2017. 02. 15.

Def. Statisztikai mező: $(\Omega, \mathcal{A}, P)$, ahol $(\Omega, \mathcal{A})$ mérhető tér, P valószínűségi eloszlások egy családja, gyakran $P = \{P_\vartheta : \vartheta \in \Theta\}$, ahol Θ a paramétertér, és $\vartheta_1 \neq \vartheta_2 \Rightarrow P_{\vartheta_1} \neq P_{\vartheta_2}$ (identifikálhatóság).

Cél: meghatározni az „igazi" eloszlást, ill. a hozzá tartozó paramétert, esetleg valamilyen függvényét.

Def. Statisztikai minta: $X : (\Omega, \mathcal{A}) \to (\mathcal{X}, \mathcal{B})$ val. változó

$\longrightarrow$ spec. eset: $X = (X_1, \dots, X_n)$, ahol X_i ftlen azonos eo. v.v.

Ekkor X n elemű (független) minta.

Ez kipróbálás: elvégzünk egy kísérletet n-szer (mintavétel), és ebből szeretnénk következtetni.

De nem csak ilyen létezik, pl. sztoch. folyamatok.

Legyen $Q_\vartheta = P_\vartheta \circ X^{-1}$ az X eloszlása $(\mathcal{X}, \mathcal{B})$-n, ha a paraméter épp ϑ. Ezzel $(\mathcal{X}, \mathcal{B}, Q = \{Q_\vartheta : \vartheta \in \Theta\})$ is stat. mező.

$(\mathcal{X}, \mathcal{B})$ neve a mintatér. $\longrightarrow$ Ennek alapján csak a Q_ϑ eloszlásokra lehet köv.

A továbbiakban csak a stat mezőn dolgozunk.

Def. Statisztika: $T : (\mathcal{X}, \mathcal{B}) \to (\mathcal{Y}, \mathcal{C})$, a minta egy függvénye

Gyakran a $T(X) = T \circ X : \Omega \to \mathcal{Y}$ fv.-t nevezzük statisztikának.

A rétfele értelmezés nem okoz félreértést.

$Q(B) = P(X \in B)$. A mintavételt n-szer megismételjük:

n elemű ftlen mintát kapunk: $X_1, \dots, X_n$

A relatív gyakoriság közelítést ad:

$$Q_n^*(B) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \in B)$$

NSZET: $Q_n^*(B) \longrightarrow Q(B)$ 1 val.

CHT: $\sqrt{n}\left(Q_n^*(B) - Q(B)\right) \xrightarrow{d} N(0, P(B)\cdot(1-P(B)))$,

ahol $\xrightarrow{d}$ az eloszlásbeli konvergencia, N pedig szórásnégyzettel (és nem szórással) paraméterezett, azaz a konvergencia $\sqrt{n}$ nagyságrendű.

$\mathcal{B} \ni B \longmapsto Q_n^*(B)$ véletlen val. mértéket, neve tapasztalati eloszlás.

A tap. eo. diszkrét: $\frac{1}{n}$ súlyokat tesz az egyes $X_1, \dots, X_n$ megfigyelésekre.

Formálisan: $Q_n^* = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$, ahol δ_c a c-re koncentrált elfajult eloszlás (Dirac-mértéket)

Legyen φ az eloszlásokon értelmezett fv.

Becsülni szeretnénk a $\varphi(Q)$ értéket.

Ha φ értelmezhető a Q_n^* tapasztalati eloszláson, akkor $\varphi(Q)$ tapasztalati becslése $\hat{\varphi}(X) = \varphi(Q_n^*)$

<u>Példa.</u> Egy- vagy többdimenziós eloszlások (Borelben)

$\varphi :=$ várható érték

Tap. becslés: $\varphi(Q_n^*) = \frac{X_1 + \dots + X_n}{n} \overset{\text{jel}}{=} \overline{X}$ mintaátlag.

<u>Példa.</u> 1-dim., $\varphi :=$ szórásnégyzet.

Tap. becslés: $s_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2 =$

tap. szórásnégyzet

$= \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \overline{X}^2$

Ez minden eo.-ra igaz tf., ezért tap. eo.-ra is, ezért nem kell külön bizonyítani.

<u>Megj.</u> Korrigált tap. szórásnégyzet: $s_n^{*2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 = \frac{n}{n-1} s_n^2$,

ez nyilvánvaló torzítatlan, míg s_n^2 nem. (majd)

<u>Példa.</u> Több dimenzióban a kovarianciamátrix:

$$S_n = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})(X_i - \overline{X})^T, \quad \text{tap. kovarianciamátrix.}$$

ahol $X_1, \dots, X_n$ k-dim vektorok, S_n $k \times k$-as

<u>Példa.</u>

Legyen $\underline{X} = (X_1, \dots, X_n)$ n elemű ftlen valós értékű minta.

$\underline{X} \longmapsto \underline{X}^* = (X_1^*, \dots, X_n^*)$ rendezett minta: $X_1^* \le \dots \le X_n^*$.

Ez is egy statisztikát ad.

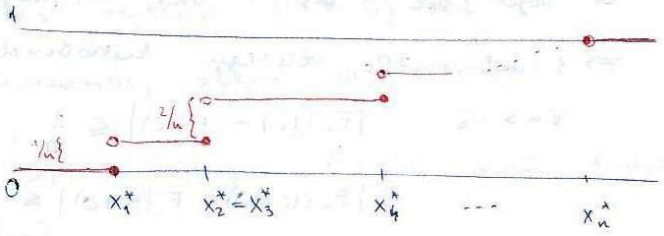
A tap. eo. tulajdonsága $\underline{X}^*$ fvénye: $Q_n^*(B) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \in B) = \frac{1}{n} \sum_{i=1}^{n} I(X_i^* \in B)$

Így vissza az indicátorokra már nem az eo. ftlenek!

Példa. φ az eofv.: $Q\ (\mathbb{Q}) = F(\cdot)$, $F(t) = P(X_1 < t) = Q((-\infty, t))$

Tap. eofv.: $F_n(t) = Q_n^*((-\infty, t)) = \frac{1}{n} \sum_{i=1}^{n} I(X_i < t)$

Ez lépcsős:



$$F_n(t) = \begin{cases} 0 & t \le x_1^* \\ i/n & x_i^* \le t \le x_{i+1}^* \qquad i = 1, 2, \dots, n-1 \\ 1 & x_n^* < t \end{cases}$$

(Ez sűrűségfüggvény nem megy: a tap. eo. diszkrét, és ezért nincs sűrűsége.)

$\forall t$: $F_n(t) \to F(t)$ 1 val. (itt a 0-mértékű halmaz t-függő).

Igaz, de erősebb és nem nyilvánvaló, hogy F_n pontonként tart 1 val.: 1 val $\forall t$ $F_n(t) \to F(t)$.

Glivenko–Cantelli-tétel (A statisztika alaptétele)

$\Delta_n := \sup_t |F_n(t) - F(t)|$.

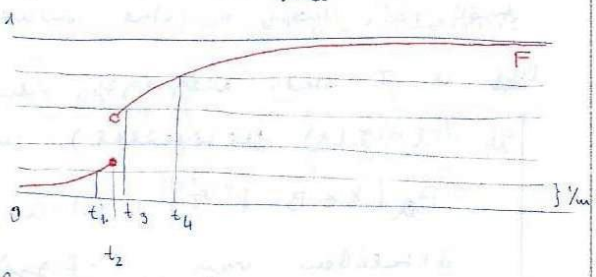
Ekkor $\Delta_n \to 0$ 1 val, azaz a tap. eofv. 1 val. tart az igazihoz.

megj. Ez azt fejezi ki, hogy a világ megismerhető: elég sok megfigyeléssel a jelenség leírható; tetszőlegesen közel lehet kerülni a megismeréshez.

megj. Δ_n val. változó: a balfolytonosság miatt $\Delta_n = \sup_{t \in \mathbb{Q}} |F_n(t) - F(t)|$.

B: Legyen $\varepsilon > 0$.

Legyen $m > \frac{2}{\varepsilon}$, és osszuk $[0,1]$-et m egyenlő részre.



$t_1 \le \dots \le t_{m-1}$: $F(t_i) \le \frac{i}{m} \le F(t_i + 0)$,

ilyenek léteznek, pl.

$$t_i := \sup\left\{t : F(t) \le \frac{i}{m}\right\} = \max\left\{t : F(t) \le \frac{i}{m}\right\}$$

NSzET: $F_n(t_i) \to F(t_i)$ 1 val. $\forall i = 1, \dots, m-1$

$$F_n(t_i) = Q_n^*((-\infty, t_i]) \longrightarrow Q((-\infty, t_i]) = F(t_i + 0)$$

Ez véges sok $\Rightarrow$ 1 val. egyszerre teljesülnek.

$\Rightarrow$ 1 val. $\exists n_0$ véletlen küszöbindex, hogy

$$\forall n > n_0: \quad |F_n(t_i) - F(t_i)| \le \frac{1}{m}, \qquad \forall i = 1, \dots, m-1$$

$$|F_n(t_i + 0) - F(t_i + 0)| \le \frac{1}{m}$$

Ezen az m eseményen, ha $n > n_0$:

- legyen $t_i < t \le t_{i+1} \Rightarrow F_n(t) - F(t) \le F_n(t_{i+1}) - F(t_i + 0) \le$

$$\le F(t_{i+1}) + \frac{1}{m} - F(t_i + 0) \le \frac{i+1}{m} + \frac{1}{m} - \frac{i}{m} = \frac{2}{m} < \varepsilon$$

analóg módon: $F(t) - F_n(t) \le F(t_{i+1}) - F_n(t_i + 0) \le$

$$\le F(t_{i+1}) - F(t_i + 0) + \frac{1}{m} \le \frac{i+1}{m} - \frac{i}{m} + \frac{1}{m} = \frac{2}{m} < \varepsilon$$

- ha $t \le t_1$: $F_n(t) - F(t) \le F_n(t_1) \le F(t_1) + \frac{1}{m} \le \frac{2}{m} < \varepsilon$

$$F(t) - F_n(t) \le F(t_1) \le \frac{1}{m} < \frac{2}{m} < \varepsilon$$

- ha $t > t_{m-1}$: $F_n(t) - F(t) \le 1 - F(t_{m-1} + 0) \le 1 - \frac{m-1}{m} = \frac{1}{m} < \varepsilon$

$$F(t) - F_n(t) \le 1 - F_n(t_{m-1} + 0) \le 1 - F(t_{m-1} + 0) + \frac{1}{m} \le$$

$$\le \frac{2}{m} < \varepsilon.$$

$\Rightarrow \Delta_n \le \frac{2}{m} < \varepsilon.$ $\square$

A tétel igaz magasabb dimenzióban is, a bizonyítás nem megy direktben át.

$\sup_B |Q_n^*(B) - Q(B)|$-vel csinálni? Itt nem triv, hogy milyen B-ket nézzük; mindre biztos nem fog 0-hoz tartani.

Elégségesség

Arról szól, hogy a stat. minden értéket tartalmazzon.

Def. A T stat. elégséges, ha az X minta felt. eloszlása ($T = T(X)$ feltétellel) már nem függ ϑ-tól.

$P_{\vartheta}(X \in B \mid T)$ jelölés: az átlúszás jelzi, hogy általában van ϑ-függés, de itt nincs.

Sir Ronald Fisher (1890–1962), 1922

Ebben a definícióban a reguláris feltételes eloszlást értjük: ez $Q(B,\omega)$ függvény, $\forall \underline{\omega}$: $B \mapsto Q(B,\omega)$ val eloszlás és $\forall \underline{B}$: $Q(B,\cdot)$ feltételes valósége $\{X \in B\}$-nek.

Doob: ha X véges dimenziós, akkor ez létezik.

A def úgy értendő, hogy $\underline{\underline{\exists}}$ ilyen, ϑ-tól nem függő verzió; ez nem egyértelmű.

Regularitás $\Rightarrow$ $E_{\vartheta}(S(X)|T)(\cdot) = \int_{\mathcal{X}} S(x)\, Q_{\vartheta}(dx, \cdot)$

<u>Áll.</u> T elégséges $(\mathcal{X}, \mathcal{B}, \mathcal{P})$-n és $\mathcal{P}' \subset \mathcal{P}$ $\Rightarrow$ T elégséges $(\mathcal{X}, \mathcal{B}, \mathcal{P}')$-n is. 2017.02.2

<u>Áll.</u> Legyen $\mathcal{X} = \mathbb{R}^n$, $\mathcal{B}$ = Borel, $\mathcal{P} = \{\underbrace{P \times \dots \times P}_{n},\ P \text{ n-dim. val. eo.}\}$ (n elemű minta összes lehetséges eloszlása).

Ekkor az $\underline{X}^* = (X_1^*, \dots, X_n^*)$ rendezett minta elégséges.

B: Legyen $\pi \in S_n$ permutáció, és jelölje $\pi : \mathbb{R}^n \to \mathbb{R}^n$ $x = (x_1, \dots, x_n) \mapsto$ $\mapsto \pi x = (x_{\pi(1)}, \dots, x_{\pi(n)})$ transzformációt.

Ezek transzformációcsoportot alkotnak, $z \in \mathbb{R}^n$ orbitja $\langle z \rangle = \{\pi z \mid \pi \in S^n\}$

A $\sigma z : \sigma \in S_n$ felsorolásban $\langle z \rangle$ minden eleme ugyanannyiszor szerepel (ez a szám attól függ, hogy z-nek hány azonos koordinátája van). Ui. ha $z_1, z_2 \in \langle z \rangle$, akkor $\exists \tau \in S_n$, hogy $z_2 = \tau z_1$, mert az orbitok particionálnak. Ezért $\sigma z = z_1 \iff \tau\sigma z = z_2$, tehát a z-t z_1-be ill. z_2-be vivő trafók között $\sigma \longleftrightarrow \tau\sigma$ megfeleltetés.

Számítsuk ki $E(f(X) \mid X^* = z)$-t! Remélhetőleg nem ez P-ftben.

$$\int_{\{X^* \in B\}} f(X)\, dP \underset{\substack{X^* \overset{d}{=} \sigma X \\ \text{(azonos eo.)}}}{=} \int_{\{(\sigma X)^* \in B\}} f(\sigma X)\, dP = \int_{\{X^* \in B\}} f(\sigma X)\, dP \quad \forall \sigma$$

mindegy, hogy permutáljuk-e, ha utána úgyis rendezzük

$$\overset{\text{átlagolás}}{=} \int_{\{X^* \in B\}} \frac{1}{n!} \sum_{\sigma} f(\sigma X)\, dP = \int_{\{X^* \in B\}} \frac{1}{n!} \sum_{\sigma} f(\sigma X^*)\, dP$$

$$\Rightarrow E(f(X) \mid X^* = z) = \frac{1}{n!} \sum_{\sigma} f(\sigma z) = \frac{1}{|\langle z \rangle|} \sum_{x \in \langle z \rangle} f(x)$$ nem függ az X_i megfigyelések eloszlásától $\to$ X^* elégséges stat.,

behelyettesítve f = indikátort. $\square$

Megj. X feltételes eloszlása X^*-ra nézve: $\langle X^* \rangle$-on egyenletes, azaz minden sorszám egyformán valószínű.

Dominált mértékosztályok

$P, Q, \dots$ val. mértékek

$\mathcal{P}, \mathcal{Q}, \dots$ eloszláscsaládok (mértékosztályok)

Def. $\mathcal{P}$ dominált, ha $\exists \lambda$ σ-véges, hogy $P \ll \lambda$ $\forall P \in \mathcal{P}$

Ekkor $\mathcal{P}$ megadható a $dP/d\lambda$ sűrűségfüggvénnyel is.

Spec. esetek: • λ = Lebesgue-mérték $\mathbb{R}^n$-ben $\Rightarrow f_\vartheta = \frac{dP_\vartheta}{d\lambda}$ a hagyományos stb.

• $\mathcal{X}$ megszámlálható, λ = számlálómérték

$\Rightarrow f_\vartheta(x) = P_\vartheta(\{x\})$ a valószínűség

Def. $B \in \mathcal{B}$ $\mathcal{P}$-0-halmaz, ha $P(B) = 0$.

Def. $B \in \mathcal{B}$ $\mathcal{P}$-0-halmaz, ha $P(B) = 0$ $\forall P \in \mathcal{P}$.

Def. $P \sim Q \Leftrightarrow P \ll Q,\ Q \ll P \Leftrightarrow$ (B P-0-h. $\Leftrightarrow$ Q-0-h.) (ekvivalencia)

Def. $\mathcal{P} \sim \mathcal{Q}$ (ekvivalencia): a $\mathcal{P}$-0-halmazok megegyeznek a $\mathcal{Q}$-0-halmazokkal.

Tétel: (Halmos – Savage, BN) Ekvivalensek:

(1) $\mathcal{P}$ dominált

(2) $\exists \mathcal{P}_0 \subset \mathcal{P}$, $\mathcal{P}_0$ megszámlálható, $\mathcal{P} \sim \mathcal{P}_0$

(3) $\exists P_1, P_2, \dots \in \mathcal{P}$, $\exists c_1, c_2, \dots \in \mathbb{R}$, $\sum c_i = 1$, $\forall c_i > 0$,

és $P_0 = \sum_{i=1}^{\infty} c_i P_i \sim \mathcal{P}$ (kikevert)

Megj: P_0 maga is val. mérték, később az ekvivalens val. mértékeket hivatkozni vele.

Tétel. (Neyman faktorizációs tétele)

Dominált statisztikai mezőn ekvivalensek:

(1) T elégséges statisztika

(2) az ekvivalens valószínűségi mértékre vonatkozó sűrűségfüggvényekre

$f_\vartheta(x) = g_\vartheta(T(x))$ ($g_\vartheta \geq 0$, mérhető)

(3) tetszőleges domináló mértékre nézve

$f_\vartheta(x) = h(x) \cdot g_\vartheta(T(x))$ ($h \geq 0$)

B: (1) $\Rightarrow$ (2) $B \in \mathcal{B}$

$\exists P_\vartheta(B \mid T)$ közös verzió

$P_0 = \sum_{n=1}^{\infty} c_n P_n$. A közös verzió jó lesz $P_0(B|T)$-nek is.

Tudjuk, hogy $A \in \sigma(T) \Rightarrow \int_A P_{\cancel{\vartheta}}(B|T)\, dP_\vartheta = P_\vartheta(A \cap B) \quad \forall \vartheta$

$$\Rightarrow \int_A P_{\cancel{\vartheta}}(B|T)\, dP_0 = P_0(A \cap B)$$

$\dfrac{dP_\vartheta|_{\sigma(T)}}{dP_0|_{\sigma(T)}}$ a RND. $\sigma(T)$-mérhető és $= g_\vartheta(T)$, ahol g_ϑ a Doob-lemma miatt létezik.

$P_\vartheta \ll P_0 \Rightarrow P_\vartheta|_{\sigma(T)} \ll P_0|_{\sigma(T)}$

Megmutatjuk, hogy $\dfrac{dP_\vartheta}{dP_0} = g_\vartheta(T)$ is teljesül.

Legyen ui. $A \in \mathcal{B}$.

$$\int_A g_\vartheta(T)\, dP_0 = \int_{\mathcal{X}} I_A \cdot g_\vartheta(T)\, dP_0 \overset{\text{TVÉT}}{=} \int_{\mathcal{X}} E_{\cancel{\vartheta}}(I_A \cdot g_\vartheta(T) \mid T)\, dP_0 =$$

T-mérhető

$$= \int_{\mathcal{X}} E_{\cancel{\vartheta}}(I_A\, g_\vartheta(T) \mid T)\, dP_0|_{\sigma(T)} =$$

$$= \int_{\mathcal{X}} E_{\cancel{\vartheta}}(I_A|T) \cdot \underbrace{g_\vartheta(T)\, dP_0|_{\sigma(T)}}_{dP_\vartheta|_{\sigma(T)}} =$$

„még egy lépés a szájbaszájas sorozatban"

$$= \int_{\mathcal{X}} E_{\cancel{\vartheta}}(I_A|T)\, dP_\vartheta = \int_{\mathcal{X}} I_A\, dP_\vartheta = P_\vartheta(A)$$

✓

(2) $\Rightarrow$ (1), $P_0(B|T)$ jó lesz közös verziónak: elég az kell, hogy $\forall A \in \sigma(T)\ \forall \vartheta \in \Theta$: $\int_A P_0(B|T)\, dP_\vartheta = P_\vartheta(A \cap B)$.

$$\int_A P_0(B|T)\, dP_\vartheta = \int_A E_0(I_B|T)\, \underbrace{g_\vartheta(T)}_{\text{mérh.}}\, dP_0 =$$

$$= \int_A E_0(I_B\, g_\vartheta(T) \mid T)\, dP_0 =$$

$$= \int_A g_\vartheta(T)\, I_B\, dP_0 = \int_{A \cap B} g_\vartheta(T)\, dP_0 = P_\vartheta(A \cap B)$$

✓

(2) $\Rightarrow$ (3) Láncszabály: $\lambda \ll \mu \ll \nu \Rightarrow \dfrac{d\lambda}{d\nu} = \dfrac{d\lambda}{d\mu} \cdot \dfrac{d\mu}{d\nu}$

Emiatt $P \ll \lambda \Rightarrow P_0 \ll \lambda \Rightarrow f_\vartheta = \dfrac{dP_\vartheta}{d\lambda} = \dfrac{dP_\vartheta}{dP_0} \cdot \dfrac{dP_0}{d\lambda} = g_\vartheta(T) \cdot h$ ✓

(3) $\Rightarrow$ (2): $P_0 = \sum c_n P_n \Rightarrow \dfrac{dP_\vartheta}{dP_0}(x) = \dfrac{dP_\vartheta/d\lambda}{dP_0/d\lambda}(x) = \dfrac{f_\vartheta(x)}{\sum c_n f_n(x)} \overset{(3)}{=} \dfrac{g_\vartheta(T(x)) \cdot \cancel{h(x)}}{\sum c_n g_n(T(x)) \cdot \cancel{h(x)}} =$

$= \tilde{g}_\vartheta(T(x))$. ✓

ahol h 0, az P_0- és P-nullmértékű.

□

Nem tipikus, hogy elégséges stat kis dimenziós, bár a neveretesellne ez igaz, de azok mind egy spec. családhoz tartoznak.

