$\xi_1, \xi_2, \ldots$ T.F.R.; azonos eloszlásúak

statisztikai minta $(\xi_1, \ldots, \xi_n)$

n a minta terjedelme

(vagy n elemű minta)

A MINTA VÁRHATÓ ÉRTÉKE:

$$\bar{\xi} \equiv \bar{\xi}_n \equiv \frac{1}{n} \sum_{i=1}^{n} \xi_i$$

A MINTA SZÓRÁSNÉGYZETE:

$$\sigma^2 \equiv \sigma_n^2 \equiv \frac{1}{n} \sum_{i=1}^{n} (\xi_i - \bar{\xi})^2$$

A MINTA KORRIGÁLT SZÓRÁSNÉGYZETE:

$$\sigma^{*2} \equiv \sigma_n^{*2} \equiv \frac{1}{n-1} \sum_{i=1}^{n} (\xi_i - \bar{\xi})^2$$

$$\sigma^2 = \frac{n-1}{n} \sigma^{*2}, \quad \sigma^{*2} = \frac{n}{n-1} \sigma^2$$

STATISZTIKAI BECSLÉSEK

$\mathbb{R}^1$-beli ELOSZLÁSOK PARAMÉTEREIRE

A SZÁMTANI KÖZÉP

$$\bar{x} \equiv \bar{x}_n \equiv \frac{1}{n} \sum_{i=1}^{n} x_i$$

AZ ÁTLAGOS NÉGYZETES ELTÉRÉS

$$s^2 \equiv s_n^2 \equiv \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

A KORRIGÁLT NÉGYZETES ELTÉRÉS

$$s^{*2} \equiv s_n^{*2} \equiv \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$s^2 = \frac{n-1}{n} s^{*2}, \quad s^{*2} = \frac{n}{n-1} s^2$$

A STEINER AZONOSSÁG (=-ság): $\forall c \in \mathbb{R}$-re

$$\sigma^2 \equiv \frac{1}{n} \sum_{i=1}^{n} (\xi_i - \bar{\xi})^2 = \frac{1}{n} \sum_{i=1}^{n} (\xi_i - c)^2 + \left[\frac{1}{n} \sum_{i=1}^{n} (\xi_i - c)\right]^2$$

STEINER AZONOSSÁG $\forall c \in \mathbb{R}$ esetén

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i-\bar{\xi})^2 = \frac{1}{n}\sum_{i=1}^{n}(\xi_i-c)^2 - \left(\frac{\sum_1^n(\xi_i-c)}{n}\right)^2$$

Biz. $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}\left((\xi_i - c) + (c-\bar{\xi})\right)^2 = \frac{1}{n}\sum_1^n\left[(\xi_i-c)^2 + 2(c-\bar{\xi})(\xi_i-c) + (c-\bar{\xi})^2\right] = \frac{1}{n}\sum_{i=1}^{n}(\xi_i-c)^2 + \frac{2}{n}(c-\bar{\xi})\sum_{i=1}^{n}(\xi_i-c) + n(c-\bar{\xi})^2 =$

$$= \frac{1}{n}\sum_{i=1}^{n}(\xi_i-c)^2 + \frac{2}{n}\left(c - \frac{\sum_1^n \xi_i}{n}\right)\sum_1^n(\xi_i-c)$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\xi_i-c)^2 - \frac{2}{n}(\bar{\xi}-c)\sum_{i=1}^{n}(\xi_i-c)$$

$$\boxed{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left((\xi_i-c)-(\bar{\xi}-c)\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left\{(\xi_i-c)^2 - 2(\bar{\xi}-c)(\xi_i-c) + (\bar{\xi}-c)^2\right\} =$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\xi_i-c)^2 - \frac{2}{n}\left(\frac{\sum_1^n \xi_i}{n} - c\right)\sum_{i=1}^{n}(\xi_i-c) + \frac{1}{n}\cdot n\left(\frac{\sum_1^n \xi_i}{n} - c\right)^2 =$$

$$-\frac{2}{n}\,\frac{\sum_1^n(\xi_i-c)}{n}\cdot\sum_{i=1}^{n}(\xi_i-c) \qquad 1\left(\frac{\sum_1^n(\xi_i-c)}{n}\right)^2$$

$$= \boxed{\frac{1}{n}\sum_{i=1}^{n}(\xi_i-c)^2} \qquad -2\left(\frac{\sum_{i=1}^n(\xi_i-c)}{n}\right)^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\xi_i-c)^2 - \left(\frac{\sum_{i=1}^n(\xi_i-c)}{n}\right)^2$$

minimumként realizálódik egy elég nagy sugarú, origó középpontú, gömb belső pontjában (vagy pontjaiban) ⟨mint kiderül – pontosan egy pontjában⟩. Véve egy ilyen nagy R_0 sugarú gömböt és kihasználva, hogy $f \in C^\infty(\|\underline{y}'\| \leqq R_0)$ – adódik, hogy $\underline{y}'$-minimumpontban (pontosabban lokális szélsőértékpontban; globális maximumhelye nincs $\underline{f}$-nek – mivel $(\ast\ast\ast)$ miatt $\sup\limits_{\underline{y}' \in \mathbb{R}^{n-1}} f(\underline{y}') = +\infty$)

$$0 = f_{y_l}(\underline{y}') = -2c_{i,l} + \sum_{\substack{j \neq i \\ j=\overline{1,n}}} \left(c_{j,j} y_j^2\right)_{y_l} + \left(\sum_{\substack{j,k=1 \\ j,k \neq i,\, k \neq j}}^{n} c_{j,k} y_j y_k\right)_{y_l} =$$
$l = \overline{1,n}\ ;\ l \neq i$

$$= -2c_{i,l} + 2c_{l,l} y_l + \sum_{\substack{k=1 \\ k \neq i,\, k \neq l \\ l \neq i}}^{n} c_{l,k} y_k + \sum_{\substack{j=1 \\ j \neq i,\, j \neq l \\ l \neq i}}^{n} c_{j,l} y_j \underset{\boxed{c_{j,l} = c_{l,j} \text{ miatt}}}{=}$$

$$= -2c_{i,l} + 2\sum_{\substack{k=1 \\ k \neq i}}^{n} c_{l,k} y_k\,,$$

azaz

$$(::)\quad \sum_{\substack{k=1 \\ l,k \neq i}}^{n} c_{l,k} y_k = c_{i,l} = c_{l,i} \qquad l = \overline{1,n}\ ,\ l \neq i$$

