... és statisztika

Tartalom:

- Statisztikai alapfogalmak: minta, realizáció, alapstatisztikák
- Empirikus eloszlásfüggvény és konvergenciája
- Becslések tulajdonságai
- Intervallumbecslés, konfidenciaintervallumok
- Hipotézisvizsgálat, paraméteres próbák

Előismeretek: differenciál- és integrálszámítás főleg egy dimenzióban,
a valszám anyagrészből leginkább: nevezetes diszkrét és folytonos eloszlások, eloszlásfüggvény, sűrűségfüggvény, súlyfüggvény, függetlenség, momentumok (főleg: várható érték és szórás), CHT és változatai, nagy számok erős törvénye.
Jegyzetkiegészítés: `cs.bme.hu/~tobias/statisztika.pdf`.
A jegyzetkiegészítés az idei előadáson el nem hangzó anyagrészeket is tartalmaz.

Emlékeztető: statisztika

Sztochasztika = a véletlen törvényszerűségének tudománya.

A sztochasztika két fő ága:

- Valószínűségelmélet: matematikai modellezés és formalizmus, elméleti alapok
- Statisztika: véletlen eljárásokból származó mérési adatok elemzése és interpretációja valószínűségelméleti módszerek segítségével

Statisztika: áttekintés

Véletlen eljárásból származó mérési adatok $\rightarrow$ milyen eloszlásúak?

Gyakori szituáció: a mérési adatok eloszlása valamilyen ismeretlen paraméterrel/paraméterekkel rendelkezik. (Pl. tudjuk, hogy Poisson, de nem tudjuk, hogy mi λ értéke.)

A statisztika két fő ága:

- Becsléselmélet: ismételt mintavétel segítségével próbáljuk meg az ismeretlen paramétert megbecsülni. A becslés lehet
  - pontbecslés $\rightarrow$ a becslésünk értéke egy valós szám (vagy n paraméter esetén egy $\mathbb{R}^n$-beli vektor), a paraméter valamilyen értelemben „legjobb közelítése".
  - intervallumbecslés/konfidenciaintervallumok $\rightarrow$ a becslésünk értéke egy intervallum, amibe az ismeretlen paraméter legalább egy bizonyos (előre megadott) valószínűséggel beleesik.
- Hipotézisvizsgálat: szintén ismételt mintavétel segítségével valamilyen állítást szeretnénk tesztelni $\rightarrow$ „van-e elég bizonyítékunk arra, hogy az állítást elfogadjuk/elutasítsuk"? $\rightarrow$ Statisztikai próbák, ebből a tárgyból: u-próba.

(Példák ismeretlen paraméterre, mintavételre, hipotézisre: 13. fej. eleje)

Minta és realizáció

(13.1.1. Def.) Legyenek $X_1, \ldots, X_n$ független, azonos eloszlású változók, amelyek peremeloszlása nem feltétlenül ismert. Ekkor az $\mathbf{X} = (X_1, \ldots, X_n)$ val. vektorváltozót n **elemű független, azonos eloszlású mintának** (röviden n **elemű fae. mintának**) nevezzük.

Jelölések: Legyen $\mathbf{X} = (X_1, \ldots, X_n)$ 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 $\vartheta \in \Theta$-ra,

(1) abban az esetben, **ha ϑ a valódi paraméter**, egy A esemény valószínűségét $\mathbb{P}_\vartheta(A)$-val jelöljük,

(2) az X_1 (vagy X_2, X_3 stb.) valószínűségi változó ϑ paraméterhez tartozó eloszlásfüggvényét F_ϑ-val jelöljük, azaz $F_\vartheta(x) = \mathbb{P}_\vartheta(X_1 < x)$, $x \in \mathbb{R}$,

(3) ha X_1-nek ezen paraméter esetén létezik sűrűségfüggvénye, akkor ezt f_ϑ-val jelöljük. Vagyis:

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

Minta és realizáció

(13.1.1. Def.) Legyenek $X_1, \ldots, X_n$ független, azonos eloszlású változók, amelyek peremeloszlása nem feltétlenül ismert. Ekkor az $\mathbf{X} = (X_1, \ldots, X_n)$ val. vektorváltozót n **elemű független, azonos eloszlású mintának** (röviden n **elemű fae. mintának**) nevezzük.

Jelölések: Legyen $\mathbf{X} = (X_1, \ldots, X_n)$ 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 $\vartheta \in \Theta$-ra,

(1) abban az esetben, **ha ϑ a valódi paraméter**, egy A esemény valószínűségét $\mathbb{P}_\vartheta(A)$-val jelöljük,

(2) az X_1 (vagy X_2, X_3 stb.) valószínűségi változó ϑ paraméterhez tartozó eloszlásfüggvényét F_ϑ-val jelöljük, azaz $F_\vartheta(x) = \mathbb{P}_\vartheta(X_1 < x)$, $x \in \mathbb{R}$,

(3) ha X_1-nek ezen paraméter esetén létezik sűrűségfüggvénye, akkor ezt f_ϑ-val jelöljük,

(4) 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 $p_\vartheta(x) = \mathbb{P}_\vartheta(X = x)$.

Minta és realizáció

- (13.1.3. Példa: f_ϑ ill. p_ϑ az előbbi példákban)
- (13.1.4. Példa: az ismeretlen paraméterrel rendelkező eloszlású val. változó értékkészlete függhet a paramétertől)

Célunk a *realizáció* precíz definiálása $\rightarrow$ ehhez először az adott paraméterhez tartozó *lényeges értékek halmazát* definiáljuk.

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

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

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

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

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

Minta és realizáció

- (13.1.3. Példa: f_ϑ ill. p_ϑ az előbbi példákban)
- (13.1.4. Példa: az ismeretlen paraméterrel rendelkező eloszlású val. változó értékkészlete függhet a paramétertől)

Célunk a *realizáció* precíz definiálása $\rightarrow$ ehhez először az adott paraméterhez tartozó *lényeges értékek halmazát* definiáljuk.

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

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

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

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

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

Minta és realizáció

(13.1.6. Példa: lényeges értékek halmaza a korábbi példákban)

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

(13.1.8. Példa: realizációk az 13.1.6. Példa esetén)

Statisztikák

Egy n elemű fae. minta egy statisztikájának a mintaelemek egy olyan függvényét nevezzük, amely szimmetrikus, azaz ,,minden mintaelemtől ugyanúgy függ'':

(13.2.1. Def.) Legyen $\mathbf{X} = (X_1, \ldots, X_n)$ egy n elemű fae. minta. Ha $T\colon \mathbb{R}^n \to \mathbb{R}$ szimmetrikus függvény, azaz

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

minden $x_1, \ldots, x_n \in \mathbb{R}$ és $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ permutáció esetén teljesül, akkor a $T(\mathbf{X}) = T(X_1, \ldots, X_n)$ valószínűségi változót az $X_1, \ldots, X_n$ egy **statisztikájának** nevezzük.

Emlékeztető: $\pi\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ permutáció, ha kölcsönösen egyértelmű függvény (azaz bijekció), lásd 1. diasor.

Megjegyzés: Tavalyelőtt a nagy Θ helyett kis θ-t használtam, de ez zavarba ejtő jelölésnek bizonyult, ezért lecseréltem.

Régi vizsgákban, feladatsorokon a θ felel meg a Θ-nak.

Alapstatisztikák: mintaátlag

Az első alapstatisztika mintaátlag, amit már a középiskolából is ismerünk:

Legyen $\mathbf{X} = (X_1, \ldots, X_n)$ egy n elemű fae. minta. Ekkor az $\overline{X_n} = \frac{X_1 + \ldots + X_n}{n}$ **mintaátlag** az $\mathbf{X}$ egy statisztikája. Ha $\mathbf{x} = (x_1, \ldots, x_n)$ az $\mathbf{X} = (X_1, \ldots, X_n)$ egy realizációja, akkor realizáció átlagát $\overline{x_n}$-nel fogjuk jelölni: $\overline{x_n} = \frac{x_1 + \ldots + x_n}{n}$.

Az eddigiek 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:

$$\mathbb{E}\left[\overline{X_n}\right] = \frac{1}{n}(\mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n)) = \mathbb{E}(X_1).$$

További tulajdonságai $\rightarrow$ lásd később.

Alapstatisztikák: korrigált empirikus szórásnégyzet

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

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

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

- „empirikus" = „tapasztalati".
- Világos: $(S_n^*)^2$ és S_n^* statisztikái **X**-nek.
- Miért $n-1$-gyel osztunk (és nem pl. n-nel)? $\rightarrow$ Köv. állítás.

(13.2.4. Állítás.) Legyen $\mathbf{X} = (X_1, \ldots, X_n)$ 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)

(13.2.5. Megjegyzés: miért nevezzük korrigáltnak)

Rendezett minták, empirikus medián, módusz

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

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

Ekkor $(X_1^*, X_2^*, \ldots, X_n^*)$-ot **rendezett mintá**nak nevezzük.

(13.2.7. Példa: nyolcszori kockadobás; ismétlődő elemek kezelése)
Vigyázat: $X_1^*, \ldots, X_n^*$ általában nem függetlenek és nem is azonos eloszlásúak!

Az empirikus medián a rendezett minta középső eleme, ha páratlan elemszámú a minta, ha pedig páros elemszámú, akkor a két középső átlaga:

(13.2.8. Def.) Az $\mathbf{X} = (X_1, \ldots, X_n)$ n elemű fae. minta **empirikus mediánj**án az $m_{n,\mathbf{X}} = X_{k+1}^*$ rendezett mintaelemet értjük, ha $n = 2k+1$, és $m_{n,\mathbf{X}} = \frac{X_k^* + X_{k+1}^*}{2}$-et, ha $n = 2k$. (13.2.9. Példa)
Mi a (nem empirikus) *medián*? $\rightarrow$ Lásd hamarosan.
Módusz: a leggyakoribb érték a mintában. Ha több ilyen érték van, akkor mindegyiket módusznak tekintjük.

Rendezett minták, empirikus medián, módusz

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

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

Ekkor $(X_1^*, X_2^*, \ldots, X_n^*)$-ot **rendezett mintá**nak nevezzük.

(13.2.7. Példa: nyolcszori kockadobás; ismétlődő elemek kezelése)
Vigyázat: $X_1^*, \ldots, X_n^*$ általában nem függetlenek és nem is azonos eloszlásúak!

Az empirikus medián a rendezett minta középső eleme, ha páratlan elemszámú a minta, ha pedig páros elemszámú, akkor a két középső átlaga:

(13.2.8. Def.) Az $\mathbf{X} = (X_1, \ldots, X_n)$ n elemű fae. minta **empirikus mediánj**án az $m_{n,\mathbf{X}} = X_{k+1}^*$ rendezett mintaelemet értjük, ha $n = 2k+1$, és $m_{n,\mathbf{X}} = \frac{X_k^* + X_{k+1}^*}{2}$-et, ha $n = 2k$. (13.2.9. Példa)
Mi a (nem empirikus) *medián*? $\rightarrow$ Lásd hamarosan.
Módusz: a leggyakoribb érték a mintában. Ha több ilyen érték van, akkor mindegyiket módusznak tekintjük.