Példa. Eltolásparaméteres Cauchy: $\frac{1}{\pi((x-\vartheta)^2+1)}$, ezt nehéz farktorálni.

Def. T minimális elégséges stat., ha elégséges és $\forall S$ elégséges statra $\exists g: T = g(S)$, azaz T tömörítése S-nek. (P-m.m.)

Áll. (BN) $\mathcal{X} = \mathbb{R}^n$

(1) $\left(\lambda\text{-m.m. } x,y: \quad T(x) = T(y) \Rightarrow \frac{f_\vartheta(x)}{f_\vartheta(y)} \text{ nem függ } \vartheta\text{-tól}\right) \Leftrightarrow T$ egs.

(2) $\left(\lambda\text{.m.m. } x,y: \quad T(x) = T(y) \Leftrightarrow \frac{f_\vartheta(x)}{f_\vartheta(y)} \text{ nem függ } \vartheta\text{-tól.}\right) \Rightarrow T$ egs.

Fischer-féle információs mátrix (Fischer-információ) 1925

Def. Likelihood-függvény $\vartheta \mapsto f_\vartheta(X)$ véletlen fv.

Def. Log-likelihood-fv. $\log f_\vartheta(X) = l(\vartheta)$

Ha a minta n elemű ftem, akkor az együttes eloszlás log-likelihoodja $L(\vartheta)$, és X_i log-likelihoodja $l_i(\vartheta)$

Ha X_i sűrűsége λ-ra $f \Rightarrow \underline{X} = (X_1, \dots, X_n)$ $\lambda^n = \lambda \times \dots \times \lambda$-ra vett sűrűsége $\prod_{i=1}^n f(x_i)$,

$\Rightarrow L(\vartheta) = \sum_{i=1}^n l_i(\vartheta)$, az $l_i(\vartheta)$ val. változók ftem azonos eloszlásúak.

mert ftemek sűrűségei

mert az X_i-k azonos eloszlásúak

Jelölés. $\partial = \partial_\vartheta$ a ϑ szerinti parc. deriválás.

Tfh. $\exists \partial l(\vartheta)$, ahol $\Theta \subset \mathbb{R}^p$ konvex nyílt.

Ekkor $\partial l(\vartheta)$ p-dim. sorvektor.

Def. Fischer-int.: $\underset{\in \mathbb{R}^p}{I(\vartheta)} = E_\vartheta\left(\partial l(\vartheta)^T \cdot \partial l(\vartheta)\right)$, ha értelmes.

Tulajdonságok.

$I(\vartheta)$ szimmetrikus, pozitív szemidefinit (mert $a \in \mathbb{R}^p$-re

$$a^T I(\vartheta) \cdot a = E_\vartheta\left(a^T \partial l(\vartheta)^T \partial l(\vartheta) \cdot a\right) = E_\vartheta\left((\partial l(\vartheta) \cdot a)^2\right) \geq 0)$$

$$I(\vartheta) = \int \frac{(\partial f_\vartheta)^T \cdot \partial f_\vartheta}{f_\vartheta^2} \cdot f_\vartheta \cdot d\lambda = 4 \int \left(\partial \sqrt{f_\vartheta}\right)^T \cdot \partial \sqrt{f_\vartheta} \, d\lambda$$

$$(\sqrt{f})' = \frac{f'}{2\sqrt{f}}$$

Gyenge regularitási feltételek (R)

① $\vartheta \longmapsto \sqrt{f_\vartheta(x)}$ folyt diff λ-m.m. x-re

② $I(\vartheta)$ létezik, folytonos ϑ-ban és invertálható

Tétel (BN) Tfh. a gy. r. felt. teljesülnek, azaz (R). Legyen $T: \mathcal{X} \to \mathbb{R}^k$ lokálisan korlátos második momentummal, azaz a $\vartheta \longmapsto E_\vartheta \|T\|^2$ fv. legyen korlátos bármely ϑ egy részében. Ekkor $E_\vartheta(T)$ folyt. diffható ϑ szerint és be lehet deriválni a várható értéket, azaz $E_\vartheta(T) = \int_{\mathcal{X}} T(x) f_\vartheta(x)\, d\lambda$-ra

$$\partial E_\vartheta(T) = \int_{\mathcal{X}} T(x)\, \partial f_\vartheta(x)\, d\lambda = \int_{\mathcal{X}} T(x) \underbrace{\frac{\partial f_\vartheta(x)}{f_\vartheta(x)}}_{\partial l(x)} f_\vartheta(x)\, d\lambda =$$

$$= E_\vartheta\left(T\, \partial l(\vartheta)\right) \quad k \times p\text{-es mátrix.}$$

Megj. Speciálisan $T \equiv 1$-re $0 = E_\vartheta(\partial l(\vartheta))$, ezért (R) $\Rightarrow$

$\Rightarrow I(\vartheta) = \Sigma_\vartheta(\partial l^T)$, $\quad \partial E_\vartheta(T) = \mathrm{cov}_\vartheta(T, \partial l^T)$

Megj. (R)-ből minelig csak a bederiválhatóságot használjuk, csak azt nehéz ellenőrzni.

Megj. A Fischer-információ nem függ a domináló mértéktől. 2017.03.01.

Ha ui. μ az eredeti val. mérték, λ pedig tetsz. domináló mérték, jel. $f_\vartheta = \frac{dP_\vartheta}{d\lambda}$, $g_\vartheta = \frac{dP_\vartheta}{d\mu}$, ekkor persze $\mu \ll \lambda$, így a láncszabály szerint $f_\vartheta = g_\vartheta \cdot \frac{d\mu}{d\lambda} = g_\vartheta \cdot h$ (h ϑ-ftlen.)

$\Rightarrow \log f_\vartheta = \log h + \log g_\vartheta \Rightarrow \partial \log f_\vartheta = \partial \log g_\vartheta$ ✓

$\Rightarrow$ (R) sem függ ϑ-tól

Tétel (1) Függetlenek információja összeadódik:

X minta eloszlásaira teljesül (R)
Y —"— —"— —"— (R)
$\Big\}$ $\forall \vartheta$-ra függetlenek

$\Rightarrow$ az együttes eloszlások családjára is teljesül (R) és

$$I_{(X,Y)}(\vartheta) = I_X(\vartheta) + I_Y(\vartheta) \quad \forall \vartheta$$

Köv. n elemű ftlen mintára: ha egy mintaelem eloszlására teljesül (R) $\Rightarrow$ a teljes mintára is, és $I_n(\vartheta) = n \cdot I_1(\vartheta)$, ahol $I_n(\vartheta)$ az n elemű minta információja.

(2) Statisztika információja nem nagyobb, mint a mintáé:

ha X eloszlásaira teljesül (R), és $T = T(X)$ eloszlásaira is teljesül (R), akkor $\forall \vartheta$: $I_T(\vartheta) \leq I_X(\vartheta)$ (azaz $I_X - I_T$ poz. nemdef.)

(3) Elégséges statisztika megőrzi az információt:

ha X eloszlásaira (R), és T elégséges $\Rightarrow$ T eloszlásaira (R), és $I_T(\vartheta) = I_X(\vartheta)$ $\forall \vartheta$.

„Az ördög a részletekben van elásva."

(4) Tfh. az eloszlások mind ekvivalensek.

Ekkor (3) megfordítható: ha X és T eloszlásaira (R) és $I_T(\vartheta) = I_X(\vartheta)$ $\forall \vartheta$ $\Rightarrow$ T elégséges.

B: (1) X minta értékei az $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ mérh. térből vesz, eloszlásai P_ϑ, $\vartheta \in \Theta$; míg Y az $(\mathcal{Y}, \mathcal{B}_{\mathcal{Y}})$-ből, eo. Q_ϑ, $\vartheta \in \Theta$

$\Rightarrow$ (X, Y) az $(\mathcal{X} \times \mathcal{Y}, \mathcal{B}_{\mathcal{X}} \times \mathcal{B}_{\mathcal{Y}})$-ből, eo. $P_\vartheta \times Q_\vartheta$, $\vartheta \in \Theta$

Ha $P_\vartheta \ll \lambda$, $Q_\vartheta \ll \mu$ $\Rightarrow$ $P_\vartheta \times Q_\vartheta \ll \lambda \times \mu$.

Jel. $f_\vartheta = \frac{dP_\vartheta}{d\lambda}$, $g_\vartheta = \frac{dQ_\vartheta}{d\mu}$ $\Rightarrow$ $\frac{d(P_\vartheta \times Q_\vartheta)}{d(\lambda \times \mu)}(x,y) = f_\vartheta(x) \cdot g_\vartheta(y) = h_\vartheta(x,y)$

(Ez egyszerű mértékelméleti állítás.)

$$\partial \log h_\vartheta(x,y) = \partial \log f_\vartheta(x) + \partial \log g_\vartheta(y)$$ (itt nagybetűk kellenének!)

várható értékük (R) miatt 0

$\Rightarrow$ $E_\vartheta(\partial \log h_\vartheta(X,Y)) = 0$ $\rightarrow$ $\underset{(X,Y)}{I(\vartheta)} = \Sigma_\vartheta(\partial \log h_\vartheta(X,Y)^T) =$

$= \Sigma_\vartheta(\partial \log f_\vartheta(X)) + \Sigma_\vartheta(\partial \log g_\vartheta(Y))$ a függetlenség miatt.

$= I_X(\vartheta) + I_Y(\vartheta)$

Az együttes eo. regularitása:

① $\sqrt{h_\vartheta(x,y)} = \sqrt{f_\vartheta(x)} \cdot \sqrt{g_\vartheta(y)}$ $\lambda \times \mu$ - m.m. (x,y)-ra folyt. diff.ható.

② $I_{(X,Y)}(\vartheta)$ folytonos ϑ-ban és poz. def., mert ilyenek összege. ✓

(2) <u>Lemma</u>. Jel. P_ϑ X eloszlását, $f_\vartheta = \frac{dP_\vartheta}{d\lambda}$. Ekkor $T(X)$ eloszlása

$$Q_\vartheta = P_\vartheta \circ T^{-1} \ll \mu = \lambda \circ T^{-1}, \quad \text{és} \quad \frac{dQ_\vartheta}{d\mu} =: g_\vartheta.$$

$$\Rightarrow E_\vartheta[\partial \log f_\vartheta(X) | T] = \partial \log g_\vartheta(T)$$

Bl: legyen B mérhető halmaz.

$$P_\vartheta(T \in B) = P_\vartheta(T^{-1}(B)) = E_\vartheta(\underbrace{I_{T^{-1}(B)}}_{\text{indikátor}}) \Rightarrow \text{folytdif. (R) miatt}$$

$$\Rightarrow \partial P_\vartheta(T^{-1}(B)) = \int_{\mathcal{X}} I_{T^{-1}(B)}\, \partial f_\vartheta \, d\lambda \underset{\substack{\uparrow \\ \text{mint korábban}}}{=} E_\vartheta(I_{T^{-1}(B)}\, \partial \log f_\vartheta(X)) =$$

$$= \int_{T^{-1}(B)} \partial \log f_\vartheta \, d\lambda$$

helyettes. integr.

$$P_\vartheta(T^{-1}(B)) = Q_\vartheta(B) \quad \text{folytdiff} \quad \Rightarrow \partial Q_\vartheta(B) = \int_B \partial \log g_\vartheta \, d(\underbrace{\lambda \circ T^{-1}}_{\mu}) =$$

$$= \int_{T^{-1}(B)} (\partial \log g_\vartheta \circ T)\, d\lambda.$$

$\boxed{1}$

<u>Lemma</u>. (A teljes szórásnégyzet tétele)

A feltételes kovarianciamátrix (szórásnégyzet): $\Sigma(X|\mathcal{F}) = E\left([X - E(X|\mathcal{F})]\cdot[X - E(X|\mathcal{F})]^T \middle| \mathcal{F}\right)$

$= E(XX^T|\mathcal{F}) - E(X|\mathcal{F})\cdot E(X|\mathcal{F})^T$.

Ezért $\Sigma(X) = E(\Sigma(X|\mathcal{F})) + \Sigma(E(X|\mathcal{F}))$

Bl: Definíció jól látható: $\Sigma(X|\mathcal{F}) = \Sigma(X - EX|\mathcal{F})$, mert EX $\mathcal{F}$-mérhető.

$\Rightarrow \Sigma(X|\mathcal{F}) = \Sigma(X - EX|\mathcal{F}) = E\left([X-EX]\cdot[X-EX]^T|\mathcal{F}\right) - [E(X|\mathcal{F}) - EX]\cdot[E(X|\mathcal{F}) - EX]^T$

$\Rightarrow E\,\Sigma(X|\mathcal{F}) = \Sigma(X) - \Sigma(E(X|\mathcal{F}))$

$\boxed{1}$

$\Rightarrow I_X(\vartheta) = \Sigma_\vartheta(\partial \log f_\vartheta(X)^T) \geq \Sigma_\vartheta(\partial \log g_\vartheta(T)) = I_T(\vartheta)$

$\partial \log f_\vartheta(X)^T$, $\mathcal{F} = \sigma(T)$-re lemmával

$\boxed{2}$

(3) T elégséges $\Rightarrow f_\vartheta(X) = h(X)\cdot g_\vartheta(T(X)) \Rightarrow \partial \log f_\vartheta(X) = \partial \log g_\vartheta(T(X))$

$\Rightarrow I_X(\vartheta) = E_\vartheta\left([\partial \log f_\vartheta(X)]^T[\partial \log f_\vartheta(X)]\right) = E_\vartheta\left([\partial \log g_\vartheta(T(X))]^T\cdot[\partial \log g_\vartheta(T(X))]\right) = I_T(\vartheta)$

✓

A regularitás domináló mértéktől-független $\Rightarrow$ ftn. λ^{at} ekvivalens mértékre, ezért $f_\vartheta(x) = g_\vartheta(T(x))$.

λ-m.m. x-re $\vartheta \mapsto \sqrt{f_\vartheta(x)}$ folytdiff. $\Rightarrow \lambda\circ T^{-1}$-m.m. t-re $\vartheta \mapsto \sqrt{g_\vartheta(t)}$ folytdiff, és az információ is megegyezik.

$\boxed{3}$

$\Rightarrow T(X)$-re is (R).

(4) $I_X(\vartheta) = I_T(\vartheta)$-hoz az kell, hogy $E_\vartheta(\Sigma_\vartheta(\partial \log f_\vartheta(x)^T|T)) = 0_{p\times p}$

$\Rightarrow \Sigma_\vartheta(\partial \log f_\vartheta(x)^T|T) = 0$ $\underbrace{P_\vartheta\text{-m.m.}}$, mert poz. szemidef. és v.értéke 0.

P-m.m., mert $\forall$ mértéket ekvivalens.

Ftn. $f_\vartheta(x) > 0$ $\forall\vartheta\ \forall x$, amire nem, azt kidobhatjuk (nullmértékű)

$\Rightarrow \partial \log f_\vartheta(X) = \varphi_\vartheta(T(x))$ $\forall\vartheta$, $\underbrace{\text{m.m. } x\text{-re}}$

$x \in N_\vartheta$, ahol N_ϑ P-0-halmaz

$\downarrow$

ϑ-függő!

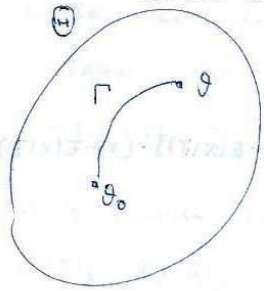
Legyen most $\Theta^* \subseteq \Theta$ megszámlálható, sűrű halmaz.

N_0: $x \notin N_0$-ra $\vartheta \mapsto \log f_\vartheta(x)$ folyt. diffható, (N_0 is P-0)

~~ha $\vartheta \notin \Theta^*$~~

Tfh. $x \notin \underbrace{\left(\bigcup_{\vartheta^* \in \Theta^*} N_{\vartheta^*}\right) \cup N_0}_{P-0} \Rightarrow \partial \log f_\vartheta(x) = \lim_{\substack{\vartheta^* \to \vartheta \\ \vartheta^* \in \Theta^*}} \varphi_{\vartheta^*}(T(x)) = \tilde{\varphi}_\vartheta(T(x))$.

Tehát m.m. x-re $\forall \vartheta$: $\partial \log f_\vartheta(x) = \tilde{\varphi}_\vartheta(T(x))$



$\vartheta_0 \in \Theta$ rögz., $\Gamma \subset \Theta$ görbe, Γ ϑ_0-t köti össze ϑ-val, rektifikálható. Integráljunk:

$$\Rightarrow \log f_\vartheta(x) - \log f_{\vartheta_0}(x) = \int_\Gamma \tilde{\varphi}_t(T(x))\, dt$$

$\longrightarrow$ ez $T(x)$ fv-ye.

$$\Rightarrow f_\vartheta(x) = \underbrace{f_{\vartheta_0}(x)}_{h(x)} \cdot \underbrace{e^{\int \dots}}_{g_\vartheta(x)}$$

faktorizációs tétel $\Rightarrow$ T elégséges

$\boxed{4}$ $\square$

Átparaméterezés

$\mathcal{P} = \{P_\vartheta : \vartheta \in \Theta\} \subset \mathbb{R}^p$, $h: G \to \Theta$ injektív és diffható, $G \subset \mathbb{R}^p$ nyílt.

$\mathcal{P}' := \{P'_\tau = P_{h(\tau)} : \tau \in G\}$ ϑ a régi, τ az új paraméter.

$(\mathcal{X}, \mathcal{B}, \mathcal{P}')$-ben a τ-ra vonatkozó információ legyen $\mathcal{I}(\tau)$.

a régi információ $I(\vartheta)$. Mi köztük a kapcsolat?

$$\cancel{\partial_\tau} \quad \frac{dP'_\tau}{d\lambda} = g_\tau, \quad \frac{dP_\vartheta}{d\lambda} = f_\vartheta, \quad g_\vartheta = f_{h(\tau)}$$

$$\Rightarrow \underbrace{\partial_\tau \log g_\tau}_{q\text{-dim sor}} = \underbrace{\partial_\vartheta \log f_\vartheta \,|_{\vartheta = h(\tau)}}_{p\text{-dim sor}} \underbrace{h'(\tau)}_{p \times q\text{-dim mátrix}}$$

diád-képzés $\Rightarrow$ $\mathcal{I}(\tau) = h'(\tau)^T \cdot I(h(\tau)) \cdot h'(\tau)$

Ez pl. jó $N(\cdot,\cdot)$ kétféle paraméterezése közti áttérésre: σ vagy σ^2.

Erős regularitási feltételek (RR)

(R) + $\exists \underbrace{\partial^2 \ell}_{p\times p}$ folytonos, és ha $M(X) = \sup_\vartheta \|\partial^2 \log f_\vartheta(X)\|^2$, akkor

$\sup_\vartheta E_\vartheta(M(X)) < +\infty$ legyen.

(Ezeket elég lenne lokálisan megfogalmazni, de most elég globálisan.)

Áll. (ZR) teljesülése esetén:
(BN)

(1) $0 = \int_{\mathcal{X}} \partial^2 f_\vartheta \, d\lambda$ (kétszer bederiválhatóság)

(2) $I(\vartheta) = -E_\vartheta\left(\partial^2 \ell(\vartheta)\right)$

Becsléselmélet

Pontbecslésekre fogunk vizsgálni. (vagy intervallumbecslés is)

$g: \Theta \to \mathbb{R}^k$, $g(\vartheta)$-t szeretnénk becsülni

Def. $g(\vartheta)$ becslése $T(X)$, ahol $T: \mathcal{X} \to \mathbb{R}^k$, azaz olyan kell, hogy ugyanolyan képezetű legyen.

Def. T torzítatlan, ha $\forall \vartheta$: $E_\vartheta(T) = g(\vartheta)$

Def. T torzítása $b_T(\vartheta) = E_\vartheta(T) - g(\vartheta)$ (bias)

Def. Veszteségfüggvény: $w: \mathbb{R}^k \times \Theta \longrightarrow \begin{cases} \mathbb{R} & \text{(általában)} \\ \mathbb{R}^{k\times k} & \text{(kivételes, de néponként)} \end{cases}$

Ha a paraméter ϑ, és $g(\vartheta)$-t $y \in \mathbb{R}^k$-val becsüljük, a veszteségünk $w(y, \vartheta)$.

(1) $w(g(\vartheta), \vartheta) = 0$

(2) $w(y, \vartheta) \geq 0$ (pozitív szemidef., ha mátrix)

(3) w az első változójában konvex

Példa. Négyzetes veszteségfv.: a leggyakrabban használt, két jelentése is van. 2017.03.08.

→ skalárértékű: $v_1(y) = \|y\|^2 \in \mathbb{R}$

→ mátrixértékű: $v_2(y) = y \cdot y^T \in \mathbb{R}^{k\times k}$

$v_1 = \mathrm{Tr}\, v_2$ ⇒ elég v_2-re bizonyítani a def. feltételeit. (Poz. szemidef. nyoma nemnegatív.)

(1) és (2) világos.

(3): $y_1 \neq y_2$, $0 < \alpha < 1$: $\alpha \cdot v_2(y_1) + (1-\alpha) v_2(y_2) - v_2(\alpha y_1 + (1-\alpha) \cdot y_2) =$

$= \alpha \cdot y_1 y_1^T + (1-\alpha) y_2 y_2^T - \left(\alpha^2 y_1 y_1^T + (1-\alpha)^2 y_2 y_2^T + \alpha \cdot (1-\alpha) \cdot (y_1 y_2^T + y_2 y_1^T)\right) =$

$= \alpha \cdot (1-\alpha) \cdot (y_1 y_1^T + y_2 y_2^T - y_1 y_2^T - y_2 y_1^T) = \alpha \cdot (1-\alpha) \cdot v_2(y_1 - y_2) \geq 0$

⇒ szigorúan is konvex

Def. Legyen $T: \mathcal{X} \to \mathbb{R}^k$ a $g(\vartheta)$ becslése, $g: \Theta \to \mathbb{R}^k$.