A $(::)$ egyenletrendszer determinánsa a $\Gamma_{i,i} \succ 0$ (pozitív definit) mátrix $\Gamma_{i,i}$ előjeles $((-1)^{i+i} = 1 > 0)$ aldeterminánsa, ami így pozitív (spec. nem nulla) – s ezért $(::)$-nak pontosan egy $\underline{y}' = (y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_n)$ megoldása van. E pont $(\underline{y}')$ eleve <u>globális minimumhely</u> (nemcsak lokális) – ami logikailag is világos; de ~~másképpen~~ is adódik, nevezetesen abból, hogy az $\underline{n-1}$ változós $\sum\limits_{j,k=1;\, j \neq i}^{n} c_{j,k} y_j y_k$ fgv. HESSE mátrixa: $\Gamma_{i,i} \succ 0$ (azaz pozitív definit) !!!

A $-\dfrac{\Gamma_{i,j}}{\Gamma_{i,i}}$ regressiós együtthatók <u>statisztikai becslése</u>

a $c_{k,l} := M[(\xi_k - m_k)(\xi_l - m_l)]$ kovariancia együtthatók, ill./vagy az $r_{k,l} = c_{k,l}/D_k \cdot D_l$ korrelációs együtthatók statisztikai becslésével történhet.

Speciálisan két V.V ξ, η esetén (ha $D(\xi), D(\eta) > 0$) a (ξ, η)-ra vonatkozó $(\xi_1, \eta_1), \ldots, (\xi_N, \eta_N)$ minta esetén $(N \gg 1)$ $c := C(\xi, \eta) := M[(\xi - M(\xi))(\eta - M(\eta))]$ becslésére a

$$(1)\ \hat{c}_N = \hat{c} = \frac{1}{N-1} \sum_{i=1}^{N} (\xi_i - \bar{\xi})(\eta_i - \bar{\eta})$$ kifejezést szokás használni,

ahol
$$\bar{\xi} := \frac{1}{N} \sum_{i=1}^{N} \xi_i, \quad \bar{\eta} := \frac{1}{N} \sum_{i=1}^{N} \eta_i .$$

Hasonlóképpen (az (1)-beli $1/N-1$ tényezővel való egyszerűsítés után) $r := R(\xi, \eta) := \dfrac{C(\xi, \eta)}{D(\xi) D(\eta)}$ becslésére

$$\hat{r}_N = \hat{r} = \frac{\sum_{i=1}^{N} (\xi_i - \bar{\xi})(\eta_i - \bar{\eta})}{\sqrt{\sum_{i=1}^{N} (\xi_i - \bar{\xi})^2 \cdot \sum_{i=1}^{N} (\eta_i - \bar{\eta})^2}}$$

adódik.

<u>TÉTEL.</u> $\forall x \in \mathbb{R}$ fix esetén $\Phi_n(x,\omega) \xrightarrow[n\to\infty]{P} \mathcal{F}(x)$.

<u>Bizonyítás.</u> Legyen $A_i := \{\omega \in \Omega \mid \xi_i(\omega) < x\}$ $i \in \mathbb{N}$.
Mivel $\xi_1, \dots, \xi_n, \dots$ T.F.R., azért az $A_1, \dots, A_n, \dots$ eseményrendszer T.F.E.R. és

$(\bullet)$ $P(A_i) = \mathcal{F}_{\xi_i}(x) = \mathcal{F}(x)$ $i \in \mathbb{N}$.

A $\xi_{A_1}, \dots, \xi_{A_n}, \dots$ indikátorváltozók rendszere szintén T.F.R. és $(\bullet)$ miatt azonos eloszlásúak a $\xi_{A_1}, \xi_{A_2}, \dots$ változók és

$M(\xi_{A_i}) = P(A_i) = \mathcal{F}(x)$ $i \in \mathbb{N}$; $D(\xi_{A_i}) = \sqrt{\mathcal{F}(x)(1-\mathcal{F}(x))}$ $i \in \mathbb{N}$.

Így a CSEBISEV-tétel szerint

$$\frac{\xi_{A_1}(\omega) + \dots + \xi_{A_n}(\omega)}{n} = \Phi_n(x,\omega) \xrightarrow[n\to\infty]{P} \mathcal{F}(x) \quad \bullet$$