Empirikus eloszlásfüggvény

Emlékeztető: egy (egydimenziós) val. változó eloszlását az eloszlásfüggvénye egyértelműen meghatározza.
Tegyük fel, hogy nem tudjuk, hogy a fae. mérési adataink milyen típusú eloszláshoz tartoznak (azaz nemcsak egy paraméter, hanem a teljes eloszlás ismeretlen), viszont képesek vagyunk nagy elemszámú mintát venni.

Kérdés: Hogyan tudjuk a mintaelemek eloszlásfüggvényét a minta alapján közelíteni/becsülni? Erre szolgál az empirikus eloszlásfüggvény.

Empirikus eloszlásfüggvény

(13.3.1. Def.) Legyen $n \in \mathbb{N}$ és legyenek $X_1, \dots, X_n$ fae. valószínűségi változók. Az $\mathbb{R} \to \mathbb{R}$,

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

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

Alternatív (és talán emészthetőbb) megfogalmazás az $X_1^*, \dots, X_n^*$ rendezett minta segítségével:

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

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

(Össznépi gyakorlat: 12.15. feladat – az empirikus eloszlásfüggvény maga is egy (diszkrét) eloszlásfüggvény)

Az empirikus eloszlásfüggvény tulajdonságai

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

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

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

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

Azaz az empirikus eloszlásfüggvény „jó becslése a valódi eloszlásfüggvénynek”: várható értéke a valódi eloszlásfüggvény, szórása 0-hoz tart, sőt az emp. eloszlásfüggvény 1 valószínűséggel pontonként a valódi eloszlásfüggvényhez tart.
Erről többet – más szavakkal – később.

(Állítás bizonyítása: (1) és (2) azért igaz, mert $nF_n^*(x) \sim \mathrm{Bin}(F(x), 1/n)$; (3) a nagy számok erős törvénye alapján, lásd előadás/jegyzetkiegészítés)

Glivenko–Cantelli-tétel (kiegészítő anyag)

Láttuk: $\mathbb{P}(\lim_{n\to\infty} F_n^*(x) = F(x)) = 1$, $\forall x \in \mathbb{R}$.

Ez még nem zárja ki azt, hogy $x \to \infty$ vagy $x \to -\infty$ esetén a konvergencia egyre lassabb legyen. (Ha ez így lenne, akkor az empirikus eloszlásfüggvényt nehezen lehetne a ,,valódi eloszlásfüggvény kirajzolására" használni.)
Azonban ez nincs így, mert az alábbi tétel szerint F_n^* 1 valószínűséggel az egész számegyenesen egyenletesen konvergál F-hez.

Glivenko–Cantelli-tétel: Legyen $n \in \mathbb{N}$ és legyenek $X_1, \ldots, X_n$ 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

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

(Megjegyzés: sűrűségfüggvény empirikus közelítése, empirikus sűrűséghisztogram)

Glivenko–Cantelli-tétel (kiegészítő anyag)

Láttuk: $\mathbb{P}(\lim_{n\to\infty} F_n^*(x) = F(x)) = 1$, $\forall x \in \mathbb{R}$.

Ez még nem zárja ki azt, hogy $x \to \infty$ vagy $x \to -\infty$ esetén a konvergencia egyre lassabb legyen. (Ha ez így lenne, akkor az empirikus eloszlásfüggvényt nehezen lehetne a ,,valódi eloszlásfüggvény kirajzolására" használni.)
Azonban ez nincs így, mert az alábbi tétel szerint F_n^* 1 valószínűséggel az egész számegyenesen egyenletesen konvergál F-hez.

Glivenko–Cantelli-tétel: Legyen $n \in \mathbb{N}$ és legyenek $X_1, \ldots, X_n$ 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

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

(Megjegyzés: sűrűségfüggvény empirikus közelítése, empirikus sűrűséghisztogram)

Becsléselmélet

Szituáció: $X_1, \ldots, X_n$ fae. mintaelemek, eloszlásuk egy $\vartheta \in \Theta$ paramétertől függ, és az ismeretlen ϑ paraméterre vagy annak valamilyen $\psi(\vartheta)$ függvényére a mintaelemek valamilyen statisztikája alapján minél jobb becslést akarunk adni.

Két fő változat:

- pontbecslések $\rightarrow$ a becslés egy konkrét valós szám/vektor lesz („a legjobb tippünk a paraméterre")
- intervallumbecslések, konfidenciaintervallumok $\rightarrow$ a becslés egy intervallum lesz, amibe az ismeretlen paraméter legalább egy adott valószínűséggel (pl. 0,95, 0,99) beleesik.

Idő hiányában a pontbecslésekről szóló rész csak ízelítő lesz, egy-két módszert ismerünk csak meg részletesebben $\rightarrow$ bővebb ismeretekért lásd a SZIT MSc-s Matematikai statisztika tárgyát.

Példa pontbecslésre

Korábbi példa: A Magyarországon vasúti útátjárókban havonta bekövetkező balesetek számai (nagyjából) fae. Poisson(ϑ)-eloszlású val. változókkal közelíthetők, ahol $\vartheta > 0$ ismeretlen.
Veszünk egy n alapú $X_1, \ldots, X_n$ mintát (n hónapig megfigyeljük a balesetek számát).
A minta milyen statisztikájával becsüljük ϑ-t?

Ötlet: Mivel a Poisson(ϑ)-eloszlásnak ϑ éppen a várható értéke, az $T(X_1, \ldots, X_n) = \overline{X_n} = \frac{1}{n}(X_1 + \ldots + X_n)$ mintaátlag jó becslés lesz.
Milyen tulajdonságokkal rendelkezik ez a becslés?

Jelölje $\mathbb{E}_\vartheta$ a $\mathbb{P}_\vartheta$ valószínűség szerinti várható értéket. Láttuk, hogy minden $\vartheta \in \Theta := (0, \infty)$ esetén (ϑ nélkül, de ϑ-val ugyanúgy igaz):

$$\mathbb{E}_\vartheta(\overline{X_n}) = \mathbb{E}_\vartheta(X_1), \tag{2}$$

ami konkrétan a Poisson(ϑ)-eloszlás esetén ϑ-val egyenlő.
Az (2) egyenlet mindig igaz, ha $\mathbb{E}_\vartheta(|X_1|) < \infty$ minden $\vartheta \in \Theta$-ra.

Példa pontbecslésre

Korábbi példa: A Magyarországon vasúti útátjárókban havonta bekövetkező balesetek számai (nagyjából) fae. Poisson(ϑ)-eloszlású val. változókkal közelíthetők, ahol $\vartheta > 0$ ismeretlen.
Veszünk egy n alapú $X_1, \ldots, X_n$ mintát (n hónapig megfigyeljük a balesetek számát).
A minta milyen statisztikájával becsüljük ϑ-t?

Ötlet: Mivel a Poisson(ϑ)-eloszlásnak ϑ éppen a várható értéke, az $T(X_1, \ldots, X_n) = \overline{X_n} = \frac{1}{n}(X_1 + \ldots + X_n)$ mintaátlag jó becslés lesz.
Milyen tulajdonságokkal rendelkezik ez a becslés?

Jelölje $\mathbb{E}_\vartheta$ a $\mathbb{P}_\vartheta$ valószínűség szerinti várható értéket. Láttuk, hogy minden $\vartheta \in \Theta := (0, \infty)$ esetén (ϑ nélkül, de ϑ-val ugyanúgy igaz):

$$\mathbb{E}_\vartheta(\overline{X_n}) = \mathbb{E}_\vartheta(X_1), \tag{2}$$

ami konkrétan a Poisson(ϑ)-eloszlás esetén ϑ-val egyenlő.
Az (2) egyenlet mindig igaz, ha $\mathbb{E}_\vartheta(|X_1|) < \infty$ minden $\vartheta \in \Theta$-ra.

Becslés torzítatlansága

Láttuk:

$$\mathbb{E}_\vartheta(\underbrace{\overline{X_n}}_{=T(X_1,\ldots,X_n)\text{ (statisztika)}}) = \underbrace{\mathbb{E}_\vartheta(X_1)}_{=\psi(\vartheta)\text{ (paraméterfüggvény)}}, \qquad \forall \vartheta \in \underbrace{(0,\infty)}_{=\Theta}.$$

(Lásd 14.1.3. Példa.)

Általában ha egy $(X_1, \ldots, X_n)$ fae. minta esetén egy $T(X_1, \ldots, X_n)$ statisztikára és egy $\vartheta \mapsto \psi(\vartheta)$ paraméterfüggvényre teljesül, hogy

$$\mathbb{E}_\vartheta(T(X_1, \ldots, X_n)) = \psi(\vartheta), \qquad \forall \vartheta \in \Theta,$$

akkor $T(X_1, \ldots, X_n)$-t torzítatlan becslésnek nevezzük a $\psi(\vartheta)$ paraméterfüggvényre.

Jelentés: Akármi is a valódi ϑ paraméter, a torzítatlan becslés várható értékben $\psi(\vartheta)$-val egyenlő.

Másik példa torzítatlanságra

(14.1.4. Példa.) Egy másik lehetséges ψ függvény a

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

(ahol $\mathbb{D}^2_\vartheta(X_1)$ jelöli X_1 szórásnégyzetét, amennyiben ϑ a valódi paraméter), ha $\mathbb{E}_\vartheta(X_1^2)$ minden $\vartheta \in \Theta$ esetén véges. Ebben az esetben a

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

korrigált empirikus szórásnégyzet torzítatlan becslése a

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

paraméterfüggvénynek.

Ugyanis korábban láttuk:

$$\mathbb{E}_\vartheta(\underbrace{(S_n^*)^2}_{=T(X_1,\ldots,X_n)\ \text{(statisztika)}}) = \underbrace{\mathbb{D}^2_\vartheta(X_1)}_{=\psi(\vartheta)\ \text{(paraméterfüggvény)}}, \qquad \forall \vartheta \in \Theta.$$

Aszimptotikus torzítatlanság

Láttuk: $$\mathbb{E}_\vartheta\Big(\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X_n})^2\Big)=\mathbb{D}^2_\vartheta(X_1),\qquad \forall\vartheta\in\Theta.$$

Tekintsünk egy másik becslést $\mathbb{D}^2_\vartheta(X_1)$-re:

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

ez a minta (nem korrigált) empirikus szórásnégyzete. A várható érték linearitása miatt minden $\vartheta\in\Theta$-ra

$$\mathbb{E}_\vartheta(S_n^2)=\mathbb{E}_\vartheta\Big(\frac{n-1}{n}\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X_n})^2\Big)=\frac{n-1}{n}\mathbb{E}_\vartheta\big((S_n^*)^2\big)=\frac{n-1}{n}\mathbb{D}^2_\vartheta(X_1).$$

Ez adott n-re nem torzítatlan becslése $\psi(\vartheta)=\mathbb{D}^2_\vartheta(X_1)$-nek.

