

Matematikai statisztika jegyzet

2021. tavasz

Tartalomjegyzék

1	2020.02.08.	1
1.1	Ismétlés valószínűségyszámításból	1
1.2	Statisztikai feladatok	1
1.3	Fogalmak	2
1.4	Dominált mértékcsalád	2
1.4.1	Példák	3
1.5	Elégséges statisztika, elégséges szigma-algebra	4
1.6	Feltételes várható érték, feltételes valószínűség	4
1.6.1	Feltételes eloszlás reguláris változata	4
1.6.2	A rendezett minta mindig elégséges	5
1.7	Feltétel elégségességre	5
1.8	Példák	6
1.8.1	Indikátor minta	6
1.8.2	Normális minta	6
2	2021.02.15.	7
2.1	Emlékeztető	7
2.2	Neymann faktorizációs tétel	7
2.3	Halmos–Savage tétel	8
2.4	Sigma-algebra teljesség tétele	10
2.5	Minimális elégségesség	10
2.6	Feltétel minimális elégségességre	11
2.7	Minimális elégséges statisztika jellemzése	12
2.8	Példák	13
2.8.1	Példa, egyenletes eloszláscsalád	13
2.8.2	Példa, Cauchy eloszlás eltolásparaméteres családja	13
2.8.3	Példa, Indikátor minta	14
3	2021.02.22.	15
3.1	Ismétlés. Elégséges és minimális elégséges statisztika	15
3.2	Teljesség	15
3.2.1	Példák	16
3.3	Momentum generáló függvény	17
3.4	Elégséges feltétel teljességre	17
3.5	Exponenciális család	18
3.5.1	Exponenciális család teljessége	18
3.5.2	Példák exponenciális családra	18
3.6	Basu tétele	19
3.6.1	Példák kísérő statisztikákra	20
3.7	Rao-Blackwell-Kolmogorov tétel	20
3.7.1	Példák	21

4	2021.03.01.	22
4.1	Becsülés, torzítatlanság, hatásosság	22
4.1.1	Alsó korlát torzítatlan becslés szórásnégyzetére	22
4.1.2	Példa. Indikátor minta	22
4.2	Cramer–Rao-egyenlőtlenség, információs határ	23
4.2.1	Cramer–Rao egyenlőtlenség, speciális esetek	23
4.3	Fisher információ	24
4.3.1	Formulák	24
4.3.2	Fisher információ exponenciális családban	24
4.3.3	Momentum generáló függvény, deriválhatóság	25
4.3.4	Példa	25
4.3.5	Átparaméterezés és Fisher információ	26
4.3.6	Átparaméterezés és információs határ	26
4.3.7	Példák	26
4.4	Gyenge regularitási feltétel	27
4.4.1	Bederiválhatóság gyenge regularitási feltétel mellett	28
4.4.2	$E_t(TI'(t))$ folytonos	28
4.4.3	Technikai lemma	28
4.4.4	A gyenge regularitási feltétel következményei	30
5	2021.03.08.	31
5.1	Fisher információ folyt.	31
5.1.1	Gyenge regularitási feltétel exponenciális családban	31
5.1.2	További példák	31
5.1.3	Független minták Fisher információja összeadódik	32
5.1.4	$S(X)$ Fisher információja nem lehet X információ tartalmánál nagyobb	32
5.1.5	Elégséges statisztika Fisher információja	32
5.1.6	Fisher információ és elégséges statisztika	33
5.1.7	Összefoglalás	33
5.2	Becslési módszerek	33
5.2.1	Momentum módszer indikátor mintára	33
5.2.2	Maximum likelihood elv indikátor mintára	34
5.2.3	Bayes módszer indikátor mintára	34
5.2.4	Tapasztalati eloszlás	35
5.2.5	Tapasztalati becslések	35
5.2.6	Momentum módszer	35
5.3	Glivenko–Cantelli tétel	36
5.3.1	Glivenko–Cantelli másképp, vázlat	37
6	2021.03.22.	39
6.1	Maximum likelihood becslés	39
6.1.1	Keverék eloszlásra	39
6.1.2	A paraméter egy függvényére	39
6.1.3	Exponenciális családban	40
6.1.4	Többdimenziós normális eloszlás	41
6.1.5	Lemma ellenőrzése	43
6.1.6	Példa, normális eloszlás	44
6.1.7	Példa, az aszimptotikus kovariancia mátrix lehet szinguláris	44

7	2021.03.29.	45
7.1	Általánosítás	45
7.1.1	konzisztencia	45
7.1.2	aszimptotikus normalitás	45
7.2	Erős regularitási feltétel	46
7.2.1	Erős regularitási feltétel következményei	46
7.3	Maximum likelihood becslés aszimptotikus tulajdonságai	47
7.3.1	Példa, keverék eloszlás	48
7.4	EM algoritmus	49
7.5	Becslések tulajdonságai	49
7.6	Bayes becslés	50
7.6.1	Aszimptotikus normalitás	51
7.6.2	Lemma bizonyítása	52
7.7	Minimax becslés indikátor mintából	53
8	2021.04.12.	54
8.1	Hipotézis vizsgálati feladatok	54
8.1.1	Példák	54
8.1.2	Fogalmak	54
8.2	Hiba lehetőségek	55
8.3	Véletlenített próba, próbafüggvény, terjedelem, szint, erő	55
8.4	Statisztikai próbák jellemzői	56
8.5	Próbák konstruálása	56
8.6	Neyman–Pearson lemma	57
8.6.1	Példa	58
8.7	Általánosítás: kritikus érték aszimptotikája	58
8.8	Kullback-Leibler divergencia	59
8.9	Két apróság	59
8.10	Erő aszimptotikája	60
9	2021.04.19.	61
9.1	Többdimenziós normális eloszlás, emlékeztető	61
9.1.1	Sűrűségfüggvény	61
9.1.2	függetlenség és korrelálatlanság	62
9.1.3	Következmények	62
9.2	Khínégyzet eloszlás, vetületek eloszlása	62
9.3	Fisher-Bartlett tétel	63
9.4	Fisher-Bartlett tétel megfordítása	63
9.5	t eloszlás	64
9.6	Klasszikus próbák	65
9.6.1	u próba	65
9.6.2	t próba	65
9.6.3	Kétmintás u-próba	66
9.6.4	Kétmintás t-próba	66
9.6.5	F-próba a szórások egyezésére	67
9.7	Diszkrét illeszkedés vizsgálat, khínégyzet-próbák	67
9.7.1	Speciális eset $r=2$	68
9.7.2	T_n határeloszlása	68

10	2021.04.26.	70
10.1	Diszkrét illeszkedés vizsgálat, khínégyzet-próbák	70
10.1.1	Speciális eset $r = 2$	70
10.1.2	T_n határeloszlása	71
10.2	Emlékeztető	71
10.3	Likelihood hányados határeloszlása	72
10.4	Diszkrét illeszkedés vizsgálat, T_n határeloszlása	73
10.5	Függetlenség vizsgálat khínégyzet-próbával	73
10.5.1	T_n határeloszlása	74
10.6	Homogenitás vizsgálat khínégyzet-próbával	74
10.7	Becsléses illeszkedés vizsgálat	75
10.8	Megjegyzések khínégyzet próbákhoz	75
11	2021.05.03.	76
11.1	Folytonos illeszkedés vizsgálat	76
11.1.1	Kolmogorov–Szmirnov statisztika eloszlásmentessége	76
11.1.2	Cramer–von Mises statisztika eloszlásmentes	76
11.2	Skálázott próbastatisztikák határeloszlása	77
11.2.1	Brown híd, mese	78
11.2.2	Skálázott Cramer–von Mises statisztika határeloszlása	78
11.2.3	Skálázott Kolmogorov–Szmirnov statisztika határeloszlása	79
11.3	Klasszikus próbák optimalitása	80
11.3.1	Emlékeztető. Neyman–Pearson lemma	80
11.4	Az u próba	80
11.4.1	Egyoldali ellenhipotézis mellett, az u -próba egyenletesen legerősebb	80
11.4.2	μ nem egyenlő nullára nincs egyenletesen legerősebb próba	81
11.4.3	az u próba egyenletes torzítatlansága	81
11.5	Egyenletesen legerősebb próbák zavaró paraméter mellett	82
11.5.1	(T, S) feltételes eloszlásai	83
11.5.2	A t -próba egyenletesen legerősebb a torzítatlan próbák között	84
12	2021.05.10	85
12.1	Lineáris regresszió, lineáris modell	85
12.1.1	Példa (lineáris regresszió)	85
12.2	Lineáris modell	86
12.3	bétakalap, és bétakalap négyzetének eloszlásai	86
12.4	Hipotézis vizsgálat normális lineáris modellben	86
12.5	Példa	87
12.6	ANOVA táblázat R-ben	87
12.7	Wald próba	88
12.8	Lineáris becslés, becsülhető függvény	88
12.9	Gauss–Markov tétel	89
12.10	Kereszt-validáció	89
12.10.1	Példa	90
12.11	Regularizáció, Ridge regresszió és LASSO	90

1. 2020.02.08.

Tudnivalók

- Tematika, irodalomjegyzék az előadás Canvas oldalán.
- Számonkérés: szóbeli vizsga, a vizsgaidőszak elején lesz egy írásbeli vizsga, amin legfeljebb középest lehet szerezni.
- A slide-ok az előadás Canvas oldalán elérhetőek lesznek az óra előtt.

1.1 Ismétlés valószínűesszámitásból

- Sűrűségfüggvény transzformációs formula:

$$f_{\phi(X)}(t) = f_X(\phi^{-1}(t)) |\det J\phi^{-1}|(t)$$

- Konvolúciós formula: X, Y függetlenek, létezik f_X , ekkor létezik f_{X+Y} és

$$f_{X+Y}(t) = \mathbb{E}(f_X(t-Y)) = \int f_X(t-y)Q_Y(dy) = \int f_X(t-y)f_Y(y)dy, \quad \text{ha } \exists f_Y.$$

- Nagy számok törvénye: X_1, X_2, \dots iid, $\mathbb{E}(X_1)$ létezik, ekkor

$$\frac{1}{n} \sum_{i=1}^n X_i \rightarrow \mathbb{E}(X_1), \quad \text{egy valószínűséggel.}$$

- Centrális határeloszlás tétel: X_1, X_2, \dots iid, $\mathbb{E}(X_1) = m \in \mathbb{R}$, $\mathbb{D}^2 X_1 = \sigma^2 \in (0, \infty)$, ekkor

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - m}{\sigma} \xrightarrow{d} N(0, 1)$$

vagy

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (X_i - m) \xrightarrow{d} N(0, \sigma^2), \quad \sigma^2 = 0 \text{ esetén is!}$$

1.2 Statisztikai feladatok

„Következtetések levonása a megfigyelésekből. Összefüggések feltárása az adatokban.”

- Pontbecslések
pl. Közvélemény-kutatás, a válaszok alapján szeretnénk egy termék kedveltségét megállapítani.
Általánosabban zajos megfigyeléseink (méréseink) vannak egy ismeretlen mennyiségre, amit szeretnénk meghatározni.
- Hipotézisvizsgálat
pl. Gyógyszerkutatás, ígéretesnek tűnő hatóanyagok vizsgálunk. A kísérlet eredményeiből szeretnénk arra következtetni, vajon hatásos-e a gyógyszerjelölt.

- Konfidencia halmaz.

pl. Közvélemény-kutatásnál nem csak egy adott termék népszerűsége érdekel minket, hanem egy olyan lehetőleg minél kisebb intervallum, ami a terméket ténylegesen kedvelők arányát nagy valószínűséggel tartalmazza.

- „Összefüggések keresése”, klasszifikáció, regresszió.

Minden megfigyelés egy (x_i, y_i) pár, ahol y_i az $f(x_i)$ zajos megfigyelése. f ismeretlen, ezt szeretnénk meghatározni.

pl. Közvélemény kutatásnál, azon kívül, hogy a terméket kedveli-e a megkérdezett, feljegyziünk egyéb adatokat is: életkor, jövedelmi helyzet, családi állapot, stb. és az érdekel minket, hogy ezek hogyan befolyásolják az adott termék kedveltségét.