Legyenek ugyanis $\xi_1, \xi_2, \ldots$ teljesen független, azonos $\mathcal{P}(\in \mathcal{C})$ -eloszlásu valószinüségi változók, és legyen $\varepsilon\,(>0)$ tetszőleges. A jelölés egyszerüsitése érdekében vezessük be a $\hat{T}_n = T_n(\xi_1, \ldots, \xi_n)$ $(n=1,2,\ldots)$ jelölést. Ekkor a feltevések szerint van olyan n_0, hogy $n > n_0$ esetén $|M(\hat{T}_n) - t(\mathcal{P})| < \varepsilon/2$. Igy $n > n_0$ esetén $(|\hat{T}_n - t(\mathcal{P})| > \varepsilon) \subseteq (|\hat{T}_n - M(\hat{T}_n)| > \varepsilon/2)$, amiből a Csebisev-egyenlőség alkalmazásával kapjuk, hogy

$$\mathcal{P}(|\hat{T}_n - t(\mathcal{P})| > \varepsilon) \leqq \mathcal{P}(|\hat{T}_n - M(\hat{T}_n)| > \varepsilon/2) \leqq 4\,\frac{D^2(\hat{T}_n)}{\varepsilon^2} \to 0 \quad (n \to \infty),$$

Ennek a segédtételnek az alapján nyilvánvaló, hogy az átlagos négyzetes eltérések $\{s_n^2\}$ sorozata az összes véges negyedik momentummal rendelkező valószinüségi eloszlások osztálya esetén a szórásnégyzetnek konzisztens becslése.

Megemlitjük, hogy a konzisztens becslés és az erősen konzisztens becslés fogalma természetes módon általánositható arra az esetre is, amikor a paraméter a $\vec{t}(\mathcal{P}) = (t_1(\mathcal{P}), \ldots, t_\ell(\mathcal{P}))$ vektor. Ekkor a $\mathcal{C}$ osztály $\vec{t}$ paraméterének konzisztens, ill. erősen konzisztens becslésén olyan $\{\vec{T}^{(n)}(x_1, \ldots, x_n)\}_{n=1}^{\infty}$ $(\vec{T}^{(n)}(x_1, \ldots, x_n) = (T_1^{(n)}(x_1, \ldots, x_n), \ldots, T_\ell^{(n)}(x_1, \ldots, x_n));\ n=1,2,\ldots)$ sorozatot értünk, ahol minden i -re $(1 \leqq i \leqq \ell)$ $\{T_i^{(n)}(x_1, \ldots, x_n)\}$ a $\mathcal{C}$ osztály t_i paraméterének konzisztens, ill. erősen konzisztens becslése.

kell

~~II. 3. Hatásos becslések. Ebben a paragrafusban is az előző paragrafusok jelöléseit használjuk. Legyen n természetes szám, és tekintsük a~~

$$\inf_T \mathcal{D}(T(\xi_1,\dots,\xi_n))$$
$$\| $$
$$D(\mathcal{P}) = D_n(\mathcal{P}) = \inf_T \left[\int_{R_n} (T(x_1,\dots,x_n) - t(\mathcal{P}))^2 d\mathcal{P}^n\right]^{1/2} \quad (\mathcal{P} \in \mathcal{C})$$

függvényt, ahol az infimum a $\mathcal{C}$ osztály t paraméterének összes n -változós torzitatlan $T(x_1,\dots,x_n)$ becslésére képezendő. Tegyük fel, hogy $D(\mathcal{P}) < \infty \ (\mathcal{P} \in \mathcal{C})$. Ezután, ha $T(x_1,\dots,x_n)$ a $\mathcal{C}$ osztály t paraméterének torzitatlan becslése, képezzük az

$$E(T,\mathcal{P}) = E_n(T,\mathcal{P}) = D_n(\mathcal{P}) \left\{ \int_{R_n} (T(x_1,\dots,x_n) - t(\mathcal{P}))^2 d\mathcal{P}^n \right\}^{-1/2} =$$
$$= D_n(\mathcal{P}) / \mathcal{D}(T(\xi_1,\dots,\xi_n))$$

hányadost, amelyet <u>a</u> T <u>becslés</u> $\mathcal{C}$ -<u>re</u> <u>vonatkozó hatásfokának</u> nevezzük. Nyilvánvaló, hogy $0 \leqq E(T,\mathcal{P}) \leqq 1 \quad (\mathcal{P} \in \mathcal{C})$.
Ha $T_1(x_1,\dots,x_n)$ és $T_2(x_1,\dots,x_n)$ a $\mathcal{C}$ osztály t paraméterének torzitatlan becslései, és

$$E(T_1;\mathcal{P}) \geqq E(T_2;\mathcal{P}) \quad (\mathcal{P} \in \mathcal{C}),$$

akkor azt mondjuk, hogy T_1 <u>a</u> $\mathcal{C}$ <u>osztály</u> t <u>paraméterének</u> T_2 <u>-nél hatásosabb becslése.</u> Ez annyit jelent, hogy bármely $\mathcal{P}(\in \mathcal{C})$ eloszlású sokaságból vett $\xi_1,\dots,\xi_n$ minta esetén

$$D(T_1(\xi_1,\dots,\xi_n)) \leqq D(T_2(\xi_1,\dots,\xi_n)) .$$

Legyen $T(x_1,\dots,x_n)$ a $\mathcal{C}$ osztály t paraméterének torzitatlan becslése. Ha $\mathcal{D}(T(\xi_1,\dots,\xi_n)) = D(\mathcal{P}) \quad (\mathcal{P} \in \mathcal{C})$, ~~azaz~~

$$\left[\int_{R_n} (T(x_1,\dots,x_n) - t(\mathcal{P}))^2 d\mathcal{P}^n\right]^{1/2} = D(\mathcal{P}) \quad (\mathcal{P} \in \mathcal{C}),$$

akkor azt mondjuk, hogy T <u>a</u> $\mathcal{C}$ <u>osztály</u> t <u>paraméterének hatásos becslése</u>. Ez viszont annyit jelent, hogy a $\mathcal{C}$ osztály t

paraméterének bármely $\overline{T}(x_1,\dots,x_n)$ torzitatlan becslése, és bármely $\xi_1,\dots,\xi_n$ $\mathcal{P}(\in\mathcal{C})$ -eloszlásu sokaságból minta esetén

$$D(T(\xi_1,\dots,\xi_n)) \leqq D(\overline{T}(\xi_1,\dots,\xi_n)).$$

A hatásos becslés tehát a minimális szórásu becslés.