Azonban aszimptotikusan torzítatlan, ami definíció szerint azt jelenti, hogy

$$\lim_{n\to\infty}\mathbb{E}_\vartheta(T(X_1,\dots,X_n))=\psi(\vartheta),\qquad \forall\vartheta\in\Theta.$$

Vissza a Poisson-eloszlásos példára

A Poisson-eloszlásos (vasúti balesetes) példában minden $\vartheta \in \Theta$ esetén $\mathbb{E}_\vartheta(X_1) = \mathbb{D}^2_\vartheta(X_1) = \vartheta$.

Tehát a $\psi(\vartheta) = \vartheta$ paraméterfüggvényre nemcsak $\overline{X_n}$, hanem $(S_n^*)^2$ is torzítatlan becslés, valamint S_n^2 is aszimptotikusan torzítatlan.

Más eloszlástípusok esetén ez nem feltétlenül teljesül!

Becslések egyéb tulajdonságai (kitekintés)

Ha egy becslés torzítatlan, átlagosan a becsülendő $\psi(\vartheta)$ paraméterfüggvény értékét veszi fel.

Cél: A torzítatlan becslések közül olyat választani, ami tipikusan is közel van $\psi(\vartheta)$-hoz.

- Adott n mintaelemszám esetén több becslés közül ugyanarra a $\psi(\vartheta)$-ra válasszuk a lehető legkisebb szórásút.
  Két (torzítatlan) becslés közül az egyiket legalább annyira hatásosnak nevezzük, mint a másikat, ha minden $\vartheta \in \Theta$-ra legfeljebb akkora a szórása, mint a másiknak.
  Egy torzítatlan becslést hatásosnak nevezünk, ha legalább annyira hatásos, mint bármely más torzítatlan becslés (ugyanarra a $\psi(\vartheta)$-ra).
  Kérdés: Létezik-e hatásos becslés? Hogyan találjuk meg?
  $\rightarrow$ Ezekkel itt nem foglalkozunk, lásd az MSc-s és PhD-s matematikai statisztika tárgyakat.
- Válasszunk olyan becslést, amely nagy n esetén tetszőlegesen megközelíti a becsült paramétert. (Ha ez nem teljesül, nem is érdemes nagy elemszámú mintát venni.)

Erős konzisztencia

Az $(X_1, \ldots, X_n)$ minta egy $T(X_1, \ldots, X_n)$ statisztikáját a $\psi(\vartheta)$ paraméterfüggvény egy erősen konzisztens becslésének nevezzük, ha bármely $\vartheta \in \Theta$ esetén

$$\mathbb{P}_\vartheta\big(\lim_{n\to\infty} T(X_1, \ldots, X_n) = \psi(\vartheta)\big) = 1.$$

Szavakkal: bármelyik ϑ is az igazi paraméter, végtelenhez tartó mintaméret esetén a becslés a becsült paraméterfüggvényhez tart.

(14.1.3. Példa folyt.) A mintaátlag erősen konzisztens becslése is $\psi(\vartheta) = \mathbb{E}_\vartheta(X_1)$-nek (ha $\mathbb{E}_\vartheta(|X_1|) < \infty$ minden $\vartheta \in \Theta$-ra teljesül, mint pl. a vasúti balesetes példában).
(Bizonyítás: lásd előadás/jegyzet $\to$ nagy számok erős törvénye)

(Torzítatlanság, aszimpt. torzítatlanság, erős konzisztencia definíciója rendszerezve: 14.1.1. Definíció $\to$ idén nem kérdezzük írásbelin)

Maximum likelihood elv (kitekintés)

Adott egy ismeretlen $\vartheta \in \Theta$ paraméterrel rendelkező eloszlás $\rightarrow$ $(X_1, \ldots, X_n)$ minta, amelynek eloszlása ettől a paramétertől függ.
Kérdés: A minta egy $(x_1, \ldots, x_n)$ realizációja alapján mit tippelnénk ϑ értékére?
Pl.: geometriai eloszlás ismeretlen ϑ paraméterrel. n-szer ismételjük azt a kísérletet, hogy az utcán talált pénzérmét az első fejig dobáljuk $\rightarrow$ dobások számai az egyes kísérletekben: $x_1, \ldots, x_n$.
$\Theta = (0, 1)$.

Egy lehetséges becslést ϑ-ra az úgynevezett maximum likelihood módszer ad. Ezt idén nem formalizáljuk és nem részletezzük (részletek $\rightarrow$ jegyzetkiegészítés), de ezt a példát megnézzük.
Részletek az MSc-s Matematikai statisztika tárgyban.
Egy másik lehetséges elv a momentumok módszere $\rightarrow$ erre egyelőre példát sem nézünk, lásd jegyzetkiegészítés.

Maximum likelihood becslés elve: Becslésnek válasszuk azt a $\vartheta_* \in \Theta$ lehetséges paramétert, amelyiknél a kapott $(x_1, \ldots, x_n)$ realizáció a lehető legvalószínűbb.

Maximum likelihood elv (kitekintés)

Adott egy ismeretlen $\vartheta \in \Theta$ paraméterrel rendelkező eloszlás $\rightarrow$ $(X_1, \ldots, X_n)$ minta, amelynek eloszlása ettől a paramétertől függ.
Kérdés: A minta egy $(x_1, \ldots, x_n)$ realizációja alapján mit tippelnénk ϑ értékére?
Pl.: geometriai eloszlás ismeretlen ϑ paraméterrel. n-szer ismételjük azt a kísérletet, hogy az utcán talált pénzérmét az első fejig dobáljuk $\rightarrow$ dobások számai az egyes kísérletekben: $x_1, \ldots, x_n$.
$\Theta = (0, 1)$.

Egy lehetséges becslést ϑ-ra az úgynevezett maximum likelihood módszer ad. Ezt idén nem formalizáljuk és nem részletezzük (részletek $\rightarrow$ jegyzetkiegészítés), de ezt a példát megnézzük.
Részletek az MSc-s Matematikai statisztika tárgyban.
Egy másik lehetséges elv a momentumok módszere $\rightarrow$ erre egyelőre példát sem nézünk, lásd jegyzetkiegészítés.

Maximum likelihood becslés elve: Becslésnek válasszuk azt a $\vartheta_* \in \Theta$ lehetséges paramétert, amelyiknél a kapott $(x_1, \ldots, x_n)$ realizáció a lehető legvalószínűbb.

Maximum likelihood elv (kitekintés)

Tegyük fel, hogy $\vartheta \in \Theta$ a valódi paraméter.
Ekkor az X_i-k geometriai eloszlásúak ϑ paraméterrel és függetlenek.
Így annak a valószínűsége, hogy X_i értéke éppen a kapott x_i:

$$\mathbb{P}_\vartheta(X_i = x_i) = \vartheta(1-\vartheta)^{x_i - 1}$$

(ez a geometriai eloszlásnál $\mathbb{P}(X_i = k)$ szokásos képlete $p = \vartheta$-val és $k = x_i$-vel).

A függetlenség miatt

$$\begin{aligned}\mathbb{P}_\vartheta(X_1 = x_1, \ldots, X_n = x_n) &= \mathbb{P}_\vartheta(X_1 = x_1) \cdot \ldots \cdot \mathbb{P}_\vartheta(X_n = x_n) \\ &= \prod_{i=1}^{n} (\vartheta(1-\vartheta)^{x_i - 1}) \\ &= \vartheta^n \prod_{i=1}^{n} (1-\vartheta)^{x_i - 1}.\end{aligned}$$

Itt $x_1, \ldots, x_n$ adottak és a kapott kifejezést ϑ-ban szeretnénk maximalizálni.

Maximum likelihood elv (kitekintés)

Tegyük fel, hogy $\vartheta \in \Theta$ a valódi paraméter.
Ekkor az X_i-k geometriai eloszlásúak ϑ paraméterrel és függetlenek.
Így annak a valószínűsége, hogy X_i értéke éppen a kapott x_i:

$$\mathbb{P}_\vartheta(X_i = x_i) = \vartheta(1-\vartheta)^{x_i-1}$$

(ez a geometriai eloszlásnál $\mathbb{P}(X_i = k)$ szokásos képlete $p = \vartheta$-val és $k = x_i$-vel).
A függetlenség miatt

$$\begin{aligned}\mathbb{P}_\vartheta(X_1 = x_1, \ldots, X_n = x_n) &= \mathbb{P}_\vartheta(X_1 = x_1) \cdot \ldots \cdot \mathbb{P}_\vartheta(X_n = x_n) \\ &= \prod_{i=1}^{n} (\vartheta(1-\vartheta)^{x_i-1}) \\ &= \vartheta^n \prod_{i=1}^{n} (1-\vartheta)^{x_i-1}.\end{aligned}$$

Itt $x_1, \ldots, x_n$ adottak és a kapott kifejezést ϑ-ban szeretnénk maximalizálni.

Maximum likelihood elv (kitekintés)

$$\mathbb{P}_\vartheta(X_1 = x_1, \ldots, X_n = x_n) = \vartheta^n \prod_{i=1}^{n} (1-\vartheta)^{x_i - 1}.$$

Ezt szeretnénk ϑ-ban maximalizálni.

Deriváljuk ϑ szerint? $\to$ Egyszerűbb, ha a logaritmusát vesszük és azt deriváljuk (a maximumhely ugyanott lesz).

$$\ln \mathbb{P}_\vartheta(X_1 = x_1, \ldots, X_n = x_n) = n \ln \vartheta + \sum_{i=1}^{n} (x_i - 1) \ln(1-\vartheta),$$

$$\frac{\partial}{\partial \vartheta} \ln \mathbb{P}_\vartheta(X_1 = x_1, \ldots, X_n = x_n) = \frac{n}{\vartheta} - \sum_{i=1}^{n} \frac{x_i - 1}{1-\vartheta}.$$

Elemi átalakítások után: a derivált 0 $\Leftrightarrow \vartheta = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\overline{x}_n}$.

$\vartheta_* = \frac{1}{\overline{x}_n}$ lesz a minta alapján a maximum likelihood becslésünk (ellenőrizhető, hogy ez valóban maximumhely).

Folytonos esetben hasonló a módszer, súlyfüggvény helyett sűrűségfüggvénnyel (példák: jegyzetkiegészítés/MSc-s tárgy).

Maximum likelihood elv (kitekintés)

$$\mathbb{P}_\vartheta(X_1 = x_1, \ldots, X_n = x_n) = \vartheta^n \prod_{i=1}^{n} (1-\vartheta)^{x_i - 1}.$$

Ezt szeretnénk ϑ-ban maximalizálni.

Deriváljuk ϑ szerint? $\to$ Egyszerűbb, ha a logaritmusát vesszük és azt deriváljuk (a maximumhely ugyanott lesz).

$$\ln \mathbb{P}_\vartheta(X_1 = x_1, \ldots, X_n = x_n) = n \ln \vartheta + \sum_{i=1}^{n} (x_i - 1) \ln(1-\vartheta),$$

$$\frac{\partial}{\partial \vartheta} \ln \mathbb{P}_\vartheta(X_1 = x_1, \ldots, X_n = x_n) = \frac{n}{\vartheta} - \sum_{i=1}^{n} \frac{x_i - 1}{1-\vartheta}.$$

Elemi átalakítások után: a derivált 0 $\Leftrightarrow \vartheta = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\overline{x}_n}$.