T rizikófüggvénye $R_T: \Theta \to \begin{cases} \mathbb{R} \\ \mathbb{R}^{k \times k} \end{cases}$, $\quad R_T = E_\vartheta(w(T, \vartheta))$

Példa. Mátrix értékű négyzetes rizikó:

$$R_T(\vartheta) = E_\vartheta\left([T - g(\vartheta)] \cdot [T - g(\vartheta)]^T\right) = \Sigma_\vartheta(T) + b_T(\vartheta) \cdot b_T(\vartheta)^T$$

Speciálisan torzítatlan becslés rizikója $R_T(\vartheta) = \Sigma_\vartheta(T)$

Def. Legyenek T_1, T_2 a $g(\vartheta)$ becslései; T_1 egyenletesen jobb (értsd: nem rosszabb), ha $R_{T_1}(\vartheta) \leq R_{T_2}(\vartheta) \quad \forall \vartheta \in \Theta$

Ha $R_{T_1}(\vartheta) = R_{T_2}(\vartheta) \quad \forall \vartheta \in \Theta \Rightarrow T_1$ és T_2 ekvivalens becslések.

Megj. Mátrixokra nehézkesen; speciálisan ez azt jelenti, hogy ez csak visszavezetés.

Def. Legyen $\mathcal{D}$ a $g(\vartheta)$ becsléseinek egy osztálya (nem feltétlenül az összes becslés). $T \in \mathcal{D}$ optimális becslés, ha $\forall T' \in \mathcal{D}$: T egy. jobb T'-nél.

A $T \in \mathcal{D}$ megengedhető (admisszibilis), ha $\forall T' \in \mathcal{D}$: ha T és T' összehasonlítható, akkor T egyenletesen jobb T'-nél.

Def. Skalárértékű w veszteségfv esetén $T \in \mathcal{D}$ minimax becslés, ha $\forall T' \in \mathcal{D}$: $\sup_\vartheta R_T(\vartheta) \leq \sup_\vartheta R_{T'}(\vartheta)$. (A pesszimista ember becslése.)

Torzítatlan becslés és négyzetes veszteségfv. esetén

egyenletesen jobb = hatásosabb

optimális = hatásos (efficiens) = UMVU

Áll. 1) T optimális $\Rightarrow$ minimax és megengedhető.

2) Ha $\mathcal{D}$ konvex és w első vált. szig. konvex, akkor az optimális becslés egyértelmű $\mathcal{P}$-m.m.

B: 1) triviális

2) T_1, T_2 optimálisak, $T := \frac{T_1 + T_2}{2} \in \mathcal{D}$. (a konvexitás csak ehhez kell)

$$R_T(\vartheta) = E_\vartheta\left(w\left(\frac{T_1 + T_2}{2}, \vartheta\right)\right) \leq E_\vartheta\left(\frac{w(T_1, \vartheta) + w(T_2, \vartheta)}{2}\right) =$$

$$= \frac{R_{T_1}(\vartheta) + R_{T_2}(\vartheta)}{2} = R_{T_1}(\vartheta)$$

Optimalitás $\Rightarrow R_{T_1}(\vartheta) = R_T(\vartheta)$, egyenlőség áll.

$$\Rightarrow w\left(\frac{T_1 + T_2}{2}, \vartheta\right) = \frac{w(T_1, \vartheta) + w(T_2, \vartheta)}{2} \quad P_\vartheta\text{-m.m. } \forall \vartheta$$

$$\Rightarrow T_1 = T_2 \quad \underbrace{P_\vartheta\text{-m.m. } \forall \vartheta}_{P_\vartheta\text{-m.m.}}$$

$\square$

Köv. A hatásos becslés egyértelmű.

B: A torzítatlan becslők osztálya konvex, a négyzetfüggvény sig. konvex. □

Tétel. (Blackwell – Rao, B. 1945, R. 1947, Kolmogorov 1950)

Legyen T a $g(\vartheta)$ becslése, S elégséges statisztika. $\tilde{T} := E_{\vartheta}(T|S)$ (S szerint)

1) $b_{\tilde{T}}(\vartheta) = b_T(\vartheta) \quad \forall \vartheta \in \Theta$

2) $R_{\tilde{T}}(\vartheta) \leq R_T(\vartheta) \quad \forall \vartheta \in \Theta$ (egyenletesen jobb)

Ennyiből becsülni csak az elégséges statisztika függvényeivel érdemes.

B: 1) Teljes várható érték tétele: $E_{\vartheta}(\tilde{T}) = E_{\vartheta}(T) \quad \forall \vartheta \in \Theta$

2) Feltételes Jensen-egyenlőtlenség:

$$R_{\tilde{T}}(\vartheta) = E_{\vartheta}\left(w(E_{\vartheta}(T|S), \vartheta)\right) \leq E_{\vartheta}\left(E_{\vartheta}\left(w(T, \vartheta)\,|\,S\right)\right) \overset{\text{F.V.É.}}{=} E_{\vartheta}\left(w(T, \vartheta)\right) = R_T(\vartheta).$$ □

Def. A T statisztika teljes, ha tetszőleges h függvénye esetén

$E_{\vartheta}(h(T)) = 0 \Rightarrow h(T) = 0$ P.m.m. (A másik irányú nyíl persze igaz.)

Van hasonló tétel, mint elégségesnél a faktorizáció.

Áll. 1) Ha T teljes és $E_{\vartheta}(h_1(T)) = E_{\vartheta}(h_2(T)) \quad \forall \vartheta \in \Theta \rightarrow h_1(T) = h_2(T)$ P-m.m. Ez becslések esetén ekvivalens $b_{h_1(T)}(\vartheta) = b_{h_2(T)}(\vartheta)$-val.

Következmény: teljes statisztika esetén legfeljebb egy torzítatlan becslés van.

2) Ha T torzítatlan, S pedig teljes és elégséges, akkor $\tilde{T}$ hatásos.

B: 1) triviális

2) Tetsz. T torzításból kiindulva $\tilde{T}$ mindig ugyanaz lesz. □

Ez az eljárás a Blackwellizálás.

1. Keressünk $g(\vartheta)$-ra „minél egyszerűbb" T torzítatlan becslést. (ha van)
2. Keressünk teljes és elégséges S statisztikát. (ha van)
3. $\tilde{T} = E_{\vartheta}(T|S)$ hatásos.

T választásától $\tilde{T}$ nem függ, de a számolás bonyolultsága függ a választástól.

Tétel. (Cramér – Rao) (R) legyen teljesülve, és legyen T a $g(\vartheta)$ torzítatlan becslése kor. korlátos másodrendű momentumnak. Tudjuk, hogy ekkor g folytdiff, deriváltja legyen $G(\vartheta)$, $k \times p$-s mátrix.

Ekkor $R_T(\vartheta) = \Sigma_{\vartheta}(T) \geq G(\vartheta) \cdot I(\vartheta)^{-1} \cdot G(\vartheta)^T$, a jobb oldal neve információs határ.

$\downarrow$ információs egyenlőtlenség.

Megj. Ha $\forall \vartheta \in \Theta$-ra egyenlőség áll $\rightarrow$ T hatásos. Fordítva nem igaz.

Megj. Tfh. a minta n elemű független. $\rightarrow$ $I(\vartheta) = I_n(\vartheta) = n \cdot I_1(\vartheta)$,

és $\Sigma_\vartheta(T) \geq \frac{1}{n} \cdot G(\vartheta) I_1(\vartheta)^{-1} \cdot G(\vartheta)^T$, mivel $(I_n(\vartheta))^{-1} = \frac{1}{n} \cdot I_1(\vartheta)^{-1}$.

Ez azt mondja, hogy bár G nemint a világ megismerhető, de csak lassan.

B_T: $0 \leq \Sigma_\vartheta(T - G I^{-1} \cdot \partial \ell^T) = \mathrm{cov}_\vartheta(T - G I^{-1} \partial \ell^T, T - G I^{-1} \partial \ell^T) =$

$$= \underbrace{\mathrm{cov}_\vartheta(T, T)}_{\Sigma(T)} - \underbrace{\mathrm{cov}_\vartheta(T, \partial \ell^T)}_{G} \cdot I^{-1} \cdot G^T + G I^{-1} \underbrace{\mathrm{cov}_\vartheta(\partial \ell^T, \partial \ell^T)}_{\Sigma(\partial \ell^T) = I} I^{-1} G^T - G I^{-1} \underbrace{\mathrm{cov}_\vartheta(\partial \ell^T, T)}_{G^T}$$

$$= \Sigma(T) - G I^{-1} G^T$$

$\square$

Köv. Legyen g diffható, T nem felt. torzítatlan becslése g-nek, T második momentuma lét. korlátos. Ekkor $b_T(\vartheta)$ is diffható, jel. $B(\vartheta) = \partial b_T(\vartheta)$, ezt tudjuk.

$\Rightarrow$ A négyzetes rizikóra $R_T(\vartheta) \geq (G(\vartheta) + B(\vartheta)) \cdot I(\vartheta)^{-1} \cdot (G(\vartheta) + B(\vartheta))^T +$
$+ b_T(\vartheta) \cdot b_T(\vartheta)^T$.

B: A CR-egyenlőtlenséggel $R_T(\vartheta) = \Sigma_\vartheta(T) + b_T(\vartheta) \cdot b_T(\vartheta)^T$.

$\Sigma_\vartheta(T)$-re az előző tétel g helyett $g + b_T$-vel.

$\square$

Átparaméterezés: régi paraméter $\vartheta \in \Theta$, $\tau \in \mathcal{G}$. Legyen $\tau = h(\vartheta)$, ahol $h: \Theta \to \mathcal{G}$ diffeomorfizmus.

$\Rightarrow$ $\mathcal{P}'_\tau = \mathcal{P}_{h(\tau)}$

τ-ban az új információs határ:

$g(\vartheta)$ helyett $g(h(\tau))$, deriváltja $G(h(\tau)) \cdot H(\tau)$, ahol $H = h'$, $p \times p$-es mátrix.

$$[G(h(\tau)) \cdot H(\tau)] \cdot \underbrace{[H(\tau)^T \cdot I(h(\tau)) \cdot H(\tau)]}_{J(\tau)}^{-1} \cdot [G(h(\tau)) \cdot H(\tau)]^T =$$

$$= G(h) \cdot \cancel{H \cdot H^{-1}} \cdot I \cdot (h)^{-1} \cdot \cancel{(H^T)^{-1}} \cdot H^T \cdot G^T(h) = G(h) \cdot I(h)^{-1} \cdot G(h)^T,$$

azaz az információs határ nem változik átparaméterezéssel.

Def. Legyenek $Y_1, \dots, Y_k$ független $N(0,1)$ val. változók, $Y = (Y_1, \dots, Y_k)^T$.

Az Y eloszlása k-dim. std. normális eo.

Legyen $A \in \mathbb{R}^{n \times k}$, $a \in \mathbb{R}^n$, $X = AY + a$. Ekkor X eloszlása n-dim. normális eo.

Áll. 1) $EX = a$, $\Sigma(X) = AA^T$. (Mert $EY = 0$, $\Sigma(Y) = \mathbb{E}_k$ egységmátrix.)

2) Ha X n-dim normális, $B \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ $\Rightarrow$ $BX + b$ normális, azaz a normális eloszlások lineárisan zártak.

Speciálisan ha X n-dim normális, akkor $\forall X_i$ marginális is normális, de ez megfordítva nem igaz.

3) $\forall m \in \mathbb{R}^n$ $\forall \Sigma \geqq 0$, $\Sigma \in \mathbb{R}^{n \times n}$, $\exists X$ normális, hogy $EX = m$, $\Sigma(X) = \Sigma$

(Ui. Y n-dim std. normálisra, $a = m$, $A = \Sigma^{1/2}$)

Áll. Ha $\Sigma \gneqq 0$ $\Rightarrow$ az eloszlásnak van sűrűségfüggvénye, és $f(x) = (2\pi)^{-n/2} \cdot (\det \Sigma)^{-1/2} \cdot \exp\left(-\frac{1}{2}(x-m)^T \Sigma^{-1} (x-m)\right)$.

Megj. Ez a Σ nem hat. meg X-et, meg A mérete is sokféle lehet úgy, hogy $AA^T = \Sigma$ legyen.

Megj. Ha $\operatorname{rang} \Sigma = r < n$ $\Rightarrow$ X 1 val. egy r-dim affin altérbe esik, ilyenkor az n-dim Lebesgue-mértéke mérve nem folytonos X. De lehet az aff. altérben is Lebesgue-mértéket definiálni, és azzal már lesz sűr.

Megj. $\Sigma = U \Lambda U^T$, ahol Λ diagonális, az átlóban a $\lambda_1, \dots, \lambda_n$ sajátértékek állnak, $\lambda_i \geqq 0$ (hiszen $\Sigma \geqq 0$), és U ortonormált mátrix, azaz sorai ortonormált rendszert alkotnak (és oszlopai is), azaz $U^T = U^{-1}$, U oszlopai a λ-knak megfelelő ortonormált sajátvektorok.

$\Sigma^{1/2} = U \Lambda^{1/2} U^T$, ahol

$\Lambda = \operatorname{diag}(\lambda_1^{1/2}, \dots, \lambda_n^{1/2})$.

B: A továbbiakban $\mu = m$. 2017. 03. 22.

$X = AY + \mu$, $\Sigma = AA^T$, $A \in \mathbb{R}^{n \times k}$

Először tfh. $k = n$.

$\Rightarrow$ Y sűrűsége $f_Y(y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{y_i^2}{2}} = (2\pi)^{-n/2} \exp\left(-\frac{1}{2}\|y\|^2\right)$

$\Rightarrow$ X sűrűsége sűrűtős formulával: az inverz $Y = A^{-1}(X - \mu)$, a Jacobi $|\det A^{-1}| = \frac{1}{|\det A|} = \frac{1}{\sqrt{\det \Sigma}}$,

$$f_X(x) = f_Y\left(A^{-1}(x-\mu)\right) \cdot \frac{1}{\sqrt{\det \Sigma}} =$$

$$= (2\pi)^{-n/2} (\det \Sigma)^{-1/2} \cdot \exp\left(-\frac{1}{2} \cdot \overbrace{(x-\mu)^T \underbrace{(A^{-1})^T A^{-1}}_{(AA^T)^{-1} = \Sigma^{-1}} (x-\mu)}^{\|A^{-1}(x-\mu)\|}\right)$$

Legyen $k > n$ $\rightarrow$ elég $\mu = 0$-ra bizonyítani, az csak eltolásparaméter.

(Az előbb is lehetett volna.)

Jelölje A-t most A_1.

$$\begin{array}{c|c} & k \\ \hline n & A_1 \\ \hline k-n & A_2 \end{array}$$

Legyen A_2 $(k-n)\times k$-as mátrix, a sorai lin. ftlenek és ortogonálisak A_1 soraira, azaz $A_1 A_2^T = 0$

$\Rightarrow$ $\operatorname{rang} A = \operatorname{rang}\left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] = k$

$X = AY = \left[\begin{array}{c} X_1 \\ \hline X_2 \end{array}\right]$ $\quad X_1 = A_1 Y$ $\quad n$ dim.

$\quad\quad\quad\quad\quad\quad X_2 = A_2 Y$ $\quad k-n$ dim.

X sűrűségfvénye, azaz X_1, X_2 együttes sfvénye:

$$f(x_1, x_2) = (2\pi)^{-k/2} (\det \Sigma)^{-1/2} \exp\left(-\tfrac{1}{2} x^T \Sigma^{-1} x\right)$$

$$\Sigma = AA^T = \left[\begin{array}{c|c} A_1A_1^T & A_1A_2^T \\ \hline A_2A_1^T & A_2A_2^T \end{array}\right] = \operatorname{diag}\left(A_1A_1^T,\ A_2A_2^T\right)$$

$$\Sigma^{-1} = \operatorname{diag}\left((A_1A_1^T)^{-1},\ (A_2A_2^T)^{-1}\right)$$

$$x^T \Sigma^{-1} x = x_1^T (A_1A_1^T)^{-1} x_1 + x_2^T (A_2A_2^T)^{-1} x_2$$

$$\det \Sigma = \det A_1A_1^T \cdot \det A_2A_2^T$$

$$\Rightarrow f(x_1, x_2) = (2\pi)^{-n/2} (\det A_1A_1^T)^{-1/2} \cdot \exp\left(-\tfrac{1}{2} x_1^T (A_1A_1^T)^{-1} x_1\right) \cdot$$
$$\cdot (2\pi)^{-(k-n)/2} (\det A_2A_2^T)^{-1/2} \cdot \exp\left(-\tfrac{1}{2} x_2^T (A_2A_2^T)^{-1} x_2\right)$$

$\Rightarrow$ X_1, X_2 függetlenek és a szét nem osztó konstansnak a megfelelő sűrűségfüggvény

De ha az egyik stv., a másik is az:

Válasszuk A_2 sorait úgy, hogy az A_1 sorai által kifeszített altér ortogonális kiegészítő alterében ortonormált bázist adjanak, ez megtehető.

$\Rightarrow$ $A_2 A_2^T = \mathbb{E}_{k-n}$ $\Rightarrow$ $f(x_1, x_2)$ második tényezője a $k-n$ dimenziós std. normális eo. sfvénye

$A_1 A_1^T = \Sigma(X_1)$ $\Rightarrow$ szorzatuk, amit bizonyítani akartunk. $\square$

<u>Áll.</u> A többdim. normális eo.-t μ és Σ meghatározza.

B: Ha $\Sigma > 0$ $\Rightarrow$ $\exists$ sfvény, ami μ-től és Σ-tól függ csak, és meghatározza az eloszlást.

Ha $\operatorname{rang} \Sigma = r < n$ $\Rightarrow$ X-nek van r olyan koordinátája, hogy a többi ezek lineáris fvénye (gyak.)

Ezen r koordináta koo. mátrixa invertálható (hiszen ha nem lenne az, akkor r-nél kevesebb koordinátával is kifejezhető lenne minden, de akkor $\operatorname{rang} \Sigma < r$ lenne ↯)

Ezért ennek az r koordinátával már van sűrűsége, ezek eloszlását meghatározta μ és Σ $\Rightarrow$ a többieket is. □

Jel. $N_n(\mu, \Sigma)$ az n-dim normális e.o. (n kiírása nem kell, ha μ és Σ dimenziója adott, hiszen az part n).

Áll. X_1, X_2 együttes eloszlása normális: $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \begin{matrix} \ell \\ n-\ell \end{matrix}$, $EX = \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

$\Rightarrow \Sigma(X) = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right]$, ahol $\Sigma_{11} = \Sigma(X_1)$, $\Sigma_{22} = \Sigma(X_2)$, $\Sigma_{12} = \operatorname{cov}(X_1, X_2)$

Tfh. $\Sigma > 0$. Ekkor

(1) $\operatorname{cov}(X_1, X_2) \Leftrightarrow X_1$ és X_2 függetlenek

(2) $X_1 - \Sigma_{12}\Sigma_{22}^{-1} X_2$ és X_2 ftlenek

(3) $X_1 | X_2$ feltételes e.o. $N_\ell(m_{1\cdot 2}, \Sigma_{11\cdot 2})$, ahol

$$m_{1\cdot 2} = E(X_1 | X_2) = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2) \quad \text{(lineáris } X_2\text{-ben)},$$

$$\Sigma_{11\cdot 2} = \Sigma(X_1 | X_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \quad \text{(konstans)}$$

B: (1) $\Sigma > 0 \Rightarrow \exists$ stv. és mint láttuk, normalitás kellett, ha

$A_1 A_2^T = \operatorname{cov}(A_1 Y, A_2 Y) = \Sigma_{12} = 0$.

(A másik irány mindig igaz.)

Megj. (1) akkor is igaz, ha Σ nem invertálható: X_1-ből és X_2-ből kiválasztandók megfelelő koordinátáit, mint az előző állításban.

(2) $\begin{bmatrix} X_1 - \Sigma_{12}\Sigma_{22}^{-1} X_2 \\ X_2 \end{bmatrix} = \left[\begin{array}{c|c} E_\ell & -\Sigma_{12}\Sigma_{22}^{-1} \\ \hline 0 & E_{n-\ell} \end{array}\right] X \Rightarrow$ együttesen normális.

Elég tehát belátni, hogy $\operatorname{cov}(X_1 - \Sigma_{12}\Sigma_{22}^{-1} X_2, X_2) = 0$

$$= \underbrace{\operatorname{cov}(X_1, X_2)}_{\Sigma_{12}} - \Sigma_{12}\Sigma_{22}^{-1} \underbrace{\operatorname{cov}(X_2, X_2)}_{\Sigma_{22}} = 0 \quad ✓$$

(3) $X_1 - \Sigma_{12}\Sigma_{22}^{-1} X_2 \mid X_2$ felt. eo. megegyezik a közönséges elmélettel (2) miatt

$$X_1 - \Sigma_{12}\Sigma_{22}^{-1} X_2 \sim N_\ell\Big(\overbrace{\mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2}^{m_{1\cdot 2}}\Big)$$

$$\Sigma(X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2) = \operatorname{cov}(X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2, X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2) =$$

$$= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \underbrace{(\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^T}_{\text{szimmetrikus}} + \Sigma_{12}\cancel{\Sigma_{22}^{-1}}\cdot\cancel{\Sigma_{22}}\cdot\Sigma_{22}^{-1}\Sigma_{12} =$$

$$= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12} = \Sigma_{11\cdot 2}$$

Ebből az áll. köv. □

Aszimptotikus Cramér – Rao:

<u>Tétel</u>: (Bahadur) (BN) Tfh. (RR). Legyen $g(\vartheta)$ folyt. diffható, $g: \Theta \to \mathbb{R}^k$. $T_n = T_n(X_1, \dots, X_n)$ becslése $g(\vartheta)$-nak.

Tfh. $\sqrt{n}\,(T_n - g(\vartheta)) \xrightarrow{d} N_k(0, \underbrace{V(\vartheta)}_{\text{variancia (kov. mátrix)}})$ $\forall \vartheta \in \Theta$

Ekkor $V(\vartheta) \geq \underbrace{G(\vartheta)\, I_1(\vartheta)^{-1} G(\vartheta)^T}_{\text{információs határ}}$ Lebesgue-m.m. ϑ-ra. $(G' = g)$

<u>Mj.</u> Minden ϑ-ra az állítás nem igaz.

<u>Def.</u> Ha valamely ϑ-ra $V(\vartheta) < G(\vartheta)\, I_1(\vartheta)^{-1} G(\vartheta)^T$, akkor a (T_n) sorozat ϑ-ban szupereffiziens (túlhatásos).

