon}{2}. \quad (15.3)$$

A $\Phi: \mathbb{R} \rightarrow (0, 1)$ eloszlásfüggvény egész \mathbb{R} -en szigorúan monoton nő, tehát létezik a $\Phi^{-1}: (0, 1) \rightarrow \mathbb{R}$ inverzfüggvény, amely szintén szigorúan monoton nő. Így tehát (15.3) mindkét oldalára Φ^{-1} -t alkalmazva azt kapjuk, hogy

$$\frac{r_\varepsilon}{\sigma} \sqrt{n} = \Phi^{-1}(1 - \varepsilon/2),$$

vagyis

$$r_\varepsilon = \frac{\sigma}{\sqrt{n}} \Phi^{-1}(1 - \varepsilon/2).$$

A végeredmény megadásához vezessük be $\delta \in (0, 1)$ esetén az

$$u_\delta = \Phi^{-1}(1 - \delta) \quad (15.4)$$

jelölést! Vegyük észre, hogy ez éppen a standard normális eloszlás $1 - \delta$ -kvantilise, hiszen ha $X \sim N(0; 1)$, akkor

$$\mathbb{P}(X < u_\delta) = \Phi(u_\delta) = \Phi(\Phi^{-1}(1 - \delta)) = 1 - \delta.$$

Így tehát

$$r_\varepsilon = \frac{\sigma}{\sqrt{n}} u_{\varepsilon/2},$$

és a keresett konfidenciaintervallum

$$\left[\bar{X}_n - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \bar{X}_n + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \right].$$

Eredményeinket a következő állításban foglaljuk össze.

15.2.2. Állítás. *Legyenek X_1, \dots, X_n fae. $N(\mu; \sigma^2)$ eloszlású valószínűségi változók, ahol a μ várható érték ismeretlen és a $\sigma^2 > 0$ szórásnégyzet ismert, és jelölje \bar{X}_n a mintához tartozó mintaátlagot. Ekkor $\varepsilon \in (0, 1)$ esetén a μ várható értékre egy $1 - \varepsilon$ szintű konfidenciaintervallum a következő:*

$$[T_1(\mathbf{X}), T_2(\mathbf{X})] = \left[\bar{X}_n - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \bar{X}_n + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \right],$$

ahol $u_{\varepsilon/2}$ a standard normális eloszlás $1 - \varepsilon/2$ -kvantilise.

15.2.3. *Megjegyzés.* Figyeljük meg, hogy

1. minél nagyobb a minta elemszáma, annál kevésbé hosszú konfidenciaintervallum szükséges (vö. nagy számok törvénye),
2. minél nagyobb a szórási, annál hosszabb konfidenciaintervallum szükséges (nagy szórási esetén „laposabb”, jobban szétterülő a sűrűségfüggvény),
3. minél kisebb az ε hibaszint, annál hosszabb konfidenciaintervallum szükséges (mert $1 - \varepsilon$ valószínűséggel kell μ -nek a konfidenciaintervallumba esnie, ha μ a valódi paraméter).

Az olvasóban joggal merülhet fel a kérdés: hogyan kell a normális eloszlás szórásnégyzetére konfidenciaintervallumot szerkeszteni, ha a várható érték ismert? Illetve hogyan lehet a normális eloszlás várható értékére akkor is konfidenciaintervallumot megadni, ha a szórási ismeretlen? A fejezet hátralévő részében erről lesz szó. Először a 15.3. fejezetben bevezetünk két újabb nevezetes eloszláscsaládot, a χ^2 - és a Student-eloszlásokat, és megismerjük ezek kapcsolatát a normális eloszlású fae. mintákkal. A Student-eloszlás a 16. fejezetben, a hipotézisvizsgálatnál is fontos szerephez fog jutni. Az itt ismertetett állításokat nem bizonyítjuk, a bizonyítások iránt érdeklődő olvasó számára a mérnök- és gazdaságinformatikus MSc képzés Matematika statisztika tárgyát (vagy bármelyik matematikus BSc képzés megfelelő tárgyát) ajánljuk. Ezek után a 15.4 fejezetben az ismert várható értékű normális eloszlás szórási, a 15.5 fejezetben pedig az ismeretlen szórási normális eloszlás várható értékére vonatkozó konfidenciaintervallum konstrukcióját ismertetjük.

15.3. χ^2 -eloszlás, Student-eloszlás és tulajdonságai

15.3.1. Definíció. Legyenek X_1, \dots, X_n fae. $N(0; 1)$ eloszlású valószínűségi változók. Ekkor $\sum_{i=1}^n X_i^2$ eloszlását n szabadságfokú χ^2 -eloszlásnak¹³ (ejtsd: khí-négyzet-eloszlásnak) nevezzük. Jelölés: $X \sim \chi^2(n)$.

¹³A χ^2 -eloszlással kapcsolatos részek a 2023/24-es tanévben mind kiegészítő anyagok.

15.3.2. Állítás. Legyen egy n pozitív egész szám. Ha az X valószínűségi változó n szabadságfokú χ^2 -eloszlású, akkor a sűrűségfüggvénye

$$f_n(x) = \begin{cases} \frac{x^{n/2-1} e^{-x/2}}{2^{n/2} \Gamma(n/2)}, & \text{ha } x > 0, \\ 0, & \text{ha } x \leq 0. \end{cases}$$

Itt Γ az úgynevezett **Euler-féle Gamma-függvény**, amelyre $\Gamma(n/2)$ rekurzívan a $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$ és $\Gamma(r+1) = r\Gamma(r)$, $r \in [0, \infty)$ egyenletekből számítható ki.

15.3.3. *Megjegyzés.* 1. Vegyük észre, hogy $n \geq 1$ esetén

$$\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = \dots = (n-1)(n-2)\dots \cdot 1\Gamma(1) = (n-1)!$$

A Γ függvény folytonos (sőt analitikus) módon kiterjeszthető az $\mathbb{R} \setminus \{0, -1, -2, \dots\}$ halmazra, ez tehát a faktoriális egyfajta kiterjesztése majdnem minden valós számra.

2. Megfigyelhetjük azt is, hogy $n = 2$ esetén

$$f_n(x) = \begin{cases} \frac{e^{-x/2}}{2}, & \text{ha } x > 0, \\ 0, & \text{ha } x \leq 0, \end{cases}$$

vagyis a 2 szabadságfokú χ^2 -eloszlás megegyezik a $\lambda = 1/2$ paraméterű exponenciális eloszlással. Általában igaz (de itt nem bizonyítjuk), hogy ha n páros, akkor az n szabadságfokú χ^2 -eloszlás megegyezik $n/2$ db független $\text{Exp}(1/2)$ eloszlású valószínűségi változó összegének eloszlásával.

15.3.4. Definíció. Legyen n egy pozitív egész szám. Az X valószínűségi változó eloszlását n **szabadságfokú t -eloszlásnak**, más szóval n **szabadságfokú Student-eloszlásnak** nevezzük, ha a sűrűségfüggvénye

$$f_n(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}, \quad x \in \mathbb{R},$$

ahol Γ ismét az Euler-féle Gamma-függvény. Jelölés: $X \sim t(n)$.

15.3.5. *Megjegyzés.* 1. A „Student” név az eloszlást felfedező William Sealy Gosset angol statisztikus áll-
neve volt.

2. Az 1 szabadságfokú,

$$f_1(x) = \frac{1}{\pi} \frac{1}{1+x^2}$$

sűrűségfüggvényű t -eloszlást ismertebb néven *Cauchy-eloszlásnak* is nevezik, ez a valószínűségi számítás számos területén fontos szerepet játszik. A Cauchy-eloszlású X valószínűségi változó érdekes tulajdonsága, hogy $\mathbb{E}[|X|] = \infty$, és emiatt az $\mathbb{E}[X]$ várható érték nem létezik. (A várható érték nem definiálható még $+\infty$ -ként vagy $-\infty$ -ként sem, mert a sűrűségfüggvény a 0-ra szimmetrikus.) Általában az n szabadságfokú t -eloszlásnak az első $n-1$ momentuma létezik, az n -edik már nem.

15.3.6. Állítás. Legyenek X_1, \dots, X_n független és $N(\mu; \sigma^2)$ eloszlású valószínűségi változók. Jelölje \overline{X}_n a hozzájuk tartozó mintaátlagot és $S_n^* = \sqrt{(S_n^*)^2}$ a korrigált empirikus szórásukat. Ekkor a

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{S_n^*}$$

valószínűségi változó t -eloszlású $n-1$ szabadságfokkal.

15.4. Konfidenciaintervallum normális eloszlás szórásnégyzetére ismert várható érték esetén (a 2023/24-es tanévben kiegészítő anyag)

Az ismert μ várható értékű és ismeretlen σ^2 szórásnégyzetű normális eloszlás szórásnégyzetére vonatkozó konfidenciaintervallum meghatározása az alábbi tételre épül, amelyet itt nem bizonyítunk:

15.4.1. Tétel (Lukács tétele, részlet). *Legyenek $X_1, \dots, X_n \sim N(\mu; \sigma^2)$ fae. valószínűségi változók. Ekkor a*

$$\frac{(n-1)(S_n^*)^2}{\sigma^2}$$

valószínűségi változó χ^2 -eloszlású $n-1$ szabadságfokkal.

Ez alapján most levezetjük a szórásnégyzetre vonatkozó konfidenciaintervallum képletét.