$\vartheta_* = \frac{1}{\overline{x}_n}$ lesz a minta alapján a maximum likelihood becslésünk (ellenőrizhető, hogy ez valóban maximumhely).

Folytonos esetben hasonló a módszer, súlyfüggvény helyett sűrűségfüggvénnyel (példák: jegyzetkiegészítés/MSc-s tárgy).

Konfidenciaintervallumok

Eddig: *pontbecslések* (átlag/szórásnégyzet becslései, maximum likelihood stb.) $\rightarrow$ a „legjobb tippünk” az ismeretlen paraméter(ek)re egy valós szám (vagy vektor).

Most: *intervallumbecslés/konfidenciaintervallumok* $\rightarrow$ nem egyetlen szám lesz a becslésünk, hanem egy intervallum,

- amelybe a keresett paraméter egy bizonyos előre megadott valószínűséggel (pl. legalább 95%, pontosan 99% stb.) beleesik,
- és amely az előbbi tulajdonságot teljesítő összes intervallum közül valamilyen értelemben optimális (pl. a legrövidebb).

Informatikai példa: szerverválaszidő.

Konfidenciaintervallumok

(15.1.1. Definíció.) Legyen $\mathbf{X} = (X_1, \ldots, X_n)$ egy n elemű fae. minta, ahol a mintaelemek eloszlása egy $\vartheta \in \Theta \subseteq \mathbb{R}^d$ paramétertől függ, legyen $\psi\colon \mathbb{R}^d \to \mathbb{R}$ 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

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

minden $\vartheta \in \Theta$ esetén.
Megjegyzések:

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. Pontosan $1 - \varepsilon$ szintű konfidenciaintervallum általában csak akkor adható meg, ha $X_1, \ldots, X_n$ peremeloszlása folytonos. Diszkrét esetben általában *legalább* $1 - \varepsilon$ *szintű konfidenciaintervallum*okról szokás beszélni, ahol a definícióban „$= 1 - \varepsilon$" helyett „$\geq 1 - \varepsilon$" szerepel.

Kvantilisek: bevezetés

Azt mondjuk, hogy a $[T_1(\mathbf{X}), T_2(\mathbf{X})]$ intervallum (pontosan) $1-\varepsilon$ **szintű konfidenciaintervallum** a $\psi(\vartheta)$ paraméterfüggvényre, ha

$$\mathbb{P}_\vartheta\big(T_1(\mathbf{X}) \le \psi(\vartheta) \le T_2(\mathbf{X})\big) = 1-\varepsilon, \qquad \forall \vartheta \in \Theta.$$

$T_1(\mathbf{X})$ és $T_2(\mathbf{X})$ meghatározásához szükséges fogalom: folytonos eloszlások *kvantilisei*.

Legyen X folytonos val. változó $\to F_X$ eloszlásfv., f_X sűrfv.
Tudjuk: F_X folytonos függvény.
Az egyszerűség kedvéért tegyük fel, hogy f_X pontosan egy I nyílt intervallumon nem 0. Ilyen az összes nevezetes folytonos eloszlás:

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

Ilyen esetben $\mathrm{Ran}(X) = I$. F_X szigorúan monoton nő I-n.
„I alsó végpontjáig" (ha az véges) konstans 0, „I felső végpontjától" (ha az véges) konstans 1.
Létezik az $F_X^{-1}\colon (0,1) \to I$ inverzfüggvény, szintén folytonos és szig. mon. növő.

Kvantilisek és medián

(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-**kvantilisé**nek, az $\frac{1}{2}$-kvantilist ($y = 1/2$) **medián**nak nevezzük.

- Tehát ha x az X y-kvantilise, akkor $\mathbb{P}(X < x) = y$.
- Miért nem jó ez a definíció diszkrét X esetén? Pl. kockadobás 0,2-kvantilise?!
- Mi a probléma folytonos esetben, ha a sűrűségfüggvény pl. két távoli diszjunkt intervallum unióján nem nulla?
- A medián az 5. feladatsor 7. e*) feladatában már szerepelt.
- Normális, exponenciális, egyenletes eloszlás mediánja, összehasonlítása a várható értékkel $\rightarrow$ gyakorlaton.
- Emlékeztető: az empirikus medián páratlan elemszám esetén a minta középső eleme, párosnál a két középső elem átlaga. Láttuk: az F_n^* empirikus eloszlásfüggvény maga is egy eloszlásfüggvény, a hozzá tartozó eloszlás diszkrét.

Kvantilisek és medián

(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-**kvantilisé**nek, az $\frac{1}{2}$-kvantilist ($y = 1/2$) **medián**nak nevezzük.

- Tehát ha x az X y-kvantilise, akkor $\mathbb{P}(X < x) = y$.
- A medián az 5. feladatsor 7. e*) feladatában már szerepelt.
- Normális, exponenciális, egyenletes eloszlás mediánja, összehasonlítása a várható értékkel $\rightarrow$ gyakorlaton.
- Emlékeztető: az empirikus medián páratlan elemszám esetén a minta középső eleme, párosnál a két középső elem átlaga.
  Láttuk: az F_n^* empirikus eloszlásfüggvény maga is egy eloszlásfüggvény, a hozzá tartozó eloszlás diszkrét.
  Az empirikus eloszlásfüggvény az empirikus medián környékén $1/2$ körül van (a részletekbe itt nem megyünk bele). Ilyen értelemben analógja az empirikus medián a mediánnak.

Konfidenciaintervallum normális eloszlás várható értékére ismert szórásnégyzet esetén

Feladat. Legyen $X \sim \mathrm{N}(\mu; \sigma^2)$, ahol a $\mu \in \mathbb{R}$ várható érték ismeretlen, de a $\sigma > 0$ szórás ismert. Adott $\varepsilon \in (0, 1)$ esetén szerkesszünk $1 - \varepsilon$ szintű konfidenciaintervallumot μ-re úgy, hogy annak hossza a lehető legkisebb legyen!

Mintaelemek sűrűségfüggvénye, ha μ a valódi paraméter:

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

Ez az x pont μ-től való távolsága függvényében szig. mon. csökken. Ezért adott ε esetén a legrövidebb $1 - \varepsilon$ szintű konfidenciaintervallumot akkor fogjuk kapni, ha azt az $\overline{X_n}$ mintaátlag körül szimmetrikusan választjuk, vagyis

$$[\overline{X_n} - r_\varepsilon, \overline{X_n} + r_\varepsilon]$$

alakban, úgy, hogy $\mathbb{P}_\mu(\mu \in [\overline{X_n} - r_\varepsilon, \overline{X_n} + r_\varepsilon]) = 1 - \varepsilon$ (3)

minden $\mu \in \mathbb{R}$ esetén teljesüljön.

Cél: r_ε (ε-tól függő) értékének meghatározása.

Konfidenciaintervallum μ-re ismert σ^2 esetén

$\mathbb{P}_\mu(\mu \in [\overline{X_n} - r_\varepsilon, \overline{X_n} + r_\varepsilon]) = \mathbb{P}_\mu(\overline{X_n} - r_\varepsilon \le \mu \le \overline{X_n} + r_\varepsilon) = 1 - \varepsilon$, $r_\varepsilon = ?$

A mintaátlag standardizálása. A mintaátlag eloszlása:

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

(Független normálisok összege normális és a paramétereik összeadódnak, normális lin. transzformáltja normális. + $\mathbb{E}, \mathbb{D}^2$ azonosságai.) Ezért

$$\frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\overline{X_n} - \mu)}{\sigma} \sim \mathrm{N}(0;1). \tag{4}$$

Itt álljunk meg egy pillanatra. A CHT pont azt mondja ki, hogy

$$\frac{\overline{X_n} - \mathbb{E}(\overline{X_n})}{\mathbb{D}(\overline{X_n})} = \frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\overline{X_n} - \mu)}{\sigma} \overset{n\to\infty}{\longrightarrow} \mathrm{N}(0;1) \text{ eloszlásban.}$$

Normális eloszlású val. változók esetén tehát a CHT nemcsak eloszlásbeli konvergenciával, hanem egyenlőséggel teljesül!

Konfidenciaintervallum μ-re ismert σ^2 esetén

$\mathbb{P}_\mu(\mu \in [\overline{X_n} - r_\varepsilon, \overline{X_n} + r_\varepsilon]) = \mathbb{P}_\mu(\overline{X_n} - r_\varepsilon \le \mu \le \overline{X_n} + r_\varepsilon) = 1 - \varepsilon$, $r_\varepsilon = ?$

A mintaátlag standardizálása. A mintaátlag eloszlása:

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

(Független normálisok összege normális és a paramétereik összeadódnak, normális lin. transzformáltja normális. + $\mathbb{E}, \mathbb{D}^2$ azonosságai.) Ezért

$$\frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\overline{X_n} - \mu)}{\sigma} \sim \mathrm{N}(0;1). \tag{4}$$

Fejezzük ki most a fenti egyenletet az $\frac{\overline{X_n}-\mu}{\sigma}\sqrt{n}$ valószínűségi változó segítségével:

$$\mathbb{P}_\mu(\overline{X_n} - r_\varepsilon \le \mu \le \overline{X_n} + r_\varepsilon) = 1 - \varepsilon \;\Leftrightarrow\; \mathbb{P}_\mu(-r_\varepsilon \le \overline{X_n} - \mu \le r_\varepsilon) = 1 - \varepsilon$$

$$\Leftrightarrow\; \mathbb{P}_\mu\Big(-\frac{r_\varepsilon}{\sigma}\sqrt{n} \le \frac{\overline{X_n} - \mu}{\sigma}\sqrt{n} \le \frac{r_\varepsilon}{\sigma}\sqrt{n}\Big) = 1 - \varepsilon.$$

Konfidenciaintervallum μ-re ismert σ^2 esetén

Cél: $\mathbb{P}_\mu\Big(-\frac{r_\varepsilon}{\sigma}\sqrt{n} \le \frac{\overline{X_n}-\mu}{\sigma}\sqrt{n} \le \frac{r_\varepsilon}{\sigma}\sqrt{n}\Big) = 1-\varepsilon$, $r_\varepsilon = ?$.

Mivel $\frac{\overline{X_n}-\mu}{\sigma}\sqrt{n} \sim \mathrm{N}(0;1)$, ezért [emlékeztető: $\Phi(-x) = 1-\Phi(x)$]

$$\mathbb{P}_\mu\Big(-\frac{r_\varepsilon}{\sigma}\sqrt{n} \le \frac{\overline{X_n}-\mu}{\sigma}\sqrt{n} \le \frac{r_\varepsilon}{\sigma}\sqrt{n}\Big) = \Phi(\frac{r_\varepsilon}{\sigma}\sqrt{n}) - \Phi(-\frac{r_\varepsilon}{\sigma}\sqrt{n}) = 2\Phi(\frac{r_\varepsilon}{\sigma}\sqrt{n}) - 1.$$

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

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

(ahol $\Phi^{-1}(1-\frac{\varepsilon}{2})$-t a szokásos módon kiolvashatjuk a normális eloszlás táblázatából), tehát

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

Visszahelyettesítve a konfidenciaintervallum:
$[\overline{X_n} - \Phi^{-1}(1-\varepsilon/2)\frac{\sigma}{\sqrt{n}}, \overline{X_n} + \Phi^{-1}(1-\varepsilon/2)\frac{\sigma}{\sqrt{n}}]$.