A tétel szerint tehát a szupereffiziensek nem lehetnek túl sokan.

<u>Példa.</u> (BN) $N(\vartheta, 1)$ $\vartheta \in \mathbb{R}$ paraméter, $g(\vartheta) = \vartheta$-t becsüljük

$\overline{X}$ torzítatlan, sőt hatásos becslés, és eléri az inf. határt.

$I_1(\vartheta) = 1 \Rightarrow$ az aszimptotikus inf. határ 1.

$$T_n(\underline{X}) = \begin{cases} \overline{X}, & \text{ha } |\overline{X}| \geq \alpha_n \\ c\cdot\overline{X}, & \text{ha } |\overline{X}| < \alpha_n, \end{cases} \quad \text{ahol } 0 \leq c < 1$$

összehúzás (shrinkage) $\alpha_n > 0$; $\alpha_n \to 0$, $\sqrt{n}\,\alpha_n \to \infty$,

pl. $\alpha_n = \frac{1}{\sqrt[4]{n}}$

Ekkor $\sqrt{n}\cdot(T_n - \vartheta) \xrightarrow{d} N_1(0, V(\vartheta))$,

ahol $V(0) = c^2$, $V(\vartheta) = 1$ $(\vartheta \neq 0)$

$\to$ (T_n) túlhatásos $\vartheta = 0$-ban

Persze T_n nem lehet torzítatlan.

Becslési módszerek

Volt: tapasztalati becslések.

Momentumok módszere: $\Theta \in \mathbb{R}^p$.

Általánosított MM-ben $h: \mathcal{X} \to \mathbb{R}^p$ fv-t veszünk, és $\varphi(\vartheta) := E_\vartheta(h(X_1))$

Olyan h kell, amire φ invertíbilis.

$$\vartheta = \varphi^{-1}\left(E_\vartheta\left(h(X_1)\right)\right) \Rightarrow \hat{\vartheta}_n := \varphi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} h(X_i)\right)$$

Voltaképp ez is tapasztalati becslés.

NSZET $\Rightarrow$ ha φ^{-1} folytonos, akkor ez erősen konzisztens becslés, azaz 1 valószínűséggel $\hat{\vartheta}_n \to \vartheta \quad \forall \vartheta$.

Klasszikus MM: h koordinátái hatványfvek

Ha $\mathcal{X} = \mathbb{R}$, p paraméter van $\Rightarrow$ kiszámítunk p momentumot.

$\mu_{i_1}, \dots, \mu_{i_p}$ momentumok, ahol $\mu_i = \mu_i(\vartheta) = E_\vartheta(X_1^i)$

$\vartheta = \varphi^{-1}(\mu_{i_1}, \dots, \mu_{i_p}) \Rightarrow \hat{\vartheta}_n = \varphi^{-1}(\hat{\mu}_{i_1}, \dots, \hat{\mu}_{i_p})$, ahol

$$\hat{\mu}_i = \frac{1}{n}\sum_{k=1}^{n} X_k^i$$

Pl. ha az eo. szimmetrikus, akkor az 1. momentum 0, nem érdemes azt venni.

Maximum-likelihood-becslés

X minta, $f_\vartheta(X)$ likelihood-fv., (λ-ra nézve)

$$\hat{\vartheta} = \hat{\vartheta}(X) = \arg\max_{\vartheta} f_\vartheta(X)$$

$\hat{\vartheta}$ nem biztos, hogy létezik, és nem feltétlenül egyértelmű

Ha $\vartheta \mapsto f_\vartheta(X)$ diffható $\Rightarrow$ ennek $\partial f_\vartheta(X)$ zérushelyét nézzük

n elemű minta esetén az együttes likelihood $\prod_{i=1}^{n} f_\vartheta(X_i)$, ennek ugyanott van maximuma, ahol a logaritmusának:

$$L(\vartheta) = \sum_{i=1}^{n} \log f_\vartheta(X_i) = \sum_{i=1}^{n} l_i(\vartheta),$$ ez kellemesebb deriválni.

$\partial L(\vartheta) = 0$ likelihood-egyenlet

Tulajdonságok

(1.) Nem függ a domináló mértéktől (a sfv. függ!)

Ha μ az ekvivalens val. mérték, és λ tetsz. dom. mérték,

akkor $\mu \ll \lambda$ (hiszt λ dominálja az egész $\mathcal{P}$-t)

$$\Rightarrow \frac{dP_\vartheta}{d\lambda} = \frac{dP_\vartheta}{d\mu} \cdot \underbrace{\frac{d\mu}{d\lambda}}_{\vartheta\text{-tól ftlen}}$$

$\Rightarrow$ a max. hely ugyanaz lesz (de az érték nem!)

$$\operatorname{argmax} \frac{dP_\vartheta}{d\mu} = \operatorname{argmax} \frac{dP_\vartheta}{d\lambda}$$

(2.) Ha $\exists$ ML-becslés, akkor olyan is van, mely az elégséges stat. fvénye.

A maxhelyek halmaza az elégséges stat fvénye:

ha μ ekvivalens mérték, akkor $f_\vartheta(x) = g_\vartheta(T(x))$ (T eqv.)

Példa. $U[\vartheta, \vartheta+1]$-ben ML ϑ-ra:

maxhelyek halmaza:

$X_n^* - 1 \leq \vartheta \leq X_1^*$, az X_1^*, X_n^* elégséges

$\Rightarrow \hat{\vartheta}_1 = X_1^*$ jó, $\hat{\vartheta}_2 = X_1^* - (X_1^* - X_n^* + 1) \cdot \frac{|\bar{X}|}{1+|\bar{X}|}$ nem jó $n \geq 3$-ra,

mert akkor már nem csak min és max van a mintában

2017.03.29. Legyen $g: \Theta \to \Gamma$.

Def. $\gamma \in \Gamma$-ra $\tilde{f}_\gamma := \sup_\vartheta \{ f_\vartheta \mid g(\vartheta) = \gamma \}$ Zehnar féle indukált likelihood-fv.

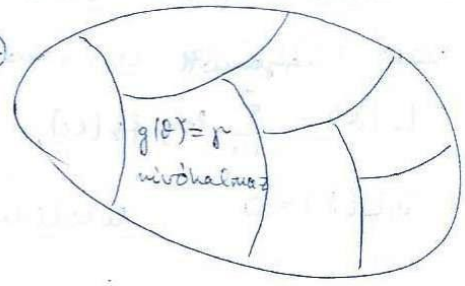
$g(\vartheta)$ ML-becslése $\hat{g}(X) = \operatorname{argmax}_\gamma \tilde{f}_\gamma(X)$

Áll. $\hat{\vartheta}$ ML-becslése ϑ-nak $\Rightarrow g(\hat{\vartheta})$ ML-becslése $g(\vartheta)$-nak.

B: $\sup_\vartheta f_\vartheta = \sup_\gamma \tilde{f}_\gamma \geq \tilde{f}_{g(\hat{\vartheta})} \geq f_{\hat{\vartheta}} =$

$= \sup_\vartheta f_\vartheta$

$\Rightarrow \tilde{f}_{g(\hat{\vartheta})} = \sup_\gamma \tilde{f}_\gamma$, és ez kellett.



$\square$

Emiatt elég a paraméter ML-becslését vizsgálni,

Milyen asz. tulajdonságai vannak a ML-becsléseknek?

<u>Tétel</u> (RR) $\Rightarrow$ a paraméter ML-becslése erősen konzisztens, és

(1) 1 val. $\forall$ elég nagy n-re a likelihood-egyenletnek van gyöke, sőt van olyan $\hat{\vartheta}_n$ gyök, mely a likelihood-fv. lok. maximumhelye, és amely erősen konzisztens, tehát $\hat{\vartheta}_n \to \vartheta$ P_ϑ-m.m. $\forall \vartheta$.

(2) Ennek a $\hat{\vartheta}_n$-nek $\sqrt{n}\,(\hat{\vartheta}_n - \vartheta) \xrightarrow{d} N_p(0, I(\vartheta)^{-1})$

<u>Mj.:</u> A Bahadur-tétel szerint $\hat{\vartheta}_n$ asszimptotikusan optimális: az asz. kovarianciamátrix eleni az info. határt.

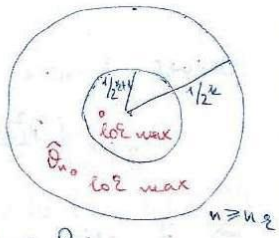
B: (1) Megmutatjuk, hogy $\forall$ elég kicsi $\varepsilon > 0$-ra 1 val. $\forall$ elég nagy n-re

$$\sup_{\|h\|=\varepsilon} \frac{1}{n} L(\vartheta + h) < \frac{1}{n} L(\vartheta) \qquad (*)$$

Ebből mi köv., hogy a gömbben van maximum, és az nem lehet a felszínen $\Rightarrow$ az ε sugarú gömb belsejében van $L(\cdot)$ maximuma, ami lok. maximum, jelölje $\hat{\vartheta}_n$, és ekkor $\partial L(\hat{\vartheta}_n) = 0$.

Megadhatók 1 val. $n_1 < n_2 < \dots$ véletlen számokat, hogy $n \geq n_k$ esetén már a ϑ körüli $\frac{1}{2^k}$ sugarú gömbben van $\hat{\vartheta}_n$.

$\Rightarrow \hat{\vartheta}_n \to \vartheta$ 1 val.



Igazoljuk tehát $(*)$-ot!

$$\frac{1}{n} L(\vartheta + h) = \frac{1}{n} L(\vartheta) + \frac{1}{n} \partial L(\vartheta)\, h + \frac{1}{2n} h^T \partial^2 L(\vartheta)\, h + R_n$$

Taylor-sor, $|R_n| \leq \frac{\|h\|^2}{2} \sup_{\|\vartheta' - \vartheta\| \leq \|h\|} \left\| \frac{\partial^2 L(\vartheta')}{n} - \frac{\partial^2 L(\vartheta)}{n} \right\|$ (Taylor-formula)

$$= \frac{\varepsilon^2}{2} \cdot \sup_{\|\vartheta' - \vartheta\| \leq \varepsilon} \underbrace{\left\| \frac{\partial^2 L(\vartheta')}{n} - \frac{\partial^2 L(\vartheta)}{n} \right\|}_{\text{a gömbben a 2. deriváltak ossz.illációja}}$$

H a norma az operátornorma (de ez mindegy mindenkor).

$$\sup_{\|\vartheta' - \vartheta\| \leq \varepsilon} \left\| \frac{1}{n} \partial^2 L(\vartheta') - \frac{1}{n} \partial^2 L(\vartheta) \right\| = \sup_{\|\vartheta' - \vartheta\| \leq \varepsilon} \left\| \frac{1}{n} \sum_{i=1}^n \left(\partial^2 l_i(\vartheta') - \partial^2 l_i(\vartheta) \right) \right\| \leq$$

$L(\vartheta) = \sum_{i=1}^n l_i(\vartheta)$

ftem, azonos eo.

$$\leq \frac{1}{n} \sum_{i=1}^n \underbrace{\sup_{\|\vartheta' - \vartheta\| \leq \varepsilon} \| \partial^2 l_i(\vartheta') - \partial^2 l_i(\vartheta) \|}_{\text{ftem., azonos eo.}}$$

NSZET miatt ez konvergál

i helyett 1 írható

$$\longrightarrow E_\vartheta\left(\sup_{\|\vartheta'-\vartheta\|\le\varepsilon} \|\partial^2 l_i(\vartheta') - \partial^2 l_i(\vartheta)\|\right) =: \eta(\varepsilon)$$

$\lim_{\varepsilon\searrow 0} \eta(\varepsilon) = 0.$ Mert $\partial^2 l_1$ folytonos (RR) miatt, ezért a sup $\to 0$ 1 val., és van integrál majoráns: $2\sup_\vartheta \|\partial^2 l_1(\vartheta)\|$ (különbség normájának becslése), ez (RR) miatt integrál. Dominált konv.-tétel.

$\Rightarrow$ Ha n elég nagy, akkor $\sup_h |R_n| < \varepsilon^2\cdot\eta(\varepsilon)$ becslés a maradéktagra

$$\frac{1}{n}\partial L(\vartheta) = \frac{1}{n}\sum_{i=1}^n \partial l_i(\vartheta) \xrightarrow[\text{NSZET}]{} \underset{(RR)}{E_\vartheta}\left(\partial l_1(\vartheta)\right) = 0$$

$$\Rightarrow \sup_{\|h\|=\varepsilon} \frac{1}{n}\partial L(\vartheta)\cdot h = o(\varepsilon),$$ ezért elég nagy n-re $< \varepsilon^2\cdot\eta(\varepsilon)$,

skalárszorzat

tehát nagyon a becslés jó az 1. tagra is.

most a 2. tag:

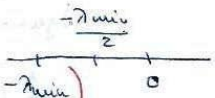
$$\sup_{\|h\|=\varepsilon} h^T \frac{\partial^2 L(\vartheta)}{n} h \le \varepsilon^2\cdot\lambda_{max}\left(\frac{\partial^2 L(\vartheta)}{n}\right)$$

a max. sérülése az argumentumnak

$$\frac{\partial^2 L(\vartheta)}{n} = \frac{1}{n}\sum_{i=1}^n \partial^2 l_i(\vartheta) \xrightarrow[\text{NSZET}]{} \underset{(RR)}{E_\vartheta}\left(\partial^2 l_1(\vartheta)\right) = -I(\vartheta)$$

Mivel λ_{max} folytonos fvénye a mátrixnak, adódik, hogy

$$\lambda_{max}\left(\frac{\partial^2 L(\vartheta)}{n}\right) \longrightarrow \lambda_{max}\left(-I(\vartheta)\right) = -\lambda_{min}\left(\underbrace{I(\vartheta)}_{\text{poz. def.}}\right) < 0$$



$\Rightarrow \forall$ elég nagy n-re $\sup_{\|h\|=\varepsilon} \frac{1}{2} h^T \frac{\partial^2 L(\vartheta)}{n} h < -\frac{\varepsilon^2}{4}\lambda_{min}(I(\vartheta))$

ide erős elég nagy n-re

$$\Rightarrow \sup_{\|h\|=\varepsilon}\left(\frac{1}{n}L(\vartheta+h) - \frac{1}{n}L(\vartheta)\right) < \varepsilon^2\left(2\eta(\varepsilon) - \frac{1}{4}\lambda_{min}\right) < 0,$$ ha

ε elég kicsiny, azaz ha $\eta(\varepsilon) < \frac{1}{8}\lambda_{min}(I(\vartheta))$

(2) Tudjuk, hogy $\hat\vartheta_n \to \vartheta$, $\sqrt{n}(\hat\vartheta_n - \vartheta) = h_n$ -re $\frac{h_n}{\sqrt{n}} \to 0.$

$$0 = \frac{1}{\sqrt{n}} \partial L(\hat{\vartheta}_n)^T \underset{\text{sorfejtés}}{=} \frac{1}{\sqrt{n}} \partial L(\vartheta)^T + \underbrace{\frac{1}{\sqrt{n}} \partial^2 L(\vartheta)(\hat{\vartheta}_n - \vartheta)}_{\frac{\partial^2 L(\vartheta)}{n} \cdot h_n} + R_n^2, \quad \text{ahol}$$

$$\|R_n^2\| \leq \|h_n\| \cdot \sup_{\|\vartheta^* - \vartheta\| \leq \|h_n\| \cdot \frac{1}{\sqrt{n}} = \|\hat{\vartheta}_n - \vartheta\|} \left\| \frac{1}{n} \partial^2 L(\vartheta^*) - \frac{1}{n} \partial^2 L(\vartheta) \right\|$$

Ha n elég nagy, akkor $\|R_n^2\| \leq \|h_n\| \cdot \underbrace{\sup_{\|\vartheta^* - \vartheta\| \leq \varepsilon} \left\| \frac{1}{n} \partial^2 L(\vartheta^*) - \frac{1}{n} \partial^2 L(\vartheta) \right\|}_{< 2 \cdot \eta(\varepsilon), \text{ ha } n \text{ elég nagy}}$

$\Rightarrow \|R_n^2\| \leq \|h_n\| \cdot 2 \cdot \eta(\varepsilon) \quad \forall \varepsilon$ van ε-hoz, amitől ε igaz.

De $\eta(\varepsilon)$ bármilyen kicsi lehet $\rightarrow \|R_n^2\| = o(\|h_n\|)$.

$\frac{1}{n} \partial^2 L(\vartheta) \longrightarrow -I(\vartheta)$ (ezt már láttuk)

$$h_n = \underbrace{\left[-\frac{1}{n} \partial^2 L(\vartheta) \right]^{-1}}_{\xrightarrow{n \to \infty} I(\vartheta)^{-1}} \cdot \left[\frac{1}{\sqrt{n}} \cdot \partial L(\vartheta)^T + \underbrace{R_n^2}_{o(h_n)} \right]$$

$$\frac{1}{\sqrt{n}} \cdot \partial L(\vartheta)^T = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \underbrace{\partial l_i(\vartheta)^T}_{\text{független, azon. eo. véletlen vektorok}}$$

$E_\vartheta(\partial l_i(\vartheta)^T) = 0$ (RR) miatt, $\mathbb{D}_\vartheta(\partial l_i(\vartheta)^T) = I(\vartheta)$

Többdimenziós CHT: $\underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \partial l_i(\vartheta)^T}_{\frac{1}{\sqrt{n}} \partial L(\vartheta)^T} \xrightarrow{d} N_p(0, I(\vartheta))$

$\Rightarrow$ h_n a Cramér–Sverdrup-lemma többdimenziós vektora szerint eo. konv.: $h_n \xrightarrow{d} I(\vartheta)^{-1} \cdot N_p(0, I(\vartheta)) = N_p(0, I(\vartheta)^{-1})$ $\square$

<u>Megj.</u>: Ha még teljesül a

$$(*) \quad \lim_{\delta \searrow 0} \liminf_{n \to \infty} P_\vartheta \left(\sup_{\|a\| \geq \varepsilon} \left(\frac{1}{n} \cdot L(\vartheta + a) - \frac{1}{n} L(\vartheta) \right) < -\delta \right) = 1$$

feltétel, akkor a tétel <u>a</u> ML-becslésre is igaz.

Ez egy teljesen ellenőrizhetetlen feltétel.

Bayes-becslések

Az eddigiek a frekventista hozzáállás.

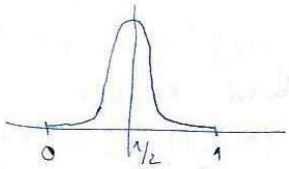
Bayes-féle hozzáállás: a ϑ paraméter is val. változó, értékeit t-vel jelöljük, hogy megkülönböztessük a v.v.-t az értékétől.

Van egy előzetes (a priori) információ a keresett eloszlásról: ← SZUBJEKTÍV!

ϑ eloszlása Q val. mérték a $\Theta \subset \mathbb{R}^p$ Borel-halmazain.

Úgy képzeljük, hogy ez tartalmaz $\forall$ előzetes infót.

Pl. pénzérme valószínűségénél



stb.

Minta: X. Most P_t az X felt. eloszlása a $\vartheta = t$ feltétel mellett. $P_t(B) = P(X \in B \mid \vartheta = t)$.

X (feltétel nélküli) eloszlása az ún. prediktív eloszlás

A TVT-vel $P_Q(B) = P(X \in B) = \int_\Theta P_t(B)\, Q(dt)$

keverékeloszlás

ϑ feltételes eloszlása $X = x$ mellett:

$Q^*(\cdot \mid x)$ a posteriori eloszlás (tapasztalat utáni)

<u>Bayes-tétel.</u> P-m.m. $x \in X$ esetén $Q^*(\cdot \mid x) \ll Q$ és a
(nem Bayes bz.)

$$\frac{dQ^*(\cdot \mid x)}{dQ}(t) = \frac{f_t(x)}{f_Q(x)},$$

ahol P_t sűrűsége f_t (valamilyen λ-ra),

P_Q sűrűsége f_Q (nagyonnan a λ-ra).

$\Rightarrow f_Q(x) = \int_\Theta f_t(x)\, Q(dt)$. a fentiek szerint.

Mj. Spec. ha Q-nak van (pl. Lebesgue) sűrűsége: $q(t)$, akkor $Q^*(\cdot \mid x)$-nek is, és annak a sűrűsége $q^*(t \mid x) = \dfrac{f_t(x)\, q(t)}{f_Q(x)}$.

Mj. Ha $\mathcal{X}$ és Ω diszkrét, és a sűrűségfv.-eket a számlálómértékre vonatkozóan nézzük, akkor visszakapjuk a közönséges Bayes-tételt.

Stat. mező helyett egyetlen val. mezőnk van, X és ϑ a két val. vektor. Ezek együttes eloszlása az $\mathcal{X} \times \Omega$ mintatéren (a mintat. σ-algebrával): $B \subset \mathcal{X}$, $C \subset \Theta$ mérhetőre $\mathbb{P}(B \times C) =$

$$= \int_C P_t(B)\, Q(dt)$$

ez nem ugyanaz, mint a P

Ez kiterjed egyértelműen a mintat. σ-algebrára $\rightarrow$ általánosított mintatmértéket.

<u>Def.</u> $g(\vartheta)$ Bayes-becslése az a (illetve: egy olyan) $T(X)$ statisztika, melyre $\mathbb{E}\left(\| T(X) - g(\vartheta) \|^2\right)$ minimális.

Tudjuk, hogy van minimalizáló T és ez egyértelmű: ez épp a feltételes várható érték, hiszen itt L^2-távolság minimumáról van szó, és a felt. v.é. épp egy L^2-vetítés.

$\Rightarrow$ a $\hat{g}(X)$ Bayes-becslésre $\hat{g}(X) = \mathbb{E}(g(\vartheta) \mid X)$, azaz $\hat{g}(X)$ a $g(\vartheta)$ a posteriori várható értéke, azaz

$$\hat{g}(X) = \int_\Theta g(t)\, Q^*(dt \mid x) = \frac{\int_\Theta g(t)\, f_t(x)\, Q(dt)}{\underbrace{\int_\Theta f_t(x)\, Q(dt)}_{f_Q(x)}}$$

$\Rightarrow \hat{g}$ az a posteriori várható érték.