15.4.2. *Példa.* A várható értékre való konfidenciaintervallumhoz képest most az a jelentős különbség, hogy a szórásnégyzet csak nemnegatív lehet. Az $1 - \varepsilon$ szintű konfidenciaintervallumok közül a lehető legkisebb hosszúságú $[a_\varepsilon, b_\varepsilon]$ ($0 < a_\varepsilon < b_\varepsilon$) konfidenciaintervallumot keressük. Bizonyítás nélkül felhasználjuk, hogy a legrövidebb konfidenciaintervallumot akkor kapjuk, ha annak a valószínűsége, hogy a valódi σ^2 paraméter kisebb, mint az alsó a_ε határ, ugyanakkora, mint azé, hogy σ^2 nagyobb, mint a felső b_ε határ, vagyis pontosan $\varepsilon/2$:

$$\mathbb{P}_{\sigma^2}(0 < \sigma^2 < a_\varepsilon) = \mathbb{P}_{\sigma^2}(\sigma^2 > b_\varepsilon) = \frac{\varepsilon}{2}.$$

Ezt átírhatjuk úgy is, hogy

$$\mathbb{P}_{\sigma^2}\left(\frac{(n-1)(S_n^*)^2}{\sigma^2} > \frac{(n-1)(S_n^*)^2}{a_\varepsilon}\right) = \mathbb{P}_{\sigma^2}\left(\frac{(n-1)(S_n^*)^2}{\sigma^2} < \frac{(n-1)(S_n^*)^2}{b_\varepsilon}\right) = \frac{\varepsilon}{2}. \quad (15.5)$$

Jelölje $\alpha \in (0, 1)$ esetén $\chi_{n-1, \alpha}^2$ az $n-1$ szabadságfokú χ^2 -eloszlás $1 - \alpha$ -kvantilisét. Lukács tételének fenti részlete alapján $\frac{(n-1)(S_n^*)^2}{\sigma^2} \sim \chi^2(n-1)$. Az a szám, amelynél ez a valószínűségi változó pontosan $\varepsilon/2$ valószínűséggel nagyobb, éppen $\chi_{n-1, \varepsilon/2}^2$, az a szám pedig, amelynél ez a valószínűségi változó pont $\varepsilon/2$ valószínűséggel kisebb, vagyis pont $1 - \varepsilon/2$ valószínűséggel nagyobb, éppen $\chi_{n-1, 1-\varepsilon/2}^2$. Ezt felhasználva a (15.5) egyenletben szereplő egyenlőségek éppen akkor teljesülnek, ha

$$\frac{(n-1)(S_n^*)^2}{a_\varepsilon} = \chi_{n-1, \varepsilon/2}^2 \quad \text{és} \quad \frac{(n-1)(S_n^*)^2}{b_\varepsilon} = \chi_{n-1, 1-\varepsilon/2}^2.$$

Nagy örömmre a kapott egyenletek már nem függenek az ismeretlen σ^2 paramétertől. Átrendezés után azt kapjuk, hogy $(a_\varepsilon, b_\varepsilon)$ csak a mintától függ, így valóban egy konfidenciaintervallum:

$$[a_\varepsilon, b_\varepsilon] = \left[\frac{(n-1)(S_n^*)^2}{\chi_{n-1, \varepsilon/2}^2}, \frac{(n-1)(S_n^*)^2}{\chi_{n-1, 1-\varepsilon/2}^2} \right].$$

A példa tanulságait az alábbi állítás foglalja össze.

15.4.3. Állítás. *Legyenek X_1, \dots, X_n fae. $N(\mu; \sigma^2)$ eloszlású valószínűségi változók, ahol a μ várható érték ismert és a $\sigma^2 > 0$ szórásnégyzet ismeretlen, és jelölje $(S_n^*)^2$ a mintához tartozó korrigált empirikus szórásnégyzetet. Ekkor $\varepsilon \in (0, 1)$ esetén a σ^2 szórásnégyzetre egy $1 - \varepsilon$ szintű konfidenciaintervallum a következő:*

$$[T_1(\mathbf{X}), T_2(\mathbf{X})] = \left[\frac{(n-1)(S_n^*)^2}{\chi_{n-1, \varepsilon/2}^2}, \frac{(n-1)(S_n^*)^2}{\chi_{n-1, 1-\varepsilon/2}^2} \right],$$

ahol $\alpha \in (0, 1)$ esetén $\chi_{n-1, \alpha}^2$ az $n-1$ szabadságfokú χ^2 -eloszlás $1 - \alpha$ -kvantilise.

15.5. Konfidenciaintervallum normális eloszlás várható értékére ismeretlen szórás esetén

A 15.3.6. Állításban érdemes egy dolgot megfigyelni. Az ismert σ szórású normális eloszlás μ várható értékére vett konfidenciaintervallum meghatározásakor láttuk, hogy

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0; 1)$$

(lásd (15.2)) ami elengedhetetlen volt a konfidenciaintervallum megadásához. Az állításban azt látjuk, hogy ha a σ szórást a S_n^* korrigált empirikus szórásra cseréljük, akkor $n - 1$ szabadságfokú t -eloszlást kapunk:

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n^*} \sim t(n - 1).$$

Miért egyszerű ez? Azért, mert az utóbbi egyenletben egyáltalán nem szerepel a σ szórás, vagyis az egyenlet bármilyen μ várható értékű normális eloszlás esetén igaz, a szórás értékétől függetlenül! Ezt felhasználva ha ismeretlen szórású normális eloszlás várható értékére kell $1 - \varepsilon$ szintű konfidenciaintervallumot szerkesztenünk, azt a 15.2.1. Példában megadott eljárás szerint tehetjük meg, azzal a különbséggel, hogy a szórást a korrigált empirikus szórásra, a standard normális eloszlás $1 - \varepsilon/2$ -kvantilisét pedig az $n - 1$ szabadságfokú t -eloszlás $1 - \varepsilon/2$ -kvantilisére kell cserélnünk. Ezzel megkapjuk a 15.2.2. Állítás analógiáját az ismeretlen szórás esetére:

15.5.1. Állítás. *Legyenek X_1, \dots, X_n fae. $N(\mu; \sigma^2)$ eloszlású valószínűségi változók, ahol a μ várható érték és a $\sigma^2 > 0$ szórásnégyzet ismeretlen, és jelölje \bar{X}_n a mintához tartozó mintaátlagot, S_n^* pedig a mintához tartozó korrigált empirikus szórást. Ekkor $\varepsilon \in (0, 1)$ esetén a μ várható értékre egy $1 - \varepsilon$ szintű konfidenciaintervallum a következő:*

$$[T_1(\mathbf{X}), T_2(\mathbf{X})] = \left[\bar{X}_n - \frac{S_n^* t_{n-1, \varepsilon/2}}{\sqrt{n}}, \bar{X}_n + \frac{S_n^* t_{n-1, \varepsilon/2}}{\sqrt{n}} \right],$$

ahol $\alpha \in (0, 1)$ esetén $t_{n-1, \alpha}$ az $n - 1$ szabadságfokú t -eloszlás $1 - \alpha$ -kvantilisét jelöli.¹⁴

16. Hipotézisvizsgálati módszerek

16.1. Bevezetés

A becslélmélet mellett matematikai statisztika másik nagy ága a *hipotézisvizsgálat*, ahol valamilyen előzetes hipotézisünket (azaz feltevésünket) akarjuk valamilyen mérés, illetve adatelemzés alapján igazolni vagy cáfolni. Ilyen hipotézis lehet például:¹⁵

- A mintaelemek várható értéke 2.
- A mintaelemek várható értéke legfeljebb 2.
- A mintaelemek várható értéke megegyezik egy adott másik minta elemeinek várható értékével.
- A mintaelemek normális eloszlásúak.
- A realizációk függetlenek valamely más, kapcsolódó mérés eredményeitől.

¹⁴A t -eloszlás nagy szabadságfok esetén „közel van” a standard normális eloszláshoz. Precízebben fogalmazva: bármely $\alpha \in (0, 1)$ esetén a t -eloszlás α -kvantilise a standard normális eloszlás α -kvantiliséhez tart, ha a szabadsági fok a végtelenhez tart. Ezért nagy mintaelemszám (kb. $n \geq 30$), azaz nagy szabadsági fok esetén szabad a t -eloszlás α -kvantilise helyett a standard normális eloszlás α -kvantilisével dolgozni. Nagy szabadsági fok esetén sem szabad viszont a korrigált empirikus szórás és a valódi szórás szerepét a konfidenciaintervallumokra, illetve a hipotézisvizsgálatok témakörénél a próbaisztatikákra vonatkozó képletekben összekeverni.

¹⁵Példa gyanánt az itt felsorolt hipotéziseken kívül természetesen gondolhatunk a 13 fejezet elején ismertetett példa-hipotézisekre is.

Célunk ebben a fejezetben az első három típusba tartozó hipotézisek vizsgálata. A továbbiak iránt érdeklődőknek megint csak a mérnök- és gazdaságinformatikus MSc képzés Matematikai statisztika tárgyát ajánljuk.

A mintavétel után számítások segítségével megvizsgáljuk, hogy a mérési eredmények a hipotézisünknek ellentmondanak-e, azaz hogy a hipotézisünket feltéve a mérési eredmények nagyon valószínűtlenek-e. Ha igen, akkor a hipotézist elutasítjuk. Általános esetben a hipotézisvizsgálat (röviden *próba*) az alábbi lépéssorozat végrehajtását jelenti:

1. A H_0 -lal jelölt *nullhipotézis* formalizálása.
2. A mintára vonatkozó *elfogadási tartomány* és *kritikus tartomány* megalkotása.
3. A kísérlet végrehajtása, amelyben az x_1, \dots, x_n realizációt (adatpontokat) kapjuk.
4. Annak megvizsgálása, hogy az adatpontok az elfogadási tartományba esnek-e. Ha igen, H_0 -t elfogadjuk. Ellenkező esetben H_0 -t elutasítjuk.

Az alábbiakban megadjuk az imént említett fogalmaknak – és néhány továbbiaknak – a definícióját.

16.1.1. Definíció. Legyen $\mathbf{X} = (X_1, \dots, X_n)$ egy fae. minta az $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mezőn, ahol a mintaelemek peremeloszlása ismeretlen. Legyen $\mathbf{x} = (x_1, \dots, x_n)$ az $\mathbf{X} = (X_1, \dots, X_n)$ egy realizációja.

- (1) Egy H_0 **nullhipotézis** egy olyan kijelentés, amelynek igazsága csak az X_1, \dots, X_n együttes eloszlásától függ. A H_1 **ellenhipotézis** a H_0 tagadása: $H_1 = \neg H_0$. Az adott próbához tartozó **alternatíva**¹⁶ a

$$H_0 \text{ vs. } H_1.$$

- (2) Egy \mathfrak{X}_e **elfogadási tartomány** az \mathbb{R}^n egy részhalmaza. Az \mathfrak{X}_e elfogadási tartományhoz tartozó \mathfrak{X}_k **kritikus tartomány** az \mathfrak{X}_e komplementere, azaz $\mathfrak{X}_k = \mathbb{R}^n \setminus \mathfrak{X}_e$. A \mathfrak{X}_k kritikus tartománnyal definiált **(statisztikai) próba** az az eljárás, amely során H_0 -t elfogadjuk, ha $\mathbf{X} \in \mathfrak{X}_e$ és H_0 -t elutasítjuk (azaz H_1 -et fogadjuk el), ha $\mathbf{X} \in \mathfrak{X}_k$.