Konfidenciaintervallum μ-re ismert σ^2 esetén

A konfidenciaintervallum:

$$[\overline{X_n} - \Phi^{-1}(1-\varepsilon/2)\frac{\sigma}{\sqrt{n}}, \overline{X_n} + \Phi^{-1}(1-\varepsilon/2)\frac{\sigma}{\sqrt{n}}].$$

Ezzel készen vagyunk, ezt még kicsit cicomázzuk – bevezetünk egy olyan jelölést, ami később a hipotézisvizsgálatnál is hasznos lesz:

$$u_\delta := \Phi^{-1}(1-\delta), \qquad \delta \in (0,1).$$

Vegyük észre, hogy u_δ az $\mathrm{N}(0;1)$ eloszlás $1-\delta$-kvantilise, hiszen az eloszlásfüggvény Φ.
A konfidenciaintervallum tehát $[\overline{X_n} - u_{\varepsilon/2}\frac{\sigma}{\sqrt{n}}, \overline{X_n} + u_{\varepsilon/2}\frac{\sigma}{\sqrt{n}}]$.

Konfidenciaintervallum μ-re ismert σ^2 esetén

Összefoglalva:

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

$$\big[T_1(\mathbf{X}), T_2(\mathbf{X})\big] = \big[\overline{X_n} - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \overline{X_n} + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}\big],$$

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

Konfidenciaintervallum μ-re ismert σ^2 esetén

$$\big[T_1(\mathbf{X}), T_2(\mathbf{X})\big] = \big[\overline{X_n} - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \overline{X_n} + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}\big].$$

Figyeljük meg, hogy

1. minél nagyobb a minta elemszáma, annál kevésbé hosszú konfidenciaintervallum szükséges (vö. nagy számok törvénye),
2. minél nagyobb a szórás, annál hosszabb konfidenciaintervallum szükséges (nagy szórás 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).

Konkrét ilyen példák $\rightarrow$ gyakorlaton (13.1. (b) feladat.)

Konfidenciaintervallumok: kitekintés

Létezik ismert képlet konfidenciaintervallumra σ^2-re ismert és ismeretlen μ esetén is, valamint μ-re ismeretlen σ^2 esetén is.
Ezek a konfidenciaintervallumok hasonló elvek mentén határozhatóak meg, de nem a normális eloszlás táblázat segítségével.
Kell hozzájuk két újabb nevezetes eloszláscsalád, a χ^2-eloszlások és a t-eloszlások.
Idén ezeket nem részletezzük, lásd a jegyzetkiegészítést, illetve az MSc-s Matematikai statisztika tárgyat.

Hipotézisvizsgálat

Hipotézisvizsgálat $\rightarrow$ valamilyen előzetes hipotézisünket (azaz feltevésünket) akarjuk valamilyen mérés, illetve adatelemzés alapján igazolni vagy cáfolni.

Ilyen hipotézis lehet például:

- 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 szórása 8.
- A mintaelemek normális eloszlásúak.
- A realizációk függetlenek valamely más, kapcsolódó mérés eredményeitől.
- A mintaelemek ugyanolyan eloszlásúak, mint egy másik gyárban hasonló méréssel kapott minta elemei.

Ebben a bevezető statisztikai részben főleg az első háromról lesz szó, az MSc-s Matematikai statisztikai tárgyban a többiről is.

Statisztikai próbák

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 hipotézisvizsgálati eljárás (röviden statisztikai próba) az alábbi lépéssorozat végrehajtását jelenti:

1. A H_0-lal jelölt *nullhipotézis* formalizálása.
2. A mintára vonatkozó *elfogadási tartomány* és annak komplementere, a *kritikus tartomány* megalkotása.
3. A kísérlet végrehajtása, amelyben az $x_1, \ldots, x_n$ 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.

* Egyes statisztikai iskolák interpretációja szerint H_0-t soha nem lehet "elfogadni", csak „nincs elég bizonyítékunk ahhoz, hogy elutasítsuk". Ebből a tárgyból ehhez nem ragaszkodunk, H_0-t el szabad fogadni.

Statisztikai próbák

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 hipotézisvizsgálati eljárás (röviden statisztikai próba) az alábbi lépéssorozat végrehajtását jelenti:

1. A H_0-lal jelölt *nullhipotézis* formalizálása.
2. A mintára vonatkozó *elfogadási tartomány* és annak komplementere, a *kritikus tartomány* megalkotása.
3. A kísérlet végrehajtása, amelyben az $x_1, \ldots, x_n$ 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.**

** Léteznek ún. *szekvenciális próbák* is, ahol a 4. pontnak ,,nem meggyőző köztes eredmény" esetén lehet olyan kimenetele is, hogy újabb mintát kell venni $\rightarrow$ ciklus indul újra 3.-tól. Pl.: Wald–Wolfowitz-próba.

Paraméteres és nemparaméteres próbák

Ebből a tárgyból alapvetően paraméteres próbákkal foglalkozunk $\rightarrow$ az eloszlás típusa ismert (pl. normális), van egy vagy több ismeretlen paraméter, ami(k)re a hipotézis vonatkozik.
Most: u-próba, több változatban $\rightarrow$ normális eloszlás várható értékére vonatkozó próbák ismert szórás esetén.
Azért ezt tanuljuk, mert a gyakorlatban ez az egyik leggyakoribb próbatípus és számos próba ehhez hasonló elven működik.
Más paraméteres próbatípusok: az MSc-s Matematikai statisztika tárgyban $\rightarrow$ kitekintés később.

MSc-s tárgy $\rightarrow$ nemparaméteres próbák is, amik a konkrét eloszlástípus ismerete nélkül is használhatóak:

- függetlenségvizsgálat,
- illeszkedésvizsgálat (egy bizonyos adott eloszlást követ-e a minta?),
- homogenitásvizsgálat (két minta peremeloszlásai azonosak-e?).

Pl. χ^2-próba, Kolmogorov–Szmirnov-próba.

Hipotézisvizsgálati alapfogalmak

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

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

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

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

[1] vs.: ejtsd ,,versus" (latinul ,,ellen").

Hipotézisvizsgálati alapfogalmak (folytatás)

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

$$\mathbb{P}\big((X_1, \ldots, X_n) \in \mathfrak{X}_{\mathrm{k}} \big| H_0 \text{ igaz}\big) = \alpha.$$

(4) **Elsőfajú hibá**nak 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).

Egy α terjedelmű próba elsőfajú hibájának valószínűsége éppen a terjedelem:

$$\mathbb{P}(H_0\text{-t elutasítjuk}|H_0 \text{ igaz}) = \mathbb{P}\big((X_1, \ldots, X_n) \in \mathfrak{X}_{\mathrm{k}} \big| H_0 \text{ igaz}\big) = \alpha.$$

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

Interpretáció – Bolla Marianna nyomán

Matematikailag nem precíz, de a szemléltetéshez segíthet:

(16.1.3. Megjegyzés.) A H_0 nullhipotézist gyakran úgy fogalmazzuk meg, hogy annak teljesülése esetén ismert legyen az $X_1, \ldots, X_n$ mintaelemek eloszlása.

Pl. utcán talált pénzérme $\rightarrow$ 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.
Nem jó nullhipotézis például a $\widehat{H}_0$: ,,a pénzérme nem szabályos" kijelentés. Mi ekkor a fae. indikátorok paramétere? Terjedelem?

Interpretáció 2.

(16.1.4. Megjegyzés.) A nullhipotézist sokszor ,,az ártatlanság vélelmében" fogalmazzuk meg.
Pl. a pénzérmét tipikusan azért vizsgáljuk meg, mert arra gyanakodunk, hogy cinkelt, de mégis ezen állítás tagadását tesszük meg H_0-nak. Ekkor

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

A legtöbb ismert próbában csak az elsőfajú hiba valószínűségét, vagyis a terjedelmet tudjuk kontrollálni: előre megadunk egy α értéket $\rightarrow$ a próba konstrukciója alapján a terjedelem éppen α lesz.
Érezhető (a próbák ismerete nélkül is): 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.

Elsőfajú és másodfajú hiba: példa

16.1.5. Példa: Egy új gyógyszert szeretnénk forgalomba hozni. Ekkor H_0: „a gyógyszer hatástalan vagy káros” vs. H_1: „a gyógyszer hatásos” jó alternatíva.

Itt: elsőfajú hiba $\Leftrightarrow$ 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.

Másodfajú hiba $\Leftrightarrow$ 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.

Ha csak az elsőfajú hiba valószínűsége kontrollálható, akkor H_0-t a fenti módon kell megválasztani.

Vissza az interpretációhoz

Láttuk:
Pl. utcán talált pénzérme $\rightarrow$ 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.
Nem jó nullhipotézis például a $\widehat{H}_0$: ,,a pénzérme nem szabályos" kijelentés. Mi ekkor a fae. indikátorok paramétere? Terjedelem?

Egész jó nullhipotézis a $\widetilde{H}_0$: ,,a pénzérme legalább $1/2$ valószínűséggel mutat fejet" kijelentés.
Igaz ugyan, hogy szintén nem tudjuk, hogy ha $\widetilde{H}_0$ igaz, akkor mi az indikátorok paramétere.
Azonban tegyük fel, hogy arra gyanakszunk, hogy a pénzérme túl kicsi valószínűséggel mutat fejet. Ekkor ha $\widetilde{H}_0$-t az ,,ártatlanság vélelmében" fogalmazzuk meg, akkor pl. intuitíve $\widetilde{H}_0$-t el lehet utasítani, ha nagyon ritkán dobunk fejet.
Ugyanis ha $\widetilde{H}_0$ igaz, legalább olyan gyakran dobunk fejet, mintha a pénzérme szabályos lenne.

Kétoldali, egymintás u-próba: példa/levezetés

(Bolla Marianna példája nyomán.)
A sarki péknél mindennap 1 db egykilós kenyeret veszünk. Az utóbbi időben mintha kisebbek lennének a kenyerek $\rightarrow$ tényleg 1 kilósak várható értékben?

Feltesszük: a kenyér tömege (kg-ban mérve) normális eloszlású $\sigma = 0{,}02$ kg szórással, tehát $\mathrm{N}(\mu; 0{,}02^2)$ eloszlású. Kérdés: $\mu \stackrel{?}{=} \mu_0$.

Alternatíva („a H_0 nullhipotézis megválasztása az ártatlanság vélelmében"):

$$H_0 \colon \mu = \mu_0 \qquad \text{vs.} \qquad H_1 \colon \mu \neq \mu_0,$$

ahol esetünkben $\mu_0 = 1$ kg.