Megjegyezzük, hogy hatásos becslés nem mindig van, viszont a hatásos becslés unicitása könnyen megmutatható.

Legyenek $T_1(x_1,\dots,x_n)$ és $T_2(x_1,\dots,x_n)$ a $\mathcal{C}$ osztály t. paraméterének hatásos becslései, akkor bármely $\mathcal{P}(\in\mathcal{C})$ -re $T_1(x_1,\dots,x_n)=T_2(x_1,\dots,x_n)$ a $\mathcal{P}^n$ mértékre vonatkozóan majdnem mindenütt teljesül.

mal. stil. — mal. stil.

A feltevések szerint

$$D(\mathcal{P}) = \int_{R_n}(T_1(x_1,\dots,x_n)-t(\mathcal{P}))^2 d\mathcal{P}^n = \int_{R_1}(T_2(x_1,\dots,x_n)-t(\mathcal{P}))^2 d\mathcal{P}^n \quad (\mathcal{P}\in\mathcal{C}).$$

Ha valamely $\mathcal{P}(\in\mathcal{C})$ esetén $D(\mathcal{P})=0$, akkor ebből nyilvánvaló, hogy $T_1(x_1,\dots,x_n)=T_2(x_1,\dots,x_n)$ a $\mathcal{P}^n$ mértékre vonatkozóan majdnem mindenütt fennáll.

Tegyük fel, hogy valamely $\mathcal{P}(\in\mathcal{C})$ esetén $D(\mathcal{P})>0$. Mivel $(T_1+T_2)/2$ is a $\mathcal{C}$ osztály t paraméterének torzitatlan becslése, igy a Bunyakovszkij-Schwarz egyenlőtlenség alkalmazásával kapjuk, hogy

$$(2)\quad D^2(\mathcal{P}) \leqq \int_{R_n}\left(\frac{T_1+T_2}{2}-t(\mathcal{P})\right)^2 d\mathcal{P}^n = \int_{R_n}\left(\frac{T_1-t(\mathcal{P})}{2}+\frac{T_2-t(\mathcal{P})}{2}\right)^2 d\mathcal{P}^n =$$

$$=\frac{1}{4}\int_{R_n}(T_1-t(\mathcal{P}))^2 d\mathcal{P}^n+\frac{1}{4}\int_{R_n}(T_2-t(\mathcal{P}))^2 d\mathcal{P}^n+\frac{1}{2}\int_{R_n}(T_1-t(\mathcal{P}))(T_2-t(\mathcal{P}))\,d\mathcal{P}^n \leqq$$

$$\leq \frac{1}{2} D^2(\mathcal{P}) + \frac{1}{2} \sqrt{\int\limits_{R_n} (T_1 - t(\mathcal{P}))^2 d\mathcal{P}^n \int\limits_{R_n} (T_2 - t(\mathcal{P}))^2 d\mathcal{P}^n} = D^2(\mathcal{P}).$$

Ezek az egyenlőtlenségek viszont csak ugy teljesülhetnek, ha

$$\int\limits_{R_n} (T_1 - t(\mathcal{P}))(T_2 - t(\mathcal{P})) d\mathcal{P}^n = \sqrt{\int\limits_{R_n} (T_1 - t(\mathcal{P}))^2 d\mathcal{P}^n \cdot \int\limits_{R_n} (T_2 - t(\mathcal{P}))^2 d\mathcal{P}^n}.$$

Tudjuk, hogy a Bunyakovszkij-Schwarz-egyenlőtlenségben egyenlőség csak ugy lehet, ha valamely c állandóval

$$T_1(x_1, \dots, x_n) - t(\mathcal{P}) = c\,(T_2(x_1, \dots, x_n) - t(\mathcal{P}))$$

a $\mathcal{P}^n$ mértékre vonatkozóan majdnem mindenütt fennáll. Mivel a feltevés szerint

$$D^2(\mathcal{P}) = \int\limits_{R_n} (T_1 - t(\mathcal{P}))^2 d\mathcal{P}^n = c^2 \int\limits_{R_n} (T_2 - t(\mathcal{P}))^2 d\mathcal{P}^n = c^2 D^2(\mathcal{P}),$$

ezért $c = 1$, vagy $c = -1$. Mivel (2) szerint

$$0 < D^2(\mathcal{P})/2 \leq c \int\limits_{R_n} (T_2 - t(\mathcal{P}))^2 d\mathcal{P}^n = c\, D^2(\mathcal{P}),$$

igy $c = 1$. Ebből nyilvánvaló, hogy most is a $\mathcal{P}^n$ mértékre vonatkozóan majdnem mindenütt $T_1(x_1, \dots, x_n) = T_2(x_1, \dots, x_n)$.

A hatásosság fogalmát nemcsak az összes torzitatlan becslések osztályára, hanem a torzitatlan becslések valamely szükebb osztályára vonatkozóan is definiálhatjuk.