<u>Áll.</u> A Bayes-becslés az elégs. stat. fv-ye 2017.04.05.

B: Már az a posteriori becslés is az egs. stat. fv-ye a fakt. tétel miatt:

$$f_t(x) = g_t(T(x)) \cdot h(x) \qquad (T \text{ egs.})$$

$$\Rightarrow f_Q(x) = h(x) \cdot \int_\Theta g_t(T(x))\, Q(dt)$$

$$\Rightarrow \frac{dQ^*(\cdot \mid x)}{dQ}(t) = \frac{f_t(x)}{f_Q(x)} = \frac{g_t(T(x))}{\int_\Theta g_s(T(x))\, Q(ds)}$$

$\square$

<u>Egyértelműség</u> $\hat{g}(X) = \mathbb{E}(g(\vartheta) \mid X)$ P_Q-m.m. egyértelmű

Ahol: $P_Q(B) = \int_\Theta P_t(B)\, Q(dt) \Rightarrow P_Q$-m.m. $= Q$-m.m. t-re P_t-m.m.

$\mathcal{P}$-m.m. $= \forall t$-re P_t-m.m.

Ezek nem mindig ugyanazok, de ugyanaz igen (Q-függő)

Megengedhető-e? Legyen $\hat{g}(X)$ Bayes-becslés, T egyenletesen jobb nála

Legyen S tetszőleges becslés, $\hat{g}(X)$ minimalizálja az

$R_S(Q) = \mathbb{E}\left(\|S(X) - g(\vartheta)\|^2\right)$ ún. a priori rizikót, ezt a minimumot bayesi rizikónak nevezik.

TVÉT: $$R_S(Q) = \mathbb{E}\left(\mathbb{E}\left(\|S(X) - g(\vartheta)\|^2 \mid \vartheta\right)\right) =$$

$$= \int_\Theta \mathbb{E}\left(\|S(X) - g(t)\|^2 \mid \vartheta = t\right) Q(dt)$$

$$= \int_\Theta \underbrace{\mathbb{E}_t\left(\|S(X) - g(t)\|^2\right)}_{R_S(t) \text{ négyzetes rizikófv.}} Q(dt),$$

tehát $$R_S(Q) = \int_\Theta R_S(t)\, Q(dt)$$

Ez filozóficilag is elfogadhatóvá teszi a Bayes-becslést: összehasonlíthatóvá tesz fügvényeket. (de a priori becslés meljérhitelére persze marad.)

Visszatérve a megengedhetőséghez: $R_T(t) \leq R_{\hat{g}}(t)$

$\Rightarrow R_T(Q) \leq R_{\hat{g}}(t)$. $<$ nem lehet, ezért $=$

$\Rightarrow$ T is Bayes-becslés

Ezért ha Bayes-becslés egyértelmű, akkor megengedhető.

Torzítatlanság

Áll: A Bayes-becslés — triv. esetből eltekintve — nem torzítatlan.

Ellenpélda: diszjunkt tartójúeloszlásnál a minta megmondja, hogy pont melyik eo.

A trivialitást így fogalmazhatjuk meg:

Q-m.m. t_1, t_2 esetén $g(t_1) \neq g(t_2) \Rightarrow P_{t_1} \perp P_{t_2}$

B: $\hat{g}(X) = \mathbb{E}(g(\vartheta) \mid X)$ Bayes-becslés, tfh. torzítatlan,

azaz $\forall t$ $\mathbb{E}_t(\hat{g}(X)) = g(t)$, vagyis $\mathbb{E}(\hat{g}(X) \mid \vartheta) = g(\vartheta)$ 1 val.

$$\mathbb{E}\left(\|\hat{g}(X) - g(\vartheta)\|^2\right) = \mathbb{E}\left(\|\hat{g}(X)\|^2\right) + \mathbb{E}\left(\|g(\vartheta)\|^2\right) - 2\,\mathbb{E}\left(\hat{g}(X)^T g(\vartheta)\right)$$

$$\mathbb{E}\left(\hat{g}(X)^T g(\vartheta)\right) \overset{\text{TVÉT}}{=} \mathbb{E}\left(\mathbb{E}\left(\hat{g}(X)^T g(\vartheta) \mid X\right)\right) =$$

$$= \mathbb{E}\left(\hat{g}(X)^T \cdot \mathbb{E}(g(\vartheta) \mid X)\right) =$$

$$= \mathbb{E}\left(\hat{g}(X)^T \cdot \hat{g}(X)\right) = \mathbb{E}\left(\|\hat{g}(X)\|^2\right)$$

$$\mathbb{E}\left(g(X)^T g(\vartheta)\right) \overset{\text{TVÉT}}{=} \mathbb{E}\left(\mathbb{E}\left(\hat{g}(X)^T g(\vartheta) \mid \vartheta\right)\right) =$$

$$= \mathbb{E}\left(\mathbb{E}\left(\hat{g}(X)^T \mid \vartheta\right) \cdot g(\vartheta)\right) \underset{\text{torzítatlanság}}{=} \mathbb{E}\left(g(\vartheta)^T \cdot g(\vartheta)\right) = \mathbb{E}\left(\|g(\vartheta)\|^2\right)$$

$\Rightarrow E\left(\|\hat{g}(X) - g(\theta)\|^2\right) = 0 \quad \Rightarrow \quad \hat{g}(X) = g(\theta) \quad \mathbb{P}$-m.m.

$\Rightarrow$ Q-m.m. t esetén $\hat{g}(x) = g(t)$ $\quad P_t$-m.m. x-re.

$\Rightarrow$ ha $g(t_1) \neq g(t_2)$, akkor $P_{t_1} \perp P_{t_2}$

<u>Aszimptotikus viselkedés</u>

<u>Áll.</u> (BN) (RR) + (*) és az a priori eloszlás Lebesgue-abszt. folyt. Θ-n, $q(t) > 0$, $\int_\Theta \|t\| \cdot q(t)\, dt < \infty$. Akkor a $\hat{\vartheta}_n$ Bayes-becslésre (n elemű mintából számolt):

$$\sqrt{n} \cdot \left(\hat{\vartheta}_n - \vartheta\right)\Big|_{\vartheta = t \text{ feltétel}} \xrightarrow{d} N\left(0, I_1(t)^{-1}\right)$$

megfelelő dimenziós normális

igaz m.m. t-re.

<u>Áll.</u> Ha a $\hat{g}(X)$ Bayes-becslés várható vesztesége konstans, akkor a Bayes-becslés minimax.

B: Legyen T tetszőleges becslés.

Bayes minimalitása az a priori várhatót

$$\sup_t R_{\hat{g}}(t) = \int_\Theta R_{\hat{g}}(t)\, Q(dt) \leq \int_\Theta R_T(t)\, Q(dt) \leq \sup_t R_T(t)$$

Hipotézisvizsgálat

$\mathcal{P} = \{ P_\vartheta : \vartheta \in \Theta \}$

Hipotézis: $\mathcal{P}_0 \subsetneq \mathcal{P}$, $\mathcal{P}_0 \neq \emptyset$

Ekvivalens: $\Theta_0 \subsetneq \Theta$, $\Theta_0 \neq \emptyset$ -ra $\mathcal{P}_0 = \{ P_\vartheta : \vartheta \in \Theta_0 \}$.

Nullhipotézis: $\vartheta \in \Theta_0$, azaz $P \in \mathcal{P}_0$. Jel.: H_0 : Ez szépíszelő az elképzelésünket.

Ellenhipotézis: $\vartheta \in \Theta_1 = \Theta \setminus \Theta_0$, azaz $P \in \mathcal{P}_1 = \mathcal{P} \setminus \mathcal{P}_0$, Jel: H_1.

Mind a null-, mind az ellenhipotézist meg kell mondani.

Minta alapján döntünk: elfogadjuk vagy elvetjük H_0-t.

$\mathcal{X} = \mathcal{X}_0 \amalg \mathcal{X}_1$, $\mathcal{X}_0$ az elfogadási tartomány: $X \in \mathcal{X}_0$-ra elfogadjuk H_0-t,

$\mathcal{X}_1$ a kritikus tartomány: $X \in \mathcal{X}_1$-re elvetjük H_0-t.

A hibás döntés típusai:

1) H_0 igaz, de elvetettük $\longrightarrow$ elsőfajú hiba

Ezt kis esélyel szeretjük elkövetni, mert H_0 az, amit elképzeltünk.

2) H_0 hamis, de elfogadtuk $\longrightarrow$ másodfajú hiba

Ennek se örülünk, de inkább ezt engedjük.

A két hiba valószínűségének csökkentése csak egymás rovására megy.

A döntést hívjuk (statisztikai) próbának.

$(\mathcal{X}_0, \mathcal{X}_1)$ próba esetén $\psi(\vartheta) = P_\vartheta$ (elvetjük H_0-t) $= P_\vartheta(\mathcal{X}_1)$ eltolásos fv.

$\vartheta \in \Theta_0$-ra $\psi(\vartheta)$ az elsőfajú hiba valószínűsége.

$\vartheta \in \Theta_1$-re $\psi(\vartheta)$ a jó döntés valószínűsége, $1 - \psi(\vartheta)$ a másodfajú hiba valsége. Ez a próba ereje ϑ-ban.

$\psi|_{\Theta_1}$ erőfüggvény

A próba terjedelme $\sup_{\Theta_0} \psi(\vartheta)$.
(pontos terjedelem)

Ennek nagyon felső korlát lenni, pl. $\alpha = 0{,}05$

A próba mérete: 1 — terjedelem, gyakran %-ban.

Próbák összehasonlítása:

φ_1, φ_2 két próba, ψ_1, ψ_2 erőfvüket

Ekkor φ_1 egyenletesen erősebb φ_2-nél, ha $\psi_1(\vartheta) \geq \psi_2(\vartheta)$ $\forall \vartheta \in \Theta_1$.
(nem gyengébb)

A φ_1 és φ_2 ekvivalensek, ha az eltolásos fvüket $\forall \vartheta \in \Theta$-ra megegyezik.

Legyen $\mathcal{D}$ próbák egy osztálya. (pl. $\leq \alpha$ terjedelmű próbák)

$\varphi \in \mathcal{D}$ egyenletesen legerősebb (optimális), ha egyenletesen erősebb minden $\mathcal{D}$-beli próbánál.

Egy φ próba torzítatlan, ha φ ereje $\geq$ φ terjedelme

Olyan próbát akarunk csinálni, amit, ha csak lehet, elfogadjuk H_0-t, el akarjuk kerülni az elsőfajú hibát. Akkor kell elvetni H_0-t, ha az eltérés már szignifikáns. Ilyenkor az információ abban van, hogy elvetjük.

Általánosítás: véletlenített (randomizált) próbák.

Próba: $\varphi : \mathcal{X} \longrightarrow [0,1]$ statisztika (próbafv.)

Ha X a minta, $\varphi(X)$ mondja meg, milyen valószínűséggel vetjük el H_0-t.

Mire jó ez? Ettől nem lesz több információ. Ez arra használható, hogy a teljes megengedett terjedelmet elérhessük. (diszkrét eset)

Az eddigi próbák determinisztikusak, ezeknél $\varphi = I(\mathcal{X}_1)$ indikátor volt.

Elutasítás fv.: $\psi(\vartheta) = P_\vartheta(\text{elvetés}) = E_\vartheta(\underbrace{P_\vartheta(\text{elvetés} \mid X)}_{\varphi(X)})$,

tehát $\psi(\vartheta) = E_\vartheta(\varphi(X))$

Innen rögtön minden fenti definíció működik az általánosított esetben is.

Áll. Minden próbához $\exists$ vele ekvivalens próba, mely az elégséges stat függvénye.

B: $\varphi'(X) := E_\vartheta(\varphi \mid T)$ próbafv. (mert a várh. érték is $[0,1]$-beli lesz), ez a T egy. stat fvye., $E_\vartheta(\varphi') = E_\vartheta(\varphi)$ $\forall \vartheta$. $\square$

A próbák gyakran olyan alakúak, hogy van egy T egyelőre próbastatisztika és $c = c_\alpha$ krit. érték, és elvetjük H_0-t, ha $T(X) > c_\alpha$. $T(X)$ méri, hogy mennyire tértünk el H_0-tól.

Def. A p-érték annak a H_0 melletti valsége, hogy $T(X)$ legalább akkora, mint amit kaptunk. Ha ez $< \alpha$, akkor elvetjük H_0-t.

Def. Egy hipotézis (H_0 vagy H_1) egyszerű, ha egyetlen eloszlás van benne. Különben összetett.

Legegyszerűbb feladat: egyszerű H_0 és egyszerű H_1.

$\longrightarrow \mathcal{P} = \{P_0, P_1\}$, $H_0 = \{P_0\}$, $H_1 = \{P_1\}$.

Ez dominált, pl. $P_0 + P_1$ dominál. $\longrightarrow$ λ-ra véve a sűrűségek f_0, f_1.

Tel. $T(x) = \dfrac{f_1(x)}{f_0(x)}$ likelihood-hányados-statisztika.

Lemma. (Neyman–Pearson)

(1) Legyen $\varphi = \begin{cases} 1, & \text{ha } T(X) > c \\ \text{tetsz.}, & \text{ha } T(x) = c, \\ 0, & \text{ha } T(X) < c \end{cases}$ ahol a tetszőleges $[0,1]$-ben értékű.

Ez a próba legerősebb a nála nem nagyobb terjedelműek között.

(2) $\forall \alpha \in (0,1)$ $\exists c_\alpha \geq 0$, $p_\alpha \in [0,1]$, hogy a

$$\varphi(T) = \begin{cases} 1, & \text{ha } T > c_\alpha \\ p_\alpha, & \text{ha } T = c_\alpha \\ 0, & \text{ha } T < c_\alpha \end{cases}$$

próba terjedelme α.

(3) Tfh. φ' legerősebb a legfeljebb α terjedelműek között.

Ekkor m.m. x-re olyan, mint (1)-ben.

Továbbá φ' terjedelme pontosan α, vagy az ereje 1.

2017.04.19.

B: (1) Jelölje φ terjedelmét α, erejét ψ

legyen φ' próba, $\alpha' \leq \alpha$, ereje ψ'.

$$\psi - \psi' = \int_{\mathcal{X}} \varphi f_1 \, d\lambda - \int_{\mathcal{X}} \varphi' f_1 \, d\lambda = \int_{\mathcal{X}} (\varphi - \varphi') f_1 \, d\lambda = *$$

$\mathcal{X} = \mathcal{X}_+ \cup \mathcal{X}_0 \cup \mathcal{X}_-$, ahol $\mathcal{X}_+ = \{x \in \mathcal{X} \mid f_1(x) > c f_0(x)\}$

$\mathcal{X}_0 = \{x \in \mathcal{X} \mid f_1(x) = c f_0(x)\}$

$\mathcal{X}_- = \{x \in \mathcal{X} \mid f_1(x) < c f_0(x)\}$

$$* = \int_{\mathcal{X}_+} + \int_{\mathcal{X}_0} + \int_{\mathcal{X}_-} \geqslant$$

$$\overset{1.}{\geqslant} \int_{\mathcal{X}_+} (1-\varphi') \overset{c f_0}{\cancel{f_1}} \, d\lambda + \int_{\mathcal{X}_0} (\varphi - \varphi') c f_0 \, d\lambda + \int_{\mathcal{X}_-} (0 - \varphi') c \cdot f_0 \, d\lambda =$$

$$= c \int_{\mathcal{X}} (\varphi - \varphi') f_0 \, d\lambda = c \cdot (\alpha - \alpha') \overset{2.}{\geqslant} 0 \quad \Rightarrow \varphi \text{ ereje nagyobb-egyenlő}$$

✓

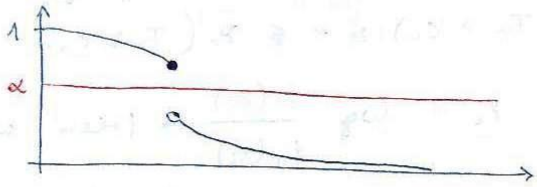
(2) $t \mapsto P_0\left(\frac{f_1(x)}{f_0(x)} \geqslant t\right)$ $\quad t=0$-ban 1

$t \to \infty$-ben 0

a nevező 1 val pozitív (v.v. a sűrűségben)

balról folytonos

monoton fogyó



$$c := \sup\left\{ t \,\middle|\, P_0\Big(\underbrace{\frac{f_1(x)}{f_0(x)}}_{T} \geqslant t\Big) > \alpha \right\}$$

Ha most $P_0(T=c) > 0 \longrightarrow$ legyen $p := \dfrac{\alpha - P_0(T>c)}{\underbrace{P_0(T\geqslant c)}_{>\alpha} - P_0(T>c)} \in [0,1]$,

máskülönben legyen $p := 0$.

Ekkor $E_0(\varphi) = \underbrace{P_0(T>c) + P_0(T=c)\,p}_{=\alpha}$ az elvetés valsége mindkét esetben. ✓

(3) Tfh. φ' legf. α terjedelmű optimális (azaz legerősebb) próba. Legyen φ a (2)-ben kapott pontosan α terjedelmű próba; ez (1) miatt legerősebb a legf. α terjedelműek között.

$\Rightarrow \psi = \psi'$

Az (1) bizonyításában lévő egyenlőtlenségsorban mindenütt egyenlőség kell álljon.

1. egyenlőtlenség:
$\Rightarrow$ $\mathcal{X}_+$-on λ-m.m. $\varphi' = 1$,
$\mathcal{X}_-$-on λ-m.m. $\varphi' = 0$.
$\Rightarrow$ φ' m.m. x-re olyan, mint az (1)-beli képlet mondja.

2. egyenlőtlenség: $\alpha = \alpha'$ vagy $c = 0$

$\downarrow$ a terjedelem α ✓

$\downarrow$ H_1 mellett $P_1(f_1(x) > 0) = 1 \to 1$ val. elvetünk. De $\frac{f_1(x)}{f_0(x)} > 0$. $\Rightarrow$ az erő 1.

✓ $\square$

Aszimptotikus elemzés

n elemű minta, $T_n = \dfrac{\prod_{i=1}^{n} f_1(X_i)}{\prod_{i=1}^{n} f_0(X_i)}$, c_n krit. érték

$$P_0(T_n > c_n) \leq \alpha \leq P_0(T_n \geq c_n)$$

Jel. $Y_i = \log \dfrac{f_0(X_i)}{f_1(X_i)}$ → fteen azonos eo.

felül negatív logaritmust veszünk

$$P_0\left(\sum_{i=1}^{n} Y_i < -\log c_n\right) \leq \alpha \leq P_0\left(\sum_{i=1}^{n} Y_i \leq -\log c_n\right)$$

Def.

Tfh. $D_0^2(Y_1) = \sigma^2 < \infty$

$$E_0(Y_i) = \int_{\mathcal{X}} \log \frac{f_0}{f_1} f_0 \, d\lambda \overset{\text{jel}}{=} D(P_0 \| P_1)$$

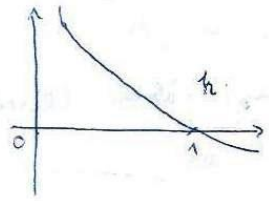
Et P_0 és P_1 (információs) divergenciája,

a relatív entrópia,

ill. P_0 és P_1 Kullback–Leibler-távolsága. (Ezek mind elnevezések)

Állítás. Fennáll, hogy $D(P_0 \| P_1) \geq 0$ és $0 \iff P_0 = P_1$

B. $h(x) = -\log x$ szig. konvex,



$$D(P_0 \| P_1) = E_0\left(h\left(\frac{f_1(X)}{f_0(X)}\right)\right) \overset{\text{Jensen}}{\geq}$$

$$\geq h\left(E_0\left(\frac{f_1(X)}{f_0(X)}\right)\right) =$$

$$= h\left(\int_{\{f_0>0\}} \frac{f_1}{f_0} f_0 \, d\lambda\right) = h\Big(\underbrace{\int_{\{f_0>0\}} f_1 \, d\lambda}_{\leq 1}\Big) \geq 0$$

Egyenlőség $\iff \dfrac{f_1(X)}{f_0(X)} =$ konstans P_0-m.m., $f_1(x) = c \cdot f_0(x)$,

$$\int_{\{f_0>0\}} \frac{f_1}{f_0} f_0 \, d\lambda = 1 \Rightarrow c = 1 \Rightarrow f_1 = f_0.$$

□

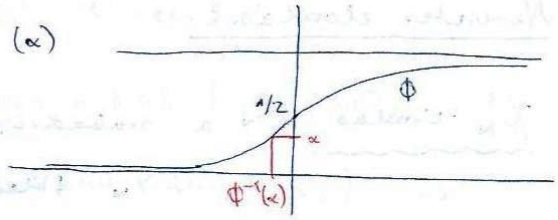
$$P_0\left(\underbrace{\frac{\sum_{i=1}^{n} Y_i - n\, D(P_0 \| P_1)}{\sqrt{n}\,\sigma}}_{\xrightarrow{d}\, N(0,1)} < \frac{-\log c_n - n\, D(P_0 \| P_1)}{\sqrt{n}\,\sigma}\right) \leq \alpha \leq P(\cdots \leq \cdots)$$

Ezt zárójel standardizálással.

Emiatt:

$$P\left(\frac{\sum Y_i - n\, D(P_0 \| P_1)}{\sqrt{n}\,\sigma} \underset{(\leq)}{<} t_n\right) \longrightarrow \Phi(t) \iff t_n \longrightarrow t$$

$$\Rightarrow \quad \frac{-\log C_n - n\,D(P_0 \| P_1)}{\sqrt{n}\,\sigma} \longrightarrow \Phi^{-1}(\alpha)$$



ugy. $\Phi(-t) = 1 - \Phi(t)$

$\Rightarrow \Phi^{-1}(\alpha) = -\Phi^{-1}(1-\alpha)$

$$\Longleftrightarrow \quad \frac{-\log C_n - n\,D(P_0 \| P_1)}{\sqrt{n}\,\sigma} = -\Phi^{-1}(1-\alpha) + o(1)$$

Áll.

$$C_n = \exp\left(-n\,D(P_0 \| P_1) + \sqrt{n}\,\sigma\,\Phi^{-1}(1-\alpha) + o(\sqrt{n})\right),$$ és

bármilyen ilyen alakú c_n sorozattal a próba terjedelme $n \to \infty$ mellett $\longrightarrow \alpha$.