Előirányzott terjedelem: $\varepsilon = 0{,}05$,
vagyis azt szeretnénk, hogy a próba során legfeljebb 0,05 valószínűséggel utasítsuk el H_0-t, amennyiben H_0 igaz.

Kétoldali, egymintás u-próba: példa/levezetés

$$H_0 \colon \mu = \mu_0 \qquad \text{vs.} \qquad H_1 \colon \mu \neq \mu_0$$

Mintát veszünk: $n = 25$ napon keresztül az aznap vásárolt kenyeret mindig mérlegre tesszük.

Kapott $\mathbf{x} = (x_1, \ldots, x_{25})$ realizáció átlaga: $\overline{x_n} = \overline{x_{25}} = 0{,}98$ kg.

Cél: olyan próbát konstruálni, aminek a terjedelme pontosan $\varepsilon = 0{,}05$, azaz olyan $\mathfrak{X}_\mathrm{e}$ elfogadási tartományt és $\mathfrak{X}_\mathrm{k}$ kritikus tartományt keresünk, amelyekre

$$\mathbb{P}(H_0\text{-t elutasítjuk} | H_0 \text{ igaz}) = \varepsilon = 0{,}05.$$

Az $\mathfrak{X}_\mathrm{e}$ elfogadási tartományt keressük olyan alakban, hogy H_0-t akkor fogadjuk el, ha μ_0 egy $[\overline{X_n} - h, \overline{X_n} + h]$ alakú intervallumba esik, ahol $h > 0$.

Cél: h-t úgy megválasztani, hogy $\mu = \mu_0 = 1$ esetén éppen ε valószínűséggel essen a minta a kritikus tartományba $\Leftrightarrow$ $1 - \varepsilon$ valószínűséggel essen az elfogadási tartományba $\Leftrightarrow$ teljesüljön, hogy

$$\mathbb{P}_{\mu_0}\left(\mu_0 \in [\overline{X_n} - h, \overline{X_n} + h]\right) = 1 - \varepsilon.$$

Kétoldali, egymintás u-próba: példa/levezetés

$$H_0\colon \mu = \mu_0 \qquad \text{vs.} \qquad H_1\colon \mu \neq \mu_0$$

Mintát veszünk: $n = 25$ napon keresztül az aznap vásárolt kenyeret mindig mérlegre tesszük.
Kapott $\mathbf{x} = (x_1, \ldots, x_{25})$ realizáció átlaga: $\overline{x_n} = \overline{x_{25}} = 0{,}98$ kg.

Cél: olyan próbát konstruálni, aminek a terjedelme pontosan $\varepsilon = 0{,}05$, azaz olyan $\mathfrak{X}_\mathrm{e}$ elfogadási tartományt és $\mathfrak{X}_\mathrm{k}$ kritikus tartományt keresünk, amelyekre

$$\mathbb{P}(H_0\text{-t elutasítjuk}|H_0 \text{ igaz}) = \varepsilon = 0{,}05.$$

Az $\mathfrak{X}_\mathrm{e}$ elfogadási tartományt keressük olyan alakban, hogy H_0-t akkor fogadjuk el, ha μ_0 egy $[\overline{X_n} - h, \overline{X_n} + h]$ alakú intervallumba esik, ahol $h > 0$.
Cél: h-t úgy megválasztani, hogy $\mu = \mu_0 = 1$ esetén éppen ε valószínűséggel essen a minta a kritikus tartományba
$\Leftrightarrow$ $1 - \varepsilon$ valószínűséggel essen az elfogadási tartományba
$\Leftrightarrow$ teljesüljön, hogy

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

Kétoldali, egymintás u-próba: példa/levezetés

Az eddigiek alapján $[\overline{X_n} - h, \overline{X_n} + h]$-ra teljesülnie kell annak, hogy $\mathbb{P}_{\mu_0}\big(\mu_0 \in [\overline{X_n} - h, \overline{X_n} + h]\big) = 1 - \varepsilon$.

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.

Láttuk: egy ilyen konfidenciaintervallum

$$\Big[\overline{X_n} - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}, \overline{X_n} + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}\Big],$$

amiből tehát azt kapjuk, hogy $h = \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}$. Vagyis H_0-t pontosan akkor fogadjuk el, ha

$$\overline{X_n} - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \le \mu_0 \le \overline{X_n} + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \quad \Leftrightarrow \quad \mu_0 - \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}} \le \overline{X_n} \le \mu_0 + \frac{\sigma u_{\varepsilon/2}}{\sqrt{n}}$$

$$\Leftrightarrow \quad -u_{\varepsilon/2} \le \frac{\overline{X_n} - \mu_0}{\sigma}\sqrt{n} \le u_{\varepsilon/2}.$$

Az elfogadási tartományunk emiatt

$$\mathfrak{X}_{\mathrm{e}} = \Big\{\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \colon -u_{\varepsilon/2} \le \frac{\overline{x_n} - \mu_0}{\sigma}\sqrt{n} \le u_{\varepsilon/2}\Big\}.$$

Kétoldali, egymintás u-próba: példa/levezetés

Az elfogadási tartományunk

$$\mathfrak{X}_{\mathrm{e}} = \left\{\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \colon\ -u_{\varepsilon/2} \leq \frac{\overline{x_n} - \mu_0}{\sigma}\sqrt{n} \leq u_{\varepsilon/2}\right\}.$$

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 $\sigma = 0{,}02$. Ez alapján tehát $u(\mathbf{x}) = \frac{0{,}98 - 1}{0{,}02}\sqrt{25} = -5$.
Normális eloszlás táblázata $\Rightarrow u_{\varepsilon/2} = \Phi^{-1}(0{,}975) \approx 1{,}9600$.
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 fenn, H_0-t elutasítjuk, vagyis kellő bizonyítékot találtunk arra, hogy a kenyerek várható értékben nem egykilósak.

Kétoldali, egymintás u-próba: példa/levezetés

Az elfogadási tartományunk

$$\mathfrak{X}_{\mathrm{e}} = \left\{\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \colon\ -u_{\varepsilon/2} \leq \frac{\overline{x_n} - \mu_0}{\sigma}\sqrt{n} \leq u_{\varepsilon/2}\right\}.$$

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 $\sigma = 0{,}02$. Ez alapján tehát $u(\mathbf{x}) = \frac{0{,}98 - 1}{0{,}02}\sqrt{25} = -5$.
Normális eloszlás táblázata $\Rightarrow u_{\varepsilon/2} = \Phi^{-1}(0{,}975) \approx 1{,}9600$.
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 fenn, H_0-t elutasítjuk, vagyis kellő bizonyítékot találtunk arra, hogy a kenyerek várható értékben nem egykilósak.

Kétoldali, egymintás u-próba: eljárás

Feltevések: az $x_1, \ldots, x_n$ adatpontok fae. normális eloszlású $X_1, \ldots, X_n$ valószínűségi változók realizációi, amelyek μ várható értéke ismeretlen, de $\sigma > 0$ szórásuk ismert.

Általános eljárás (nem feltétlenül kétoldali) egymintás u-próbánál:

1. A H_0 nullhipotézis felállítása.
2. Az előirányzott $\varepsilon \in (0, 1)$ terjedelem kiválasztása.
3. Az $x_1, \ldots, x_n$ adatpontok kinyerése.
4. Az elfogadási tartomány meghatározása (függ a próba altípusától).
5. A próbastatisztika $u(\mathbf{x}) = \frac{\overline{x_n} - \mu_0}{\sigma}\sqrt{n}$ 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).

Kétoldali eset. Nullhipotézis: $H_0 \colon \mu = \mu_0$.

Releváns kvantilis: $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$.

Elfogadási tartomány: $\mathfrak{X}_{\mathrm{e}} = \{(x_1, \ldots, x_n) \colon |u(\mathbf{x})| \leq u_{\varepsilon/2}\}$, tehát:

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

Egyoldali, egymintás u-próba: példa/levezetés

Térjünk vissza a kenyérvásárlás példájához.
Jogos kérdés: miért azt tettük meg H_1-nek, hogy $\mu \neq \mu_0$?
Ha a mintaátlag kisebb, mint 1 (pl. 0,98 kg), akkor józan ésszel arra gyanakszunk, hogy a kenyér súlyának várható értéke kisebb, mint 1 kg. Arra nem, hogy nagyobb.

Egy másik lehetséges alternatíva (ahol μ_0 továbbra is 1)

$$H_0 \colon \mu \geq \mu_0 \qquad \text{vs.} \qquad H_1 \colon \mu < \mu_0. \tag{5}$$

Kérdés: ha tudjuk, hogy H_0 teljesül, akkor a mintaelemek eloszlása is ismert?
Szigorú értelemben nem. H_0 teljesülése esetén lehetnek a mintaelemek például $\mathrm{N}(\mu_0; \sigma^2)$ és $\mathrm{N}(\mu_0 + 113; \sigma^2)$ 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).
Hasonló, mint az egész jó nullhipotézis a pénzérmés példánál.

Egyoldali, egymintás u-próba: példa/levezetés

Térjünk vissza a kenyérvásárlás példájához.
Jogos kérdés: miért azt tettük meg H_1-nek, hogy $\mu \neq \mu_0$?
Ha a mintaátlag kisebb, mint 1 (pl. 0,98 kg), akkor józan ésszel arra gyanakszunk, hogy a kenyér súlyának várható értéke kisebb, mint 1 kg. Arra nem, hogy nagyobb.

Egy másik lehetséges alternatíva (ahol μ_0 továbbra is 1)

$$H_0 \colon \mu \geq \mu_0 \qquad \text{vs.} \qquad H_1 \colon \mu < \mu_0. \tag{5}$$

Kérdés: ha tudjuk, hogy H_0 teljesül, akkor a mintaelemek eloszlása is ismert?
Szigorú értelemben nem. H_0 teljesülése esetén lehetnek a mintaelemek például $\mathrm{N}(\mu_0; \sigma^2)$ és $\mathrm{N}(\mu_0 + 113; \sigma^2)$ 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).
Hasonló, mint az egész jó nullhipotézis a pénzérmés példánál.

Egyoldali, egymintás u-próba: példa/levezetés

$$H_0\colon \mu \geq \mu_0 \qquad \text{vs.} \qquad H_1\colon \mu < \mu_0,$$

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.

A próbastatisztika most is $u = \frac{\overline{x_n}-\mu_0}{\sigma}\sqrt{n}$ lesz, de H_0-t most akkor utasítjuk el, ha u értéke nagyon kicsi.

Elfogadási tartomány: jobbról végtelen, és $\mu = \mu_0$ esetén $1-\varepsilon$ valószínűséggel esik μ_0 az $\mathfrak{X}_\mathrm{e}$ elfogadási tartományba, tehát ε valószínűséggel a tőle balra eső $\mathfrak{X}_\mathrm{k}$ kritikus tartományba:

$$\mathfrak{X}_\mathrm{k} = \{\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n\colon u(\mathbf{x}) < -u_\varepsilon\},$$

ahol $u_\varepsilon = \Phi^{-1}(1-\varepsilon)$, ahogy eddig.