1.3 Fogalmak

- Minta: X_1, X_2, \dots, X_n független azonos eloszlású megfigyelések.

Általában nem kellene, hogy függetlenek és azonos eloszlásúak legyenek, de csak ezt az esetet nézzük ezen az órán.

n a mintaelemszám.

- Mintatér: minta lehetséges értékei. Ha skalár megfigyeléseink vannak, akkor \mathbb{R}^n része. Jelölés \mathfrak{X} . Ez játssza a lehetséges kimenetek halmazát.

- Statisztikai mező: $(\Omega, \mathcal{A}, \mathbb{P})$

Általában $\Omega = \mathfrak{X}$.

$\mathbb{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ valószínűségi mértékek családja \mathcal{A} -n. Θ neve paramétertér.

Az a feltételezésünk, hogy a megfigyelések $\mathbb{P}^* \in \mathcal{P}$ -ből származnak.

- Statisztika: mérhető függvény a mintatéren. Ha $X = (X_1, \dots, X_n)$ a minta, akkor S -et és $S(X)$ -et azonosítjuk. Ez egyébként is teljesül, ha $\Omega = \mathfrak{X}$ és $X_i(x) = x_i$.

Példa Közvélemény-kutatás.

- Minden megfigyelés vagy 0, vagy 1. Nulla ha nem kedveli a terméket, és 1 ha igen.
- Mintatér $\mathfrak{X} = \{0, 1\}^n$. n a megfigyelések száma.
- Ha minden megkérdezett, egymástól függetlenül és találmra választjuk a teljes populációból, akkor X_1, \dots, X_n független θ paraméterű indikátorok, ahol $\theta \in [0, 1]$ az ismeretlen paraméter.
- Paramétertér $\Theta = [0, 1]$
- $\mathbb{P}_\theta(X = x) = \prod_i \mathbb{P}_\theta(X_i = x_i) = \prod_i \theta^{x_i} (1 - \theta)^{1-x_i}$.
- $S(x) = \sum_i x_i$ a mintaösszeg, egy példa statisztikára. $S(X) = \sum_i X_i$.

1.4 Dominált mértékcsalád

Definíció 1. \mathcal{P} *dominált mértékcsalád* az (Ω, \mathcal{A}) mérhető téren, ha létezik olyan λ σ -véges mérték amire nézve mindegyik $\mathbb{P} \in \mathcal{P}$ abszolút folytonos. λ neve *domináló mérték*.

Emlékeztető.

- $\mathbb{P} \ll \lambda$ azt jelenti, hogy a λ nullmértékű halmazok \mathbb{P} nullmértékűek is.
- $\mathbb{P} \ll \lambda$ szükséges ahhoz, hogy létezzen Z (≥ 0 , mérhető), amivel $\mathbb{P}(A) = \int_A Z d\lambda$ teljesül. A Radon-Nikodym tétel szerint elégséges is.
- Jelölés: $Z = \frac{d\mathbb{P}}{d\lambda}$ a \mathbb{P} Radon-Nikodym deriváltja λ -ra nézve.
- Mértékcseré integrálban:

$$\mathbb{P} \ll \lambda, \quad \implies \quad \int h d\mathbb{P} = \int h \frac{d\mathbb{P}}{d\lambda} d\lambda$$

- Láncre szabály. Ha $\mathbb{P}_1 \ll \mathbb{P}_0 \ll \lambda$, akkor

$$\frac{d\mathbb{P}_1}{d\lambda} = \frac{d\mathbb{P}_1}{d\mathbb{P}_0} \frac{d\mathbb{P}_0}{d\lambda}$$

1.4.1 Példák

- Ha a mintaelemek közös eloszlása abszolút folytonos a Lebesgue mértékre nézve, akkor az n elemű minta együttes eloszlása abszolút folytonos az n dimenziós Lebesgue mértékre nézve.
- Példa. X_1, \dots, X_n normális minta ismeretlen m várható értékkel és $\sigma^2 > 0$ szórásnégyzettel. Ekkor $\mathfrak{X} = \mathbb{R}^n$, $\Theta = \mathbb{R} \times (0, \infty)$, $\theta = (m, \sigma)$, a domináló mérték λ az n dimenziós Lebesgue mérték, az eloszláscsalád tagjait a sűrűségfüggvényükkel lehet megadni:

$$f_\theta(x) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_i (x_i - m)^2 \right\}$$

- Ha mintaelemek közös eloszlása diszkrét, akkor a mintatér megszámlálható és a mintaelemek együttes eloszlását a $\mathbb{P}_\theta(X = x)$ valószínűségekkel lehet leírni. Ez nem más mint a számlálómértékre vonatkozó Radon-Nikodym derivált.
- Példa: Közvélemény-kutatás, azaz n elemű indikátor minta.
Mintatér: $\mathfrak{X} = \{0, 1\}^n$, paraméterter: $\Theta = [0, 1]$, minta együttes sűrűségfüggvénye a számláló mértékre nézve:

$$f_\theta(x) = \mathbb{P}_\theta(X = x) = \prod_i \theta^{x_i} (1 - \theta)^{(1-x_i)}, \quad \theta \in \Theta = [0, 1]$$

Nem minden mértékcsalád dominált és nem minden dominált családra igaz, hogy vagy a számlálómérték, vagy a Lebesgue mérték dominálja.

- Exponenciális eloszlású élettartamokat figyelünk meg ismeretlen c értéknél levágva. Azaz $\Theta = \{(\lambda, c) : \lambda, c > 0\}$, X_i eloszlása \mathbb{P}_θ mellett azonos $Z \wedge c$ eloszlásával, ahol $Z \sim \exp(\lambda)$. HF. Ez az eloszláscsalád nem dominált.
- Előző példa, de $c = 1$ ismert. Ekkor $\mathfrak{X} = [0, 1]^n$ és ha $\mu = \text{Leb} + \delta_1$, akkor $\lambda = \mu^{\otimes n}$ domináló mérték. A minta sűrűségfüggvénye:

$$f_\theta(x) = \prod_i \left(\theta e^{-\theta x} \mathbf{1}_{(0 < x_i < 1)} + e^{-\theta} \mathbf{1}_{(x_i = 1)} \right), \quad \theta > 0$$

ugyanis

$$\mathbb{P}_\theta(X_i \in H) = \mathbb{P}_\theta(Z \wedge 1 \in H) = \mathbb{P}_\theta(Z \in H \cap (0, 1)) + \mathbb{P}_\theta(Z \geq 1) \mathbf{1}_{(1 \in H)}$$

1.5 Elégséges statisztika, elégséges σ -algebra

$(\Omega, \mathcal{A}, \mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\})$ statisztikai mező. $\Omega = \mathfrak{X}$.

Definíció 2. $S : \mathfrak{X} \rightarrow \mathbb{R}^d$ *elégséges statisztika*, ha minden $A \in \mathcal{B}(\mathfrak{X})$ -re $\mathbb{P}_\theta(X \in A | S)$ -nek létezik θ -tól független, közös változata.

$\mathcal{S} \subset \mathcal{A}$ *elégséges σ -algebra*, ha minden $A \in \mathcal{B}(\mathfrak{X})$ -re $\mathbb{P}_\theta(X \in A | \mathcal{S})$ -nek létezik θ -tól független, közös változata.

Motiváció:

- Cél: A minta alapján \mathbb{P}^* meghatározása, vagy közelítése.
- Mikor nem hordoz plusz információt a generáló eloszlásról a minta az S statisztikához képest?
Ha a minta feltételes eloszlása az eloszláscsalád minden tagjára ugyanaz, azaz S elégséges.
- Ha S elégséges, akkor elég $S(X)$ -et megjegyezni a mintából.

1.6 Feltételes várható érték, feltételes valószínűség

Definíció 3. $X \in L^1$ valószínűségi változó, \mathcal{F} σ -algebra. $\mathbb{E}(X | \mathcal{F})$ olyan \mathcal{F} mérhető, integrálható változó, ami teljesíti a parciális átlagolási tulajdonságot:

$$\mathbb{E}(X \mathbf{1}_A) = \mathbb{E}(\mathbb{E}(X | \mathcal{F}) \mathbf{1}_A), \quad \text{minden } A \in \mathcal{F}\text{-re.}$$

- Ha $\mathcal{F} = \sigma(Y)$, akkor $\mathbb{E}(X | \mathcal{F}) = \mathbb{E}(X | Y) = g(Y)$ alakú, ahol g Borel mérhető.
- Ez valójában egy Radon-Nikodym derivált, $\mu(A) = \mathbb{E}(X \mathbf{1}_A)$ véges előjeles mérték \mathcal{F} -n, ami $\mathbb{P}|_{\mathcal{F}}$ -re nézve abszolút folytonos. $\mathbb{E}(X | \mathcal{F}) = \frac{d\mu}{d\mathbb{P}|_{\mathcal{F}}}$.
- Szemléletes jelentés $g(y) = \lim_{\varepsilon \rightarrow 0} \mathbb{E}(X | |Y - y| < \varepsilon)$ (Y eloszlása szerint majdnem minden y -ra).
- Ha Y diszkrét, $\mathbb{P}(Y = y) > 0$, akkor $g(y) = \mathbb{E}(X | Y = y)$. Ezt a jelölést szokás használni, a g függvényre, akkor is ha Y nem diszkrét.
- Ha $X \in L^2$, akkor $\mathbb{E}(X | \mathcal{F}) \in L^2$ és $X - \mathbb{E}(X | \mathcal{F})$ merőleges $L^2(\mathcal{F})$ -re. $\mathbb{E}(X | \mathcal{F})$ nem más, mint X vetülete $L^2(\mathcal{F})$ -re.
- $\mathbb{P}(A | \mathcal{F})$ egy másik jelölés $\mathbb{E}(\mathbf{1}_A | \mathcal{F})$ -re.

1.6.1 Feltételes eloszlás reguláris változata

Definíció 4. X valószínűségi változó, \mathcal{F} σ -algebra. $Q : \mathcal{B}(\mathbb{R}) \times \Omega \rightarrow [0, 1]$ az X feltételes eloszlásának reguláris változata, ha

- $\omega \mapsto Q(A, \omega)$, a $\mathbb{P}(X \in A | \mathcal{F})$ egy változata.
- $A \mapsto Q(A, \omega)$ valószínűségi mérték $\mathcal{B}(\mathbb{R})$ -en, minden ω -ra.

Ha X vektorváltozó, vagy általánosabban (S, \mathcal{S}) értékű, akkor $\mathcal{B}(\mathbb{R})$ \mathcal{S} -sel helyettesítendő.

Állítás 1. Ha X feltételes eloszlásának létezik reguláris változata Q , $h(X) \in L^1$, akkor $\mathbb{E}(h(X) | \mathcal{F}) = \int h(x) Q(dx, \cdot)$

Tétel 1 (bizonyítás nélkül). Ha X teljes szeparábilis metrikus térbe képez (pl. \mathbb{R}, \mathbb{R}^n), akkor tetszőleges feltétel mellett létezik a feltételes eloszlás reguláris változata.

1.6.2 A rendezett minta mindig elégséges

Definíció 5. X_1, \dots, X_n minta, $X_1^* = \min_i X_i$ a legkisebb megfigyelés, X_2^* a második legkisebb megfigyelés, ..., $X_n^* = \max_i X_i$ a legnagyobb megfigyelés. $X^* = (X_1^*, \dots, X_n^*)$ neve **rendezett minta**. Formulával:

$$X_k^* = \max\{\min_{i \in I} X_i : I \subset \{1, \dots, n\}, |I| = n - (k - 1)\}$$

Állítás 2. $(X, X^*) \stackrel{p}{=} (\Pi X^*, X^*)$, ahol Π az $\{1, \dots, n\}$ indexhalmaz X -től független véletlen permutációja ($\mathbb{P}(\Pi = \pi) = 1/n!$ minden $\pi \in S_n$ -re) és

$$\pi X^* = (X_{\pi^{-1}(1)}^*, \dots, X_{\pi^{-1}(n)}^*).$$

Az állítás alapján

$$\mathbb{P}(X \in H \mid X^*) = \mathbb{P}(\Pi X^* \in H \mid X^*) = \mathbb{P}(\Pi x \in H) \Big|_{x=X^*} = \frac{1}{n!} \sum_{\pi \in S_n} \mathbf{1}_{(\pi x \in H)} \Big|_{x=X^*}$$

Általánosabban: $\mathbb{E}(T(X) \mid X^*) = \mathbb{E}(T(\Pi X^*) \mid X^*) = \frac{1}{n!} \sum_{\pi} T(\pi X^*)$.