Legyen ugyanis $\mathcal{T}$ a $\mathcal{C}$ osztály t paramétere n változós torzitatlan becsléseinek valamely nem üres osztálya. $\mathcal{P} \in \mathcal{C}$ esetén tekintsük a

$$D(\mathcal{P}; \mathcal{T}) = D_n(\mathcal{P}; \mathcal{T}) = \inf_{T \in \mathcal{T}} \left[\int\limits_{R_n} (T(x_1, \dots, x_n) - t(\mathcal{P}))^2 d\mathcal{P}^n \right]^{1/2}$$

függvényt; tegyük fel, hogy $\mathcal{D}(\mathcal{P};\mathcal{T}) < \infty$ $(\mathcal{P}\in\mathcal{C})$.
Ha most $T(x_1,\dots,x_n)\in\mathcal{T}$ olyan becslés, amelyre $\mathcal{D}(T(\xi_1,\dots,\xi_n)) \equiv$

$$\equiv\left[\int_{R_n}(T-t(\mathcal{P}))^2\,d\mathcal{P}^n\right]^{1/2} = \mathcal{D}(\mathcal{P};\mathcal{T}) \qquad (\mathcal{P}\in\mathcal{C})$$

teljesül, akkor azt mondjuk, hogy $T(x_1,\dots,x_n)$ <u>a $\mathcal{C}$ osztály t paraméterének a $\mathcal{T}$ osztályra vonatkozóan hatásos becslése</u>.

Ennek az ujabb fogalomnak az illusztrálására egy további megjegyzést teszünk.

Lineáris közép alatt értünk valamely $\bar{x}^a = \bar{x}_n^a = \sum_{i=1}^{n} a_i x_i$ alaku függvényt, ahol $a_1,\dots,a_n$ adott konstansok, amelyekre $\sum_{i=1}^{n} a_i = 1$.

<u>Bármely lineáris közép a véges várható értékü valószinüségi eloszlások osztálya esetén a várható értéknek torzitatlan becslése</u>.

Valóban, ha $\xi_1,\dots,\xi_n$ statisztikai minta, ahol $m = M(\xi_i)$ $(i=1,\dots,n)$ létezik, és $\bar{x}_n^a$ lineáris közép, akkor

$$M\left(\sum_{i=1}^{n} a_i \xi_i\right) = \sum_{i=1}^{n} a_i M(\xi_i) = m\sum_{i=1}^{n} a_i = m.$$

Érvényes a következő állitás.

<u>Az $\bar{x}_n$ számtani közép a véges szórásu valószinüségi eloszlások osztálya esetén az összes $\bar{x}_n^a$ lineáris közepek osztályára vonatkozóan a várható értéknek hatásos becslése</u>.

Legyen ugyanis $\bar{x}_n^a = \sum_{i=1}^{n} a_i x_i$ lineáris közép, és $\xi_1,\dots,\xi_n$ tetszőleges minta, amely elemeinek közös D szórása véges. Ekkor a Cauchy-egyenlőtlenség alkalmazásával nyerjük, hogy

mul. vált.

$$D^2\Big(\sum_{i=1}^{n} a_i\xi_i\Big)=\sum_{i=1}^{n} a_i^2 D^2(\xi_i)=D^2\sum_{i=1}^{n} a_i^2 \geq \frac{D^2}{n}\Big(\sum_{i=1}^{n} a_i\Big)^2=\frac{D^2}{n}=D^2(\bar{\xi}_n),$$

mul. vált.

amivel az állítást bebizonyitottuk.

Torzitatlan becslésekre vonatkozik az alábbi nevezetes tétel is, amelyet csak speciális esetre fogalmazunk meg.

Cramér-Rao-egyenlőtlenség. Legyen $\mathcal{C}=\{\mathcal{P}_\theta : \theta\in\Theta\}$ R_1-beli folytonos valószinüségi eloszlások egyparaméteres osztálya, ahol Θ véges, vagy végtelen nyitott intervallum; a $\mathcal{P}_\theta$ eloszlás sürüségfüggvényét jelöljük $f(x;\theta)$ -val. Tegyük fel, hogy van olyan $E\ (\subseteq R_1)$ halmaz, hogy minden $\theta\ (\in\Theta)$ -ra $\mathcal{P}_\theta(E)=0$, továbbá bármely $x\notin E$, $\theta\in\Theta$ esetén $f(x;\theta)>0$, és $\dfrac{\partial f(x;\theta)}{\partial\theta}$ létezik. Legyen

$$h(\underline{x};\theta)=h(x_1,\dots,x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta).$$

Tegyük fel, hogy

$$(3)\qquad I_\theta^2=\int_{R_n}\left(\frac{\partial \log h(\underline{x};\theta)}{\partial\theta}\right)^2 h(\underline{x};\theta)\,d\underline{x}<\infty \quad (\theta\in\Theta),$$

és valahányszor $g(\underline{x})=g(x_1,\dots,x_n)$ olyan Borel-mérhető függvény, amelyre

$$(4)\qquad \int_{R_n} g^2(\underline{x})\,h(\underline{x};\theta)\,d\underline{x}<\infty \quad (\theta\in\Theta)$$

Tekintsük a Cramér-Rao-egyenlőtlenségnek egy speciális esetét. Tegyük fel, hogy a $\xi_1,\dots,\xi_n$ statisztikai minta elemei $N(m, D^2)$ -eloszlásuak, ahol $-\infty < m < \infty$ és D rögzitett. Ekkor $f(x;m) = \frac{1}{\sqrt{2\pi}D} e^{-\frac{(x-m)^2}{2D^2}}$, és

$$h(x_1,\dots,x_n;m) = \frac{1}{(2\pi)^{\frac{n}{2}} D^n} e^{-\frac{1}{2D^2}\sum_{i=1}^{n}(x_i-m)^2}.$$

Könnyen belátható, hogy ebben az esetben f -re és h -ra teljesülnek az előző tétel feltevései. Ezért ha $g_0(x_1,\dots,x_n)$ a $\mathcal{C} = \{N(m,D^2) : m \in R_1, D \text{ fix}\}$ osztály m paraméterének torzitatlan becslése, akkor

$$D^2(g_0(\xi_1,\dots,\xi_n)) \geqq 1/I_m^2 \quad (m \in R_1),$$

ahol

$$I_m^2 = \int_{R_n}\left(\frac{\partial \log h(\underline{x};m)}{\partial m}\right)^2 h(\underline{x};m)\,d\underline{x} = M\left(\left(\frac{\partial \log h(\xi_1,\dots,\xi_n;m)}{\partial m}\right)^2\right).$$