- (3) A \mathfrak{X}_k kritikus tartománnyal definiált próba **terjedelme** a $0 < \alpha < 1$ szám, ha

$$\mathbb{P}((X_1, \dots, X_n) \in \mathfrak{X}_k | H_0 \text{ igaz}) = \alpha.$$

- (4) **Elsőfajú hibának** nevezzük azt, amikor egy próba eljárása során H_0 -t elutasítjuk, pedig H_0 igaz. **Másodfajú hibának** nevezzük azt, amikor az eljárás során H_0 -t elfogadjuk, pedig H_1 igaz (azaz H_0 hamis).

16.1.2. *Megjegyzés.* Egy α terjedelmű próba elsőfajú hibájának valószínűsége éppen a terjedelem:

$$\mathbb{P}(H_0\text{-t elutasítjuk} | H_0 \text{ igaz}) = \mathbb{P}((X_1, \dots, X_n) \in \mathfrak{X}_k | H_0 \text{ igaz}) = \alpha.$$

A másodfajú hiba valószínűségének kiszámítására viszont nem mindig áll rendelkezésre explicit módszer, amint azt konkrét példákban látni fogjuk.

Konkrét próbák működési elvének megértésében segíthetnek a következő általános, matematikailag nem precíz megfontolások és példák¹⁷.

16.1.3. *Megjegyzés.* A H_0 nullhipotézist gyakran úgy fogalmazzuk meg, hogy annak teljesülése esetén ismert legyen az X_1, \dots, X_n mintaelemek eloszlása. Ilyen értelemben például az utcán talált pénzérme esetén jó nullhipotézis a H_0 : „a pénzérme szabályos” kijelentés, mert ekkor annak ismeretében, hogy H_0 igaz, hogy a mintaelemek fae. $1/2$ paraméterű indikátorváltozók. Ezzel szemben nem jó nullhipotézis például a H'_0 : „a pénzérme nem szabályos” kijelentés, mert ekkor H'_0 ismeretében csak azt tudjuk, hogy a mintaelemek valamilyen, $1/2$ -től különböző paraméterrel rendelkező fae. indikátorváltozók, ami nem ad elegendő információt a terjedelem kiszámításához.

¹⁶vs.: ejtsd „versus” (latinul „ellen”).

¹⁷Ezek Bolla Marianna 2012-es *Matematikai statisztika* előadásáról származnak.

16.1.4. *Megjegyzés.* Sokszor helyes az az interpretáció, hogy a nullhipotézist „az ártatlanság vélelmében” fogalmazzuk meg. Például a pénzérme vizsgálatára tipikusan azért kerül sor, mert arra gyanakodunk, hogy a pénzérme cinkelt, de mégis ezen állítás tagadását tesszük meg H_0 -nak. Ebben az értelemben

- az elsőfajú hiba úgy interpretálható, mint az ártatlan elítélése,
- a másodfajú hiba úgy interpretálható, mint a bűnös felmentése.

(Még egyszer hangsúlyozzuk, hogy ezek matematikailag nem precíz, sőt nem is igazán formalizálható állítások.) A legtöbb ismert próbában csak az elsőfajú hiba valószínűségét, vagyis a terjedelmet tudjuk kontrollálni: előre megadunk egy α értéket, és a próba konstrukciója alapján a terjedelem éppen α lesz. Viszont érezhető, – anélkül is, hogy bármit tudnánk konkrét próbákról – hogy ha az előirányzott terjedelmet nagyon kicsinek választjuk, akkor a másodfajú hiba valószínűsége megnő. („Ha biztosra akarunk menni, hogy ártatlanokat ne ítéljünk el, akkor bűnösöket is gyakrabban fogunk felmenteni.”) A gyakorlatban $\varepsilon = 0.05$ és $\varepsilon = 0.01$ a leggyakrabban választott terjedelemértékek.

16.1.5. *Példa.* Ha például egy új gyógyszert szeretnénk forgalomba hozni, akkor H_0 : „a gyógyszer hatástalan vagy káros” vs. H_1 : „a gyógyszer hatásos” jó alternatíva, mert ekkor az elsőfajú hiba azt a scenáriót jelenti, hogy a mintavétel alapján egy hatástalan vagy káros gyógyszert hatásosnak mondunk (és ez alapján vélhetően forgalomba hozzuk). Mivel ez az elsőfajú hiba, ennek valószínűségét tudjuk kontrollálni, tetszőlegesen alacsony $1 - \alpha$ szint alatt tarthatjuk, a másodfajú hiba valószínűségének megnövekedésének árán. A másodfajú hiba itt az az eshetőség, hogy a gyógyszer hatásos, de a mintavétel alapján mégis hatástalannak vagy károsnak nyilvánítjuk (és ezért bizonyára nem vezetjük be). Természetesen a másodfajú hiba bekövetkezése sem kedvező (pl. gazdasági szempontból), de nem jár olyan fatális következményekkel, mint az elsőfajú hiba. Ezért ha olyan hipotézisvizsgálati módszert használunk a gyógyszer tesztelésére, amelyben csak az elsőfajú hiba valószínűsége kontrollálható, akkor H_0 -t a fenti módon kell megválasztani.

16.2. u -próba

A fejezet hátralévő részében ismertetjük az u -próba és a t -próba több aletétét. Ezek a normális eloszlás várható értékével kapcsolatos próbák ismert, illetve ismeretlen szórás esetén. Látni fogjuk, hogy ezek a próbák számos ponton kapcsolódnak a konfidenciaintervallumok témájához (ami már az u és t jelölések alapján is sejthető). A fejezet ezen részével kapcsolatban érdemes mindjárt az elején tisztázni, hogy az u - és t -próbák csak normális eloszlású minták esetén működnek. A centrális határeloszlás-tétel miatt sokféle mintáról feltehető, hogy (jó közelítéssel) normális eloszlású, de azért számos olyan alkalmazás is van, amikor ez nem teljesül. Például a mérnök- és gazdaságinformatikus MSc képzés Matematikai statisztika tárgyának anyagában szerepelnek ún. *nemparaméteres próbák* is, amelyek a normális eloszlást nem tételezik fel és nagyon sokfajta eloszlásra alkalmazhatóak. A nemparaméteres próbák legismertebb példái közé tartozik a χ^2 -próba és a *Kolmogorov–Smirnov-próba*.

Az u -próba kétoldali, egymintás változatát egy példán vezetjük be, majd áttekinthető formában összefoglaljuk az eljárást.

16.2.1. *Példa* (Kétoldali, egymintás u -próba: kenyérvásárlás). A sarki péknél mindennap egy kilós kenyeret veszünk, és a kenyerek az utóbbi napokban kisebbnek tűnnek a korábban megszokottnál, ezért felmerül a gyanúnk, hogy újabban várható értékben is 1 kg-nál kisebb tömegű kenyereket ad el nekünk a pék. Feltesszük, hogy a vásárolt kenyerek kilogrammban mért tömegei egymástól függetlenek és normális eloszlásúak ismeretlen μ várható értékkel és ismert $\sigma = 0.02$ kilogramm szórással (tehát $\sigma^2 = 0.0004$ kg²). Ezért felállítjuk a

$$H_0: \mu = \mu_0 \quad \text{vs.} \quad H_1: \mu \neq \mu_0$$

alternatívát, ahol esetünkben $\mu_0 = 1$. Terjedelemként $\varepsilon = 0.05$ -öt irányzunk elő, vagyis azt szeretnénk, hogy a próba során legfeljebb 0.05 valószínűséggel utasítsuk el H_0 -t, amennyiben H_0 igaz.

Ezután mintát veszünk: $n = 25$ napon keresztül az aznap vásárolt kenyeret mindig mérlegre tesszük. Egy olyan $\mathbf{x} = (x_1, \dots, x_{25})$ realizációt kapunk, amelyre az átlagos tömeg $\bar{x}_n = \bar{x}_{25} = 0.98$ kg. A minta alapján

szeretnénk egy olyan próbát konstruálni, amelynek terjedelme pontosan $\varepsilon = 0.05$, azaz az \mathfrak{X}_e elfogadási tartományt és az \mathfrak{X}_k kritikus tartományt úgy akarjuk megválasztani, hogy

$$\mathbb{P}(H_0\text{-t elutasítjuk}|H_0) = \varepsilon = 0.05$$

teljesüljön.

Az elfogadási tartományt keressük

$$\mathfrak{X}_e = [\overline{X}_n - h, \overline{X}_n + h]$$

alakban, ahol $h > 0$. Ekkor h -t úgy kívánjuk megválasztani, hogy $\mu = \mu_0 = 1$ esetén (az ennek megfelelő valószínűséget \mathbb{P}_{μ_0} -al jelöljük) éppen ε valószínűséggel essen $\mu = \mu_0$ a kritikus tartományba, azaz $1 - \varepsilon$ valószínűséggel essen az elfogadási tartományba, vagyis

$$\mathbb{P}_{\mu_0}(\mu_0 \in [\overline{X}_n - h, \overline{X}_n + h]) = 1 - \varepsilon$$

teljesüljön. Azaz $[\overline{X}_n - h, \overline{X}_n + h]$ egy \overline{X}_n körül szimmetrikus, $1 - \varepsilon$ szintű konfidenciaintervallum kell, hogy legyen a μ_0 paraméterre. Azt viszont már a 15.2. alfejezetből tudjuk, hogy egy ilyen konfidenciaintervallum

$$\left[\overline{X}_n - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \overline{X}_n + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \right],$$

amiből tehát azt kapjuk, hogy

$$h = \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}.$$

Vagyis H_0 -t pontosan akkor fogadjuk el, ha

$$\begin{aligned} \overline{X}_n - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \leq \mu_0 \leq \overline{X}_n + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} &\Leftrightarrow \mu_0 - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \leq \overline{X}_n \leq \mu_0 + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \\ &\Leftrightarrow -u_{\varepsilon/2} \leq \frac{\overline{X}_n - \mu_0}{\sigma} \sqrt{n} \leq u_{\varepsilon/2}. \end{aligned}$$

Az elfogadási tartományunk emiatt

$$\mathfrak{X}_e = \left\{ \mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n : -u_{\varepsilon/2} \leq \frac{\overline{x}_n - \mu_0}{\sigma} \sqrt{n} \leq u_{\varepsilon/2} \right\}.$$

Ha tehát az $u(\mathbf{X}) = \frac{\overline{X}_n - \mu_0}{\sigma} \sqrt{n}$ próbastatisztika $u(\mathbf{x}) = \frac{\overline{x}_n - \mu_0}{\sigma} \sqrt{n}$ realizációja $-u_{\varepsilon/2}$ és $u_{\varepsilon/2}$ közé esik, akkor H_0 -t elfogadjuk, különben pedig elutasítjuk (azaz H_1 -et fogadjuk el).