$X = (X_1, \dots, X_n)$ minta, $X^* = (X_1^* \leq X_2^* \leq \dots \leq X_n^*)$ a rendezett minta

- Π az X -től független véletlen permutáció, ekkor $X \stackrel{p}{=} \Pi X$

$$\mathbb{P}(\Pi X \in H) = \mathbb{E}(\mathbb{P}(\Pi X \in H \mid \Pi)) = \mathbb{E}(\mathbb{P}(\pi X \in H) \Big|_{\pi=\Pi})$$

de $\pi X \stackrel{p}{=} X$ minden rögzített permutációra és így $\mathbb{P}(\Pi X \in H \mid \Pi) = \mathbb{P}(X \in H)$.

- $(X, X^*) \stackrel{p}{=} (\Pi X, (\Pi X)^*) = (\Pi X, X^*)$.
- $X = \rho X^*$ alkalmas $\sigma(X)$ mérhető ρ permutációval és $\Pi X = (\Pi \rho) X^*$.
- $\Pi \rho$, X -től független véletlen permutáció, mert

$$\mathbb{P}(\Pi \rho = \pi \mid X) = \mathbb{P}(\Pi = \pi \rho^{-1} \mid X) = \frac{1}{n!}, \quad \text{minden } \pi \in S_n\text{-re.}$$

- $(X, X^*) \stackrel{p}{=} (\Pi X, X^*) = (\Pi \rho X^*, X^*) \stackrel{p}{=} (\Pi X^*, X^*)$.

1.7 Feltétel elégségessége

Állítás 3. \mathcal{P} dominált mértékcsalád, a λ domináló mérték valószínűségi mérték.

Ha $\frac{d\mathbb{P}}{d\lambda} \sigma(S)$ mérhető (pontosabban létezik $\sigma(S)$ mérhető változata) minden $\mathbb{P} \in \mathcal{P}$ -re, akkor S elégséges statisztika.

Megfordítás is igaz, bizonyítás később.

Tetszőleges $\mathbb{P} \in \mathcal{P}$ -re $\frac{d\mathbb{P}}{d\lambda} = f(S)$ és

$$\begin{aligned} \mathbb{P}(X \in A, S \in H) &= \mathbb{E}_{\lambda}(\mathbf{1}_{(X \in A)} \mathbf{1}_{(S \in H)} f(S)) \\ &= \mathbb{E}_{\lambda}(\mathbb{E}_{\lambda}(\mathbf{1}_{(X \in A)} \mid S) \mathbf{1}_{(S \in H)} f(S)) = \mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\lambda}(\mathbf{1}_{(X \in A)} \mid S) \mathbf{1}_{(S \in H)}) \end{aligned}$$

Azaz $\mathbb{E}_{\lambda}(\mathbf{1}_{(X \in A)} \mid S)$ a $\mathbb{P}(X \in A \mid S)$, $\mathbb{P} \in \mathcal{P}$ feltételes valószínűségek közös változata.

Valójában a Bayes szabályt használtuk a feltételes várható érték kiszámítására:

$$\mathbb{E}_{\mathbb{P}}(Y \mid \mathcal{F}) = \frac{\mathbb{E}_{\lambda}(Y \frac{d\mathbb{P}}{d\lambda} \mid \mathcal{F})}{\mathbb{E}_{\lambda}(\frac{d\mathbb{P}}{d\lambda} \mid \mathcal{F})}$$

Ha a λ domináló mérték véges, akkor $\tilde{\lambda} = \frac{1}{\lambda(\Omega)} \lambda$ valószínűségi mérték és az $\frac{d\mathbb{P}}{d\lambda} = \frac{d\mathbb{P}}{d\tilde{\lambda}} \cdot \frac{d\tilde{\lambda}}{d\lambda} = \lambda(\Omega) \frac{d\mathbb{P}}{d\tilde{\lambda}}$ pontosan akkor S mérhető, ha $\frac{d\mathbb{P}}{d\tilde{\lambda}}$ az.

1.8 Példák

1.8.1 Indikátor minta

X_1, \dots, X_n , n elemű indikátor minta ismeretlen $\theta \in [0, 1]$ paraméterrel.

- $\mathfrak{X} = \{0, 1\}^n$. Domináló mérték λ : a számláló mérték.

•

$$f_\theta(x) = \mathbb{P}_\theta(X = x) = \prod_i \mathbb{P}_\theta(X_i = x_i) = \prod_i \theta^{x_i} (1 - \theta)^{1-x_i} = \theta^{\sum x_i} (1 - \theta)^{n - \sum x_i}$$

- λ véges mérték

$$f_\theta(x) = g_\theta(S(x)), \quad S(x) = \sum x_i, \quad g_\theta(s) = \theta^s (1 - \theta)^{n-s},$$

$S = \sum X_i$, azaz **mintaösszeg**, elégséges statisztika.

Hogyan adható meg, a $\mathbb{P}_\theta(X \in A \mid S)$ feltételes eloszlás?

- S diszkrét változó. S lehetséges értékei $0, 1, \dots, n$, S binomiális n renddel és θ paraméterrel.

•

$$\mathbb{P}_\theta(X = x \mid S = s) = \frac{\mathbb{P}_\theta(X = x, S = s)}{\mathbb{P}_\theta(S = s)} = \begin{cases} 0 & \sum x_i \neq s \\ \frac{\theta^s (1-\theta)^{n-s}}{\binom{n}{s} \theta^s (1-\theta)^{n-s}} = \frac{1}{\binom{n}{s}} & \sum x_i = s \end{cases}$$

X S -re vonatkozó feltételes eloszlása:

n S elemű részhalmazai közül választunk egyet taláalomra, legyen ez I , $j \in I$ -re $Y_j = 1$ egyébként nulla.

X feltételes eloszlása S -re nézve azonos Y eloszlásával, azaz minden θ -ra ugyanaz.

1.8.2 $N(\theta, 1)$ minta

X_1, \dots, X_n , n elemű minta ismeretlen $\theta \in \mathbb{R}$ várható értékű és egységnyi szórású normális eloszlásból.

- $\mathfrak{X} = \mathbb{R}^n$, domináló mérték λ lehet a Lebesgue mérték, de lehet a X eloszlása $\theta = 0$ mellett.
- A sűrűségfüggvény a Lebesgue mértékre nézve

$$f_\theta(x) = \prod_i \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i - \theta)^2} = f_0(x) \prod_i e^{x_i \theta - \frac{1}{2}\theta^2} = f_0(x) \exp\left\{\theta \sum_i x_i - \frac{n}{2}\theta^2\right\}$$

•

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}_0}(x) = \frac{f_\theta(x)}{f_0(x)} = \exp\left\{\theta S(x) - \frac{n}{2}\theta^2\right\}, \quad S(x) = \sum_i x_i$$

$S = \sum X_i$, azaz **mintaösszeg**, elégséges statisztika.

Hogyan írható le a $\mathbb{P}_\theta(X \in A \mid S)$ feltételes eloszlás?

- $\bar{X} = \frac{1}{n} \sum X_i$ a mintaátlag.
- $(X_1 - \bar{X}, \dots, X_n - \bar{X})$ eloszlása nem függ θ -tól és később kiszámoljuk, hogy független S -től így $e = (1, \dots, 1)$ jelöléssel

$$\mathbb{P}_\theta(X \in H \mid S) = \mathbb{P}_\theta(X - \bar{X}e + xe \in H) \big|_{x=S/n} = \mathbb{P}_0(X - \bar{X}e + xe \in H) \big|_{x=S/n}$$

2. 2021.02.15.

2.1 Emlékeztető

$(\Omega, \mathcal{A}, \mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\})$ statisztikai mező. $\Omega = \mathfrak{X}$.

Definíció 6. $S : \mathfrak{X} \rightarrow \mathbb{R}^d$ **elégseges statisztika**, ha minden $A \in \mathcal{B}(\mathfrak{X})$ -re $\mathbb{P}_\theta(X \in A | S)$ -nek létezik θ -tól független, közös változata.

$\mathcal{S} \subset \mathcal{A}$ **elégseges σ -algebra**, ha minden $A \in \mathcal{B}(\mathfrak{X})$ -re $\mathbb{P}_\theta(X \in A | \mathcal{S})$ -nek létezik θ -tól független, közös változata.

Tétel 2 (Neymann faktorizációs tétel). $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ dominált mértékcsalád az $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ minta téren, λ domináló mértékkel.

Az S statisztika pontosan akkor elégseges, ha az $f_\theta = \frac{d\mathbb{P}_\theta}{d\lambda}$ sűrűségek

$$f_\theta(x) = g_\theta(S(x))h(x) \quad (*)$$

alakban is megadhatóak, alkalmas g_θ, h mérhető függvényekkel.

- Először $\lambda = \mathbb{P}_0 \in \mathcal{P}$ -ra ellenőrizzük, hogy S pontosan akkor elégseges, ha a $\frac{d\mathbb{P}_\theta}{d\mathbb{P}_0} = g_\theta \circ S$ alakú, azaz a sűrűségeknek van S mérhető változata.
- \mathcal{P}' a \mathcal{P} konvex keverékeiből álló család

$$\mathcal{P}' = \left\{ \sum_i c_i \mathbb{P}_i : c_i \geq 0, \sum_i c_i = 1, \mathbb{P}_i \in \mathcal{P} \right\}$$

Megmutatjuk, hogy S pontosan akkor elégseges \mathcal{P} -re nézve, ha \mathcal{P}' -re az, továbbá létezik $\mathbb{P}_0 \in \mathcal{P}'$ ami, \mathcal{P}' -t dominálja (Halmos–Savage tétel) és $(*)$ pontosan akkor teljesül \mathcal{P} -re, ha \mathcal{P}' -re. .

- Ha λ tetszőleges domináló mérték, akkor

$$S \text{ egs.} \implies \frac{d\mathbb{P}_\theta}{d\lambda} = \frac{d\mathbb{P}_\theta}{d\mathbb{P}_0} \frac{d\mathbb{P}_0}{d\lambda} = (g_\theta \circ S)h, \quad (*) \implies \frac{d\mathbb{P}_\theta}{d\mathbb{P}_0} = \frac{d\mathbb{P}_\theta}{d\lambda} : \frac{d\mathbb{P}_0}{d\lambda} = \frac{(g_\theta \circ S)h}{(g_0 \circ S)h} = \tilde{g}_\theta \circ S \implies S \text{ egs..}$$

2.2 Neymann faktorizációs tétel, $\lambda = \mathbb{P}_0$

Tétel 3. $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ dominált mértékcsalád az $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ minta téren, $\mathbb{P}_0 \in \mathcal{P}$ domináló mértékkel.

Az S statisztika pontosan akkor elégseges, ha az $f_\theta = \frac{d\mathbb{P}_\theta}{d\mathbb{P}_0}$ sűrűségeknek van $\sigma(S)$ mérhető változata.

A \mathcal{S} σ -algebra pontosan akkor elégseges, ha az $f_\theta = \frac{d\mathbb{P}_\theta}{d\mathbb{P}_0}$ sűrűségeknek van \mathcal{S} mérhető változata.

- \mathcal{S} elégseges σ -algebra, ekkor $\mathbb{P}_\theta(X \in H | \mathcal{S}) = \mathbb{P}_0(X \in H | \mathcal{S})$.
- Teljes valószínűség tétellel

$$\mathbb{P}_\theta(X \in H) = \mathbb{E}_\theta(\mathbb{P}_\theta(X \in H | \mathcal{S})) = \mathbb{E}_0\left(\frac{d\mathbb{P}_\theta|_{\mathcal{S}}}{d\mathbb{P}_0|_{\mathcal{S}}} \mathbb{E}_0(\mathbf{1}_{(X \in H)} | \mathcal{S})\right) = \mathbb{E}_0\left(\mathbf{1}_{(X \in H)} \frac{d\mathbb{P}_\theta|_{\mathcal{S}}}{d\mathbb{P}_0|_{\mathcal{S}}}\right)$$

- Azt kaptuk, hogy $\frac{d\mathbb{P}_\theta|_{\mathcal{S}}}{d\mathbb{P}_0|_{\mathcal{S}}}$, ami \mathcal{S} mérhető, a $\frac{d\mathbb{P}_\theta}{d\mathbb{P}_0}$ Radon-Nikodym derivált egy változata.