Egyoldali, egymintás u-próba: példa, előnyök, változatok

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 $H_0\colon \mu \le \mu_0$ vs. $H_1\colon \mu > \mu_0$

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

$$\mathfrak{X}_\mathrm{k} = \{\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n \colon u(\mathbf{x}) > u_\varepsilon\}$$

kritikus tartományt kapjuk. H_0-t elutasítjuk $\Leftrightarrow$ $u(\mathbf{x})$ nagyon nagy.

Kétoldali, egymintás u-próba: eljárás

Feltevések: az $x_1, \ldots, x_n$ adatpontok fae. normális eloszlású $X_1, \ldots, X_n$ valószínűségi változók realizációi, amelyek μ várható értéke ismeretlen, de $\sigma > 0$ szórásuk ismert.

Általános eljárás (nem feltétlenül egyoldali) egymintás u-próbánál:

1. A H_0 nullhipotézis felállítása.
2. Az előirányzott $\varepsilon \in (0, 1)$ terjedelem kiválasztása.
3. Az $x_1, \ldots, x_n$ adatpontok kinyerése.
4. Az elfogadási tartomány meghatározása (függ a próba altípusától).
5. A próbastatisztika $u(\mathbf{x}) = \frac{\overline{x_n} - \mu_0}{\sigma}\sqrt{n}$ 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).

Egyoldali eset ($H_0\colon \mu \geq \mu_0$ változat). Nullhipotézis: $H_0\colon \mu \geq \mu_0$.

Releváns kvantilis: $-u_\varepsilon = -\Phi^{-1}(1 - \varepsilon)$.

Elfogadási tartomány: $\mathfrak{X}_\mathrm{e} = \{(x_1, \ldots, x_n)\colon u(\mathbf{x}) \geq -u_\varepsilon\}$, tehát:

- ha $u(\mathbf{x}) < -u_\varepsilon$, H_0-t elutasítjuk,
- különben H_0-t elfogadjuk.

Kétoldali, egymintás u-próba: eljárás

Feltevések: az $x_1, \ldots, x_n$ adatpontok fae. normális eloszlású $X_1, \ldots, X_n$ valószínűségi változók realizációi, amelyek μ várható értéke ismeretlen, de $\sigma > 0$ szórásuk ismert.

Általános eljárás (nem feltétlenül egyoldali) egymintás u-próbánál:

1. A H_0 nullhipotézis felállítása.
2. Az előirányzott $\varepsilon \in (0, 1)$ terjedelem kiválasztása.
3. Az $x_1, \ldots, x_n$ adatpontok kinyerése.
4. Az elfogadási tartomány meghatározása (függ a próba altípusától).
5. A próbastatisztika $u(\mathbf{x}) = \frac{\overline{x_n} - \mu_0}{\sigma}\sqrt{n}$ 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).

Egyoldali eset ($H_0\colon \mu \leq \mu_0$ változat). Nullhipotézis: $H_0\colon \mu \leq \mu_0$.

Releváns kvantilis: $u_\varepsilon = \Phi^{-1}(1 - \varepsilon)$.

Elfogadási tartomány: $\mathfrak{X}_\mathrm{e} = \{(x_1, \ldots, x_n)\colon u(\mathbf{x}) \leq u_\varepsilon\}$, tehát:

- ha $u(\mathbf{x}) > u_\varepsilon$, H_0-t elutasítjuk,
- különben H_0-t elfogadjuk.

Próba ereje, az u-próba optimalitása (kieg. anyag)

Láttuk: a kétoldali, egymintás u-próba terjedelme, vagyis elsőfajú hibájának valószínűsége legfeljebb az előzetesen választott ε lesz.
Kérdés: Mit mondhatunk a másodfajú hibáról?
Ha tudjuk, hogy H_1 teljesül $\Rightarrow$ mintaelemek eloszlása??
Vegyünk most egy konkrét $\mu_1 \neq \mu_0$ 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 hiba képlete ez lenne:

$$\mathbb{P}(H_0\text{-t elfogadjuk}|\, \mu = \mu_1) = \mathbb{P}\big(u(\mathbf{X}) \in (-u_{\varepsilon/2}, u_{\varepsilon/2})\big|\, X_1 \sim \mathrm{N}(\mu_1; \sigma^2)\big).$$

Ebben az esetben a próba **erejé**nek 1 mínusz a másodfajú hiba valószínűségét nevezzük. Az erő tehát annak a valószínűsége, hogy H_0-t valóban elutasítjuk, ha μ_1 a valódi paraméter.
Az erő itt tehát csak konkrét H_1-beli paraméter esetén definiálható.
Minél távolabb μ_1 a μ_0 értéktől, annál kisebb lesz a másodfajú hiba valószínűsége, vagyis annál nagyobb lesz a próba ereje.

Egyenletesen legerősebb próbák (kieg. anyag)

Miért a (kétoldali, egymintás) u-próbát alkalmazzuk a gyakorlatban és nem valamilyen más, szintén ε terjedelmű próbát ugyanerre 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.

Vagyis: bárhogy is választunk egy H_1-beli μ_1 paramétert, $\mu = \mu_1$ esetén a másodfajú hiba akkor lesz a legkisebb (vagyis az erő a legnagyobb), ha az u-próbát alkalmazzuk.

(Ezt nem bizonyítjuk, sőt ki sem mondjuk formálisan.
Megjegyzés: ez a **Neyman–Pearson-lemma** következménye.
A Neyman–Pearson-lemma az egyenletesen legerősebb próba létezését mondja ki és konstrukcióját adja meg általános esetben $\rightarrow$ lásd BME-TTK-s matematikai statisztika tárgyak.)

Ilyen értelemben egyenletesen legerősebb próba az összes u- és t-próba, amit az előadáson tárgyalunk, mindegyik a maga alternatívájára vonatkozóan, a maga feltevései mellett.

Az egymintás u-próba alkalmazásai

Az u-próba esetenként olyan esetekben is alkalmazható, ahol a mintaelemek ugyan nem normális eloszlásúak, de a CHT/de Moivre–Laplace-tétel miatt azzal közelíthetőek.

(Példa: u-próba binomiális eloszlásra, 16.2.4. Példa)

(Házi feladat: ugyanerre a példára megfelelő kétoldali próba?)

p-érték

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.

Fontos mennyiség: az a terjedelem, ahol H_0-t még éppen elfogadjuk.

Ezt szokás a próba p-**értéké**nek (angolul *p-value*) nevezni. Minél kisebb a p-érték, annál jogosabb a H_0 nullhipotézis elutasítása.

A pénzérmés 16.2.4. példában ez az érték $\varepsilon = 0{,}01$ és $\varepsilon = 0{,}05$ között van. Ebben a példában (egyoldali, egymintás u-próba) a p-érték éppen 1 mínusz Φ-nek a próbastatisztikában vett értéke:

$$u(\mathbf{x}) = u_\varepsilon \;\Leftrightarrow\; \mathbb{P}(u(\mathbf{X}) > u(\mathbf{x})) = \varepsilon \;\Leftrightarrow\; \varepsilon = 1-\Phi(u(\mathbf{x})) = 1-\Phi(2) \approx 0{,}0228.$$

Vagyis a próba p-értéke kb. 2,28%. Maximum ekkora terjedelemnél 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, jogos H_0 elutasítása.

p-érték

p-érték használata $\rightarrow$ alternatív szemlélet a terjedelemhez képest. Itt nem rögzítjük a mintavétel előtt a terjedelmet, hanem a minta alapján adjuk meg azt az ε terjedelmet, amely pont a H_0 elfogadása és elutasítása határán áll, azaz a p-értéket.

Gyakori szóhasználat (nemcsak u-próbánál): hogy valamilyen mintára vonatkozó állítás „*szignifikáns*", ha a p-érték legfeljebb 0,05, és „*nagyon szignifikáns*", ha legfeljebb 0,01.
Ha elfogadjuk a nullhipotézist, max. 0,05, ill. 0,01 lehet az elsőfajú hiba.

Megjegyzés: a sok törölt statisztika anyagrész (maximum likelihood becslés, t-eloszlásos konfidenciaintervallum és próbák stb.) kompenzálására a p-értéket idén a korábbiaknál jobban hangsúlyozzuk, vizsgán is szerepelhet.
Régi vizsgákban nincsenek ilyen feladatok, ehelyett a gyakorlati feladatok átnézését javaslom mindenkinek.

Kétmintás próbák

Következő téma: kétmintás u-próba. $\rightarrow$ Főleg a kétoldali változatára koncentrálunk.
Ebből az egyoldali verziók nagyon hasonlóan vezethetők le, mint az egymintás esetben.

Itt két független (nem feltétlenül azonos elemszámú) fae. normális eloszlású minta lesz adott. Szeretnénk tesztelni, hogy a két minta várható értéke megegyezik-e.
Feltevés: A mintaelemek szórásait ismerjük (ezek egy-egy mintán belül természetesen egyenlőek, de a két minta között nem feltétlenül).

Kétoldali, kétmintás u-próba

Kenyeres példa $\rightarrow \varepsilon = 0{,}05$ terjedelem mellett el kell utasítanunk azt a nullhipotézist, hogy a péknél vásárolt kenyerek várható értékben egykilósak $\rightarrow$ alternatív kenyérforrás után nézünk.
Egy másik pék nagyon hasonló egykilós kenyereket gyárt.
Reklámja szerint képes olyan kenyeret gyártani, amelyek tömegének szórása csak 1 dkg (azaz $\sigma_1 = 0.01$ kg, tehát $\sigma_1^2 = 0.0001$ kg^2 a szórásnégyzet).
Most $n_1 = 10$ napon keresztül ennél a péknél vásárolunk napi egy kenyeret, és megmérjük a tömegüket.
Egy olyan $\mathbf{x} = (x_1, \ldots, x_{n_1})$ realizációt kapunk, amelynek átlaga $\overline{x_{n_1}} = 0{,}9825$ kg.
A későbbiekben hasznos jelölés a régi péknél vásárolt kenyerek tömegeire: $\mathbf{y} = (y_1, \ldots, y_{n_2})$, ahol a mintaelemszám $n_2 = 25$, a mintaátlag $\overline{y_{n_2}} = 0{,}98$, a szórás $\sigma_2 = 0{,}02$ volt.
(Feltételezzük, hogy az – akárki által gyártott – kenyerek tömegei együttesen függetlenek.)
Kérdés: ugyanaz marad-e a helyzetünk, ha az új pékhez járunk?