Mivel most

$$\frac{\partial \log h(x_1,\dots,x_n;m)}{\partial m} = -\frac{1}{D^2}\sum_{i=1}^{n}(x_i-m),$$

igy a mintaelemek függetlensége miatt

$$M\left(\left(\frac{\partial \log h(\xi_1,\dots,\xi_n;m)}{\partial m}\right)^2\right) = \frac{1}{D^4} M\left(\left(\sum_{i=1}^{n}(\xi_i-m)\right)^2\right) =$$

$$= \frac{1}{D^4} \sum_{i=1}^{n} M\left(\left(\xi_i - m\right)^2\right) = \frac{n}{D^2}.$$

Tehát

$$D^2\left(g_0(\xi_1, \dots, \xi_n)\right) \geq \frac{D^2}{n}.$$

Másrészt tudjuk, hogy

$$D^2(\bar{\xi}_n) = \frac{D^2}{n}.$$

Kaptuk tehát a következő eredményt:

mat. stat.

<u>A számtani közép a normális eloszlások osztálya esetén a várható értéknek hatásos becslése.</u> mat. stat.

II. 4. <u>Elégséges statisztikák.</u> Legyen $\mathcal{C}=\{\mathcal{P}\}$ – R_1 -beli valószinüségi eloszlások valamely nem üres osztálya. Legyen továbbá $\vec{T}(x_1, \dots, x_n) = (T_1(x_1, \dots, x_n), \dots, T_\ell(x_1, \dots, x_n))$ a $\mathcal{C}$ osztályra vonatkozó statisztika. Azt mondjuk, hogy $\vec{T}$ <u>a $\mathcal{C}$ osztály elégséges statisztikája</u>, ha bármely $\mathcal{P}(\in \mathcal{C})$ – eloszlásu sokaságból vett $\xi_1, \dots, \xi_n$ minta esetén a $\xi_1, \dots, \xi_n$ mintaelemeknek a $\hat{\vec{T}} = \vec{T}(\xi_1, \dots, \xi_n)$ valószinüségi vektorváltozóra vonatkozó feltételes eloszlása ugyanaz. Az, hogy $\vec{T}$ a $\mathcal{C}$ osztály elégséges statisztikája, ugy interpretálható, hogy ha egy $\mathcal{P}(\in \mathcal{C})$ -eloszlásu sokaságból vett $\xi_1, \dots, \xi_n$ minta esetén a $\hat{\vec{T}} = \vec{T}(\xi_1, \dots, \xi_n)$ vektort meghatároztuk, akkor a $\xi_1, \dots, \xi_n$ mintaelemek közös eloszlásukra vonatkozóan további információt nem szolgálhatnak.

Az elégséges statisztika fogalma könnyen általánosítható

$$\int_{\mathcal{U}} P((\eta_1,\dots,\eta_n)\in TA \mid \sqrt{n}\,\bar{\xi}_n = x)\frac{1}{\sqrt{2\pi}D}e^{-\frac{(x-\sqrt{n}m)^2}{2D^2}}dx \quad (\mathcal{U}\in\mathcal{B}_1).$$

Az utóbbi három egyenlőségből következik, hogy

$$P((\xi_1,\dots,\xi_n)\in A \mid \sqrt{n}\,\bar{\xi}_n = x) = \chi_{(TA)^*}(x)\,P((\eta_1,\dots,\eta_{n-1})\in (TA)_x) \quad (A\in\mathcal{B}_n).$$

Mivel a jobb oldal független az m várható értéktől, ezzel az állítást bebizonyitottuk.

II. 5. Maximum-likelihood módszer. Az előző paragrafusokban több példát mutattunk be arra, hogy a mintaelemek eloszlásának egyes jellemzőit milyen függvényekkel célszerü becsülni. Természetes módon felmerül a kérdés, hogy egy tetszőleges jellemző paraméterre hogyan adhatunk becslést. Általában használható eljárást ad az un. maximum-likelihood módszer, amelyet itt a két legfontosabb speciális esetben tárgyalunk.

Diszkrét eset. Legyen $\mathcal{C}=\{\mathcal{P}_{\vec{t}}\,;\ \vec{t}\in\mathcal{T}\}$ $(\vec{t}=(t_1,\dots,t_\ell))$ $\mathcal{R}_1$-beli diszkrét valószinüségi eloszlások ℓ-paraméteres nemüres osztálya. Tegyük fel, hogy van olyan megszámlálható X halmaz, hogy $\mathcal{P}_{\vec{t}}(X)=1$ $(\vec{t}\in\mathcal{T})$. Legyen

$$\mathcal{P}_{\vec{t}}(\{x\}) = p(x,t_1,\dots,t_\ell) \qquad (x\in X,\ \vec{t}\in\mathcal{T}).$$

Az

$$L(x_1,\dots,x_n,t_1,\dots,t_\ell)=\prod_{i=1}^{n} p(x_i,t_1,\dots,t_\ell)$$

függvényt likelihood-függvénynek nevezzük.