- A fordított irányt korábban leellenőriztük a Bayes szabály felhasználásával.

\mathcal{P} és \mathcal{P}' \mathcal{P} mértékcsalád a mintatéren

$$\mathcal{P}' = \left\{ \sum_i c_i \mathbb{P}_i : c_i \geq 0, \sum_i c_i = 1, \mathbb{P}_i \in \mathcal{P} \right\}$$

Állítás 4. S pontosan akkor elégséges \mathcal{P} -re nézve, ha \mathcal{P}' -re elégséges.

- Legyen H rögzített és $q(S) = \mathbb{P}(X \in H \mid S)$ a \mathcal{P} -re vonatkozó közös változat.
- Tetszőleges $\mathbb{P} \in \mathcal{P}'$ -re $\mathbb{P} = \sum_i c_i \mathbb{P}_i$ alakú és

$$\mathbb{P}(X \in H, S \in A) = \sum_i c_i \int \mathbb{P}_i(X \in H \mid S) \mathbf{1}_{(S \in A)} d\mathbb{P}_i = \sum_i c_i \int q(S) \mathbf{1}_{(S \in A)} d\mathbb{P}_i = \int q(S) \mathbf{1}_{(S \in A)} d\mathbb{P}$$

- $q(S)$ S mérhető és teljesíti a parciális átlagolási tulajdonságot $q(S) = \mathbb{P}(X \in H \mid S)$ minden $\mathbb{P} \in \mathcal{P}'$ -re.
- Másik irányhoz elég észrevenni, hogy ha S elégséges egy mértékcsaládra, akkor tetszőleges részcsaládra is az és $\mathcal{P} \subset \mathcal{P}'$.

\mathcal{P} és \mathcal{P}' \mathcal{P} mértékcsalád a mintatéren λ domináló mértékkel.

$$\mathcal{P}' = \left\{ \sum_i c_i \mathbb{P}_i : c_i \geq 0, \sum_i c_i = 1, \mathbb{P}_i \in \mathcal{P} \right\}$$

Ha $\mathbb{P} = \sum_i c_i \mathbb{P}_i \in \mathcal{P}'$, akkor

$$\frac{d\mathbb{P}}{d\lambda} = \sum_i c_i \frac{d\mathbb{P}_i}{d\lambda}$$

Állítás 5. λ domináló mérték \mathcal{P} -hez, S statisztika. Ha

$$\frac{d\mathbb{P}_\theta}{d\lambda} = (g_\theta \circ S)h, \quad \text{minden } \mathbb{P}_\theta \in \mathcal{P}\text{-re,}$$

akkor ugyanez igaz minden $\mathbb{P} \in \mathcal{P}'$ -re.

2.3 Halmos–Savage tétel

Tétel 4. \mathcal{P} konvex keverésre zárt, dominált mértékcsalád $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ -en.

Ekkor létezik $\mathbb{P}_0 \in \mathcal{P}$, ami dominálja \mathcal{P} -t.

- λ σ -véges domináló mérték. Feltehető, hogy véges mérték. Ugyanis, ha $\mathfrak{X} = \cup_n A_n$ és $\lambda(A_n) < \infty$, akkor $c_n > 0$, $\sum_n c_n \lambda(A_n) < \infty$ esetén

$$\lambda'(H) = \sum_n c_n \lambda(A_n \cap H) = \int_H \sum_n c_n \mathbf{1}_{A_n} d\lambda$$

véges mérték és λ, λ' ekvivalensek és λ' véges.

- λ véges domináló mérték. Legyen

$$c = \sup \left\{ \lambda(f > 0) : f = \frac{d\mathbb{P}}{d\lambda}, \mathbb{P} \in \mathcal{P} \right\}$$

- Elég megmutatni, hogy c maximum, azaz létezik $\mathbb{P}_0 \in \mathcal{P}$, amire $c = \lambda(\frac{d\mathbb{P}_0}{d\lambda} > 0)$.

Lemma 1. \mathcal{P} konvex keverésre zárt, véges λ mértékkel dominált mértékcsalád $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ -en. Ha $\mathbb{P}_0 \in \mathcal{P}$ -re

$$\lambda\left(\frac{d\mathbb{P}_0}{d\lambda} > 0\right) = c = \sup\left\{\lambda(f > 0) : f = \frac{d\mathbb{P}}{d\lambda}, \mathbb{P} \in \mathcal{P}\right\}$$

Ekkor \mathbb{P}_0 dominálja \mathcal{P} -t.

- Tetszőleges $\mathbb{P} \in \mathcal{P}$ -re $\mathbb{P}' = \frac{1}{2}(\mathbb{P} + \mathbb{P}_0) \in \mathcal{P}$ a konvex keverésre zárttság miatt.
- $f = \frac{d\mathbb{P}}{d\lambda}$, $f_0 = \frac{d\mathbb{P}_0}{d\lambda}$, $f, f_0 \geq 0$. Ekkor $2\frac{d\mathbb{P}'}{d\lambda} = f + f_0$ és

$$c \geq \lambda((f + f_0) > 0) \geq \lambda(f_0 > 0) = c, \quad \text{mert } \{f + f_0 > 0\} = \{f > 0\} \cup \{f_0 > 0\}.$$
- $\lambda(\{f > 0\} \cap \{f_0 > 0\}) = 0$ és

$$\mathbb{P}(X \in H) = \int_H f d\lambda = \int_H f \mathbf{1}_{(f_0 > 0)} d\lambda = \int_H \frac{f}{f_0} \mathbf{1}_{(f_0 > 0)} f_0 d\lambda = \int_H \frac{f}{f_0} \mathbf{1}_{(f_0 > 0)} d\mathbb{P}_0$$
- Legyen $\mathbb{P}_n \in \mathcal{P}$ olyan, amire az $f_n = \frac{d\mathbb{P}_n}{d\lambda}$ jelöléssel

$$\lambda(f_n > 0) \geq c - \frac{1}{n}$$
- $\mathbb{P}_0 = \sum_{n=1}^{\infty} 2^{-n} \mathbb{P}_n \in \mathcal{P}$, ekkor $f = \frac{d\mathbb{P}_0}{d\lambda} = \sum_n 2^{-n} f_n$ és

$$\{f > 0\} = \cup_n \{f_n > 0\}, \quad c = \sup_n \lambda(f_n > 0) \leq \lambda(f > 0) \leq c.$$

Neymann-faktorizációs tétel összefoglalás

Tétel 5. $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ dominált mértékcsalád az $(\mathfrak{X}, \mathcal{B}(\mathfrak{X}))$ minta téren, λ domináló mértékkel.

Az S statisztika pontosan akkor elégséges, ha az $f_\theta = \frac{d\mathbb{P}_\theta}{d\lambda}$ sűrűségek

$$f_\theta(x) = g_\theta(S(x))h(x)$$

alakban is megadhatóak, alkalmas g_θ, h mérhető függvényekkel.

- Feltehető, hogy \mathcal{P} zárt a konvex keverésre.
- Ha \mathcal{P} zárt a konvex keverésre, akkor a domináló mérték választható $\mathbb{P}_0 \in \mathcal{P}$ -nek is.

Lemma 2. $\mathbb{P}_0 \in \mathcal{P}$ domináló mérték \mathcal{P} -hez. Ekkor S pontosan akkor elégséges, ha $\frac{d\mathbb{P}}{d\mathbb{P}_0}$ -nak van $\sigma(S)$ mérhető változata minden $\mathbb{P} \in \mathcal{P}$ -re.

Elégségesség, összefoglalás \mathcal{P} mértékcsalád, $\mathcal{P}' = \{\sum_i c_i \mathbb{P}_i : c_i \geq 0, \sum c_i = 1, \mathbb{P}_i \in \mathcal{P}\}$.

Állítás 6 (Halmos–Savage tétel). Ha \mathcal{P} dominált, akkor van olyan $\mathbb{P}_0 \in \mathcal{P}'$, ami dominálja \mathcal{P}' -t és így \mathcal{P} -t is.

Állítás 7. S pontosan akkor elégséges statisztika, ha $\frac{d\mathbb{P}}{d\mathbb{P}_0}$ -nak létezik $\sigma(S)$ mérhető változata minden $\mathbb{P} \in \mathcal{P}$ -re.

\mathcal{S} pontosan akkor elégséges σ -algebra, ha $\frac{d\mathbb{P}}{d\mathbb{P}_0}$ -nak létezik \mathcal{S} mérhető változata minden $\mathbb{P} \in \mathcal{P}$ -re.

- $\mathcal{N}(\mathcal{P}) = \{N : \mathbb{P}(N) = 0, \forall \mathbb{P} \in \mathcal{P}\}$, $\mathcal{N}(\mathbb{P}_0) = \{N : \mathbb{P}_0(N) = 0\}$. Ekkor $\mathcal{N}(\mathcal{P}) = \mathcal{N}(\mathcal{P}') = \mathcal{N}(\mathbb{P}_0)$.
- Tetszőleges \mathcal{S} σ -algebrára, legyen $\mathcal{S}^* = \{A \circ N : A \in \mathcal{S}, N \in \mathcal{N}(\mathcal{P})\}$ a \mathcal{S} teljessé tétele, vagy teljes burka.
- Belátjuk, hogy \mathcal{S}^* σ -algebra és $\frac{d\mathbb{P}}{d\mathbb{P}_0}$ -nak pontosan akkor létezik \mathcal{S} mérhető változata ha \mathcal{S}^* mérhető.

Állítás 8. S pontosan akkor elégséges statisztika, ha $\frac{d\mathbb{P}}{d\mathbb{P}_0} \sigma(S)^*$ mérhető minden $\mathbb{P} \in \mathcal{P}$ -re.

\mathcal{S} pontosan akkor elégséges σ -algebra, ha $\frac{d\mathbb{P}}{d\mathbb{P}_0} \mathcal{S}^*$ mérhető minden $\mathbb{P} \in \mathcal{P}$ -re.

2.4 σ -algebra teljess  t tele

$A \circ B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$ az A  s B szimmetrikus differenci ja.

 ll t s 9. (Ω, \mathcal{A}) m rhet  t r. $\mathcal{S} \subset \mathcal{A}$ σ -algebra, $\mathcal{N} \subset \mathcal{A}$ (nem  res) lesz ll   s σ -uni  z rt. Ekkor $\mathcal{S}^* = \{S \circ N : S \in \mathcal{S}, N \in \mathcal{N}\}$ σ -algebra.

- $(S \circ N)^c = S^c \circ N$,
- $N = (\cup_i (S_i \circ N_i)) \circ (\cup_i S_i) \subset \cup_i N_i$, mert

$$(\cup_i S_i) \setminus (\cup_i N_i) \subset \cup_i (S_i \setminus N_i) \subset (\cup_i S_i) \circ (\cup_i N_i)$$

$\cup_i N_i \in \mathcal{N}$  s $N \subset \cup_i N_i$ miatt $N \in \mathcal{N}$.

$$\cup_i (S_i \circ N_i) = (\cup_i S_i) \circ N \in \mathcal{S}^*, \quad (B \circ (A \circ B) = A).$$

A továbbiakban $\mathcal{S}^* = \sigma(\mathcal{S} \cup \mathcal{N}(\mathcal{P}))$. Ha $\{T' \neq T\} \in \mathcal{N}(\mathcal{P})$, akkor azt mondjuk, hogy T' a T egy v ltozata.

 ll t s 10. T -nek pontosan akkor l tezik \mathcal{S} m rhet  v ltozata, ha \mathcal{S}^* m rhet .