Konkrét esetben D és σ számítható, $o(\sqrt{n})$-et elhagyjuk $\Rightarrow$ α-hoz tartó c_n.

Áll. A likelihood-hányados-próba ereje (ψ_n) exp. sebességgel tart 1-hez (tehát a próba konzisztens).

B:

$$\underbrace{1-\psi_n}_{\text{másodfajú hiba val.}} = P_1(\text{elfogadás}) \leq P_1\left(\frac{\prod_{i=1}^n f_1(X_i)}{\prod_{i=1}^n f_0(X_i)} \leq c_n\right) =$$

$$= P_1\left(\prod_{i=1}^n \underbrace{\frac{f_0(X_i)}{f_1(X_i)}}_{\text{ftemp}} \geq \frac{1}{c_n}\right) \overset{\text{Markov}}{\leq} E_1\left(\prod_{i=1}^n \frac{f_0(X_i)}{f_1(X_i)}\right) \cdot \frac{1}{1/c_n} = c_n \cdot \left(E_1\left(\frac{f_0(X_1)}{f_1(X_1)}\right)\right)^n =$$

$$= c_n \cdot \left(\int \frac{f_0}{f_1} \cdot f_1 \, d\lambda\right)^n \leq c_n \cdot 1^n = c_n = \underbrace{e^{-n\,D(P_0 \| P_1) + o(n)}}_{\to 0 \text{ exp.}}$$

$\square$

Ennyi a Neyman–Pearson-lemmáról.

Nevezetes eloszlások

χ^2_N-eloszlás: N a szabadsági fok

$Y_1, \dots, Y_N$ ftlen $N(0,1)$

$Y_1^2 + \dots + Y_N^2 \sim \chi^2_N$ (centrális) χ^2_N-eo.

$\chi^2_n * \chi^2_m = \chi^2_{n+m}$ definícióból

$\chi^2_1 = N(0,1)^2$ $\quad P(N(0,1)^2 < t) = P(-\sqrt{t} < N(0,1) < \sqrt{t}) =$

$= \Phi(\sqrt{t}) - \Phi(-\sqrt{t}) = 2\Phi(\sqrt{t}) - 1$ eofs.

Sfs.: $f_1(t) = 2\varphi(\sqrt{t}) \cdot \frac{1}{2\sqrt{t}} = \frac{1}{\sqrt{2\pi t}} e^{-t/2} \sim \Gamma_{1/2, 1/2}$

$\Rightarrow \chi^2_n = \Gamma_{n/2, 1/2}$

t-eloszlás (Student-eo.)

W. S. Gosset (1876–1937) minőségellenőrzésre fejlesztette ki, Student néven publikálta (1908)

$X \sim N(0,1)$
$Y \sim \chi^2_n$ függetlenek $\longrightarrow$ $\dfrac{X}{\sqrt{\frac{Y}{n}}}$ n szabadságfokú t-eloszlás: t_n

Áll. t_n szimmetrikus

B: $-X$ is $N(0,1)$ $\Rightarrow$ $t_n \overset{d}{=} -t_n$ □

Áll. $t_n \to N(0,1)$, ha $n \to \infty$.

B: $\frac{Y}{n} \to 1$ NSZET miatt.

Cramér–Sluckij $\Rightarrow$ a határeloszlás tart. □

$t_1 = \dfrac{N(0,1)}{|N(0,1)|}$ $\Rightarrow$ t_1 = Cauchy-eo.

F-eloszlás

$X \sim \chi^2_n$
$Y \sim \chi^2_m$ ftlenek $\Rightarrow$ $\dfrac{\frac{X}{n}}{\frac{Y}{m}}$ eloszlása (n,m) szabadságfokú F-eloszlás (Fisher)

Áll. $\dfrac{1}{F_{n,m}} = F_{m,n}$ □

Áll. $X \sim N_n(m, E_n)$, akkor a koordinátái $N(m_i, 1)$ eloszlásúak.

Ezért $\|X\|^2$ eloszlása csak $\|m\|^2$-en keresztül függ m-től.

Def. $\|X\|^2$ n szab. fokú, $\lambda = \|m\|^2$ nemcentralitási paraméterű nemcentrális χ^2-eloszlás: $\chi^2_n[\lambda]$.

Spec. $\lambda = 0$-ra ez a szükséges (centrális) χ^2-eo.

B: Legyen U ortogonális mátrix, hogy $Um = \|m\| \cdot (1, 0, \dots, 0)^T$ (U lehetséges). Ekkor $UX \sim N_n(\|m\| \cdot (1, 0, \dots, 0)^T, E_n) \Rightarrow$
$\Rightarrow \|UX\|^2 = \|X\|^2$. □

Áll. $X \sim N_n(m, \sigma^2 E_n)$. $\mathbb{R}^n = \mathcal{L}_1 \oplus \dots \oplus \mathcal{L}_k$ páronként ortogonális alterek, $\dim \mathcal{L}_i = r_i$. Az $\mathcal{L}_i$-re való merőleges vetítés P_i.
$X = X^{(1)} + \dots + X^{(k)}$ ort. felbontás, ahol $X^{(i)} = P_i(X)$. Ekkor:

(1) $X^{(i)} \sim N_n(m^{(i)}, \sigma^2 P_i)$ ftlenek

(2) $\|X^{(i)}\| \sim \sigma^2 \chi^2_{r_i}\left[\frac{\|m^{(i)}\|^2}{\sigma^2}\right]$

B: (1) $X^{(i)} = N_n(P_i m, \sigma^2 P_i P_i^T)$

P_i szimmetrikus és idempotens ($P_i = P_i^T$ és $P_i^2 = P_i$)
$\Rightarrow X^{(i)} = N_n(m^{(i)}, \sigma^2 P_i)$

$X^{(1)}, \dots, X^{(k)}$ együttesen is normális, mert X lin. transzformáltja. Ezért elég belátni, hogy korrelálatlanok.

$$\operatorname{cov}(X^{(i)}, X^{(j)}) = \operatorname{cov}(P_i X, P_j X) =$$
$$= P_i \operatorname{cov}(X, X) P_j^T = P_i \cdot \sigma^2 E_n \cdot P_j^T =$$
$$= \sigma^2 \cdot P_i P_j = \sigma^2 \cdot 0 = 0, \text{ ha } i \neq j.$$ ✓

(2) Ftt. $\sigma = 1$: az általános eset visszavezethető $\frac{X}{\sigma}$-ra. 2017.04.26.

Legyen $e_1, \dots, e_r$ ONB $\mathcal{L}_i$-ben $\Rightarrow Q = [e_1, \dots, e_r] \in \mathbb{R}^{n \times r}$
(A továbbiakban az i indexet elhagyjuk, az i betűt másra használjuk)
$\Rightarrow Q^T Q = E_r$ és $P = QQ^T = \sum_{i=1}^{r} e_i e_i^T$,

mert $\forall i$: $Pe_i = e_i$, és $a \perp \mathcal{L}$-re $Pa = \sum_{i=1}^{r} e_i \underbrace{e_i^T a}_{0} = 0$,
és a vetítést pont ezek szorzatának.

$$\|PX\|^2 = X^T \cdot P^T P \cdot X = X^T \cdot Q \underbrace{Q^T Q}_{E} Q^T X = \|Q^T X\|^2$$

$Q^T X$ is norm. eo., $Q^T X \sim N_r(Q^T m, \underbrace{Q^T Q}_{E})$

$\Rightarrow \|Q^T X\|^2 \sim \chi^2_r\left[\underbrace{\|Q^T m\|^2}_{\|Pm\|^2}\right]$, és épp ez kellett. □

Köv. (Fisher – Bartlett-tétel) $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ minta. Ekkor

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{\mu}\right), \quad s_n^{*2} \sim \frac{\sigma^2}{n-1}\chi^2_{n-1}, \text{ és ez a két stat ftlen.}$$

(F–B)

Mj. Ez a ftlenség karakterizálja is a normális eloszlást. (BN)

B: $\underline{X} = (X_1, \dots, X_n)^T \sim N(\mu e, \sigma^2 \mathbb{E})$, ahol $e = (1, \dots, 1)^T \in \mathbb{R}^n$

$\mathbb{R}^n = \mathcal{L}_1 \oplus \mathcal{L}_2$, ahol $\mathcal{L}_1 = \langle e \rangle$, $\mathcal{L}_2 = \mathcal{L}_1^\perp$,

$\dim \mathcal{L}_1 = 1$, $\dim \mathcal{L}_2 = n-1$.

$X = X^{(1)} + X^{(2)}$ ortogon. felbontás

$\mathcal{L}_1$-re való vetítés: $P = \frac{ee^T}{n}$, mert $\mathcal{L}_1$-ben $\frac{e}{\sqrt{n}}$ ONB,

így $X^{(1)} = PX = e \cdot \frac{1}{n} \cdot e^T X = e \cdot \overline{X} = \overline{X} \cdot e$

$\Rightarrow X^{(2)} = X - X^{(1)} = X - \overline{X} e$

$\Rightarrow \|X^{(2)}\|^2 = \sum_{i=1}^n (X_i - \overline{X})^2 = (n-1)\, s_n^{*2}$

$X^{(1)}$ és $X^{(2)}$ ftlen $\Rightarrow$ $\overline{X}$ és s_n^{*2} ftlen.

Továbbá $\|X^{(2)}\|^2 \sim \sigma^2 \chi^2_{n-1}$ centrális, mert: $\mu e \in \mathcal{L}_1$ a várh. érték

$\Rightarrow \mu e \perp \mathcal{L}_2 \Rightarrow (\mu e)^{(2)} = 0 \Rightarrow$ centralitás.

A normális eloszlás paramétereire vonatkozó klasszikus próbák

u-próbák (z-próbák)

Egymintás: $X_1, \dots, X_n \sim N(\mu, \sigma^2)$, σ ismert, $\mu \in \mathbb{R}$ paraméter

- $H_0: \mu = \mu_0$ $\quad H_1: \mu \neq \mu_0$ kétoldali ellenhipotézis
- $H_0: \mu = \mu_0$, vagy $H_0^>: \mu \leq \mu_0$, $\quad H_1^>: \mu > \mu_0$ - egyoldali ehpt.

$$T := \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{\overline{X} - \mu_0}{\sigma}\sqrt{n} \underset{H_0}{\sim} N(0,1)$$

H_1 esetén is normális T, de $n \to \infty$ mellett elszáll a $\pm\infty$-be

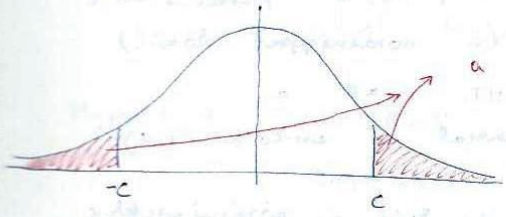
Próba: elfogad. H_0-t, ha

- egyoldali esetben: $T > c_{krit}$, ahol $c_{krit} = c_\alpha$, α a terj.

$\Rightarrow c_\alpha =$ az $N(0,1)$ eo. $(1-\alpha)$-kvantilise, tehát

$c_\alpha = \Phi^{-1}(1-\alpha)$

• kétoldali esetben: itt a tól nagy és a tól kicsi is baj,



a piros részek együtt α területűek

$\Rightarrow$ $|T| > c_{krit}$ -re vetünk el, $c_{krit} = c_{\alpha/2} = \Phi^{-1}(1-\frac{\alpha}{2})$

<u>Kétmintás</u>: $\left.\begin{array}{l} X_1, \dots, X_{n_1} \sim N(\mu_1, \sigma_1^2) \\ Y_1, \dots, Y_{n_2} \sim N(\mu_2, \sigma_2^2) \end{array}\right\}$ függetlenek $\quad \sigma_1, \sigma_2$ ismert, $\mu_1, \mu_2 \in \mathbb{R}$ paraméter

• Kétoldali: $H_0: \mu_1 = \mu_2 \qquad H_1: \mu_1 \neq \mu_2$

• Egyoldali: $H_0: \mu_1 = \mu_2 \qquad H_1': \mu_1 > \mu_2$

vagy $H_0': \mu_1 \leq \mu_2$

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \underset{H_0}{\sim} N(0,1)$$

standardizálás

Próba: mint az egymintás esetben.

• Egyoldali: elvetés, ha $T > c_\alpha = \Phi^{-1}(1-\alpha)$

• Kétoldali: elvetés, ha $|T| > c_{\alpha/2} = \Phi^{-1}(1-\frac{\alpha}{2})$

t-próbák

<u>Egymintás</u>: $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ $\quad \sigma > 0$, $\mu \in \mathbb{R}$ paraméterek

Hipotézisek mint az u-próbánál

$$T = \frac{\overline{X} - \mu_0}{\sqrt{S_n^{*2}}} \sqrt{n} = \underbrace{\left(\frac{\overline{X} - \mu_0}{\sigma} \sqrt{n}\right)}_{\underset{H_0}{\sim} N(0,1)} \cdot \frac{1}{\sqrt{\frac{S_n^{*2}}{\sigma^2}}}$$

ez csak figyelemfelhívás

$\frac{S_n^{*2}}{\sigma^2} \underset{H_0}{\sim} \frac{\chi^2_{n-1}}{n-1}$

Az első és a második tényező ftlen (FB-tétel)

$\Downarrow$

$T \underset{H_0}{\sim} t_{n-1}$

Próba: • Egyoldali: elvetés, ha $T > c_{krit}$, ahol c_{krit} a t_{n-1}-eo. $(1-\alpha)$-kvantilise, ez c_α.

• Kétoldali: elvetés, ha $|T| > c_{krit}$ (t_{n-1} szimmetrikus eo.), ahol $c_{krit} = c_{\alpha/2}$

Kétmintás: $X_1, \dots, X_{n_1} \sim N(\mu_1, \sigma^2)$
$Y_1, \dots, Y_{n_2} \sim N(\mu_2, \sigma^2)$ } ffvek, μ_1, μ_2, σ paraméterek (a szórásnégyzet közös!)

Sok esetben $\sigma_1^2 = \sigma_2^2$ természetes, pl. a CHT miatt normális eo., és ha a mérések ugyanazzal az eszközzel mérjük, akkor valószínűleg a szórás ugyanaz.
Ez nem mindig nyilvánvaló, olyankor F-próba kell a szórásnégyzetekre.

Hipotézisek: mint az u-próbánál.

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\bar{X} - \bar{Y}}{\sigma \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

u-próba alapján

$$\approx \widehat{\sigma^2} = \frac{(n_1-1)\, s_X^{*2} + (n_2-1)\, s_Y^{*2}}{n_1+n_2-1} = \frac{\sum_{i=1}^{n_1} (X_i - \bar{X})^2 + \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2}{n_1+n_2-2}$$

$$\Rightarrow T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{(n_1-1)\, s_X^{*2} + (n_2-1)\, s_Y^{*2}}{n_1+n_2-1}}} \sqrt{\frac{n_1 n_2}{n_1+n_2}} =$$

$$= \underbrace{\frac{\bar{X} - \bar{Y}}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}}_{\underset{H_0}{\sim} N(0,1)} \cdot \frac{1}{\sqrt{\frac{\hat{\sigma}^2}{\sigma^2}}}$$

$$\underset{H_0}{\sim} \frac{\chi^2_{n_1+n_2-2}}{n_1+n_2-2}$$

$$\hat{\sigma}^2 = \frac{\sigma^2 \chi^2_{n_1-1} + \sigma^2 \chi^2_{n_2-1}}{n_1+n_2-1} = \frac{\sigma^2 \chi^2_{n_1+n_2-2}}{n_1+n_2-2}$$

(függetlenek)

$\bar{X}, \bar{Y}, s_X^{*2}, s_Y^{*2}$ függetlenek $\Rightarrow$ T-ben az 1. és 2. tényező ffvek

$$\Rightarrow T \underset{H_0}{\sim} t_{n_1+n_2-2}$$

A próbák mint az egymintás esetben:

- Egyoldali: elvetünk, ha $T > c_\alpha$
- Kétoldali: elvetünk, ha $|T| > c_{\alpha/2}$,

ahol c_α a $t_{n_1+n_2-2}$-eo. $(1-\alpha)$-kvantilise

Van olyan módszer, ami kétmintásnál $\sigma_1 \neq \sigma_2$-re is működik, az csak közelítő próbát ad, nem egzakt. $\longrightarrow$ s-próba

F-próba

$$\left.\begin{array}{l} X_1, \dots, X_{n_1} \sim N(\mu_1, \sigma_1^2) \\ Y_1, \dots, Y_{n_2} \sim N(\mu_2, \sigma_2^2) \end{array}\right\} \text{ függetlenek} \qquad \mu_1, \mu_2, \sigma_1, \sigma_2 \text{ paraméterek}$$

Hipotézisek:
- Kétoldali: $H_0: \sigma_1^2 = \sigma_2^2$ $\quad H_1: \sigma_1^2 \neq \sigma_2^2$
- Egyoldali: $H_0: \sigma_1^2 = \sigma_2^2$ $\quad H_1: \sigma_1^2 > \sigma_2^2$

  vagy $H_0': \sigma_1^2 \leq \sigma_2^2$

$$T = \frac{s_X^{*2}}{s_Y^{*2}} \underset{H_0}{\sim} F_{n_1-1,\, n_2-1}$$

Próba:
- Egyoldali: elvetés, ha $T > c_{krit} = c_\alpha$, ahol c_α az $F_{n_1-1,\, n_2-1}$-eo. $(1-\alpha)$-kvantilise.
- Kétoldali: adott α-ra $0 < c_1 < c_2 < \infty$ krit. értékek, és elvetés, ha $T < c_1$ vagy $T > c_2$.

  Ha X-et és Y-t felcseréljük:

  az $(n_1-1,\, n_2-1)$-hez c_1', c_2' krit. ért.,

  az $(n_2-1,\, n_1-1)$-hez c_1'', c_2'' krit. ért.,

  és ezekre $c_1'' = \frac{1}{c_2'}$, $c_2'' = \frac{1}{c_1'}$

  A gyakorlatban kamatolt α-ra $c_1 < 1 < c_2$ fennáll. (kivételtan $T = 1$-re $s_X^{*2} = s_Y^{*2}$, ilyenkor nyilván el akarunk fogadni.)

  Gyakorlatban T-hez a nagyobb tap. négyzetet szorzás ini a számlálóba $\Rightarrow$ $1 \leq T$ automatikus, csak a felsővel kell összehasonlítani: $T > c_2$-re elvetés.

Azért, hogy az egyoldali esetben kamatolt krit. értékek a kétoldali esetben is kamatolhatók legyenek, szokás c_2-nek $c_{\alpha/2}$-t kamatolni. Ez nem indokolt matematikailag (kivéve, ha $n_1 = n_2$), de kényelmes.

Tulajdonságok

Az egymintás egyoldali u-próba egyenletesen legerősebb.

A többi csak a torzítatlanok közötti egy. legerősebb (minesen olyan, ami mindenre egy. legerősebb lenne).

Kivétel: az $\alpha/2$-es F-próba $n_1 \neq n_2$ esetén nem is torzítatlan!

(De a sorolás, ám nehezen számolható c_2-vel jó lenne.)

Ezek a próbák $n \to \infty$, ill. $n_1, n_2 \to \infty$ mellett 1-hez tartó erejűek θ_1-en, azaz konzisztens próbák.

χ^2 - próbák

Tétel. r osztály, n objektum az osztályokba sorolva (mindegyik pontosan egybe): minden objektum p_i valószínűséggel kerül az i-edik osztályba $\left(p_i > 0,\ \sum_{i=1}^{r} p_i = 1\right)$

Az i-edik osztályba N_i objektum került.

Jel. $T_n := \sum_{i=1}^{r} \frac{(N_i - np_i)^2}{np_i}$.

Ekkor $n \to \infty$ esetén $T_n \xrightarrow{d} \chi^2_{r-1}$

B: i az osztályokon, j az objektumokon megy mindig.

$I_{ij} := I$(a j-edik objektum az i-edik osztályba sorolták)

Legyen $Y_1, \dots, Y_n$ egymástól és ezektől az indikátorokból (azaz az objektumokból) független $N(0,1)$ val. változók.

$$X_{ij} := \frac{I_{ij} - p_i}{\sqrt{p_i}} + \sqrt{p_i}\, Y_j \qquad (i = 1, \dots, r,\quad j = 1, \dots, n)$$

$$X_j = (X_{1j}, \dots, X_{rj})^T \qquad (j = 1, \dots, n)$$

Ezek ftlen, azonos eo. vektorok.

Meg fogjuk mutatni: $\underset{(\text{tsuc})}{EX_j = 0}$, $\Sigma(X_j) = E_r$

2017.05.03.

$$\operatorname{cov}(X_{ij}, X_{kj}) = \operatorname{cov}\left(\frac{I_{ij} - p_i}{\sqrt{p_i}}, \frac{I_{kj} - p_k}{\sqrt{p_k}}\right) + \operatorname{cov}\left(\frac{I_{ij} - p_i}{\sqrt{p_i}}, \sqrt{p_k}\, Y_j\right) +$$

$$+ \operatorname{cov}\left(\sqrt{p_i}\, Y_j, \frac{I_{kj} - p_k}{\sqrt{p_k}}\right) + \operatorname{cov}\left(\sqrt{p_i}\, Y_j, \sqrt{p_k}\, Y_j\right) =$$

$$= \frac{\operatorname{cov}(I_{ij}, I_{kj})}{\sqrt{p_i p_k}} + 0 + 0 + \sqrt{p_i p_k}\, \underbrace{D^2(Y_j)}_{1} = *$$

Y_j ftem

I_{ij}-ből és I_{kj}-ből

$$\operatorname{cov}(I_{ij}, I_{kj}) = \underbrace{E(I_{ij} \cdot I_{kj})}_{\substack{p_i, \text{ ha } i=k \\ 0, \text{ ha } i \neq k}} - \underbrace{E(I_{ij})}_{p_i}\, \underbrace{E(I_{kj})}_{p_k}$$

$$* = \begin{cases} 1, & \text{ha } i = k \\ 0, & \text{ha } i \neq k \end{cases}$$

CHT r dimenzióban: $\frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_j \xrightarrow{d} N_r(0, E_r)$

$\rightarrow \left\| \frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_j \right\|^2 \xrightarrow{d} \chi_r^2$ normát véve.

$$\left\| \frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_j \right\|^2 = \sum_{i=1}^{r} \left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_{ij} \right)^2 = \sum_{i=1}^{r} \left[\frac{1}{\sqrt{n p_i}} \overbrace{\sum_{j=1}^{n} (I_{ij} - p_i)}^{N_i - n \cdot p_i} + \overbrace{\sqrt{\frac{p_i}{n}} \sum_{j=1}^{n} Y_j}^{\sqrt{p_i}\, Z_n^{*}} \right]^2$$

$$= \sum_{i=1}^{r} \frac{(N_i - n p_i)^2}{n p_i} + \underbrace{\sum_{i=1}^{r} p_i}_{1} \cdot Z_n^2 + 2 \cdot \frac{Z_n}{\sqrt{n}} \underbrace{\sum_{i=1}^{r} (N_i - n p_i)}_{n - n = 0} =$$

$* \; Z_n := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} Y_j \sim N(0,1)$

$$= T_n + Z_n^2.$$

Ezek ftem tagok, mert az egyikben indikátorok, a másikban Y_j-k vannak, és azok ftemek.