Tegyük fel, hogy minden $x_1,\dots,x_n$ $(x_i\in X,\ i=1,\dots,n)$

esetén van egyetlen olyan $\vec{T} = (T_1, \dots, T_\ell) \in \mathcal{T}$ $(T_i = T_i(x_1, \dots, x_n),$ $i = 1, \dots, \ell)$ vektor, amelyre

$$L(x_1, \dots, x_n, T_1, \dots, T_\ell) = \max_{(t_1, \dots, t_\ell) \in \mathcal{T}} L(x_1, \dots, x_n, t_1, \dots, t_\ell).$$

Ekkor a $\mathcal{C}$ osztály $\vec{t}$ paraméterének maximum-likelihood becslésén a $\vec{T}(x_1, \dots, x_n) = (T_1(x_1, \dots, x_n), \dots, T_\ell(x_1, \dots, x_n))$ vektorfüggvényt értjük. A definició alapján nyilvánvaló, hogy ha a $\xi_1, \dots, \xi_n$ $\mathcal{C}$-beli eloszlásu sokaságból vett minta esetén a $(\xi_1 = x_1) \cap \dots \cap (\xi_n = x_n)$ esemény következik be, akkor $\vec{T}(x_1, \dots, x_n)$ éppen az a paraméterérték, amely mellett a $(\xi_1 = x_1) \cap \dots \cap (\xi_n = x_n)$ esemény valószinüsége maximális.

Ha bármely $x_1, \dots, x_n$ $(x_i \in X, i = 1, \dots, n)$, $\vec{t} \in \mathcal{T}$ esetén $L(x_1, \dots, x_n, t_1, \dots, t_\ell) > 0$, továbbá $\mathcal{T}$ R_ℓ-beli nyitott halmaz, és bármely $x_1, \dots, x_n$ $(x_i \in X, i = 1, \dots, n)$, $\vec{t} \in \mathcal{T}$ esetén a $\dfrac{\partial L(x_1, \dots, x_n, t_1, \dots, t_\ell)}{\partial t_j}$ $(j = 1, \dots, \ell)$ parciális differenciálhányadosok léteznek, akkor $\vec{T}$ meghatározása a

$$\frac{\partial \log L(x_1, \dots, x_n, t_1, \dots, t_\ell)}{\partial t_j} = 0 \quad (j = 1, \dots, \ell)$$

un. likelihood-egyenletek megoldásával könnyebben történhet. A

$$\log L(x_1, \dots, x_n, t_1, \dots, t_\ell) = \sum_{i=1}^{n} \log p(x_i, t_1, \dots, t_\ell)$$

függvényt logaritmus-likelihood-függvénynek nevezzük.

Tekintsük a következő példát. Legyen $\mathcal{C}$ a Poisson-eloszlások osztálya, ahol $t = \lambda \ (>0)$ az eloszlás várható értéke. Most

$$p(x;\lambda) = \frac{\lambda^x}{x!} e^{-\lambda} \qquad (x = 0, 1, \ldots),$$

és

$$L(x_1, \ldots, x_n; \lambda) = \frac{\lambda^{x_1 + \ldots + x_n}}{x_1! \ldots x_n!} e^{-n\lambda}.$$

Továbbá

$$\log L(x_1, \ldots, x_n; \lambda) = \log \lambda \sum_{i=1}^{n} x_i - n\lambda - \sum_{i=1}^{n} \log(x_i!) .$$

Igy a likelihood-egyenlet

$$\frac{\partial \log L(x_1, \ldots, x_n; \lambda)}{\partial \lambda} = \frac{1}{\lambda} \sum_{i=1}^{n} x_i - n = 0 .$$

Ebből $\sum_{i=1}^{n} x_i > 0$ esetén az egyenlet megoldás

$$T = T(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i .$$

Mivel

$$\left. \frac{\partial^2 \log L(x_1, \ldots, x_n; \lambda)}{\partial \lambda^2} \right|_{\lambda = T} = - \frac{1}{\lambda^2} \sum_{i=1}^{n} x_i \Big|_{\lambda = T} = - \frac{n^2}{\sum_{i=1}^{n} x_i} < 0,$$

igy $\sum_{i=1}^{n} x_i > 0$ esetén $L(x_1, \ldots, x_n; \lambda)$ az $\frac{1}{n} \sum_{i=1}^{n} x_i$ paraméterérték mellett maximális. Ha viszont $\sum_{i=1}^{n} x_i = 0$, akkor $L(x_1, \ldots, x_n; \lambda)$ a $0 \left(= \frac{1}{n} \sum_{i=1}^{n} x_i\right)$ paraméterérték mellett maximális. Kaptuk tehát, hogy <u>a Poisson-eloszlások osztálya esetén a várható érték maximum-likelihood becslése a számtani közép</u>.

<u>Folytonos eset.</u> Legyen $\mathcal{C} = \{\mathcal{P}_{\vec{t}} : \vec{t} \in \mathcal{P}\}$ $(\vec{t} = (t_1, \dots, t_\ell))$ R_1-beli folytonos valószínüségi eloszlások ℓ-paraméteres, nem üres osztálya; $f(x; t_1, \dots, t_\ell)$ jelölje a $\mathcal{P}_{\vec{t}}$ eloszlás sürüségfüggvényét. Ebben az esetben a likelihood-függvény

$$L(x_1, \dots, x_n; t_1, \dots, t_\ell) = \prod_{i=1}^{n} f(x_i; t_1, \dots, t_\ell) \, ;$$

$\log L(x_1, \dots, x_n, t_1, \dots, t_\ell)$ pedig a <u>logaritmus-likelihood-függvény</u>.