- \Rightarrow : T' a T statisztika \mathcal{S} m rhet  v ltozata.
 $N = \{T \in H\} \circ \{T' \in H\} \subset \{T \neq T'\} \in \mathcal{N}(\mathcal{P})$  gy $N \in \mathcal{N}(\mathcal{P})$  s $\{T \in H\} = \{T' \in H\} \circ N \in \mathcal{S}^*$.
- \Leftarrow : T skal r  rt k , \mathcal{S}^* m rhet . $\{T < q\} = H_q \circ N_q$, ahol $H_q \in \mathcal{S}$  s $N_q \in \mathcal{N}(\mathcal{P})$.
Legyen $T'(x) = \inf\{q \in \mathbb{Q} : x \in H_q\}$.

$$\{T' < c\} = \cup_{q < c} H_q \in \mathcal{S} \quad \Rightarrow \quad T' \text{ } \mathcal{S} \text{ m rhet }$$

$$\{T < c\} \circ \{T' < c\} = (\cup_{q < c} H_q \circ N_q) \circ (\cup_{q < c} H_q) \subset \cup_q N_q \in \mathcal{N}(\mathcal{P}) \quad \Rightarrow \quad \{T \neq T'\} \in \mathcal{N}(\mathcal{P}).$$

Ha $T : \mathfrak{X} \rightarrow \mathbb{R}^p$, akkor a koordin ta lek pez sekre a fenti  rvel s.

2.5 Minim lis el gs ges g

Def nici  7. \mathcal{S} minim lis el gs ges σ -algebra, ha el gs ges  s tetsz leges \mathcal{S}' el gs ges σ -algebr ra: $\mathcal{S}^* \subset (\mathcal{S}')^*$.

\mathcal{S} minim lis el gs ges statisztika, ha $\sigma(\mathcal{S})$ minim lis el gs ges σ -algebra.

- Ha \mathcal{S} minim lis el gs ges  s T el gs ges, akkor $\sigma(\mathcal{S}) \subset \sigma(\mathcal{S})^* \subset \sigma(T)^*$  s $\mathbb{P}(S = g(T)) = 1$ minden $\mathbb{P} \in \mathcal{P}$ -re, alkalmas m rhet  g -vel.
- Ha \mathcal{P} domin lt, akkor $\mathcal{S} = \sigma(\{\frac{d\mathbb{P}}{d\mathbb{P}_0} : \mathbb{P} \in \mathcal{P}\})$, minim lis el gs ges σ -algebra.
 \mathcal{S}' pontosan akkor el gs ges, ha $\frac{d\mathbb{P}}{d\mathbb{P}_0} \sim (\mathcal{S}')^*$ minden $\mathbb{P} \in \mathcal{P}$ -re, amib l $\mathcal{S} \subset (\mathcal{S}')^*$.
- Ha $\mathcal{P} = \{\mathbb{P}_n : n \geq 1\}$ megs ml lhat , akkor domin lt  s van minim lis el gs ges statisztika.
 $\mathbb{P}_0 = \sum_n 2^{-n} \mathbb{P}_n$ domin l  m rt k. $\mathcal{S} = \sigma(\{\frac{d\mathbb{P}}{d\mathbb{P}_0} : \mathbb{P} \in \mathcal{P}\})$ megs ml lhat an gener lt.
 \mathcal{S} gener torrendszer $\{H_n : n \geq 1\}$, $T = \sum_n 3^{-n} 2 \mathbf{1}_{H_n}$, ekkor $\mathcal{S} = \sigma(T)$.
(HF. ellen rizni,  tmutat s: $\{\{3^{n-1}T\} > 1/2\} = H_n$).

- Ha $\mathcal{P}_0 \subset \mathcal{P}$, $\mathcal{N}(\mathcal{P}_0) = \mathcal{N}(\mathcal{P})$, továbbá T elégséges \mathcal{P} -re és minimális elégséges \mathcal{P}_0 -ra, akkor minimális elégséges \mathcal{P} -re is.
- \mathcal{S} elégséges σ -algebra \mathcal{P} -re $\implies \mathcal{S}$ elégséges. \mathcal{P}_0 -ra $\implies \sigma(T)^* \subset \mathcal{S}^*$, ahol $\mathcal{N}(\mathcal{P}_0) = \mathcal{N}(\mathcal{P})$ -vel tettük teljessé a σ -algebrákat, azaz T minimális elégséges \mathcal{P} -re.

2.6 Feltétel minimális elégségességre

$\mathfrak{X}' \subset \mathfrak{X}$ teljes mértékű, ha a komplementere $\mathcal{N}(\mathcal{P})$ eleme.

Tétel 6. $\mathcal{P} = \{\mathbb{P}_\theta = f_\theta d\lambda : \theta \in \Theta\}$ megszámlálható mértékcsalád az $(\mathfrak{X}, \mathcal{B})$ mintatéren.

Jelölés: $x \sim y$, ha $\Theta_x = \{\theta \in \Theta : f_\theta(x) \neq 0\} = \Theta_y$ és $\theta \mapsto f_\theta(x)/f_\theta(y)$ konstans $\Theta_x = \Theta_y$ -on.

1. Ha S elégséges, akkor létezik \mathfrak{X}' teljes mértékű rész úgy, hogy $x, y \in \mathfrak{X}'$ -re, $S(x) = S(y) \implies x \sim y$.
2. Ha S elégséges és $x \sim y \implies Sx = Sy$, akkor S minimális elégséges.

1. rész

- S elégséges. A Neymann faktorizációs tétel szerint $f_\theta \neq (g_\theta \circ S)h$ λ -nullmértékű és így \mathbb{P}_θ nullmértékű is minden θ -ra.
- $\mathbb{P}_\theta(h = 0) = 0$ minden θ -ra.
- Legyen $\mathfrak{X}' = \{x \in \mathfrak{X} : f_\theta(x) = g_\theta(S(x))h(x), h(x) \neq 0\}$. \mathfrak{X}' teljes mértékű.
- Ha $x, y \in \mathfrak{X}'$, akkor $h(x), h(y) \neq 0$ és $f_\theta(x) = g_\theta(S(x))h(x)$, $f_\theta(y) = g_\theta(S(y))h(y)$.
 $x, y \in \mathfrak{X}', S(x) = S(y) \implies \Theta_x = \{\theta : g_\theta(S(x))h(x) \neq 0\} = \{\theta : g_\theta(S(y))h(y) \neq 0\}$
és Θ_x -en $f_\theta(x)/f_\theta(y) = h(x)/h(y)$, ami nem függ θ -tól.

2. rész:

- \mathcal{P} megszámlálható \implies létezik T minimális elégséges statisztika és 1. alkalmazható T -re.
-

$$x, y \in \mathfrak{X}', \quad T(x) = T(y) \implies x \sim y \implies S(x) = S(y).$$

Azaz valamilyen g függvénnyel $S(x) = g(T(x))$, ha $x \in \mathfrak{X}'$, $g(T)$ az S egy változata, S „egyszerűbb” mint T , azaz $\sigma(S)^* \subset \sigma(T)^*$.

- Ugyanakkor $\sigma(T)^* \subset \sigma(S)^*$, mert S elégséges és T minimális elégséges.

Pontosítások

- $x, y \in \mathfrak{X}'$ -re $T(x) = T(y) \implies S(x) = S(y)$. Feltehető, hogy ez mindenhol teljesül.
 $\tilde{S}(x) = \text{tg}(S(x))\mathbf{1}_{(x \in \mathfrak{X}')} + 2\mathbf{1}_{(x \notin \mathfrak{X}')} , \tilde{T}(x) = \text{tg}(T(x))\mathbf{1}_{(x \in \mathfrak{X}')} + 2\mathbf{1}_{(x \notin \mathfrak{X}')} (koordinátánként):$
 $\sigma(S)^* = \sigma(\tilde{S})^*, \quad \sigma(T)^* = \sigma(\tilde{T})^*, \quad \tilde{T}(x) = \tilde{T}(y) \implies \tilde{S}(x) = \tilde{S}(y)$
- Ha S, T mérhető és $T(x) = T(y) \implies S(x) = S(y)$, akkor általában **nem** igaz, hogy $\sigma(S) \subset \sigma(T)$.

A tétel bizonyításának vége azon múlik, hogy ha $T : (\mathfrak{X}, \mathcal{B}(\mathfrak{X})) \rightarrow (\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$ mérhető, ahol \mathfrak{X} teljes szeparábilis metrikus tér, pl. \mathbb{R}^n , akkor

$$\sigma(T) = \{\{T \in H\} : H \in \mathcal{B}(\mathbb{R}^p)\} = \tilde{\sigma}(T) = \{A \in \mathcal{B}(\mathfrak{X}) : x \in A \implies [x]_T \subset A\}$$

ahol $[x]_T = \{y : Tx = Ty\}$.

- A bizonyítás végén $T(x) = T(y) \implies S(x) = S(y)$ miatt $[x]_T \subset [x]_S$ és
 $\tilde{\sigma}(S) = \{A \in \mathcal{B}(\mathfrak{X}) : x \in A \implies [x]_S \subset A\} \subset \tilde{\sigma}(T) = \{A \in \mathcal{B}(\mathfrak{X}) : x \in A \implies [x]_T \subset A\}.$

2.7 Minimális elégséges statisztika jellemzése

$\mathfrak{X}' \subset \mathfrak{X}$ teljes mértékű, ha a komplementere $\mathcal{N}(\mathcal{P})$ eleme.

Állítás 11. $\mathcal{P} = \{\mathbb{P}_\theta = f_\theta d\lambda : \theta \in \Theta\}$ megszámlálható mértékcsalád az $(\mathfrak{X}, \mathcal{B})$ mintatéren.

Jelölés: $x \sim y$, ha $\Theta_x = \{\theta \in \Theta : f_\theta(x) \neq 0\} = \Theta_y$ és $\theta \mapsto f_\theta(x)/f_\theta(y)$ konstans $\Theta_x = \Theta_y$ -on.

S pontosan akkor minimális elégséges statisztika, ha megadható olyan \mathfrak{X}' teljes mértékű része a mintatérnek, amire $x, y \in \mathfrak{X}'$ esetén $x \sim y \iff S(x) = S(y)$.

- Legyen $\mathbb{P}_0 = \sum_\theta c_\theta \mathbb{P}_\theta$, ahol $c_\theta > 0$, $\sum c_\theta = 1$. Ekkor $\mathbb{P}_0 \in \mathcal{P}'$ és domináló mérték.
- Ha S minimális elégséges, akkor $\sigma(S)^* = \sigma^*\left(\left\{\frac{d\mathbb{P}_\theta}{d\mathbb{P}_0} : \theta \in \Theta\right\}\right)$.

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}_0} = \frac{f_\theta}{h} \mathbf{1}_{(h>0)} \sim \sigma(S)^*, \quad h = \sum c_\theta f_\theta$$

- Létezik \mathfrak{X}' teljes mértékű és g_θ mérhető függvények, hogy $f_\theta(x) = g_\theta(S(x))h(x)$, ha $x \in \mathfrak{X}'$ és ezzel a választással a feltétel teljesül.
- Megfordítva, $S(x) = S(y) \implies x \sim y$ \mathfrak{X}' -n miatt S elégséges. Ha $x \sim y \implies S(x) = S(y)$ \mathfrak{X}' -en, akkor a bizonyítás végét lehet módosítani.

Miért kell megszámlálható részcsaládra áttérni?

- Egy elemű X minta az

$$f_\theta(x) = \theta e^{-2\theta|x|}, \quad \theta \in \Theta = (0, \infty)$$

sűrűségfüggvénnyel megadott eloszláscsaládból.

- $S = |X|$ elégséges.
-

$$\tilde{f}_\theta(x) = \begin{cases} \theta e^{-2\theta|x|} & x \neq \theta \\ 1 & x = \theta \end{cases}$$

- \tilde{f}_θ választás mellett akarjuk $|X|$ minimális elégségességét ellenőrizni.

$$\frac{\tilde{f}_\theta(x)}{\tilde{f}_\theta(y)} = \begin{cases} e^{-2\theta(|x|-|y|)} & x \neq \theta \neq y \\ e^{-2\theta|x|} & x \neq \theta = y \\ e^{2\theta|y|} & x = \theta \neq y \\ 1 & x = \theta = y \end{cases}$$

Milyen x, y párokra teljesül, hogy a hányados nem függ θ -tól? Csak akkor, ha $x = y$, holott $S = |X|$ elégséges.