A konkrét esetünkben $\varepsilon = 0.05$, $n = 25$, $\overline{x}_n = 0.98$, $\mu_0 = 1$ és $\sigma = 0.03$. Ez alapján tehát $u(\mathbf{x}) = \frac{0.98 - 1}{0.03} \sqrt{25} = -5$. Ugyanakkor a normális eloszlás táblázatába pillantva látjuk, hogy $u_{\varepsilon/2} = \Phi^{-1}(0.975) \approx 1.9600$. Tehát H_0 -t akkor fogadnánk el, ha az $u(\mathbf{X})$ próbastatisztika értéke -1.96 és 1.96 közé esne. Mivel ez nem áll fent, H_0 -t elutasítjuk, vagyis kellő bizonyítékot találtunk arra, hogy a kenyerek várható értékben nem egykilósak.

Az alábbiakban formalizáljuk az egymintás u -próba általános eljárását, amelynek egyes részletei a próba altípusától (egy- vagy kétoldali) függenek:

Eljárás (Egymintás u -próba). *Feltevés:* az x_1, \dots, x_n adatpontok fae. normális eloszlású X_1, \dots, X_n valószínűségi változók realizációi, amelyek μ várható értéke ismeretlen, de $\sigma > 0$ szórásuk ismert. Általános eljárás:

1. A H_0 nullhipotézis felállítása.
2. Az előírt $\varepsilon \in (0, 1)$ terjedelm kiválasztása.
3. Az x_1, \dots, x_n adatpontok kinyerése.

4. Az elfogadási tartomány meghatározása. Ez függ a próba altípusától, de mindig magába foglalja a normális eloszlás egy megfelelő kvantilisének kiszámítását (táblázatból való kiolvasását).
5. A próbastatisztika $u(\mathbf{x}) = \frac{\bar{x}_n - \mu_0}{\sigma} \sqrt{n}$ értékének meghatározása.
6. Annak eldöntése, hogy a próbastatisztika értéke az elfogadási tartományba esik-e. Ha igen, H_0 -t elfogadjuk. Ha nem, H_0 -t elutasítjuk (tehát H_1 -et fogadjuk el).

Példánk alapján a kétoldali, egymintás u -próba esetén az eljárást így pontosíthatjuk:

Kétoldali, egymintás u -próba.

- Nullhipotézis: $H_0: \mu = \mu_0$.
- Releváns kvantilis: $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$.
- Elfogadási tartomány: $\mathfrak{X}_e = \{(x_1, \dots, x_n) : |u(\mathbf{x})| \leq u_{\varepsilon/2}\}$, tehát:
 - ha $|u(\mathbf{x})| > u_{\varepsilon/2}$, H_0 -t elutasítjuk,
 - különben H_0 -t elfogadjuk.

Most előző példánknál maradvá bevezetjük az egyoldali, egymintás u -próbát is.

16.2.2. *Példa* (Egyoldali, egymintás u -próba: kenyérvásárlás). Az előző példánál felmerülhet a kérdés, hogy miért azt tettük meg nullhipotézisnek, hogy a kenyér várható értékben egykilós. Ugyanis a 0.98 kg-os mintaátlagot józan ésszel senki nem gondolná bizonyítéknak arra, hogy a kenyér tömege várható értékben *nagyobb*, mint 1 kg, hanem csak arra, hogy *kisebb*. Ennek megfelelően felállíthatunk egy másik alternatívát (megint csak „az ártatlanság vélelmében”):

$$H_0: \mu \geq \mu_0 \quad \text{vs.} \quad H_1: \mu < \mu_0, \quad (16.1)$$

ahol μ_0 változatlanul 1-gyel egyenlő.

Vajon teljesíti-e ez az alternatíva azt a feltételt, hogy ha tudjuk, hogy teljesül, akkor a mintaelemek eloszlása is ismert? Szigorú értelemben nem: H_0 teljesülése esetén lehetnek a mintaelemek például $N(\mu_0; \sigma^2)$ és $N(\mu_0 + 113; \sigma^2)$ eloszlásúak is. Azonban a H_0 teljesülése esetén egy μ_0 -nál kisebb átlagú mintát akkor kapunk a legnagyobb valószínűséggel, ha μ megegyezik μ_0 -lal (nem pedig nagyobb nála). Ezért ha egy tetszőleges A eseményre – a jelöléssel enyhén visszaélve – bevezetjük a

$$\mathbb{P}(A|H_0 \text{ igaz}) := \mathbb{P}(A | \text{a mintaelemek várható értéke } \mu_0)$$

jelölést, vagyis úgy fogjuk fel, hogy H_0 teljesülése esetén a várható érték egyenlő μ_0 -lal, akkor a kétoldali, egymintás u -próbához hasonlóan járhatunk el.

Az $u(\mathbf{x})$ próbastatisztika-realizácóink most is ugyanaz lesz, mint a kétoldali esetben, a különbség csak az lesz, hogy most csak akkor utasítjuk el H_0 -t, ha az $u = \frac{\bar{x}_n - \mu_0}{\sigma} \sqrt{n}$ értéke nagyon kicsi. Pontosabban olyan elfogadási tartományt választunk, amely jobbról végtelen, és $\mu = \mu_0$ esetén $1 - \varepsilon$ valószínűséggel esik μ_0 az \mathfrak{X}_e elfogadási tartományba, tehát ε valószínűséggel a tőle balra eső \mathfrak{X}_k kritikus tartományba:

$$\mathfrak{X}_k = \{\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n : u(\mathbf{x}) < -u_\varepsilon\},$$

ahol $u_\varepsilon = \Phi^{-1}(1 - \varepsilon)$ továbbra is a standard normális eloszlás $1 - \varepsilon$ -kvantilisét jelöli.

Példánkban $u(\mathbf{x}) = -5$ -öt kaptunk. A standard normális eloszlás $1 - \varepsilon = 0.95$ -kvantilise $\Phi^{-1}(1 - \varepsilon) = \Phi^{-1}(0.95) \approx 1.6449$. Mivel $u(\mathbf{x}) = -5 < -u_\varepsilon$, H_0 -t ismét elutasítjuk, vagyis azt kapjuk, hogy a kenyerek tömege várható értékben *kisebb*, mint 1 kg.

Az egyoldali, egymintás u -próba tehát μ_0 -nál nagyobb átlagú minta esetén soha nem utasítja el H_0 -t, viszont μ_0 -nál kisebb átlagú minták esetén szigorúbb, mint a kétoldali, egymintás próba: a példánkban ha $u(\mathbf{x})$ a -1.9600 és a -1.6449 számok közé esik, akkor az egyoldali próba már elutasítja H_0 -t, de a kétoldali még elfogadja. Ez az egyoldali próba fő előnye.

Ha viszont a

$$H_0: \mu \leq \mu_0 \quad \text{vs.} \quad H_1: \mu > \mu_0 \quad (16.2)$$

alternatívára kívánunk egyoldali, egymintás u -próbát szerkeszteni, akkor a szimmetria miatt az

$$\mathfrak{X}_k = \{\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n : u(\mathbf{x}) > u_\varepsilon\}$$

kritikus tartományt kapjuk. Most tehát csak akkor utasítjuk el H_0 -t, ha a próbastatisztika értéke nagyon nagy. Az eddigiek alapján összefoglaljuk az egyoldali, egymintás u -próba jellemzőit mindkét lehetséges nullhipotézis esetén.

Egyoldali, egymintás u -próba ($H_0: \mu \geq \mu_0$ változat).

- *Nullhipotézis:* $H_0: \mu \geq \mu_0$.
- *Releváns kvantilis:* $-u_\varepsilon = -\Phi^{-1}(1 - \varepsilon)$.
- *Elfogadási tartomány:* $\mathfrak{X}_e = \{(x_1, \dots, x_n) : u(\mathbf{x}) \geq -u_\varepsilon\}$, tehát:
 - ha $u(\mathbf{x}) < -u_\varepsilon$, H_0 -t elutasítjuk,
 - különben H_0 -t elfogadjuk.

Egyoldali, egymintás u -próba ($H_0: \mu \leq \mu_0$ változat).

- *Nullhipotézis:* $H_0: \mu \leq \mu_0$.
- *Releváns kvantilis:* $u_\varepsilon = \Phi^{-1}(1 - \varepsilon)$.
- *Elfogadási tartomány:* $\mathfrak{X}_e = \{(x_1, \dots, x_n) : u(\mathbf{x}) \leq u_\varepsilon\}$, tehát:
 - ha $u(\mathbf{x}) > u_\varepsilon$, H_0 -t elutasítjuk,
 - különben H_0 -t elfogadjuk.

16.2.3. *Megjegyzés* (Próba ereje, az u -próba optimalitása). Igazoltuk (legalább vázlatosan), hogy a kétoldali, egymintás u -próba terjedelme, vagyis elsőfajú hibájának valószínűsége legfeljebb az előzetesen választott ε paraméter lesz. Mit mondhatunk a másodfajú hibáról? Ha H_1 teljesül, akkor nem tudjuk, milyen eloszlásúak a mintaelemeink, mert normális eloszlásuk várható értéke ismeretlen. Vegyünk most egy $\mu_1 \neq \mu_0$ paramétert, ahol μ_0 a nullhipotézis szerinti várható érték. Tehát ha μ_1 a valódi paraméter, akkor H_1 igaz. Ha tudnánk, hogy μ_1 a valódi paraméter, akkor a másodfajú hibát a

$$\mathbb{P}(H_0\text{-t elfogadjuk} \mid \mu_1 \text{ a valódi paraméter}) = \mathbb{P}(u(\mathbf{X}) \in [-u_{\varepsilon/2}, u_{\varepsilon/2}] \mid X_1 \sim N(\mu_1; \sigma^2))$$

képlettel írhatnánk le, ami egy kiszámítható, konkrét érték. Ebben az esetben a próba **erejének** 1 mínusz a másodfajú hiba valószínűségét nevezzük. Az erő tehát ebben az esetben csak konkrét H_1 -beli paraméter esetén definiálható. Látható, hogy minél távolabb van μ_1 a μ_0 értéktől, annál kisebb lesz a másodfajú hiba valószínűsége, vagyis annál nagyobb lesz a próba ereje.