Ez és a χ_r^2 is konvolúció, de a konvolúció fécsoperfjában általában nem lehet egyszerűsíteni.

Legyen φ_n a T_n, f_k a χ_k^2 karakteisztikus fvénye.

Igenge kell: $\varphi_n \cdot f_1 \longrightarrow f_r = f_{r-1} f_1$ pontonként

Ha $f_1(t) \neq 0 \quad \forall t \in \mathbb{R} \rightarrow$ le lehet osztani, és az adódik, amit szeretnénk.

Definíció. Egy Q val. eo. korlátlanul osztható, ha $\forall n\ \exists Q_n$:

$Q = \underbrace{Q_n * \dots * Q_n}_{n \text{ db}}$. Kar. függvénnyel: ψ korl. osztható,

ha $\forall n\ \exists \psi_n$, hogy $\psi = \psi_n^n$ és ψ_n kar. fv.

Lemma. Korlátlanul osztható kar. fv. sehol sem nulla.

B: Ha ψ kar fv, akkor $|\psi(t)|^2 = \psi(t)\cdot\psi(-t)$ is az.

Ha ψ korl. osztható $\Rightarrow |\psi|^2 = (|\psi_n|^2)^n$, $|\psi_n|^2 = |\psi|^{\frac{2}{n}}$

$$|\psi_n(t)|^2 = |\psi(t)|^{\frac{2}{n}} \xrightarrow[n\to\infty]{} \begin{cases} 0, & \text{ha } \psi(t)=0 \\ 1, & \text{ha } \psi(t)\neq 0 \end{cases}$$

Folytonos-e a határfüggvény 0-ban?

$\psi(t)\neq 0$ a 0 egy kis. ben, mert $\psi(0)=1$ és folyt.

$\Rightarrow$ a határfv. kar. fv. (folytonossági tétel)

$\Rightarrow$ folytonos mindenütt, csak 2 értéket vehet fel és az 1-et felveszi.

$\Rightarrow \lim_{n\to\infty} |\psi_n(t)|^2 \equiv 1 \Rightarrow \psi(t)$ sehol sem 0. $\boxed{L}$

$\Rightarrow \chi_r^2$ kar. függvény sehol se 0, mert $\chi_r^2 = \Gamma_{\frac{r}{2},\frac{1}{2}}$,

és Γ korl. osztható: $\Gamma_{\alpha,\lambda} = \left(\Gamma_{\frac{\alpha}{n},\lambda}\right)^{*n}$ (n-edik konvolúció-hatvány). $\square$

Illeszkedésvizsgálat

n megfigyelés, r osztály, N_i az i-edik osztály gyakorisága

Tiszta illeszkedésvizsgálat

H_0: a vszger $(p_1, p_2, \dots, p_r)$ $\qquad H_1$: nem ez az eloszlás

$$T_n = \sum_{i=1}^{r} \frac{(N_i - np_i)^2}{np_i} \xrightarrow[H_0]{d} \chi^2_{r-1}$$

Próba: elvetjük H_0-t, ha $T_n > c_{krit} = \chi_r^2$-eo. $(1-\alpha)$-kvantilise

$$T_n = n\cdot\sum_{i=1}^{r} \frac{\left(\frac{N_i}{n} - p_i\right)^2}{p_i}$$

Látható, hogy ha H_0 nem igaz, akkor $T_n \to \infty$

$\Rightarrow$ a próba konzisztens, $\forall$ alternatívát 1-hez tartó valszínnel felismerünk

Ez ún. aszimptotikus próba: a terjedelme nem α, csak $n \to \infty$ mellett tart α-hoz.

Kell, hogy n elég nagy legyen: $\forall N_i \geq 4$ legyen; ebben nem egyértelmű a szabályban.

Ha ez nem áll fenn: osztályokat kell összevonni. Praktikusan is osztályokat érdemes összevonni.

Az összevonás csökkenti az erőt.

Becsléses illeszkedésvizsgálat

H_0: a $(p_1, \dots, p_r)$ eloszlás egy $\{p(\vartheta) : \vartheta \in \Theta\}$ osztály eleme, $\Theta \subset \mathbb{R}^s$

Fisher tétele (BN). Tfh. $s \leq r-2$, továbbá $p(\vartheta)$ ϑ-ban folyt. diff.ható és $\text{rang} \underbrace{p'(\vartheta)}_{r \times s\text{-es mátrix}} = s$ $\forall \vartheta \in \Theta$. Legyen $\hat{\vartheta}_n$ a ϑ ML-becslése.

Ekkor $$\sum_{i=1}^{r} \frac{N_i - n p_i(\hat{\vartheta}_n)}{n p_i(\hat{\vartheta}_n)} \xrightarrow[H_0]{d} \chi^2_{r-1-s}.$$

Tehát a becslésben a $p(\hat{\vartheta}_n)$ eloszlást illesztünk. χ^2-próbával, $r-1-s$ szab. fokkal.

Homogenitásvizsgálat

r osztály, két féle mintát szoktunk be

a bevalószínűségek $p_1, \dots, p_r$ ill. $q_1, \dots, q_r$,

a gyakoriságok pedig $N_1, \dots, N_r$ ill. $M_1, \dots, M_r$,

$\sum_{i=1}^{r} N_i = n$, $\sum_{i=1}^{r} M_i = m$

H_0: $\underline{p} = \underline{q}$ $\qquad H_1$: $\underline{p} \neq \underline{q}$

$N_1, \dots, N_r, M_1, \dots, M_r$ előállásának valósége (likelihood) =

$$= \underbrace{\frac{n!}{\prod_{i=1}^{r} N_i!} \cdot \frac{m!}{\prod_{i=1}^{r} M_i!}}_{\text{konstans}} \cdot \underbrace{\prod_{i=1}^{r} p_i^{N_i} \cdot \prod_{i=1}^{r} q_i^{M_i}}_{\text{érdekes rész}}$$

H_0: $\underline{p} = \underline{\vartheta}$, $\underline{q} = \underline{\vartheta}$, ahol $\vartheta = (\vartheta_1, \dots, \vartheta_{r-1}) \in \Delta$, és

$$\Delta = \left\{ z \in \mathbb{R}^{r-1} \;\middle|\; \forall z_i > 0,\ \sum_{i=1}^{r-1} z_i < 1 \right\} \quad \text{simplex}$$

Tel.: $\vartheta_r = 1 - \sum_{i=1}^{r-1} \vartheta_i$, $\quad s = r-1$

Tudjuk:

$$\sum_{i=1}^{r} \frac{(N_i - np_i)^2}{np_i} + \sum_{i=1}^{r} \frac{(M_i - nq_i)^2}{nq_i} \xrightarrow[H_0]{d} \chi^2_{(r-1)\cdot 2}, \quad f = 2r-2$$

A nab. fok $s = (r-1)$-gyel fog csökkenni — szab. fok

$$\text{likelihood} = C \cdot \prod_{i=1}^{r} \vartheta_i^{M_i + N_i},$$

ezt szeretnénk maximalizálni

log

$$\sum_{i=1}^{r} (M_i + N_i) \log \vartheta_i - t$$

kell maximalizálni

$$\frac{\partial}{\partial \vartheta_i}: \quad \frac{M_i + N_i}{\vartheta_i} - \frac{M_r + N_r}{\vartheta_r} = 0 \quad \text{legyen}$$

$$\Rightarrow \quad \forall i: \quad \frac{M_i + N_i}{\vartheta_i} = C \quad \rightarrow \quad \vartheta_i = C \cdot (M_i + N_i)$$

$$\Rightarrow \quad \underbrace{\sum_{i=1}^{r} \vartheta_i}_{1} = C \cdot (m+n) \quad \Rightarrow \quad C = \frac{1}{m+n}$$

$\Rightarrow$ a ML-becslés $\hat{\vartheta}_i = \dfrac{M_i + N_i}{m+n}$, azaz az összevont mintából a relatív gyakoriság

Ez áll $i = r$-re is, mert az összeg 1.

$$\Rightarrow \quad T_n = \sum_{i=1}^{r} \left[\frac{\left(N_i - n \cdot \frac{M_i + N_i}{m+n}\right)^2}{n \cdot \frac{M_i + N_i}{m+n}} + \frac{\left(M_i - m \cdot \frac{M_i + N_i}{m+n}\right)^2}{m \cdot \frac{M_i + N_i}{m+n}} \right]$$

Fisher alapján $T_n \xrightarrow{d} \chi^2_{(2r-2)-(r-1)} = \chi^2_{r-1}$.

Még alakítjuk kicsit T_n-et.

$$\left(N_i - n \cdot \frac{M_i + N_i}{m+n}\right)^2 = \frac{(mN_i - nM_i)^2}{(m+n)^2} =$$

$$= \frac{m^2 n^2}{(m+n)^2} \cdot \left(\frac{N_i}{n} - \frac{M_i}{m}\right)^2$$

Ugyanezt tesszük a 2. tag számlálójára is.

$$\Rightarrow T_n = \sum_{i=1}^{r} \frac{m^2 n^2}{(m+n)^2} \cdot \frac{(m+n)\left(\frac{1}{n}+\frac{1}{m}\right)}{M_i + N_i} \left(\frac{N_i}{n} - \frac{M_i}{m}\right)^2 =$$

$$= \sum_{i=1}^{r} m \cdot n \cdot \frac{\left(\frac{N_i}{n} - \frac{M_i}{m}\right)^2}{N_i + M_i}$$

próbastatisztikáro végünk χ^2_{r-1} - próbát

Függetlenségvizsgálat

n megfigyelés, 2 szempont: r ill. s osztállyal

Az (i,j) kombináció avagy cella valósége p_{ij} $(1 \le i \le r,\ 1 \le j \le s)$, gyakorisága pedig N_{ij}.

Marginálisok:
$$p_{i\cdot} = \sum_{j=1}^{s} p_{ij}, \qquad p_{\cdot j} = \sum_{i=1}^{r} p_{ij},$$
$$N_{i\cdot} = \sum_{j=1}^{s} N_{ij}, \qquad N_{\cdot j} = \sum_{i=1}^{r} N_{ij}$$

2. szempont

1. szempont	1	2	...	s	
1	N_{11}	N_{12}		N_{1s}	$N_{1\cdot}$
2	N_{21}	N_{22}		N_{2s}	$N_{2\cdot}$
⋮					
r	N_{r1}	N_{r2}		N_{rs}	$N_{r\cdot}$
	$N_{\cdot 1}$	$N_{\cdot 2}$		$N_{\cdot s}$	n

kontingenciatábla

$r \times s$ cella

H_0: a két szempont ftlen, azaz $p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$ $\forall i,j$

H_1: nem független

Ha a marginálisok adottak $\longrightarrow$ tiszta illeszkedésvizsgálat

Ha nem: becsléses illeszkedésvizsgálat $r+s-2$ paraméterre (a marginálisok), azért ennyi, mert az utolsók ezekből már adódnak.

$$(\vartheta_1, \dots, \vartheta_{r-1}) \in \Delta_{r-1}$$
$$(\tau_1, \dots, \tau_{s-1}) \in \Delta_{s-1}$$

Tel. $\vartheta_r = 1 - \sum_{i=1}^{r} \vartheta_i, \quad \tau_s = 1 - \sum_{j=1}^{s} \tau_j$

Cél H_0: $p_{ij} = \vartheta_i \tau_j$ $\quad (1 \le i \le r \cancel{-1},\ 1 \le j \le s \cancel{-1})$

Mj. ekkor $p_{i\cdot} = \sum_{j=1}^{s} \vartheta_i \tau_j = \vartheta_i \sum_{j=1}^{s} \tau_j = \vartheta_i \cdot 1$, hasonlóan $p_{\cdot j} = \tau_j$

Ez nyilván majdnem ugyanaz, csak $i=r$-re vagy $j=s$-re kérdéses.

$\underbrace{\text{mint az eredeti } H_0}$

Nem ide való! (Ettől még igaz.)

És ott is jó, pl. $p_{rj} = p_{\cdot j} \sum_{i=1}^{r-1} p_{ij} = p_{\cdot j} - \sum_{i=1}^{r-1} p_{i\cdot}\, p_{\cdot j} =$

$$= p_{\cdot j}\left(1 - \sum_{i=1}^{r-1} p_{i\cdot}\right) = p_{\cdot j} \cdot p_{r\cdot} \checkmark$$

Hasonlóan $p_{is} = p_{i\cdot} \cdot p_{\cdot s}$ ✓

Végül $p_{rs} = 1 - \sum_{(i,j)\neq(r,s)} p_{ij} = \sum_{\forall(i,j)} p_{i\cdot}\, p_{\cdot j} - \sum_{(i,j)\neq(r,s)} p_{i\cdot} \cdot p_{\cdot j} =$

$$= p_{r\cdot} \cdot p_{\cdot s} \checkmark$$

Tehát H_0-hoz elég, ha $p_{ij} = p_{i\cdot}\, p_{\cdot j}$ fennáll $i \le r-1$, $j \le s-1$-re

2017.05.10.

ML-becslés kell.

$$\text{likelihood} := C \prod_{i=1}^{r} \prod_{j=1}^{s} p_{ij}^{N_{ij}} = C \prod_{i=1}^{r} \prod_{j=1}^{s} (\vartheta_i \tau_j)^{N_{ij}} =$$

$$= C \underbrace{\prod_{i=1}^{r} \vartheta_i^{N_{i\cdot}}}_{} \cdot \underbrace{\prod_{j=1}^{s} \tau_j^{N_{\cdot j}}}_{}$$

külön maximalizáljuk a L tényezőt

A ML-becslés a relatív gyak. lesz:

$$\hat{\vartheta}_i = \frac{N_{i\cdot}}{n}, \quad \hat{\tau}_j = \frac{N_{\cdot j}}{m}, \quad \text{et } i=r\text{-re és } j=s\text{-re is igaz.}$$

$$\Rightarrow T_n = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{\left(N_{ij} - \frac{N_{i\cdot} N_{\cdot j}}{n}\right)^2}{\frac{N_{i\cdot} N_{\cdot j}}{n}} \xrightarrow[H_0]{d} \chi^2_{rs-1-(r+s-2)} = \chi^2_{(r-1)(s-1)}$$

Fisher

Erre végünk tehát χ^2-próbát $f=(r-1)(s-1)$ szabadsági fokkal.

T_n-et lehetne még szebb alakra hozni.

Speciális eset: $r = s = 2$.

Két igen-nem tulajdonság van.

$(I=1,\ N=2)$

	I	N	
I	N_{11}	N_{12}	$N_{1\cdot}$
N	N_{21}	N_{22}	$N_{2\cdot}$
	$N_{\cdot 1}$	$N_{\cdot 2}$	n

T_n 4 tagú. Választunk egy állat, ugyen a gyakorisága a, a vele egy sorban levő b, vele egy oszlopban c, a negyedik (átlós helyzetű) d.

Pl. a (2.1) állára:

c	d
a	b

$$\longrightarrow \frac{\left(a - \frac{(a+b)(a+c)}{n}\right)^2}{\frac{(a+b)(a+c)}{n}}$$

$n = a+b+c+d \longrightarrow$ a számláló: $\left[\frac{1}{n}\left(a(a+b+c+d) - (a+b)(a+c)\right)\right]^2 = \frac{(ad-bc)^2}{n^2}$

$$\Rightarrow T_n = \frac{(N_{11}N_{22} - N_{12}N_{21})^2}{n} \cdot \left(\underbrace{\frac{1}{N_{1\cdot}N_{\cdot 1}} + \frac{1}{N_{1\cdot}N_{\cdot 2}}}_{\frac{1}{N_{1\cdot}} \cdot \frac{\overbrace{N_{\cdot 2}+N_{\cdot 1}}^{n}}{N_{\cdot 1}N_{\cdot 2}}} + \underbrace{\frac{1}{N_{2\cdot}N_{\cdot 1}} + \frac{1}{N_{2\cdot}N_{\cdot 2}}}_{\frac{1}{N_{2\cdot}} \cdot \frac{n}{N_{\cdot 1}N_{\cdot 2}}}\right)$$

ez mind a h tagra ugyanaz lesz!

$$\frac{1}{N_{1\cdot}} \cdot \frac{N_{\cdot 2}+N_{\cdot 1}}{N_{\cdot 1}N_{\cdot 2}} + \frac{1}{N_{2\cdot}} \cdot \frac{n}{N_{\cdot 1}N_{\cdot 2}} = \frac{n}{N_{1\cdot}N_{2\cdot}} \cdot \frac{n}{N_{\cdot 1}N_{\cdot 2}}$$

$$T_n = \frac{(N_{11}N_{22} - N_{12}N_{21})^2}{N_{1\cdot}N_{2\cdot}N_{\cdot 1}N_{\cdot 2}} \cdot n \xrightarrow[H_0]{d} \chi_1^2$$

Egyszerűbb képlet, mint általában.

Sőt: ezzel egyoldali ellenőrzőpróbát is megy!

H_1: a két tulajdonság pozitívan korrelál (H_0: ettől tovább is)

(igen–nem esetével negatív korreláció is vizsgálható)

A pozitív korreláció egy lehetséges megfogalmazása:

$$P(2.\ I \mid 1.\ I) > P(2.\ I \mid 1.\ N)$$

$$\frac{p_{11}}{p_{11}+p_{12}} > \frac{p_{21}}{p_{21}+p_{22}}$$

$$p_{11}\,p_{22} > p_{12} \cdot p_{21}$$

$$\Longleftrightarrow \begin{vmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{vmatrix} > 0$$

$\chi_1^2 = N(0,1)^2 \Rightarrow$ Plausibilis (de nem igazolható), hogy $\sqrt{T_n} \longrightarrow |N(0,1)|$. Ennek alapján az $U_n = \sqrt{n}\,\frac{N_{11}N_{22} - N_{12}N_{21}}{N_{1\cdot}N_{2\cdot}N_{\cdot 1}N_{\cdot 2}}$ eloszlásra $U_n \xrightarrow[H_0]{d} N(0,1)$.

H_1 esetén U_n nagy (sőt: ∞-be tart: $U_n \approx \sqrt{n} \cdot \frac{n^2}{n^2} \longrightarrow \infty$)

$\rightarrow U_n$-re egyoldali u-próbát végzünk.

(Ugy. a kétoldali u-próba χ^2-próbával lesznek ekvivalens.)

Illeszkedésvizsgálat folytonos esetben

$X_1, \dots, X_n$ minta folyt. eloszlásból

H_0: $P(X_1 < t) = F(t) \quad \forall t$

egyoldl. H_1: $P(X_1 < t) \geq F(t) \quad \forall t$ és $\exists t$: $P(X_1 < t) > F(t)$

kétold. H_1: $\exists t$: $P(X_1 < t) \neq F(t)$.

Kétoldali esetben diszkretizálás:

gyakran más eleve csoportosított adatok vannak, pl. évenként

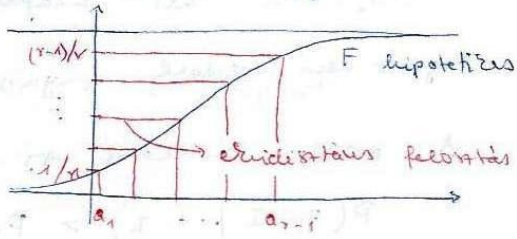
Felveszünk $a_1 < \dots < a_{r-1}$ számokat, az osztályok

$(-\infty, a_1), [a_1, a_2), \dots, [a_{r-2}, a_{r-1}), [a_{r-1}, \infty) \longrightarrow r$ db.

$F(a_1), F(a_2), \dots F(a_{r-1}), 1 - F(a_{r-1})$ valségek

χ^2-próbával ellenőrzik

Ha maguk választjuk az osztópontokat, akkor úgy célszerű, hogy az osztályok valségei minél egyenletesebbek legyenek.



Ha F_n a tapasztalati eofv.,

$$D_n^+ = \sup_t \left(F_n(t) - F(t)\right)$$

$$D_n^- = \sup_t \left(F(t) - F_n(t)\right)$$

$$D_n = \sup \left|F(t) - F_n(t)\right| = \max\{D_n^+, D_n^-\}$$

<u>Áll.</u> H_0 esetén ezek eloszlásmentesek, azaz az eloszlásuk nem függ F-től.

B: $F_n(t) - F(t)$ monoton fogyó két szomszédos megfigyelés között: $(X_i^*, X_{i+1}^*]$-on mon. csökk.

$\Rightarrow$ a sup a bal végpontban vett jobbliminek:

$$F_n(X_i^* + 0) - F(X_i^* + 0) = \frac{i}{n} - F(X_i^*)$$ Ez $i=n$-re is jó,

és $(-\infty, X_1^*]$-on a különbség ≤ 0, de ez nem játszik, mert $i=n$-re $\frac{n}{n} - F(X_i^*) \geq 0$.

$$\Rightarrow \quad D_n^+ = \max_{1 \le i \le n} \left(\frac{i}{n} - F(X_i^*) \right)$$

Hasonlóan $D_n^- = \max_{1 \le i \le n} \left(F(X_i^*) - \frac{i-1}{n} \right)$, a kiindító félegyenlőtlenség a különbség nemnegatív -, és ez nem játszik, mert $i=1$-re nemneg.