Tegyük fel, hogy van olyan $E \ (\subseteq R_n)$ halmaz, amelyre $\mathcal{P}_{\vec{t}}^{\,n}(E) = 0 \ (\vec{t} \in \mathcal{T})$ teljesül, és $(x_1, \dots, x_n) \in R_n - E$ esetén van egyetlen olyan

$$\vec{T} = (T_1, \dots, T_\ell) \in \mathcal{T} \quad (T_j = T_j(x_1, \dots, x_n); \ j = 1, \dots, \ell)$$

vektor, hogy

$$L(x_1, \dots, x_n, T_1, \dots, T_\ell) = \max_{(t_1, \dots, t_\ell) \in \mathcal{T}} L(x_1, \dots, x_n, t_1, \dots, t_\ell).$$

Ha a $T_j(x_1, \dots, x_n)$ függvények Borel-mérhetők, akkor <u>a $\mathcal{C}$ osztály $\vec{t}$ paraméterének maximum-likelihood becslésén a</u> $\vec{T}(x_1, \dots, x_n) = (T_1(x_1, \dots, x_n), \dots, T_\ell(x_1, \dots, x_n))$ vektorfüggvényt értjük.

Ha bármely $(x_1, \dots, x_n) \in R_n$, $\vec{t} \in \mathcal{T}$ esetén $L(x_1, \dots, x_n, t_1, \dots, t_\ell) > 0$, továbbá $\mathcal{T}$ R_ℓ-beli nyitott halmaz, és bármely $\vec{t} \in \mathcal{T}$, $(x_1, \dots, x_n) \in R_n$ esetén a $\dfrac{\partial L(x_1, \dots, x_n, t_1, \dots, t_\ell)}{\partial t_j}$ $(j = 1, \dots, \ell)$ parciál differenciálhányadosok léteznek, akkor $\vec{T}$ meghatározása a

$$\frac{\partial \log L(x_1, \dots, x_n, t_1, \dots, t_\ell)}{\partial t_j} = 0 \quad (j = 1, \dots, \ell)$$

likelihood-egyenletek megoldásával könnyebben történhet.

A folytonos esetre példaként tekintsük az R_1 -beli normális eloszlások osztályát. Ekkor $t_1 = m$ a várható érték, és $t_2 = D^2$ a szórásnégyzet $(m \in R_1,\ D^2 > 0)$. A likelihood-függvény most

$$L(x_1,\dots,x_n, m, D^2) = \frac{1}{(2\pi)^{\frac{n}{2}} D^n} e^{-\frac{1}{2D^2}\sum_{i=1}^{n}(x_i - m)^2},$$

és

$$\log L(x_1,\dots,x_n, m, D^2) = -\frac{1}{2D^2}\sum_{i=1}^{n}(x_i - m)^2 - \frac{n}{2}\log D^2 - \frac{n}{2}\log 2\pi.$$

A likelihood-egyenletek:

$$\frac{\partial \log L}{\partial m} = +\frac{1}{D^2}\sum_{i=1}^{n}(x_i - m) = 0, \quad \frac{\partial \log L}{\partial D^2} = -\frac{n}{2}\frac{1}{D^2} + \frac{1}{2D^4}\sum_{i=1}^{n}(x_i - m)^2 = 0.$$

$\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 \neq 0$ esetén az egyértelműen meghatározott megoldások

$$T_1 = T_1(x_1,\dots,x_n) = \bar{x}_n, \quad T_2 = T_2(x_1,\dots,x_n) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2.$$

A likelihood-függvény alakjából következik, hogy ilyen paraméterértékek esetén a likelihood-függvénynek maximuma van. Ha pedig $\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 = 0$, akkor a likelihood-egyenleteknek nincs megoldása, sőt a likelihood-becslés sem értelmezhető, azonban az ilyen $(x_1,\dots,x_n)$ pontok halmazának Lebesgue-mértéke 0 -val egyenlő. Kaptuk tehát, hogy a normális eloszlások osztálya (m, D^2) paraméterének maximum-likelihood becslése az $(\overline{x}, s^2)$ vektorfüggvény.

A maximum-likelihood becslés fogalma kézenfekvő módon általánositható k-dimenziós eloszlások esetére.

Megmutatható, hogy az eloszlásosztály eloszlásaira tett elég általános feltételek mellett a maximum-likelihood-becslés konzisztens lesz, és az illető osztály elégséges statisztikáját szolgáltatja. A torzitatlanság következménye azonban nem mindig teljesül.

II. 6. Momentum-módszer. Egy további általános becslési módszer az un. momentum-módszer. Legyen $\mathcal{C} = \{\mathcal{P}_{\vec{t}} : \vec{t} \in \mathcal{T}\}$ $(\vec{t} = (t_1, \dots, t_\ell))$ R_1-beli valószinüségi eloszlások ℓ paraméteres, nem üres osztálya. Tegyük fel, hogy az

$$m_j(\mathcal{P}_{\vec{t}}) = \int_{-\infty}^{\infty} x^j \, d\mathcal{P}_{\vec{t}} \qquad (j = 1, \dots, \ell)$$

momentumok minden $\vec{t} \in \mathcal{T}$ esetén léteznek, és a $\mathcal{T}$-n értelmezett

$$m_j(\mathcal{P}_{\vec{t}}) = m_j(t_1, \dots, t_\ell) \qquad (j = 1, \dots, \ell)$$

függvények ismeretesek. Tegyük fel továbbá, hogy van olyan $E \ (\subseteq R_n)$ halmaz, hogy $\mathcal{P}^n_{\vec{t}}(E) = 0$ $(\vec{t} \in \mathcal{T})$, és ha $(x_1, \dots, x_n) \notin E$, akkor az

$$m_j(t_1, \dots, t_\ell) = \frac{1}{n} \sum_{i=1}^{n} x_i^j \qquad (j = 1, \dots, \ell)$$

egyenletrendszernek egyetlen

$$\vec{T}_i(x_1, \dots, x_n) = (T_1(x_1, \dots, x_n), \dots, T_\ell(x_1, \dots, x_n))$$