- Ha csak megszámlálható $\Theta_0 \subset \Theta$ részcsaládot veszünk, akkor $\mathfrak{X}' = \mathbb{R} \setminus \Theta_0$ teljes mértékű, $T(x) = |x| \mathbf{1}_{(x \in \mathfrak{X}')} + x \mathbf{1}_{(x \notin \mathfrak{X}')}$. $T(X) = |X|$ majdnem mindenütt és T minimális elégséges a feltétel alapján.

2.8 Példák

2.8.1 Példa, $U(\theta, \theta + 1)$ egyenletes eloszláscsalád

- $\Theta = \mathbb{R}$, domináló mérték a Lebesgue mérték. Egy megfigyelés sűrűségfüggvénye $f_\theta(x) = \mathbf{1}_{(\theta < x < \theta+1)}$.
- Minta tér: $\mathfrak{X} = \mathbb{R}^n$, a minta sűrűségfüggvénye:

$$f_\theta(x) = \prod_i f_\theta(x_i) = \prod_i \mathbf{1}_{(\theta < x_i < \theta+1)} = \mathbf{1}_{(\theta < \min x_i, \max x_i < \theta+1)} = \mathbf{1}_{(\max x_i - 1 < \theta < \min x_i)}$$

- $X_1^* = \min X_i$, $X_n^* = \max X_i$, $S = (X_1^*, X_n^*)$ elégséges statisztika.
- Van-e egyszerűbb elégséges statisztika?
- $\Theta' = \mathbb{Q} \subset \Theta$ megszámlálható részcsalád, $\mathcal{N}(\{\mathbb{P}_\theta : \theta \in \Theta'\}) = \mathcal{N}(\{\mathbb{P}_\theta : \theta \in \Theta\})$. Így ha S minimális elégséges a szűkebb családra nézve, akkor az eredetire is az.
- $\mathfrak{X}' = \{x \in \mathfrak{X} : \max x_i - 1 < \min x_i\}$ teljes mértékű része \mathfrak{X} -nek. $x, y \in \mathfrak{X}'$ -re

$$x \sim y, \quad \text{ha} \quad \Theta_x = \{\theta : f_\theta(x) \neq 0\} = \Theta_y, \quad \theta \mapsto \frac{f_\theta(x)}{f_\theta(y)} = \frac{\mathbf{1}_{(\max x_i - 1 < \theta < \min x_i)}}{\mathbf{1}_{(\max y_i - 1 < \theta < \min y_i)}}, \quad \text{konstans } \Theta_x\text{-n}$$

Milyen $x, y \in \mathfrak{X}'$ párokra teljesül $x \sim y$?

$$\{\theta \in \mathbb{Q} : f_\theta(x) \neq 0\} = \mathbb{Q} \cap (\max x_i - 1, \min x_i) = \{\theta \in \mathbb{Q} : f_\theta(y) \neq 0\} = \mathbb{Q} \cap (\max y_i - 1, \min y_i)$$

azaz $\min x_i = \min y_i$ és $\max x_i = \max y_i$ azaz S minimális elégséges.

2.8.2 Példa, Cauchy eloszlás eltolásparaméteres családjá

- $\Theta = \mathbb{R}$, egy mintaelem sűrűségfüggvénye $f_\theta(x) = \frac{1}{\pi} \frac{1}{1+(x-\theta)^2}$.
- Minta tér: $\mathfrak{X} = \mathbb{R}^n$, a minta sűrűségfüggvénye:

$$f_\theta(x) = \prod_k f_\theta(x_k) = \frac{1}{\pi^n} \cdot \frac{1}{\prod_k (1 + (x_k - \theta)^2)}$$

A rendezett minta mindig elégséges statisztika. Nem látszik hogyan lehetne ennél egyszerűbbet megadni.

- Elég $\Theta' = \mathbb{Q}$ részcsaládot nézni.

•

$$\frac{f_\theta(x)}{f_\theta(y)} = \frac{\prod (1 + (y_k - \theta)^2)}{\prod (1 + (x_k - \theta)^2)} = \frac{p_y(\theta)}{p_x(\theta)}$$

ahol $p_x(\theta)$ $2n$ -ed fokú polinom θ -ban, 1 főegyütthatóval. Ha ez nem függ θ -tól, akkor a hányados konstans egy és $p_x(\theta) = p_y(\theta)$, $\theta \in \mathbb{Q}$, de akkor a két polinom azonos.

$p_x = p_y$ miatt azonosak a gyökök (multiplicitással együtt), p_x gyökei $\{\pm i + x_k : k = 1, \dots, n\}$, így

$$\{\pm i + x_k : k = 1, \dots, n\} = \{\pm i + y_k : k = 1, \dots, n\} \implies \{x_k : k = 1, \dots, n\} = \{y_k : k = 1, \dots, n\}$$

Azaz x és y legfeljebb a koordináták sorrendjében térhet el, a rendezett minta minimális elégséges statisztika.

2.8.3 Példa, Indikátor minta

- $\Theta = (0, 1)$, $\mathfrak{X} = \{0, 1\}^n$, $f_\theta(x) = \mathbb{P}_\theta(X = x) = \theta^s(1 - \theta)^{n-s}$, ahol $s = \sum x_i$.
- $S = \sum X_i$ elégséges. Van-e egyszerűbb elégséges statisztika?

$$\frac{f_\theta(x)}{f_\theta(y)} = \left(\frac{\theta}{1 - \theta} \right)^{\sum x_i - \sum y_i}$$

ez csak akkor nem függ θ -tól, ha $\sum x_i = \sum y_i$, azaz a mintaösszeg minimális elégséges.

- Ugyanez másképp.
 \mathcal{S} elégséges σ -algebra. $\mathbb{P}_\theta(X \in A \mid \mathcal{S})$ nem függ θ -tól, T_A jelöli a közös változatot.
- S a mintaösszeg elégséges, $\mathbb{E}_\theta(\mathbf{1}_{(X \in A)} - T_A \mid S) = h(S)$ a közös változat. Ekkor

$$\mathbb{E}_\theta(h(S)) = \mathbb{E}_\theta(\mathbf{1}_{(X \in A)} - T_A) = \mathbb{P}_\theta(A) - \mathbb{E}_\theta(\mathbb{E}_\theta(\mathbf{1}_A \mid \mathcal{S})) = 0.$$

- S binomiális eloszlású n -renddel és θ paraméterrel:

$$0 = \mathbb{E}_\theta(h(S)) = \sum_{k=0}^n h(k) \binom{n}{k} \theta^k (1 - \theta)^{n-k} \quad \text{minden } \theta \in (0, 1).$$

Így $h(k) = 0$, $k = 0, 1, \dots, n$ és $0 = h(S) = \mathbb{E}_\theta(\mathbf{1}_A - T_A \mid S)$.

- Ha $A \in \sigma(S)$, akkor $h(S) = \mathbf{1}_A - \mathbb{E}(T_A \mid S)$ és $T_A = \mathbf{1}_A$, vagyis $A \in \mathcal{S}^*$ és $\sigma(S) \subset \mathcal{S}^*$.

3. 2021.02.22.

3.1 Ismétlés. Elégséges és minimális elégséges statisztika

Definíció 8. S elégséges a $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ eloszláscsaládra nézve, ha a $\mathbb{P}_\theta(X \in A | S)$ feltételes eloszlásoknak létezik közös változata.

S minimális elégséges, ha elégséges és $\sigma(S)^* \subset \mathcal{S}^*$ tetszőleges \mathcal{S} elégséges σ -algebrára.

Tétel 7 (Neyman faktoriációs tétel). Dominált családban S pontosan akkor elégséges, ha $f_\theta = (g_\theta \circ S) \cdot h$ alakban írható.

Tétel 8 (Egyszerűsített, majdnem mindig igaz változat). \mathcal{P} dominált. S elégséges statisztika pontosan akkor minimális elégséges, ha $x \sim y \iff S(x) = S(y)$, ahol $x \sim y$, ha $\Theta_x = \{\theta : f_\theta(x) \neq 0\} = \Theta_y$ és $\theta \mapsto \frac{f_\theta(x)}{f_\theta(y)}$ konstans $\Theta_x = \Theta_y$ -on.

Példa, minimális elégségesség másképp

- $\Theta = (0, 1)$, $\mathfrak{X} = \{0, 1\}^n$, $f_\theta(x) = \mathbb{P}_\theta(X = x) = \theta^s(1 - \theta)^{n-s}$, ahol $s = \sum x_i$.
- $S = \sum X_i$ minimális elégségessége másképp.
- \mathcal{S} elégséges σ -algebra. $\mathbb{P}_\theta(X \in A | \mathcal{S})$ nem függ θ -tól, T_A jelöli a közös változatot.
- S a mintaösszeg elégséges, $\mathbb{E}_\theta(\mathbf{1}_{(X \in A)} - T_A | S) = h(S)$ a közös változat. Ekkor

$$\mathbb{E}_\theta(h(S)) = \mathbb{E}_\theta(\mathbf{1}_{(X \in A)} - T_A) = \mathbb{P}_\theta(A) - \mathbb{E}_\theta(\mathbb{E}_\theta(\mathbf{1}_A | \mathcal{S})) = 0.$$

- S binomiális eloszlású n -renddel és θ paraméterrel:

$$0 = \mathbb{E}_\theta(h(S)) = \sum_{k=0}^n h(k) \binom{n}{k} \theta^k (1 - \theta)^{n-k} \quad \text{minden } \theta \in (0, 1).$$

Így $h(k) = 0$, $k = 0, 1, \dots, n$ és $0 = h(S) = \mathbb{E}_\theta(\mathbf{1}_A - T_A | S)$.

- Ha $A \in \sigma(S)$, akkor $h(S) = \mathbf{1}_A - \mathbb{E}(T_A | S) = 0$, azaz $\mathbb{E}(\mathbb{E}(\mathbf{1}_A | \mathcal{S}) | S) = \mathbf{1}_A$.
- A feltételes várható érték képzés L^2 -ben merőleges vetítés, ha nem csökkenti az L^2 normát akkor nem is változtathat.
- $\mathbf{1}_A \sim \mathcal{S}^*$ minden $A \in \sigma(S)$ -re, vagyis $\sigma(S) \subset \mathcal{S}^*$, amiből $\sigma(S)^* \subset \mathcal{S}^*$ és $\sigma(S)$ minimális elégséges.

3.2 Teljesség

Definíció 9. $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ eloszláscsalád, S statisztika. S teljes statisztika (a \mathcal{P} eloszláscsaládra nézve), ha

$$\mathbb{E}_\theta(h(S)) = 0, \quad \text{minden } \theta \in \Theta\text{-ra} \implies \mathbb{P}_\theta(h(S) = 0) = 1, \quad \text{minden } \theta\text{-ra}$$

Azaz S teljes, ha az S függvényei között csak a lényegében nulla változóra teljesül, hogy az eloszlás család minden tagja mellett létezik és nulla a várható értéke.

S korlátosan teljes, ha az S korlátos függvényei között csak a lényegében nulla változóra teljesül, hogy az eloszlás család minden tagja mellett nulla a várható értéke.

S teljessége az S eloszlásai családjának tulajdonsága.

pl. \mathcal{P} dominált: $f_\theta = \frac{d\mathbb{P}_\theta}{d\lambda}$ a sűrűségek. Ekkor $S : \mathfrak{X} \rightarrow \mathbb{R}^p$ eloszlásainak családját $\mu = \lambda \circ S^{-1}$ dominálja:

$$0 = \mu(H) = \lambda(S \in H) \implies (\mathbb{P}_\theta \circ S^{-1})(H) = \mathbb{P}_\theta(S \in H) = \int_{S \in H} f_\theta d\lambda = 0, \quad \forall \theta.$$

Azaz, S eloszlásainak van sűrűségfüggvénye μ -re nézve $\tilde{f}_\theta = \frac{d(\mathbb{P}_\theta \circ S^{-1})}{d\mu}$. S teljes, ha

$$\left\{ h : \int_{\mathbb{R}^p} h \tilde{f}_\theta d\mu \text{ létezik és } 0 \text{ minden } \theta\text{-ra} \right\} = \left\{ h : h \tilde{f}_\theta = 0 \text{ a } \mu \text{ majdnem mindenütt minden } \theta\text{-ra} \right\}$$

Tétel 9. Ha S korlátosan teljes és elégséges a \mathcal{P} eloszláscsaládra, akkor minimális elégséges.