Kétoldali, kétmintás u-próba

Vagyis: az új péknél vásárolt kenyerek tömegének (szintén ismeretlen) μ_1 várható értéke ugyanaz-e, mint a régi péknél vásároltak tömegének μ_2 várható értéke? Ezért a

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

alternatívára kívánunk egy $\varepsilon = 0{,}05$ terjedelmű próbát konstruálni. Feltételezve, hogy az $(X_1, \ldots, X_{n_1})$ minta független az $(Y_1, \ldots, Y_{n_2})$ mintától, az $\overline{X_{n_1}}$ mintaátlag is független az $\overline{Y_{n_2}}$ mintaátlagtól. Ezért $\overline{X_{n_1}}$ és $\overline{Y_{n_2}}$ lineáris kombinációi is normális eloszlásúak. Elemi számításokkal (lásd ea.) megmutatható, hogy

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

Így a kétoldali, egymintás u-próbához hasonlóan, ha

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

akkor H_0-t elfogadjuk, különben elutasítjuk. Kritikus tartományunk:

$$\mathfrak{X}_{\mathrm{k}} = \big\{(\mathbf{x}, \mathbf{y}) = (x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}) \in \mathbb{R}^{n_1+n_2} \colon |u(\mathbf{x}, \mathbf{y})| > u_{\varepsilon/2}\big\}.$$

Kétoldali, kétmintás u-próba

Konkrét példánkban

$$u(\mathbf{x},\mathbf{y}) = \frac{\overline{x_{n_1}} - \overline{y_{n_2}}}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} = \frac{0{,}9825 - 0{,}98}{\sqrt{\frac{0.0001}{10} + \frac{0.0004}{25}}} \approx 0{,}4903.$$

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 változtat szignifikánsan a helyzetünkön...

Kétoldali, kétmintás u-próba: eljárás

Feltevések: az $x_1, \ldots, x_{n_1}$ adatpontok fae. $\mathrm{N}(\mu_1, \sigma_1^2)$ eloszlású $X_1, \ldots, X_{n_1}$ valószínűségi változók realizációi, az $y_1, \ldots, y_{n_2}$ adatpontok pedig az $X_1, \ldots, X_{n_1}$-től (együttesen) független, egymás között fae. $\mathrm{N}(\mu_2, \sigma_2^2)$ eloszlású $Y_1, \ldots, Y_{n_2}$ valószínűségi változók realizációi. A $\sigma_1, \sigma_2 > 0$ szórások ismertek, a μ_1, μ_2 várható értékek ismeretlenek. Eljárás:

1. Nullhipotézis: $H_0 \colon \mu_1 = \mu_2$ vs. $H_1 \colon \mu_1 \neq \mu_2$.
2. Az előirányzott $\varepsilon \in (0, 1)$ terjedelem kiválasztása.
3. Az $x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}$ adatpontok kinyerése.
4. Próbastatisztika:
$$u(\mathbf{x}, \mathbf{y}) = \frac{\overline{x_{n_1}} - \overline{y_{n_2}}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$
5. Elfogadási tartomány: $\mathfrak{X}_{\mathrm{e}} = \big\{(\mathbf{x}, \mathbf{y}) = (x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}) \in \mathbb{R}^{n_1+n_2} \colon |u(\mathbf{x}, \mathbf{y})| \leq u_{\varepsilon/2}\big\}$, ahol $u_{\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$.
6. Döntés: ha $|u(\mathbf{x}, \mathbf{y})| > u_{\varepsilon/2}$, H_0-t elutasítjuk, kül. elfogadjuk.

Egyoldali, kétmintás u-próba

Mi a helyzet a kétmintás u-próba egyoldali változatával?
Az egymintás esetből világos, hogy mi a lehetséges kétfajta egyoldali változat: a kritikus tartományt $u(\mathbf{x}, \mathbf{y}) < -u_\varepsilon$ vagy $u(\mathbf{x}, \mathbf{y}) > u_\varepsilon$ határozza meg.
Kevésbé egyszerű kérdés: a kenyeres példában melyik a jó? ($\mathbf{x}$, azaz 1. minta: új, $\mathbf{y}$, azaz régi: régi pék mintája.)

Az előbb az volt a nullhipotézisünk, hogy nem változott a helyzetünk az új pékkel. Egymintás esetben ennek az felel meg, hogy nem változott vagy romlott. Vagyis:

$$H_0 \colon \mu_1 \leq \mu_2 \text{ vs. } \mu_1 > \mu_2.$$

Ugyanis csak akkor szeretnénk elfogadni, hogy javult, ha biztosak vagyunk benne $\rightarrow$ az elsőfajú hiba legyen az, hogy nem javult, de a próba alapján mégis azt mondjuk, hogy igen.

Egyoldali, kétmintás u-próba

Mi a helyzet a kétmintás u-próba egyoldali változatával?
Az egymintás esetből világos, hogy mi a lehetséges kétfajta egyoldali változat: a kritikus tartományt $u(\mathbf{x}, \mathbf{y}) < -u_\varepsilon$ vagy $u(\mathbf{x}, \mathbf{y}) > u_\varepsilon$ határozza meg.
Kevésbé egyszerű kérdés: a kenyeres példában melyik a jó? ($\mathbf{x}$, azaz 1. minta: új, $\mathbf{y}$, azaz régi: régi pék mintája.)

Az előbb az volt a nullhipotézisünk, hogy nem változott a helyzetünk az új pékkel. Egymintás esetben ennek az felel meg, hogy nem változott vagy romlott. Vagyis:

$$H_0 \colon \mu_1 \leq \mu_2 \text{ vs. } \mu_1 > \mu_2.$$

Ugyanis csak akkor szeretnénk elfogadni, hogy javult, ha biztosak vagyunk benne $\rightarrow$ az elsőfajú hiba legyen az, hogy nem javult, de a próba alapján mégis azt mondjuk, hogy igen.

Egyoldali, kétmintás u-próba

A kenyeres példában tehát az alternatíva ($\mathbf{x}$, azaz 1. minta: új, $\mathbf{y}$, azaz régi: régi pék mintája.)

$$H_0\colon \mu_1 \leq \mu_2 \text{ vs. } \mu_1 > \mu_2.$$

A próbastatisztikánk ez volt:

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

Kritikus eset: ha $\overline{Y_{n_2}}$ jóval kisebb $\overline{X_{n_1}}$-nél. Ekkor $u(\mathbf{X}, \mathbf{Y})$ kicsi. Ebből látszik, hogy a megfelelő egyoldali kritikus tartomány:

$$\mathfrak{X}_\mathrm{k} = \left\{(\mathbf{x}, \mathbf{y}) = (x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}) \in \mathbb{R}^{n_1+n_2} \colon u(\mathbf{x}, \mathbf{y}) > u_\varepsilon\right\}.$$

Láttuk: $u(\mathbf{x}, \mathbf{y}) \approx 0{,}4903$. Ez kisebb, mint $u_\varepsilon = \Phi^{-1}(0{,}95) \approx 1{,}645$, tehát H_0-t elfogadjuk.
Nem javult tehát a helyzetünk az új pékkel. (Annak ellenére, hogy az egyoldali próba megint csak ,,szigorúbb" a kétoldalinál, H_0-t ismét elfogadtuk.)

Kétoldali, kétmintás u-próba: példa

(13.7. (b) feladat)
Véletlenszerűen kiválasztottuk a BME 10 gépészmérnök férfi hallgatóját, az ő átlagos testmagasságuk 179 cm volt, 2 cm korrigált empirikus szórással, valamint a BME 20 építőmérnök férfi hallgatóját, az ő esetükben az átlag 178 cm, a korrigált empirikus szórás 1 cm volt. Feltéve, hogy az egyes hallgatók testmagasságai függetlenek, normális eloszlásúak és azonos szak esetén azonos eloszlásúak is, vizsgáljuk meg a minta alapján 5% és 1% terjedelmű próba segítségével is, hogy a két szakra járók testmagasságának várható értéke megegyezik-e, ha a gépészmérnök férfi hallgatók testmagasságának szórása $\sqrt{2}$ cm, az építőmérnök férfi hallgatóké pedig 1 cm.

Adjuk meg H_0-t, H_1-et, a próbastatisztika elfogadási tartománynak megfelelő értékei halmazát, a próbastatisztika értékét és a döntést.

Kitekintés: t-próba (lásd Matematikai statisztika tárgy)

t-próba: ugyanazok a nullhipotézisek, mint az u-próbánál.

Különbség: a normális eloszlású mintaelemek szórása ismeretlen.

- Egymintás (egy-/kétoldali) t-próba: hasonló az u-próbához, csak a próbastatisztika más (σ helyére a korrigált empirikus szórás kerül), és a normális eloszlás kvantilisei helyett a megfelelő paraméterű t-eloszlás kvantilisével hasonlítjuk össze.
- Kétmintás (egyoldali/kétoldali) t-próba: alapvetően akkor működik, ha a két minta szórása ismeretlen, de egyenlő.
- Mi a teendő, ha nem tudjuk, hogy a szórások egyenlőek-e?
  - Szórások egyenlőségének ellenőrzésére is van próba: F-próba, $H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_1 : \sigma_1^2 \neq \sigma_2^2$.
  - Ha az F-próbánál H_0-t elfogadjuk, mehet a kétmintás t-próba.
  - Ha az F-próbánál H_0-t elutasítjuk: Welch-próba.
    A próbastatisztika még mindig t-eloszlású, más paraméterrel.
- Összetartozó mintás t-próba: $(x_1, y_1), \ldots, (x_n, y_n)$ adatpontok, egymás közt $x_1, \ldots, x_n$ és $y_1, \ldots, y_n$ is fae. normális eloszlású val. változók realizációi, de x_i és y_i nem függetlenek, hanem ugyanabból a mérésből származnak. Pl: n csavart mérünk, x_i az i-edik csavar hossza, y_i az átmérője.

Kitekintés: t-próba (lásd Matematikai statisztika tárgy)

t-próba: ugyanazok a nullhipotézisek, mint az u-próbánál.

Különbség: a normális eloszlású mintaelemek szórása ismeretlen.

- Egymintás (egy-/kétoldali) t-próba: hasonló az u-próbához, csak a próbastatisztika más (σ helyére a korrigált empirikus szórás kerül), és a normális eloszlás kvantilisei helyett a megfelelő paraméterű t-eloszlás kvantilisével hasonlítjuk össze.
- Kétmintás (egyoldali/kétoldali) t-próba: alapvetően akkor működik, ha a két minta szórása ismeretlen, de egyenlő.
- Mi a teendő, ha nem tudjuk, hogy a szórások egyenlőek-e?
  - Szórások egyenlőségének ellenőrzésére is van próba: F-próba, $H_0 \colon \sigma_1^2 = \sigma_2^2$ vs. $H_1 \colon \sigma_1^2 \neq \sigma_2^2$.
  - Ha az F-próbánál H_0-t elfogadjuk, mehet a kétmintás t-próba.
  - Ha az F-próbánál H_0-t elutasítjuk: Welch-próba.
    A próbastatisztika még mindig t-eloszlású, más paraméterrel.
- Összetartozó mintás t-próba: $(x_1, y_1), \ldots, (x_n, y_n)$ adatpontok, egymás közt $x_1, \ldots, x_n$ és $y_1, \ldots, y_n$ is fae. normális eloszlású val. változók realizációi, de x_i és y_i nem függetlenek, hanem ugyanabból a mérésből származnak. Pl: n csavart mérünk, x_i az i-edik csavar hossza, y_i az átmérője.