Miért a (kétoldali, egymintás) u -próbát alkalmazzuk a gyakorlatban és nem valamilyen más, szintén ε terjedelmű próbát ugyanezre az alternatívára (feltételezve, hogy a mintaelemek eloszlása normális és a σ szórásuk adott)? Azért, mert bizonyítható, hogy az u -próba az **egyenletesen legerősebb** próba valamennyi ilyen próba közül. Ez azt jelenti, hogy bárhogy is választunk egy H_1 -beli μ_1 paramétert, $\mu = \mu_1$ esetén a másodfajú hiba akkor lesz a legkisebb, ha az u -próbát alkalmazzuk. Ezt az állítást itt nem bizonyítjuk, sőt nem is mondjuk ki formálisan, csak megjegyezzük, hogy ez a **Neyman–Pearson-lemma** következménye. A Neyman–Pearson-lemma az egyenletesen legerősebb próba létezését mondja ki és konstrukcióját adja meg általános esetben, és többek között matematikus hallgatóknak szóló matematikai statisztikai tárgyaknál találkozhatunk vele (pl. a BME-TTK-n is).

Ilyen értelemben egyenletesen legerősebb próba az egyoldali, egymintás és a kétoldali, kétmintás u -próba is, valamint az összes t -próba, amit ebben a jegyzetben tárgyalunk, mindegyik a maga alternatívájára vonatkozóan.

Lássunk most egy olyan példát, ahol eredendően nem normális eloszlással van dolgunk, de a CHT miatt az u -próba mégis alkalmazható!

16.2.4. *Példa.* Az utcán talált, ismeretlen $\vartheta \in (0, 1)$ valószínűséggel fejet mutató pénzérmét $n = 100$ -szor feldobjuk, a dobás eredménye 60-szor lesz fej és 40-szer írás. Döntsük el az egyoldali, egymintás u -próba segítségével, hogy a pénzérme szabályosnak tekinthető-e, ha a próba terjedelme $\varepsilon = 0.05$, illetve $\varepsilon = 0.01$!

Az $X_i = \mathbf{1}_{\{\text{az } i. \text{ dobás eredménye fej}\}}$ fae. mintaelemeink ebben az esetben nem normális eloszlásúak. Ezek ϑ paraméterű indikátorváltozók, tehát $\mathbb{E}\vartheta[X_i] = \vartheta$ valószínűséggel mutatnak fejet, a szórásnégyzetük pedig szintén ismeretlen: $\mathbb{D}_\vartheta^2[X_i] = \vartheta(1 - \vartheta)$.

A minta mérete $n \geq 30$ esetén elég nagy ahhoz, hogy a centrális határeloszlás-tételt alkalmazzuk. Tekintsük a

$$H_0: \vartheta \leq \frac{1}{2} \quad \text{vs.} \quad H_1: \vartheta > \frac{1}{2}$$

alternatívát. Ismét igaz az, hogy egy $n\vartheta = 50$ -nél nagyobb átlagú mintát H_0 teljesülése esetén akkor kapunk a legnagyobb valószínűséggel, ha $\vartheta = \frac{1}{2}$. Ezért (a jelöléssel ismét némiképpen visszaélve) feltehetjük, hogy H_0 teljesülése esetén $\vartheta = \frac{1}{2}$. Ebben az esetben az X_i mintaelemek várható értéke $\frac{1}{2}$, szórásnégyzete $\frac{1}{4}$. Így a centrális határeloszlás-tétel miatt

$$u(X_1, \dots, X_n) = \frac{\sum_{i=1}^n X_i - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}}$$

valószínűségi változó közelítőleg $N(0; 1)$ eloszlású. Az egyoldali, egymintás u -próba szellemében ha ezen statisztika értéke a $(-\infty, u_\varepsilon]$ intervallumba esik, akkor H_0 -t elfogadjuk, különben pedig H_0 -t elutasítjuk.

Konkrét példánkban

$$u(x_1, \dots, x_n) = \frac{60 - 50}{10/2} = 2.$$

Ezt kell $\varepsilon = 0.05$ esetén $u_\varepsilon = \Phi^{-1}(1 - \varepsilon) = \Phi^{-1}(0.95) = 1.6449$ -cel, $\varepsilon = 0.01$ esetén pedig $u_\varepsilon = \Phi^{-1}(0.99) = 2.3263$ -mal összehasonlítani. Ezért $\varepsilon = 0.05$ esetén H_0 -t elutasítjuk (tehát arra következtetünk, hogy a pénzérme nem szabályos), viszont $\varepsilon = 0.01$ esetén elfogadjuk (vagyis arra a konklúzióra jutunk, hogy az érme szabályos).

16.2.5. *Megjegyzés* (p -érték). Láthatjuk, hogy ha ε -t csökkentjük, akkor H_0 -t gyakrabban fogjuk elfogadni; ez nem meglepő, hiszen ε a próba terjedelmével, vagyis az elsőfajú hiba valószínűségével egyezik meg. A fenti pénzérmés példában $\varepsilon = 0.01$ és $\varepsilon = 0.05$ között található az a legnagyobb terjedelem, amelyre H_0 -t az adott mintának megfelelő $u(\mathbf{x})$ próbastatisztika-érték esetén még elfogadjuk. Ezt szokás a próba p -értékének (az angol nyelvű irodalomban p -value) nevezni. Minél kisebb a p -érték, annál jogosabb a H_0 nullhipotézis elutasítása.

A pénzérmés példában (egyoldali, egymintás u -próba) a p -érték éppen 1 mínusz a standard normális eloszlásfüggvénynek a próbastatisztikában vett értéke:

$$u(\mathbf{x}) = u_\varepsilon \Leftrightarrow \mathbb{P}(u(\mathbf{X}) > u(\mathbf{x})) = \varepsilon \Leftrightarrow \varepsilon = 1 - \Phi(u(\mathbf{x})) = 1 - \Phi(2) \approx 1 - 0.9772 = 0.0228.$$

Vagyis a próba p -értéke kb. 2,28%. Ennél kisebb (vagy ezzel egyenlő) terjedelem esetén H_0 -t elfogadjuk, ennél nagyobb terjedelem esetén elutasítjuk.

Kétoldali u -próba esetén, például a kenyeres példában:

$$|u(\mathbf{x})| = u_{\varepsilon/2} \Leftrightarrow \varepsilon = 2(1 - \Phi(|u(\mathbf{x})|)) = 2(1 - \Phi(5)) \approx 5.733 \times 10^{-7}.$$

Vagyis a próba p -értéke rendkívül alacsony; H_0 -t csak akkor fogadjuk el, ha $\varepsilon \leq 5.733 \times 10^{-7}$.

A p -érték használata egy alternatív szemléletet jelent az előírányzott terjedelemhez képest: itt nem a mintavétel előtt határozzuk meg a próba ε terjedelmét, ez és a minta alapján döntést hozva H_0 elfogadásáról vagy elutasításáról, hanem a minta alapján határozzuk meg azt az ε terjedelmet, amely pont a H_0 elfogadása és elutasítása határán áll, azaz a p -értéket.

Gyakori szóhasználat, hogy valamilyen mintára vonatkozó állítás „szignifikáns”, ha a p -érték legfeljebb 0.05, és „nagyon szignifikáns”, ha legfeljebb 0.01. Ez a kifejezőmód szinte bármilyen hipotézisvizsgálati módszer esetén használható, és azt jelenti, hogy ha elfogadjuk a minta alapján a nullhipotézist, akkor legfeljebb 0.05, illetve 0.01 lehet az elsőfajú hiba valószínűsége (azaz a terjedelem).

Alfejezetünk végén térjünk rá a **kétmintás** u -próbara, amellyel két normális eloszlású minta várható értékének egyenlőségét tesztelhetjük. Ennek itt csak a kétoldali változatát közöljük, ebből az egyoldali változat ugyanúgy vezethető le, mint az egymintás esetben. Kezdjük ismét a kenyeres példával, illetve annak egy folytatásával.

16.2.6. *Példa* (Kétoldali, kétmintás u -próba: kenyérvásárlás). Miután megtudtuk, hogy $\varepsilon = 0.05$ terjedelem mellett el kell utasítanunk azt a nullhipotézist, hogy a péknél vásárolt kenyerek várható értékben egykilósak, alternatív kenyérforrás után nézünk. Lakóhelyünk közelében van egy másik pék, aki nagyon hasonló kinézetű, ízű és árú egykilós kenyereket gyárt, mint az első. Ez a pék azzal reklámozza magát, hogy precíziós eszközök segítségével képes olyan kenyeret gyártani, amelyek tömegének szórása csak 1 dkg (azaz $\sigma_1 = 0.01$ kg, tehát $\sigma_1^2 = 0.0001$ kg² a szórásnégyzet). Most $n_1 = 10$ napon keresztül ennél a péknél vásárolunk napi egy kenyeret, és megmérjük a tömegüket. Egy olyan $\mathbf{x} = (x_1, \dots, x_{n_1})$ realizációt kapunk, amelynek átlaga $\bar{x}_{n_1} = 0.9825$ kg. Vezessünk be egy új, a későbbiekben hasznos jelölést a régi péknél vásárolt kenyerek tömegeire: $\mathbf{y} = (y_1, \dots, y_{n_2})$, ahol a mintaelemszám $n_2 = 25$, a mintaátlag $\bar{y}_{n_2} = 0.98$, a szórás $\sigma_2 = 0.02$ volt. (Feltételezzük, hogy bármely, a két pék egyike által gyártott kenyér tömege független bármely, ugyanezen pék vagy a másik pék által gyártott kenyér tömegétől.)