- $A \in \sigma(S)$, \mathcal{S} tetszőleges elégséges σ -algebra, $T_A = \mathbb{P}_\theta(X \in A | \mathcal{S})$ közös változat.
- $h(S) = \mathbb{E}_\theta(\mathbf{1}_A - T_A | S) = \mathbf{1}_A - \mathbb{E}(T_A | S)$ korlátos statisztika, $\mathbb{E}_\theta(h(S)) = \mathbb{P}_\theta(A) - \mathbb{P}_\theta(A) = 0$.
- Teljesség miatt $h(S) = 0$, azaz $\mathbf{1}_A = \mathbb{E}(T_A | S)$ \mathcal{P} majdnem mindenütt. $0 \leq T_A \leq 1$ miatt $\mathbb{E}_\theta((\mathbf{1}_A - T_A)^2 | S) = \mathbb{E}(\mathbf{1}_A - 2\mathbf{1}_A T_A + T_A^2 | S) = \mathbb{E}(T_A^2 | S) - \mathbf{1}_A \leq \mathbb{E}(T_A | S) - \mathbf{1}_A = 0$.
- $\mathbb{P}_\theta(\mathbf{1}_A = T_A) = 1$ minden θ -ra, azaz $\mathbb{P}_\theta(A \circ \{T_A = 1\}) = 0$, így $A \in \mathcal{S}^*$ és $\sigma(S) \subset \mathcal{S}^*$.

3.2.1 Példák

Nem minden minimális elégséges statisztika korlátosan teljes és nem minden teljes statisztika elégséges

- $U(\theta, \theta + 1)$, $\theta \in \mathbb{R}$ eloszlásból származó, $n \geq 2$ elemű minta.
- $S = (X_1^*, X_n^*)$ minimális elégséges statisztika.
- $X_n^* - X_1^*$ eloszlása nem függ θ -tól és $\mathbb{E}_\theta(X_n^* - X_1^*) = \frac{n-1}{n+1}$.
- $h(S) = (X_n^* - X_1^*) \wedge 1 - \frac{n-1}{n+1}$ olyan korlátos függvénye S -nek, amire

$$\mathbb{E}_\theta(h(S)) = 0, \quad \forall \theta \in \mathbb{R}.$$

- A $T = 0$ konstans statisztika teljes tetszőleges eloszláscsaládra nézve, de csak a triviális esetben elégséges.
- HF. X $n = 2$ elemű minta $U(0, \theta)$, $\theta > 0$ eloszlásból. Ekkor $T(x) = x_1 + x_2$ teljes, de nem elégséges.

További példa teljes statisztikára

- $N(\theta, 1)$ eloszlásból származó n elemű mintánk van.
- $S = \sum X_i \sim N(n\theta, n)$ elégséges statisztika.
-

$$0 = \mathbb{E}_\theta(h(S)) = \int h(x) \frac{1}{\sqrt{2\pi n}} e^{-\frac{(x-n\theta)^2}{2n}} dx = C(\theta) \int h(x) e^{-\frac{x^2}{2n}} e^{\theta x} dx \quad \forall \theta \in \mathbb{R}$$

Tétel 10. $h : \mathbb{R} \rightarrow \mathbb{R}$ mérhető, $a < b$,

$$\int h(x) e^{\theta x} dx, \quad \text{létezik és nulla minden } \theta \in (a, b)\text{-re.}$$

Ekkor $h = 0$ Lebesgue majdnem mindenütt.

- A tétel alapján, $h(x) e^{-\frac{x^2}{2n}} = 0$ mm. és így $\mathbb{P}_\theta(h(S) = 0) = 1$ minden θ -ra
- S teljes és elégséges, ezért minimális elégséges.

3.3 Momentum generáló függvény

X p -dimenziós vektor változó, $M_X(t) = \mathbb{E}(e^{t \cdot X})$, $t \in \mathbb{R}^p$ az X momentum generáló függvénye.

Állítás 12. *Ha X momentum generáló függvénye véges az origó egy környezetében, akkor meghatározza X eloszlását.*

Bizonyítás vázlat.

- X skalár változó. $h(z) = h(x + iy) = \mathbb{E}(e^{(x+iy)X})$ a képzetes tengely körüli sávban definiált és véges. Itt deriválható (komplex értelemben), ezért a valós tengelyen felvett értékek (a momentum generáló függvény) meghatározzák a karakterisztikus függvényt, az pedig az eloszlást.
- X vektor változó. α rögzített vektor, $Y = \alpha \cdot X$ skalár változó. $\mathbb{E}(e^{tY}) = \mathbb{E}(e^{(t\alpha)X})$, azaz Y eloszlását X momentum generáló függvénye meghatározza.
- X eloszlását az $\alpha \cdot X$ alakú változók eloszlása meghatározza, ugyanis a karakterisztikus függvényekre $\phi_X(\alpha) = \phi_{\alpha \cdot X}(1)$.
- Összefoglalva, a momentum generáló függvény meghatározza az egy dimenziós vetületek eloszlását, azok pedig X eloszlását.

3.4 Elégséges feltétel teljességre

Tétel 11. *Legyen μ mérték $\mathcal{B}(\mathbb{R}^p)$ -n, $h : \mathbb{R}^p \rightarrow \mathbb{R}$ mérhető és*

$$A = \left\{ \alpha \in \mathbb{R}^p : \int_{\mathbb{R}^p} e^{\alpha \cdot x} h(x) \mu(dx) \text{ létezik és nulla} \right\}$$

Ha A -nak létezik belső pontja, akkor h μ -majdnem mindenütt nulla.

- Legyen $\alpha_0 \in \text{int } A$, és

$$\nu_{\pm}(H) = \int_H e^{\alpha_0 \cdot x} |h|_{\pm}(x) \mu(dx), \quad H \in \mathcal{B}(\mathbb{R}^p).$$

ν_+ , ν_- véges Borel mértékek, hiszen $\alpha_0 \in A$.

- Ha $\nu_+(\mathbb{R}^p) = \nu_-(\mathbb{R}^p) = 0$, akkor h μ -majdnem mindenütt nulla.
- Ha $\nu_+(\mathbb{R}^p) = \nu_-(\mathbb{R}^p) > 0$, akkor konstanssal való szorzás után feltehető, hogy ν_+ , ν_- valószínűségi mértékek. Azt akarjuk megmutatni, hogy ez az eset nem fordulhat elő.
- A két mérték momentum generáló függvénye az origó egy környezetében azonos, hiszen

$$\int e^{tx} \nu_+(dx) - \int e^{tx} \nu_-(dx) = \int e^{(t+\alpha_0)x} h(x) \mu(dx) = 0 \quad \text{ha } \alpha_0 + t \in A.$$

- Mivel a momentum generáló függvény az eloszlást meghatározza, $\nu_+ = \nu_-$ és

$$\nu_-(h > 0) = 0 = \nu_+(h > 0) = \int \mathbf{1}_{(h(x)>0)} e^{\alpha_0 \cdot x} |h|_+(x) \mu(dx) = \nu_+(\mathbb{R}^p) = 1, \quad \text{!}$$

3.5 Exponenciális család

Definíció 10. \mathcal{P} exponenciális eloszlás család, ha dominált és alkalmas paraméterezéssel $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^p$ valamint alkalmas domináló mértékkel, $T : \mathfrak{X} \rightarrow \mathbb{R}^p$ statisztikával a sűrűségfüggvények alakja

$$f_\theta(x) = \exp\{\theta \cdot T(x) - b(\theta)\}. \quad (*)$$

Tulajdonságok:

- T elégséges statisztika (Neyman faktorizációs tétel)
- Ha $\Theta \subset \mathbb{R}^p$ -ben van $p + 1$ általános helyzetű pont, akkor T minimális elégséges.
 $\theta_i \in \Theta$, $i = 0, \dots, p$ és $\theta_i - \theta_0$ lineáris burka \mathbb{R}^p .
 Ekkor, ha $\frac{f_{\theta_i}(x)}{f_{\theta_i}(y)}$ nem függ θ -tól, akkor $(\theta_i - \theta_0)(T(x) - T(y)) = 0$, $i = 1, \dots, p$ és $T(x) = T(y)$, ami a minimális elégségesség feltétele.
 Exponenciális család mindig megadható úgy is, hogy T minimális elégséges legyen. HF.
- Ha Θ belseje nem üres, akkor T teljes is (következő dián).

3.5.1 Exponenciális család teljessége

Ha Θ belseje nem üres, akkor T teljes is.

- λ a domináló mérték, amivel a sűrűség $(*)$ alakú. $\tilde{\lambda}(H) = \lambda(T \in H)$, $H \in \mathcal{B}(\mathbb{R}^p)$.
- Ha minden $\theta \in \Theta$ -ra, $\mathbb{E}_\theta(h(T))$ létezik és nulla, akkor

$$0 = \mathbb{E}_\theta(h(T)) = \int_{\mathfrak{X}} h \circ T e^{\theta T - b(\theta)} d\lambda = \int_{\mathbb{R}^p} h(t) e^{\theta t - b(\theta)} \tilde{\lambda}(dt) \implies \left\{ \theta : \int_{\mathbb{R}^p} e^{\theta t} h(t) \tilde{\lambda}(dt) = 0 \right\} = \Theta.$$

- $\text{int } \Theta \neq \emptyset \implies h d\tilde{\lambda}$ az azonosan nulla mérték. Tetszőleges $H \in \mathcal{B}(\mathbb{R}^p)$ -re, pl. $H = \{h > 0\}$ és $H = \{h < 0\}$

$$0 = \int_H h d\tilde{\lambda} = \int_{\mathfrak{X}} \mathbf{1}_{(T \in H)} h \circ T d\lambda, \quad \forall H \in \mathcal{B}(\mathbb{R}^p) \implies \lambda(h \circ T \neq 0) = 0.$$

$$\mathbb{P}_\theta(h(T) \neq 0) = 0, \text{ azaz } \mathbb{P}_\theta(h(T) = 0) = 1.$$

3.5.2 Példák exponenciális családra

- n elemű minta $\exp(\theta)$, $\theta > 0$ eloszlásból, $\mathfrak{X} = (0, \infty)^n$, domináló mérték a Lebesgue mérték.

$$f_\theta(x) = \theta^n \exp\left\{-\theta \sum x_i\right\}, \quad T(x) = -\sum x_i, \quad b(\theta) = -n \ln \theta.$$

- n elemű indikátor minta, $p \in (0, 1)$ paraméterrel, $\mathfrak{X} = \{0, 1\}^n$, domináló mérték a számláló mérték

$$\begin{aligned} f_\theta(x) &= p^{\sum x_i} (1-p)^{n-\sum x_i} = \left(\frac{p}{1-p}\right)^{\sum x_i} (1-p)^n \\ &= \exp\{\theta T(x) - b(\theta)\}, \quad \theta = \ln \frac{p}{1-p}, \quad p = \frac{1}{1+e^{-\theta}}, \quad T(x) = \sum x_i, \quad b(\theta) = n \ln(1+e^\theta) \end{aligned}$$

- $N(\mu, \sigma^2)$ eloszlásból származó n elemű minta. $\mathfrak{X} = \mathbb{R}^n$ $\mu \in \mathbb{R}$, $\sigma^2 > 0$. Sűrűségfüggvény a Lebesgue mértékre.

$$f_{\mu, \sigma^2}(x) = \frac{1}{(2\pi)^{n/2} \sigma^n} \exp\left\{-\frac{1}{2\sigma^2} \sum (x_i - \mu)^2\right\}$$

$$= \exp\left\{-\frac{1}{2\sigma^2} \sum x_i^2 + \frac{\mu}{\sigma^2} \sum x_i - \tilde{b}(\mu, \sigma^2)\right\}, \quad \theta = \left(\frac{1}{2\sigma^2}, \frac{\mu}{\sigma^2}\right), \quad T(x) = \left(-\sum x_i^2, \sum x_i\right).$$

- HF. A gyakorlaton szereplő további „szép” eloszláscsaládok: geometriai, Poisson, Gamma, Béta, stb. mind exponenciális családot alkotnak.