Ha X v.v. F eloszlásfüggvénye folytonos $\underset{\text{gyak.}}{\Rightarrow}$ $F(X)$ eloszlása $U[0,1]$

$\Rightarrow$ $F(X_1^*), \dots, F(X_n^*)$ egy $U(0,1)$-beli rendezett minta

$\Rightarrow$ eloszlásmentesség. $\square$

<u>Köv.</u> A próba: egyoldali esetben elvetés, ha $D_n^+ > c_{n,\alpha}$

kétoldali esetben elvetés, ha $D_n > c'_{n,\alpha}$ } F-független! $\square$

Ha n nagy: asimptotikus próba is végezhető, mert van határeloszlás: $y > 0$-ra

$$\lim_{n \to \infty} P\left(\sqrt{n}\, D_n^+ < y\right) = 1 - e^{-2y^2}$$

($\sqrt{n}$-es szorzó nélkül 0-hoz tartana)

$$\lim_{n \to \infty} P\left(\sqrt{n}\, D_n < y\right) = K(y) = \sum_{k=-\infty}^{+\infty} (-1)^k e^{-2k^2 y^2}$$

Kolmogorov-eloszlás, nincs zárt alak.

Erre alapozva: elvetjük H_0-t, ha egyoldalira $\sqrt{n}\, D_n^+ > \sqrt{\frac{1}{2} \log \frac{1}{\alpha}}$,

kétoldalira $\sqrt{n}\, D_n > K^{-1}(1-\alpha)$

Ezek a próbák konzisztensek.

Az egyoldalit Szmirnov-, a kétoldalit Kolmogorov-próbának nevezik.

Homogenitásvizsgálat folytonos esetben

$$\left.\begin{array}{lll} X_1, \dots, X_n \sim F & \text{folyt.} \\ Y_1, \dots, Y_m \sim G & \text{folyt.} \end{array}\right\} \text{ függetlenek}$$

H_0: $F = G$ egyo. H_1: $F(t) \geq G(t)$ $\forall t$, $\exists t$: $F(t) > G(t)$

kéto. H_1': $F \not\equiv G$.

$$D^+_{n,m} = \sup_t \left(F_n(t) - G_m(t)\right)$$

$$D^-_{n,m} = \sup_t \left(G_m(t) - F_n(t)\right)$$

$$D_{n,m} = \sup \left|G_m(t) - F_n(t)\right| = \max\left(D^+_{n,m}, D^-_{n,m}\right)$$

Áll. H_0 mellett ezek eloszlásmentesek.

B: $D^+_{n,m}$-ra a további gondolatmenettel

$$D^+_{n,m} = \max\left\{\frac{i}{n} - G_m(X_i^* + 0) \;\middle|\; 1 \leq i \leq n\right\}$$

Szimmetriaokokból:

$$D^-_{n,m} = \max\left\{\frac{i}{m} - F_n(Y_i^* + 0) \;\middle|\; 1 \leq i \leq m\right\}$$

$$G_m(X_i^* + 0) = \frac{1}{m}\sum_{j=1}^{m} I(Y_j \underset{+0 \text{ miatt}}{\leq} X_i^*) \underset{F=G}{\leq} \frac{1}{m}\sum_{j=1}^{m} I\left(G(Y_j) \leq F(X_i^*)\right)$$

Másrészt

$$G_m(X_i^* + 0) \geq \frac{1}{m}\sum_{j=1}^{m} I\left(G(Y_j) < F(X_i^*)\right)$$

Ezek 1 val megegyeznek,

mert $G(Y_j)$ és $F(X_i^*)$ függetlenek, $G(Y_j) \sim U[0,1]$,

$\left(F(X_i^*) \overset{d}{=} U_i^* \sim \text{Béta}(i, n+1-i)\right.$ is fennáll, de ez nem kell (gyak.))

$$\Rightarrow P\left(F(X_i^*) = G(Y_j)\right) = 0$$

Tehát $D^+_{n,m}$, $D^-_{n,m}$ és $D_{n,m}$ eloszlása ugyanaz, mintha minden minta $U[0,1]$-ből származna.

Aszimptotikus próba:

$$\lim_{n,m\to\infty} P\left(\sqrt{\frac{nm}{n+m}}\, D^+_{n,m} < y\right) = 1 - e^{-2y^2}$$

$$\lim_{n,m\to\infty} P\left(\sqrt{\frac{nm}{n+m}}\, D_{n,m} < y\right) = K(y)$$

Próba mint az illeszkedésvizsgálatnál.

Megj.: Nincs becslés illeszkedésvizsgálat.

Ha itt ϑ-t becsüljük, és abból illesztünk, az a határeloszlást is befolyásolja.

A nevezetes családokhoz tartozó esetekre vannak próbák, de általános nincs.

Lineáris modell

y, $x_1, x_2, \dots, x_p$ fizikai mennyiségek,

$$y = \sum_{j=1}^{p} a_j x_j \qquad a_1, \dots, a_p \text{ nem ismert együtthatók}$$

n mérést végzünk: $\quad y_i = \sum_{j=1}^{p} a_j x_{ij} + \varepsilon_i \qquad 1 \le i \le n$

$p \ll n$ $\qquad$ véletlen mérési hiba

Feltesszük, hogy a mérések során az ε_i hiba 0 várh. értékű, korrelálatlanok, azonos szórásúak.

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n \text{ véletlen vektor}, \quad a = \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} \in \mathbb{R}^p \text{ paraméternek vektor},$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \in \mathbb{R}^n \text{ véletlen hibavektor}, \quad X = \left[x_{ij} \right] \in \mathbb{R}^{n \times p} \text{ ismert eltérőmátrix}$$

korrelálatlanság

$$y = Xa + \varepsilon, \quad E(\varepsilon) = 0, \quad \Sigma(\varepsilon) = \sigma^2 \cdot \mathbb{E}_n$$

σ^2 is paraméter, de nem érdekel

(ún. zavaró paraméter, nuisance parameter)

<u>Cél</u>: a, illetve fréjeinek becslése, valamint hipotézisvizsgálat a-ra (ebben az esetben kell ε eloszlása is $\longrightarrow$ ún. normális lineáris modellben $\varepsilon \sim N_p(0, \sigma^2 \mathbb{E}_n)$)

$p \ll n \Rightarrow$ az egyenletrendszer ε miatt nem túlhatározott, hanem ellentmondó

Gauss: legkisebb négyzetek elve

$\hat{a}$ olyan, hogy $\|y - X\hat{a}\|^2$ minimális; ekkor $\hat{a}$ legkisebb négyzetes becslés (LNB)

$$\|y - X\hat{a}\|^2 = \sum_{i=1}^{n} \underbrace{\left(y_i - \sum_{j=1}^{p} x_{ij} \hat{a}_j \right)}_{\hat{\varepsilon}_i \text{ reziduális}}^2 = \sum_{i=1}^{n} \hat{\varepsilon}_i$$

$$y - Xa = \varepsilon \quad \Rightarrow \quad y - X\hat{a} = \hat{\varepsilon}$$

$\|y - Xa\|^2$: Q_R reziduális négyzetösszeg, RSS (residual sum of squares)

Világos, hogy $X\hat{a}$ az y merőleges vetülete $\operatorname{Im} X$-re. (és így $X\hat{a}$ egyértelmű)

$\hat{a}$ legkisebb négyzetes becslés $\Leftrightarrow$ $\underbrace{y - X\hat{a} \perp \operatorname{Im} X}$

$$y - X\hat{a} \in (\operatorname{Im} X)^{\perp} = \operatorname{Ker} X^T$$

Tehát $\hat{a}$ LNB $\Leftrightarrow$ $X^T(y - X\hat{a}) = 0$

$\Leftrightarrow$ $X^T X \hat{a} = X^T y$ Gauss-féle normálegyenlet

Ez $p \times p$-es egyenletrendszer. $r := \operatorname{rang}(X) = \operatorname{rang}(X^T X)$

Esetek: • $r = p$ $\Rightarrow$ $\exists!\, \hat{a}$:

• $r < p$ és ellentmondó $\longrightarrow$ ez nem lehet, mert tudjuk, hogy $\exists$ megoldás

• $r < p$ és a megoldások egy affin altér.

Teljes rangú eset: $\hat{a} = (X^T X)^{-1} X^T y$ az LNB egyértelműen.

$$E\hat{a} = (X^T X)^{-1} X^T \underbrace{Ey}_{Xa} = a$$

2017. 05. 17.

$$\Sigma(\hat{a}) = (X^T X)^{-1} X^T \underbrace{\Sigma(y)}_{\sigma^2 E_n} X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1}$$

$r < p$: a Gauss-egyenlet megoldásai $p-r$ dim aff. alteret alkotnak, az LNB nem egyértelmű

Az altér egy spec. eleme a min. normájú:

$\hat{a} = (X^T X)^{-} X^T y$, ahol $(X^T X)^{-}$ a Moore-Penrose-féle pszeudoinverz, amire $X^T X = U \Lambda U^T$, ← spektrálelőállítás ahol

$\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_p)$ sajátérték (pozit. szemidef. miatt $\forall \lambda_i \geq 0$),

U ortonormált: $U = [u_1, \dots, u_p]$, ahol $u_1, \dots, u_p$ a megfelelő ortonormált sajátvektorokból álló bázis (SONB),

$UU^T = U^T U = E_p$.

Ennek $(X^T X)^{-} = U \Lambda^{-} U^T$, ahol $\Lambda^{-} = \operatorname{diag}(\lambda_1^{-}, \dots, \lambda_p^{-})$, ahol

$$\lambda_i^{-} = \begin{cases} 0 & \lambda_i = 0 \\ 1/\lambda_i & \lambda_i > 0 \end{cases}$$

Ez valóban jó (BN), minimális normájút ad.

Mj. $r = p$ P_X jel. az $\operatorname{Im} X$-re való merőleges vetítést.

$$P_X y = X\hat{a} = X(X^T X)^{-1} X^T y \quad \forall y \in \mathbb{R}^n$$

$$\Rightarrow P_X = X(X^T X)^{-1} X^T$$

$y - X\hat{a} = \hat{\varepsilon}$ reziduálisok vektora

$\|\hat{\varepsilon}\|^2 = Q_R$ reziduális négyzetösszeg

$\mathbb{R}^n = \mathcal{L}_X \oplus \mathcal{L}_R$ ortogon. felbontás $\mathcal{L}_X = \operatorname{Im} X$

$\uparrow$ resid. altér, rá vetítés P_R

$y = y^X + y^R$ ortogon. felbontás

$y^R = P_R y = \hat{\varepsilon}$

$y^X = P_X y = X\hat{a}$

$\dim \mathcal{L}_X = r$, $\dim \mathcal{L}_R = n-r$

$s_R^2 = \dfrac{Q_R}{n-r}$ reziduális szórásnégyzet

<u>Áll.</u> s_R^2 torzítatlan becslése σ^2-nek.

B: $EQ_R = E\|P_R y\|^2 = E\Big(\underbrace{y^T \underbrace{P_R^T P_R}_{P_R} y}_{\text{skalár}}\Big) =$

$= E\big(\operatorname{tr}(y^T P_R y)\big) \underset{\operatorname{tr}(AB)=\operatorname{tr}(BA)}{=} E\big(\operatorname{tr}(P_R y y^T)\big) \underset{\operatorname{tr}\text{ lineáris}}{=}$

$= \operatorname{tr}\big(E(P_R yy^T)\big) \underset{P_R\text{ lineáris}}{=} \operatorname{tr}\big(P_R \underbrace{E(yy^T)}_{\Sigma(y)+Ey\,Ey^T}\big) =$

$= \operatorname{tr}\big(P_R(\sigma^2 E_n + Xaa^TX^T)\big) = \sigma^2 \operatorname{tr}(P_R) + 0$

$P_R Xa = 0$, mert $Xa \perp \mathcal{L}_R$

$\operatorname{tr}(P_R) = \dim \mathcal{L}_R$, mert a sajátérték összege, az idempotencia miatt csak 0 vagy 1 lehet sé, annyi 1-es van, amennyi $\mathcal{L}_R$ dimenziója.

$\Rightarrow EQ_R = \sigma^2 \operatorname{tr}(P_R) = \sigma^2 \cdot (n-r)$.

<u>Def.</u>

Legyen C $q \times p$-es mátrix, $\psi = \psi(a) = Ca$ a paraméter lineáris függvénye. Ekkor ψ becsülhető, ha van torzítatlan lineáris becslése: $T(y) = By$, ahol B $q \times n$-es.

$E(By) = \psi$

Áll. ψ becsülhető, ha C sorai benne vannak egy $\underbrace{X \text{ sorai}}_{\text{az egyes mérések körülményeit kódolja}}$ által kifeszített altérben (lin. fátistér), és $\operatorname{Im} X^T$.

B: ψ becsülhető $\iff \forall a \in \mathbb{R}^p: \quad \psi = C \cdot a = E(By) = B\,Ey = BXa.$

$\iff C = BX$

$\iff$ C sorai az X sorainak lin. kombinációi. $\square$

ez egy másik!

Tétel. (Gauss–Markov) Legyen ψ becsülhető, és legyen $\hat{a}$ (tetszőleges) LNB. Legyen $\hat{\psi} = C\hat{a}$. Ekkor $\hat{\psi}$ független $\hat{a}$ választásától, $\hat{\psi}$ torzítatlan lineáris becslés ψ-re és az ilyenek között minimális szórásnégyzetű (azaz négyzetes eltérségfüggvénye nézve optimális).

B: $\hat{\psi} = C\hat{a} = B\underbrace{X\hat{a}}_{\text{egyértelmű}} = BP_X y \implies \hat{a}$-tól függetlenség

lineáris becslés, mert y-nak lin. füvye

$E\hat{\psi} = BP_X Ey = BP_X Xa = BXa = Ca = \psi \implies$ torzítatlanság

Legyen $T = T(y) = By$ tetszőleges torzítatlan lin. becslés ψ-re.

Ekkor $C = BX$ a torzítatlanság miatt.

$$\underbrace{E\left(T(y)\cdot T(y)^T\right)}_{\text{elegendő ezt vizsgálni, mert mindkettőnél ugyanezt kell kivonni, hogy } \Sigma\text{-t kapjuk}} = E\left(By\,(By)^T\right)$$

$$By = B(P_X y + P_R y)$$

$$\to By\,(By)^T = (\hat{\psi} + BP_R y)\cdot(\hat{\psi} + BP_R y)^T = \hat{\psi}\hat{\psi}^T + \hat{\psi}(BP_R y)^T + (BP_R y)\hat{\psi}^T + (BP_R y)(BP_R y)^T$$

$$E(\underbrace{\hat{\psi}}_{BP_X y}(BP_R y)^T) = BP_X \underbrace{E(yy^T)}_{\Sigma y + Ey\cdot Ey^T = \sigma^2\cdot \mathbb{E} + (Xa)(Xa)^T} P_R B^T = **$$

$$** = \sigma^2 \cdot B \underbrace{P_X P_R}_{0} B^T + BP_X (Xa)\underbrace{(Xa)^T \cdot P_R}_{(P_R Xa)^T = 0} \cdot B^T = 0$$

Ugyanígy a másik vegyes tag vár. értéke is 0.

$$\implies \Sigma(T) - \Sigma(\hat{\psi}) = E\left(By\,(By)^T\right) - E(\hat{\psi}\hat{\psi}^T) = E\left((BP_R y)(BP_R y)^T\right) \geq 0 \quad \square$$

Def. $\hat{\varphi}$-ot is LNB-nek nevezzük.

Def. Teljes rangú esetben $\forall \varphi^{lin}$ f. becsülhető:

$$C = \underbrace{C(X^TX)^{-1}X^T}_{B}X$$

$\Rightarrow \hat{a}$ és így $C\hat{a}$ is optimális.

Normális lineáris modell

$\varepsilon \sim N_n(0, \sigma^2 E_n) \quad \rightarrow \quad y \sim N_n(Xa, \sigma^2 E_n)$

Itt $Q_R = \|P_R y\|^2 \sim \sigma^2 \cdot \chi^2_{n-r}\Big[\underbrace{\frac{1}{\sigma^2}\|P_R Ey\|^2}_{0,\text{ mert } Ey = Xa \perp \mathcal{L}_R}\Big]$

Továbbá $X\hat{a}$ ftlen $\hat{\varepsilon}$-tól és így Q_R-tól.

Ez a norm. lin. modell.

Lineáris hipotézis: $H_0: Ba = b$

$H_1: Ba \neq b$,

ahol B $q \times p$-es mátrix és $b \in \operatorname{Im} B$. (egyébként nem lineáris H_0)

Ftu. $r(B) = q$ (sorok lin ftlenek).

Feltehető, hogy $b = 0$: legyen ui. a_0 olyan, hogy $Ba_0 = b$.

(ilyen pontosan van) $\Rightarrow \tilde{a} = a - a_0$, $\tilde{y} = y - Xa_0$

$\tilde{y} = X\tilde{a} + \varepsilon$ lin. modell, és ebben $H_0: B\tilde{a} = 0$

Általános elv

Általánosított likelihood-hányados-statisztika:

$$T(y) = \frac{\sup\{f_{a,\sigma}(y) \mid a \in \mathbb{R}^p, \sigma > 0\}}{\sup\{f_{a,\sigma}(y) \mid \underbrace{Ba = 0}_{H_0}, \sigma > 0\}} \geq 1$$

Elvetjük H_0-t, ha T nagy (általánosított lrh.-próba).

(Regularitás esetén erre asimptotikus próbát lehet onnan alt.)

T helyett ennek bmely szig. növő fvénye használható.

$$f_{a,\sigma}(y) = (2\pi)^{-\frac{n}{2}} \cdot \sigma^{-n} \cdot \exp\left(-\frac{1}{2\sigma^2}\|y - Xa\|^2\right)$$

$y_1, \dots, y_n$ ftlenek σ szórással

Rögz. σ esetén a-ban maximalizálva $\|y - Xa\|^2$ minimális

$\rightarrow$ az $\hat{a}$ ML-becslés $= \hat{a}$ LNB

$$f_{\hat{a},\sigma}(y) = (2\pi)^{-\frac{n}{2}} \cdot \sigma^{-n} \cdot \exp\left(-\frac{1}{2\sigma^2}\|y - X\hat{a}\|^2\right)$$ maximalizálandó

log: $$-n \log \sigma - \frac{1}{2\sigma^2}\|y - X\hat{a}\|^2 + C$$

$\frac{\partial}{\partial\sigma} = 0$: $$-\frac{n}{\sigma} + \frac{1}{\sigma^3}\|y - X\hat{a}\| = 0 \quad \Rightarrow \quad \hat{\sigma}^2 = \frac{1}{n}\|y - X\hat{a}\|^2$$

$\Rightarrow$ $\sup f_{a,\sigma}(y) = \left(\frac{2\pi e}{n}\right)^{-\frac{n}{2}} \cdot \|y - X\hat{a}\|^{-n}$ a számláló T-ben

$\sup_{H_0} f_{a,\sigma}(y) = \left(\frac{2\pi e}{n}\right)^{-\frac{n}{2}} \cdot \|y - X\tilde{a}\|^{-n}$ a nevező T-ben,

ahol $\tilde{a}$ minimalizálja $\|y - Xa\|^2$-et H_0-on, azaz Ker B-n.

$$\Rightarrow \quad T = \frac{\|y - X\tilde{a}\|^n}{\|y - X\hat{a}\|^n}$$

$\mathbb{R}^n = \mathcal{L}_X \oplus \mathcal{L}_R$

$\mathcal{L}_X = X\mathbb{R}^p = X(\operatorname{Ker} B \oplus \operatorname{Im} B^T) = \mathcal{L}_0 \oplus \mathcal{L}_1$, ahol $\mathcal{L}_0 = X \operatorname{Ker} B = \{Xa \mid Ba = 0\}$.

$y^X = X\hat{a} = y^0 + y^1$

$\tilde{a}$: $X\tilde{a}$ az y vetülete $\mathcal{L}_0$-ra, azaz $X\tilde{a} = y^0$

$y = y^0 + y^1 + y^R$ ort. felbontás

$y - y^0 = y^1 + y^R$

$$\left.\begin{aligned} \|y - X\tilde{a}\|^2 = \|y - y^0\|^2 &\underset{\text{Pitagorasz}}{=} \|y^1\|^2 + \|y^R\|^2 \\ \|y - X\hat{a}\|^2 &= \|y^R\|^2 \end{aligned}\right\} \Rightarrow T = \left(\frac{\|y^1\|^2}{\|y^R\|^2} + 1\right)^{n/2}$$

$r_0 := \dim \mathcal{L}_0 \quad \Rightarrow \quad \dim \mathcal{L}_1 = r - r_0$

Legyen a stat. a növekvő: $F(y) = \left(T(y)^{\frac{2}{n}} - 1\right) \cdot \underbrace{\frac{n-r}{r-r_0}}$,

es sig. mon. növő fvénye T-nek.

számlálót és nevezőt is leosztjuk a dimenziójával

<u>Áll.</u> H_0 esetén $F(y) \sim F_{r-r_0,\, n-r}$

Ezért ezzel egyoldali F-próbát végezni.

B: $F(y) = \frac{s_1^2}{s_R^2}$, ahol $s_1^2 = \frac{Q_1}{r - r_0}$, $Q_1 = \|y^1\|^2$,

s_R^2 a reziduális négyzetösszeg.

$$s_R^2 \sim \frac{\sigma^2}{n-r} \chi^2_{n-r}$$

$$s_1^2 \sim \frac{\sigma^2}{r-r_0} \cdot \chi^2_{r-r_0}[\lambda], \text{ ahol } \lambda = \frac{1}{\sigma^2} \cdot \|(Xa)^1\|^2$$

$Ey = Xa$

$H_0 \Rightarrow Xa \in \mathcal{L}_0 \Rightarrow (Xa)^1 = 0 \Rightarrow \lambda = 0$

$\mathcal{L}_1 \perp \mathcal{L}_R \Rightarrow y^1$ és y^R ftlen $\Rightarrow s_1^2$ és s_R^2 ftlen, mert y^1 és y^R fv-ei.

Osztással valóban F-eloszlást kapunk. □

y^1 is szimmetrikus, ezt nem volt idő.