Ez alapján kellene eldöntenünk, hogy jobban járunk-e, ha ehhez a pékhez járunk, mint a másikhoz, vagyis hogy az új péknél vásárolt kenyerek tömegének ismeretlen μ_1 várható értéke nagyobb-e, mint a régi péknél vásároltak tömegének μ_2 várható értéke. Ezért a

$$H_0: \mu_1 = \mu_2 \quad \text{vs.} \quad H_1: \mu_1 \neq \mu_2$$

alternatívára kívánunk egy $\varepsilon = 0.95$ terjedelmű próbát konstruálni. Feltételezve, hogy az (X_1, \dots, X_{n_1}) minta független az (Y_1, \dots, Y_{n_2}) mintától, az \bar{X}_{n_1} mintaátlag is független az \bar{Y}_{n_2} mintaátlagtól. Ezért \bar{X}_{n_1} és \bar{Y}_{n_2} lineáris kombinációi is normális eloszlásúak. Elemi számításokkal (amiket itt nem részletezünk) megmutatható, hogy

$$u(\mathbf{X}, \mathbf{Y}) = \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

standard normális eloszlású. Ennek ismeretében hasonlóan járhatunk el, mint a kétoldali, egymintás u -próba esetén. Ha

$$u(\mathbf{X}, \mathbf{Y}) \in [-u_{\varepsilon/2}, u_{\varepsilon/2}] = [-\Phi^{-1}(1 - \varepsilon/2), \Phi^{-1}(1 - \varepsilon/2)],$$

akkor H_0 -t elfogadjuk, különben pedig elutasítjuk. Vagyis a kritikus tartományunk

$$\mathfrak{X}_k = \{(\mathbf{x}, \mathbf{y}) = (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \in \mathbb{R}^{n_1+n_2} : |u(\mathbf{x}, \mathbf{y})| > u_{\varepsilon/2}\}.$$

Konkrét példánkban

$$u(\mathbf{x}, \mathbf{y}) = \frac{0.9825 - 0.98}{\sqrt{\frac{0.0001}{10} + \frac{0.0004}{25}}} \approx 0,4903.$$

Mivel $|u(\mathbf{x}, \mathbf{y})| < u_{\varepsilon/2} = \Phi^{-1}(0.975) = 1.9600$, H_0 -t elfogadjuk. Vagyis várható értékben az új péknél is ugyanolyan tömegűek a kenyerek, mint a réginél, tehát a pékváltás nem segíti a helyzetünket...

Foglaljuk össze tehát a kétoldali, kétmintás u -próba eljárását:

Eljárás (Kétoldali, kétmintás u -próba). *Feltevés:* az x_1, \dots, x_{n_1} adatpontok fae. normális eloszlású X_1, \dots, X_{n_1} valószínűségi változók realizációi, amelyek $\sigma_1 > 0$ szórása ismert, μ_1 várható értéke viszont ismeretlen, az y_1, \dots, y_{n_2} adatpontok pedig az X_1, \dots, X_{n_1} -től (együttesen) független, egymás között fae. normális eloszlású Y_1, \dots, Y_{n_2} valószínűségi változók realizációi, ezek $\sigma_2 > 0$ szórása ismert, μ_2 várható értéke viszont ismeretlen. *Eljárás:*

1. Nullhipotézis: $H_0: \mu_1 = \mu_2$ vs. $H_1: \mu_1 \neq \mu_2$.
2. Az előírányzott $\varepsilon \in (0, 1)$ terjedelem kiválasztása.

3. Az $x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}$ adatpontok kinyerése.

4. Próbastatisztika:

$$u(\mathbf{x}, \mathbf{y}) = \frac{\bar{x}_{n_1} - \bar{y}_{n_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

5. Elfogadási tartomány: $\mathfrak{X}_e = \{(\mathbf{x}, \mathbf{y}) = (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \in \mathbb{R}^{n_1+n_2} : |u(\mathbf{x}, \mathbf{y})| \leq u_{\varepsilon/2}\}$, ahol $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$.

6. Döntés: ha $|u(\mathbf{x}, \mathbf{y})| > u_{\varepsilon/2}$, akkor H_0 -t elutasítjuk, különben elfogadjuk.

16.3. t -próba

Amint már említettük, a t -próba ismeretlen szórású normális eloszlással rendelkező fae. mintaelemek esetén használható. Ebből a próbatípusból is az egyoldali, egy- illetve kétmintás, valamint a kétoldali, kétmintás próbát ismertetjük ebben a jegyzetben. Ezek többé-kevésbé analóg módon viselkednek, mint az u -próba megfelelő változataival, azzal a különbséggel, a próbastatisztikákban a szórás helyett a korrigált empirikus szórás fog szerepelni, és a próbastatisztikát a standard normális eloszlás helyett egy bizonyos szabadságfokú t -eloszlás megfelelő kvantilisével kell összehasonlítani. Kétmintás esetben a próbastatisztika formája jelentősebb mértékben változik az u -próbaéhoz képest. Kezdjük most az egymintás esettel.

16.3.1. *Példa* (Kétoldali, egymintás t -próba). Legyen $\mathbf{X} = (X_1, \dots, X_n)$ fae. minta $N(\mu; \sigma^2)$ eloszlású mintaelemekkel, ahol ezúttal mind a μ várható érték, mind a σ^2 szórásnégyzet ismeretlen. Kétoldali, egymintás esetben az alternatívánk továbbra is

$$H_0: \mu = \mu_0 \quad \text{vs.} \quad H_1: \mu \neq \mu_0.$$

A (szintén kétoldali, egymintás) ε terjedelmű u -próba alapja az volt, hogy

$$u(\mathbf{X}) = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma^2} \sim N(0; 1),$$

és ez alapján ha $|u(\mathbf{X})| \leq u_{\varepsilon/2}$, akkor H_0 -t elfogadjuk, különben elutasítjuk. Ugyanakkor a 15.5 alfejezetben említettük és használtuk, hogy

$$t(\mathbf{X}) = \frac{\sqrt{n}(\bar{X}_n - \mu)}{S_n^*} \sim t(n-1),$$

ahol $S_n^* = \sqrt{(S_n^*)^2}$ a mintához tartozó korrigált empirikus szórást jelöli. Emiatt az u -próba levezetésével (lásd 16.2.1. Példa) analóg módon azt kapjuk, hogy akkor kapunk ε terjedelmű próbát, ha H_0 -t akkor fogadjuk el, ha $|t(\mathbf{X})| \leq t_{n-1, \varepsilon/2}$, ahol $t_{n-1, \varepsilon/2}$ az $n-1$ szabadságfokú t -eloszlás $1 - \varepsilon/2$ -kvantilisét jelöli, és különben H_0 -t elutasítjuk.

Hasonlóan az egyoldali, kétmintás t -próbát is az u -próba megfelelő változatával definiálhatjuk, a próbastatisztikában (amit most $t(\mathbf{X})$ -szel jelölünk) a szórást a korrigált empirikus szórásra cserélve és a normális eloszlás releváns kvantilise helyett az $n-1$ szabadságfokú t -eloszlás ugyanezen kvantilisét tekintve. Írjuk le az egymintás t -próbák eljárását az u -próbaéhoz hasonló formában:

Eljárás (Egymintás t -próba). *Feltevések:* az x_1, \dots, x_n adatpontok fae. normális eloszlású X_1, \dots, X_n valószínűségi változók realizációi, amelyek szórása és μ várható értéke ismeretlen. *Általános eljárás:*

1. A H_0 nullhipotézis felállítása.
2. Az előírányzott $\varepsilon \in (0, 1)$ terjedelem kiválasztása.
3. Az x_1, \dots, x_n adatpontok kinyerése.
4. Az $f = n - 1$ szabadsági fok meghatározása.

5. Az elfogadási tartomány meghatározása. Ez függ a próba altípusától, de mindig magába foglalja az $n - 1$ szabadságfokú t -eloszlás egy megfelelő kvantilisének kiszámítását (táblázatból való kiolasását). Ez a táblázat csak nemnegatív helyeken adja meg az $n - 1$ szabadságfokú t -eloszlás eloszlásfüggvényének értékeit, a negatív helyeken vett értékek kiszámítása a normális eloszlással analóg módon történik (mivel a t -eloszlások sűrűségfüggvénye is a 0-ra szimmetrikus).
6. A próbastatisztika $t(\mathbf{x}) = \frac{\bar{x}_n - \mu_0}{s_n^*} \sqrt{n}$ értékének meghatározása.
7. Annak eldöntése, hogy a próbastatisztika értéke az elfogadási tartományba esik-e. Ha igen, H_0 -t elfogadjuk. Ha nem, H_0 -t elutasítjuk.

Kétoldali, egymintás t -próba.

- Nullhipotézis: $H_0: \mu = \mu_0$.
- Releváns kvantilis: $t_{n-1, \varepsilon/2}$ az $f = n - 1$ szabadságfokú t -eloszlás $1 - \varepsilon/2$ -kvantilise.
- Elfogadási tartomány: $\mathfrak{X}_e = \{(x_1, \dots, x_n) : |t(\mathbf{x})| \leq t_{n-1, \varepsilon/2}\}$, tehát:
 - ha $|t(\mathbf{x})| > t_{n-1, \varepsilon/2}$, H_0 -t elutasítjuk,
 - különben H_0 -t elfogadjuk.

Egyoldali, egymintás t -próba ($H_0: \mu \geq \mu_0$ változat).

- Nullhipotézis: $H_0: \mu \geq \mu_0$.
- Releváns kvantilis: $-t_{n-1, \varepsilon}$, ahol $t_{n-1, \varepsilon}$ az $n - 1$ szabadságfokú t -eloszlás $1 - \varepsilon$ -kvantilise.
- Elfogadási tartomány: $\mathfrak{X}_e = \{(x_1, \dots, x_n) : t(\mathbf{x}) \geq -t_{n-1, \varepsilon}\}$, tehát:
 - ha $t(\mathbf{x}) < -t_{n-1, \varepsilon}$, H_0 -t elutasítjuk,
 - különben H_0 -t elfogadjuk.

Egyoldali, egymintás t -próba ($H_0: \mu \leq \mu_0$ változat).

- Nullhipotézis: $H_0: \mu \leq \mu_0$.
- Releváns kvantilis: $t_{n-1, \varepsilon}$ az $n - 1$ szabadságfokú t -eloszlás $1 - \varepsilon$ -kvantilise.
- Elfogadási tartomány: $\mathfrak{X}_e = \{(x_1, \dots, x_n) : t(\mathbf{x}) \leq t_{n-1, \varepsilon}\}$, tehát:
 - ha $t(\mathbf{x}) > t_{n-1, \varepsilon}$, H_0 -t elutasítjuk,
 - különben H_0 -t elfogadjuk.

16.3.2. *Példa* (Egymintás t -próba: kenyérvásárlás.). Miután elszomorodtunk azon, hogy az u -próbáink eredménye alapján a pékünk valóban várható értékben 1 kg-nál kisebb kenyereket ad nekünk, viszont érdemben a másik péknél sem jobb a helyzet, esziünkbe jut, hogy talán csak azért kaptuk ezeket az eredményeket, mert a szórás értékéről rossz feltételezéseink voltak. Ezért elhatározzuk, hogy az adott alternatívákra elvégezzük a t -próba megfelelő változatait is. Kezdjük ebben a példában az egymintás t -próbákkal!

Szerencsére mindennap felírtuk, hogy aznap milyen nehéz kenyeret vettünk, ezért ki tudjuk számolni a korrigált empirikus szórásokat is. Az első péknél vett kenyerek tömegéből álló, $n = 25$ elemű (x_1, \dots, x_n) mintarealizáció esetén, amelyre $\bar{x}_n = 0.98$ kg teljesült, azt kapjuk, hogy a korrigált empirikus szórás $s_{25}^* = 0.05$ kg (vagyis lényegesen nagyobb a valódi szórás korábban feltételezett 0.02 értékénél).

Kétoldali, egymintás t -próba esetén a

$$H_0: \mu = 1 \quad \text{vs.} \quad H_1: \mu \neq 1$$

alternatívára felállított $\varepsilon = 0.05$ terjedelmű próbánál a próbastatisztika értéke

$$t(\mathbf{x}) = \frac{\sqrt{25}(0.98 - 1)}{0.05} = -\frac{10}{5} = -2.$$

Ennek abszolút értékét kell az $n - 1 = 24$ szabadságfokú t -eloszlás $1 - \varepsilon/2 = 0.975$ -kvantilisével összehasonlítani. A kvantilis értéke

$$t_{24,0.025} \approx 2.0639.$$

Ez nagyobb, mint $u_{0.975} \approx 1.9600$, és valamivel nagyobb $|t(\mathbf{x})|$ -nél is. Ezért H_0 -t elfogadjuk, vagyis azzal a következtetéssel vigasztalhatjuk magunkat, hogy a kenyerek várható értékben mégis 1 kilogramm tömegűek.

Egyoldali, egymintás t -próba esetén a

$$H_0: \mu \geq 1 \quad \text{vs.} \quad H_1: \mu < 1$$

alternatívára felállított $\varepsilon = 0.05$ terjedelmű próbánál a próbastatisztika értéke megint csak $t(\mathbf{x}) = -2$. Ezt kell $-t_{24,0.05} \approx -1.7109$ -cel összehasonlítani. Mivel $t(\mathbf{x}) < -t_{24,0.05}$, H_0 -t a „szigorúbb” egyoldali próba alapján még mindig elutasítjuk, vagyis arra a következtetésre jutunk, hogy a kenyerek várható értékben 1 kg-nál kisebb tömegűek.

Végül térjünk rá a kétoldali, kétmintás t -próba esetére.

16.3.3. *Példa.* Legyenek $X_1, \dots, X_{n_1} \sim N(\mu_1; \sigma_1^2)$ és $Y_1, \dots, Y_{n_2} \sim N(\mu_2; \sigma_2^2)$ függetlenek, ahol a μ_1 és μ_2 várható értékek ismeretlenek, és tekintsük a

$$H_0: \mu_1 = \mu_2 \quad \text{vs.} \quad H_1: \mu_1 \neq \mu_2$$

alternatívát. Láttuk (bár nem bizonyítottuk), hogy

$$u(\mathbf{X}, \mathbf{Y}) = \frac{\overline{X_{n_1}} - \overline{Y_{n_2}}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0; 1).$$

Ezért ha a σ_1^2 és σ_2^2 szórásnégyzetek ismertek, akkor az $u(\mathbf{X}, \mathbf{Y})$ statisztikát használva konstruálhatunk egy ε terjedelmű próbát erre az alternatívára, ami nem más, mint a kétoldali, kétmintás u -próba.

Hogyan alkossunk ε terjedelmű próbát ugyanerre az alternatívára, ha σ_1^2 és σ_2^2 is ismeretlenek? Ebben az esetben belátható, hogy amennyiben $\sigma_1^2 = \sigma_2^2$,

$$t(\mathbf{X}, \mathbf{Y}) = \frac{\overline{X_{n_1}} - \overline{Y_{n_2}}}{\sqrt{\frac{(n_1-1)(s_X^*)^2 + (n_2-1)(s_Y^*)^2}{n_1+n_2-2}}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

t -eloszlású $n_1 + n_2 - 2$ szabadsági fokkal, ahol $(s_X^*)^2$ jelöli az \mathbf{X} mintához tartozó korrigált empirikus szórásnégyzetet és $(s_Y^*)^2$ az \mathbf{Y} -hoz tartozót (ezt sem bizonyítjuk).¹⁸ A kritikus tartományunk ebben az esetben tehát

$$\mathfrak{X}_k = \{(\mathbf{x}, \mathbf{y}) = (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \in \mathbb{R}^{n_1+n_2} : |t(\mathbf{x}, \mathbf{y})| > t_{n_1+n_2-2, \varepsilon/2}\},$$

ahol $t_{n_1+n_2-2, \varepsilon/2}$ az $n_1 + n_2 - 2$ szabadsági fokú t -eloszlás $1 - \varepsilon/2$ -kvantilise (az eddigiekkel konzisztens jelölést használva). Vagyis ha $|t(\mathbf{x}, \mathbf{y})| \leq t_{n_1+n_2-2, \varepsilon/2}$, akkor H_0 -t elfogadjuk, különben pedig H_0 -t elutasítjuk.

Ez alapján formalizálhatjuk a kétoldali, kétmintás t -próba eljárását, amely természetesen sokban hasonlít a kétoldali, kétmintás u -próbaéra. Vigyázat: a próbastatisztika változása itt nem csupán annyi, hogy a szórásnégyzeteket a korrigált empirikus megfelelőjükre cseréljük, és a próba csak akkor alkalmazható, ha a két minta ismeretlen szórással megegyeznek!

¹⁸Általában igaz az a (matematikailag nem mindig egyértelmű) ökölszabály, hogy a szabadsági fok az összes mintaelem száma mínusz a becsült paraméterek száma. Esetünkben az egymintás t -próbánál egyetlen becsült paraméterünk volt: a valódi szórással helyett a minta korrigált empirikus szórással dolgoztunk, így a szabadsági fok $n - 1$ volt. A kétmintás t -próbánál már két szórást helyettesítünk a becsült értékével, ezért kapunk $n_1 + n_2 - 2$ szabadsági fokot.

Eljárás (Kétoldali, kétmintás t -próba). *Feltevés:* az x_1, \dots, x_{n_1} adatpontok fae. normális eloszlású X_1, \dots, X_{n_1} valószínűségi változók realizációi, amelyek szórása ismeretlen, ahogy μ_1 várható értékük is, az y_1, \dots, y_{n_2} adatpontok pedig az X_1, \dots, X_{n_1} -től (együttesen) független, egymás között fae. normális eloszlású Y_1, \dots, Y_{n_2} valószínűségi változók realizációi, ezek μ_2 várható értékük sem ismert, szórásuk pedig megegyezik az X_i valószínűségi változók (ismeretlen) szórásával.¹⁹ Eljárás:

1. Nullhipotézis: $H_0: \mu_1 = \mu_2$ vs. $H_1: \mu_1 \neq \mu_2$.
2. Az előírányzott $\varepsilon \in (0, 1)$ terjedelem kiválasztása.
3. Az $x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}$ adatpontok kinyerése.
4. Az $f = n_1 + n_2 - 2$ szabadsági fok meghatározása.
5. Próbat statisztika:

$$t(\mathbf{x}, \mathbf{y}) = \frac{\bar{x}_{n_1} - \bar{y}_{n_2}}{\sqrt{\frac{(n_1-1)(s_x^*)^2 + (n_2-1)(s_y^*)^2}{n_1+n_2-2}}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}.$$

6. Elfogadási tartomány: $\mathfrak{X}_e = \{(\mathbf{x}, \mathbf{y}) = (x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}) \in \mathbb{R}^{n_1+n_2} : |t(\mathbf{x}, \mathbf{y})| \leq t_{n_1+n_2-2, \varepsilon/2}\}$, ahol $t_{n_1+n_2-2, \varepsilon/2}$ az $f = n_1 + n_2 - 2$ szabadságfokú t -eloszlás $1 - \varepsilon/2$ -kvantilise.
7. Döntés: ha $|t(\mathbf{x}, \mathbf{y})| > t_{\varepsilon/2}$, akkor H_0 -t elutasítjuk, különben elfogadjuk.

Zárásként alkalmazzuk a kétoldali, kétmintás t -próbát a kenyeres példára.

16.3.4. *Példa.* A kétmintás próbának megfelelő jelöléssel az eddigiek szerint az új péknél vásárolt 10 db kenyér adatai: $n_1 = 10$, $\bar{x}_{n_1} = 0.9825$ kg. Szerencsére gondosan felírtuk az új péknél vásárolt kenyerek tömegadatait is, ezek korrigált empirikus szórására pedig $s_x^* = 0.015$ kg adódott. A régi péknél vásárolt 25 kenyérhez tartozó adatok az eddigiek szerint: $n_2 = 25$, $\bar{y}_{n_2} = 0.98$ kg, $s_y^* = 0.05$ kg. Így a szabadsági fok $f = n_1 + n_2 - 2 = 33$. A

$$H_0: \mu_1 = \mu_2 \text{ vs. } H_1: \mu_1 \neq \mu_2$$

alternatívára – ismeretlen szórások mellett – a kétoldali, kétmintás t -próbát alkalmazzuk. A próbat statisztika értéke:

$$t(\mathbf{x}, \mathbf{y}) = \frac{\bar{x}_{n_1} - \bar{y}_{n_2}}{\sqrt{\frac{(n_1-1)(s_x^*)^2 + (n_2-1)(s_y^*)^2}{n_1+n_2-2}}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \frac{0.9825 - 0.98}{\sqrt{\frac{9 \cdot 0.015^2 + 24 \cdot 0.05^2}{33}}} \sqrt{\frac{10 \cdot 25}{10 + 25}} \approx 0.1541.$$

Ez (abszolút értékben) jóval kisebb, mint az $n_1 + n_2 - 2 = 33$ szabadságfokú t -eloszlás $1 - \varepsilon/2 = 0.975$ -kvantilise, $t_{33, 0.025} \approx 2.0345$. Ezért H_0 -t elfogadjuk, azaz (ismét) arra a következtetésre jutunk, hogy a régi és az új péknél vásárolt kenyér várható értéke megegyezik.²⁰

16.4. Megjegyzések a hipotézisvizsgálati képletgyűjtemény használatához

A vizsgán a dokumentum utolsó előtti oldalán szereplő képletgyűjtemény és az utolsó oldalon szereplő táblázatok használhatóak, ennek megfelelő részletei szerepelni fognak a dolgozaton, amennyiben van konfidenciaintervallummal vagy hipotézisvizsgálattal kapcsolatos feladat. A táblázat a mérnök- és gazdaságinformatikus MSc Matematikai statisztika tárgyából származik (kissé cenzúrázva). Az alábbi megjegyzések ismerete fontos ahhoz, hogy a képletgyűjtemény valóban hasznunkra legyen:

¹⁹Ebből a tárgyból nem foglalkozunk azzal az esettel, amikor $\sigma_1^2 = \sigma_2^2$ nem teljesül. A szórásnégyzetek egyenlőségét is lehet statisztikai próbával tesztelni, mégpedig a $H_0: \sigma_1^2 = \sigma_2^2$ vs. $H_1: \sigma_1^2 \neq \sigma_2^2$ alternatívájú F -próba segítségével. Ha az F -próba azt az eredményt adja, hogy $\sigma_1^2 = \sigma_2^2$, akkor alkalmazható a kétoldali, kétmintás t -próba. Ellenkező esetben ehelyett a *Welch-próbát* kell alkalmazni, amelynek próbat statisztikáját más módon kell kiszámolni. A próbat statisztika szintén t -eloszlású lesz, de a t -eloszlás szabadsági foka általában nem egész szám lesz. Ezek a próbák a mérnök- és gazdaságinformatikus MSc képzés Matematikai statisztika tárgyának anyagához tartoznak.

²⁰A kapott próba p -értéke rendkívül alacsony, mert a próbat statisztika $t(\mathbf{x}, \mathbf{y})$ értéke közel van a 0-hoz.

- A képletgyűjteményben a realizációkra utalva szinte mindent kisbetűvel jelölünk, kivéve Φ -t, a standard normális eloszlás eloszlásfüggvényét.
- A képletgyűjteményben szereplő konfidenciaintervallumok szintje $1 - \varepsilon$.
- Általában egy nullhipotézist akkor utasítunk el, ha a próbastatisztikánk értéke (a nullhipotézishez képest) túl extrém. H_0 teljesülése esetén kétoldali próbánál a kritikus tartományba esés valószínűsége mind a megadott kvantilistól jobbra, mind annak ellentettjétől balra is $\varepsilon/2$, míg egyoldali próba esetén a kritikus tartomány egyetlen intervallum, amelybe ε valószínűséggel esik a minta.
- A (bármilyen szabadságfokú) t -eloszlás sűrűségfüggvénye ugyanúgy szimmetrikus a 0-ra, mint a standard normális eloszlásé, ezért a t -eloszlások táblázatában is csak nemnegatív helyen vett értékek szerepelnek. Ha X Student-eloszlású tetszőleges pozitív egész paraméterrel, akkor is igazak a

$$\mathbb{P}(X < -x) = \mathbb{P}(X > x) = \mathbb{P}(X \geq x) = 1 - \mathbb{P}(X < x), \quad \mathbb{P}(|X| < x) = \mathbb{P}(-x < X < x) = 2\mathbb{P}(X < x) - 1,$$

képletek $x \geq 0$ esetén, amelyek segítségével az eloszlásfüggvény értéke negatív helyeken is meghatározható, ugyanúgy, mint a standard normális eloszlás esetén.

- A vizsgán használható hipotézisvizsgálati képletgyűjteményben (lásd ezen dokumentum végén) az egyoldali próbáknak csak a $H_0: \mu \leq \mu_0$ változata szerepel. Ha a $H_0: \mu \geq \mu_0$ nullhipotézisű próbát kell végrehajtanunk és a képletgyűjteményt szeretnénk használni, akkor a mintaelemek -1 -szeresére a $H_0: \mu \leq -\mu_0$ nullhipotézisű próbát kell alkalmazni, hiszen egy normális eloszlású valószínűségi változó -1 -szerese is normális eloszlású, változatlan szórással és -1 -szeresére változó várható értékkel, illetve egy bármilyen szabadságfokkal t -eloszlású valószínűségi változó -1 -szerese is ugyanilyen szabadságfokkal t -eloszlású.

Felhasznált irodalom

A jegyzetkiegészítés anyaga (a Mészáros Szabolcs-féle jegyzeten túl) Noemi Kurt *Stochastik für Informatiker* című tankönyvére (Springer, 2020) és Bolla Marianna (BME-TTK) 2013. tavaszi félévi *Matematikai statisztika* című előadásának anyagára épül.

Konfidenciaintervallumokhoz és hipotézisvizsgálatokhoz képletek

$$s_n^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}_n^2, (s_n^*)^2 = \frac{n}{n-1} s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2, s_n^* = \sqrt{(s_n^*)^2}.$$

u -próba

1. Kétoldali, egymintás: $u = \frac{\bar{x}_n - \mu_0}{\sigma} \sqrt{n}$, $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$,
konfidenciaintervallum μ -re:

$$\left[\bar{x}_n - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \bar{x}_n + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \right].$$

2. Egyoldali, egymintás: $u = \frac{\bar{x}_n - \mu_0}{\sigma} \sqrt{n}$, $u_{\varepsilon} = \Phi^{-1}(1 - \varepsilon)$.

3. Kétoldali, kétmintás: $u = \frac{\bar{x}_{n_1} - \bar{y}_{n_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$, $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$.

t -próba

1. Kétoldali, egymintás: $t = \frac{\bar{x}_n - \mu_0}{s_n^*} \sqrt{n}$, $t_{n-1, \varepsilon/2}$ a t_{n-1} -eloszlás $1 - \varepsilon/2$ -kvantilise,
konfidenciaintervallum μ -re:

$$\left[\bar{x}_n - \frac{s_n^* t_{n-1, \varepsilon/2}}{\sqrt{n}}, \bar{x}_n + \frac{s_n^* t_{n-1, \varepsilon/2}}{\sqrt{n}} \right].$$

2. Egyoldali, egymintás: $t = \frac{\bar{x}_n - \mu_0}{s_n^*} \sqrt{n}$, $t_{n-1, \varepsilon}$ a t_{n-1} -eloszlás $1 - \varepsilon$ -kvantilise.

3. Kétoldali, kétmintás: $t = \frac{\bar{x}_{n_1} - \bar{y}_{n_2}}{\sqrt{\frac{(n_1-1)(s_x^*)^2 + (n_2-1)(s_y^*)^2}{n_1+n_2-2}}} \sqrt{\frac{n_1 n_2}{n_1+n_2}}$, $t_{n-1, \varepsilon/2}$ a $t_{n_1+n_2-2}$ -eloszlás $1 - \varepsilon/2$ -kvantilise.

Konfidenciaintervallum σ^2 -re ismert μ esetén (a 2023/24-es tanévben kiegészítő anyag, nem vizsgaanyag)

$$\left(\frac{(n-1)(s_n^*)^2}{\chi_{n-1, \varepsilon/2}^2}, \frac{(n-1)(s_n^*)^2}{\chi_{n-1, 1-\varepsilon/2}^2} \right),$$

$\chi_{n-1, \alpha}^2$ a χ_{n-1}^2 -eloszlás $1 - \alpha$ -kvantilise.

A t -(Student-)próba kritikus értékei

f	0,1 0,05	0,05 0,025	0,02 0,01	f	0,1 0,05	0,05 0,025	0,02 0,01
1	6,314	12,71	31,82	21	1,721	2,080	2,518
2	2,920	4,303	6,965	22	1,717	2,074	2,508
3	2,353	3,182	4,541	23	1,714	2,069	2,500
4	2,132	2,776	3,747	24	1,711	2,064	2,492
5	2,015	2,571	3,365	25	1,708	2,060	2,485
6	1,943	2,447	3,143	26	1,706	2,056	2,479
7	1,895	2,365	2,998	27	1,703	2,052	2,473
8	1,860	2,306	2,896	28	1,701	2,048	2,467
9	1,833	2,262	2,821	29	1,699	2,045	2,462
10	1,812	2,228	2,764	30	1,697	2,042	2,457
11	1,796	2,201	2,718	40	1,684	2,021	2,423
12	1,782	2,179	2,681	50	1,676	2,009	2,403
13	1,771	2,160	2,650	60	1,671	2,000	2,390
14	1,761	2,145	2,624	70	1,667	1,994	2,381
15	1,753	2,131	2,602	80	1,664	1,990	2,374
16	1,746	2,120	2,583	90	1,662	1,987	2,369
17	1,740	2,110	2,567	100	1,660	1,984	2,364
18	1,734	2,101	2,552	200	1,653	1,972	2,345
19	1,729	2,093	2,539	500	1,648	1,965	2,334
20	1,725	2,086	2,528	∞	1,645	1,960	2,326

Az eloszlás szabadságfoka f , az oszlopok felett a próba terjedelmét adtuk meg, a felső érték kétoldali, az alsó egyoldali ellenhipotézisre vonatkozik.

A χ^2 -próba kritikus értékei

f	0,1	0,05	0,01	f	0,1	0,05	0,01
1	2,71	3,84	6,63	16	23,5	26,3	32,0
2	4,61	5,99	9,21	17	24,8	27,6	33,4
3	6,25	7,81	11,3	18	26,0	28,9	34,8
4	7,78	9,49	13,3	19	27,2	30,1	36,2
5	9,24	11,1	15,1	20	28,4	31,4	37,6
6	10,6	12,6	16,8	21	29,6	32,7	38,9
7	12,0	14,1	18,5	22	30,8	33,9	40,3
8	13,4	15,5	20,1	23	32,0	35,2	41,6
9	14,7	16,9	21,7	24	33,2	36,4	43,0
10	16,0	18,3	23,2	25	34,4	37,7	44,3
11	17,3	19,7	24,7	26	35,6	38,9	45,6
12	18,5	21,0	26,2	27	36,7	40,1	47,0
13	19,8	22,4	27,9	28	37,9	41,3	48,3
14	21,1	23,7	29,1	29	39,1	42,6	49,6
15	22,3	25,0	30,6	30	40,3	43,8	50,9

Az eloszlás szabadságfoka f , az oszlopok felett a próba terjedelmét adtuk meg.