Kivételek: ha a minimális elégséges statisztika a rendezett minta pl. Cauchy eltolás paraméteres család, ill. amikor az eloszlások tartója nem közös, pl. „egyenletes eloszlás”.

T nem mindig teljes

- $N(\mu, \sigma^2)$ eloszlásból származó minta $\mu = \sigma$ ismeretlen.

Átparaméterezés után $\theta(\mu, \sigma^2) = (\frac{1}{2\sigma^2}, \frac{\mu}{\sigma^2})$, $\Theta = \{(\tau, \sqrt{2\tau}) : \tau > 0\}$. $T(x) = (-\sum x_i^2, \sum x_i)$ minimális elégséges.

- $n = 2$,

$$\mathbb{E}(T_1) = \mathbb{E}(-(X_1^2 + X_2^2)) = -2(\sigma^2 + \mu^2) = -4\sigma^2,$$

$$\mathbb{E}(T_2) = \mathbb{E}((X_1 + X_2)) = 2\mu, \quad \mathbb{E}(T_2^2) = \mathbb{E}((X_1 + X_2)^2) = 2\sigma^2 + (2\mu)^2 = 6\sigma^2$$

•

$$\mathbb{E}_\theta(3T_1 + 2T_2^2) = 0$$

$\mathbb{E}_\theta(h(T)) = 0$, minden $\theta = (\tau, \sqrt{2\tau})$ paraméterre, ha $h(t_1, t_2) = 3t_1 + 2t_2^2$.

- $\mathbb{E}_\theta(h(T)) = 0$ minden θ -ra, de $\mathbb{P}_\theta(h(T) = 0) = 0$.

$$h(T(X)) = -3(X_1^2 + X_2^2) + 2(X_1 + X_2)^2 = 4X_1X_2 - (X_1^2 + X_2^2) = 3X_2^2 - (X_1 - 2X_2)^2$$

és

$$\mathbb{P}_\theta(h(T) = 0) = \mathbb{E}_\theta(\mathbb{P}_\theta(3X_2^2 - (X_1 - 2X_2)^2 = 0 \mid X_2)) = \mathbb{E}_\theta(\mathbb{P}_\theta((X_1 - 2x)^2 = 3x^2)_{|x=X_2}) = 0.$$

3.6 Basu tétele

Definíció 11. T kísérő (ancillary) statisztika (a \mathcal{P} eloszláscsaládra nézve), ha $\mathbb{P}(T \in A)$ nem függ $\mathbb{P} \in \mathcal{P}$ -től, azaz T eloszlása azonos mindegyik $\mathbb{P} \in \mathcal{P}$ mellett.

Tétel 12. Ha T kísérő statisztika, S korlátosan teljes és elégséges a \mathcal{P} eloszláscsaládra nézve, akkor T és S független mindegyik $\mathbb{P} \in \mathcal{P}$ mellett.

- $\mathbb{P}_\theta \in \mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$

$$\mathbb{P}_\theta(T \in A \mid S) - \mathbb{P}_\theta(T \in A) = h(S)$$

mert S elégséges és T kísérő.

- $\mathbb{E}_\theta(h(S)) = \mathbb{P}_\theta(T \in A) - \mathbb{P}_\theta(T \in A) = 0$.
- Mivel S korlátosan teljes, ezért $\mathbb{P}_\theta(h(S) = 0) = 1$ minden θ -ra, azaz $\mathbb{P}_\theta(T \in A \mid S) = \mathbb{P}_\theta(T \in A)$ és

$$\mathbb{P}_\theta(T \in A, S \in B) = \mathbb{E}_\theta(\mathbb{P}(T \in A \mid S) \mathbf{1}_{(S \in B)}) = \mathbb{P}_\theta(T \in A) \mathbb{P}_\theta(S \in B), \quad \text{minden } A, B\text{-re}$$

3.6.1 Példák kísérő statisztikákra

- X_1, \dots, X_n exponenciális eloszlású minta $\lambda > 0$ paraméterrel. $T_i = X_i / \sum X_i$, $T = (T_1, \dots, T_n)$. Ekkor T kísérő statisztika, ugyanis \mathbb{P}_λ alatt X ugyanolyan eloszlású, mint $\frac{1}{\lambda}Z$, ahol Z_1, \dots, Z_n független egységnyi paraméterű exponenciálisok és T kiszámításakor λ kiesik. $S = \sum X_i$ teljes és elégséges, azaz T és S független.

Sűrűségtranszformációs formulával is könnyen ellenőrizhető.

Ugyanígy kísérő T' , ha $T'_i = X_i / \|X\|_2$ és T' és S is független.

- Hasonlóan tetszőleges skálaparaméteres családban $X / \|X\|_p$ kísérő statisztika.
- $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ eloszlású minta, ahol $\mu \in \mathbb{R}$, $\sigma > 0$. Ekkor X ugyanolyan eloszlású, mint $\mu + \sigma Z$, ahol Z_1, \dots, Z_n független standard normálisok. Így

$$T(X) = \frac{X_1 - \bar{X}}{s_n(X)} \stackrel{p}{=} \frac{Z_1 - \bar{Z}}{s_n(Z)} \quad \text{kísérő statisztika, ahol } \bar{X} = \frac{1}{n} \sum_i X_i \text{ és } s_n^2(X) = \frac{1}{n} \sum_i (X_i - \bar{X})^2.$$

$S = (\sum X_i, \sum X_i^2)$ teljes és elégséges, vagyis független T -től.

Példa kísérő statisztikára, ahol Basu tétele nem alkalmazható

- X_1, \dots, X_n $n \geq 2$ elemű minta $U(\theta, \theta + 1)$ eloszlásból, $\theta \in \mathbb{R}$.
- $S = (X_1^*, X_n^*)$ minimális elégséges statisztika.
- θ eltolás paraméter, $T = X_n^* - X_1^*$ kísérő statisztika.
- T nem független S -től, mert T S függvénye és nem elfajult eloszlású.
- Nincs ellentmondás. S nem korlátosan teljes.

3.7 Rao-Blackwell-Kolmogorov tétel

Definíció 12. $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$. $g : \Theta \rightarrow \mathbb{R}$.

T statisztika a $g(\theta)$ torzítatlan becslése, ha $\mathbb{E}_\theta(T) = g(\theta)$.

T a $g(\theta)$ hatásos becslése, ha torzítatlan és tetszőleges T' torzítatlan becslésre $\mathbb{D}_\theta^2 T \leq \mathbb{D}_\theta^2 T'$.

Tétel 13. S elégséges a \mathcal{P} eloszláscsaládra és T torzítatlan $g : \Theta \rightarrow \mathbb{R}$ -re. Ekkor $\mathbb{E}_\theta(T | S) = h(S)$ közös változata torzítatlan $g(\theta)$ -ra és $\mathbb{D}_\theta^2 h(S) \leq \mathbb{D}_\theta^2 T$.

Ha az S statisztika teljes is, akkor $\mathbb{E}_\theta(T | S)$ hatásos becslés.

- S elégséges, ezért $h(S) = \mathbb{E}_\theta(T | S)$ nem függ θ -tól!
- $\mathbb{E}_\theta(h(S)) = \mathbb{E}_\theta(T) = g(\theta)$ a teljes várható érték tétel miatt, azaz $h(S)$ torzítatlan.
- $\mathbb{D}_\theta^2(h(S)) = \mathbb{E}_\theta((h(S) - g(\theta))^2) \leq \mathbb{E}_\theta((T - g(\theta))^2) = \mathbb{D}_\theta^2(T)$, mert

$$(h(S) - g(\theta))^2 = (\mathbb{E}_\theta(T | S) - g(\theta))^2 \leq \mathbb{E}((T - g(\theta))^2 | S)$$

a Jensen egyenlőtlenség szerint.

- Ha S teljes és elégséges, T, T' tetszőleges torzítatlan becslések $g(\theta)$ -ra. $h(S) = \mathbb{E}_\theta(T | S)$.
- $\mathbb{E}_\theta(T - T' | S) = h(S)$ és $\mathbb{E}_\theta(h(S)) = g(\theta) - g(\theta) = 0$ így $h(S) = 0$ és $\mathbb{E}(T' | S) = h(S)$.
- $\mathbb{D}^2 h(S) \leq \mathbb{D}^2 T'$ tetszőleges T' torzítatlan becslésre, vagyis $\mathbb{E}(T | S)$ hatásos.

Általánosítás

- $W(t, \theta)$ **veszteségfüggvény**, pl. $W(t, \theta) = |t - g(\theta)|^2$ a négyzetes veszteségfüggvény. $R_T(\theta) = \mathbb{E}_\theta(W(T, \theta))$ a T becslés **rizikója**. A torzítatlan becslések körében a négyzetes rizikó a szórásnégyzet.
- W az első változóban konvex, S teljes és elégséges, T torzítatlan, akkor $\mathbb{E}_\theta(T | S)$ közös változata minimális rizikójú a torzítatlan becslések között: $R_T \leq R_{T'}$ tetszőleges torzítatlan T' -re.

3.7.1 Példák

- X_1, \dots, X_n n elemű indikátor minta $p \in (0, 1)$ paraméterrel. $S = \sum X_i$ teljes és elégséges.
- $\mathbb{E}_p(T) = \sum_{x \in \mathfrak{X}} T(x) p^{S(x)} (1-p)^{n-S(x)}$ a p legfeljebb n -ed fokú polinomja.
- Becsüljük $g(p) = p^k$, ahol $1 \leq k \leq n$.
- $T(X) = \mathbf{1}_{(X_1=1, \dots, X_k=1)}$ torzítatlan becslés p^k -ra. $\mathbb{E}(T | S)$ hatásos.
-

$$\begin{aligned} \mathbb{E}_p(T | S = s) &= \frac{\mathbb{P}_p(T = 1, S = s)}{\mathbb{P}_p(S = s)} \\ &= \frac{\mathbb{P}_p(T = 1, \sum_{i=k+1}^n X_i = s - k)}{\mathbb{P}_p(S = s)} = \begin{cases} 0 & \text{ha } s < k \\ \frac{p^k \binom{n-k}{s-k} p^{s-k} (1-p)^{(n-k)-(s-k)}}{\binom{n}{s} p^s (1-p)^{n-s}} = \frac{s \cdots (s-(k-1))}{n \cdots (n-(k-1))} & \text{ha } k \leq s \leq n \end{cases} \end{aligned}$$

$$\mathbb{E}(T | S) = \frac{S}{n} \cdot \frac{S-1}{n-1} \cdots \frac{S-(k-1)}{n-(k-1)} \text{ hatásos becslés } p^k\text{-ra.}$$

- Egy gépről kikerülő termék jellemző mérete, például sugara $R \sim N(\mu, \sigma^2)$ eloszlású, $\mu > 0, \sigma^2 > 0$.
- A termék selejtes, ha $R > c$, ahol c adott konstans.
- n elemű X_1, \dots, X_n minta alapján szeretnénk a selejt arányt becsülni.
- $S = (\sum X_i, \sum X_i^2)$ teljes elégséges statisztika.
- $T = \mathbf{1}_{(X_1 > c)}$ torzítatlan becslés a selejt arányra.

$$\mathbb{E}(T | S) = \mathbb{P}(X_1 > c | S) = \mathbb{P}\left(\frac{X_1 - \bar{X}}{s_n(X)} > \frac{c - \bar{X}}{s_n(X)} \middle| S\right)$$

Itt $\eta = \frac{X_1 - \bar{X}}{s_n(X)}$ kísérő statisztika és S -től független $\bar{X} = \frac{1}{n} \sum X_i$, $s_n^2(X) = \frac{1}{n} \sum X_i^2 - (\frac{1}{n} \sum X_i)^2$ a feltételből kiolvasható.

$$\mathbb{E}(T | S) = \mathbb{P}\left(\eta > \frac{c - \bar{X}}{s_n(X)} \middle| S\right) = \mathbb{P}(\eta > x) \Big|_{x=\frac{c-\bar{X}}{s_n(X)}} = 1 - F_\eta\left(\frac{c - \bar{X}}{s_n(X)}\right).$$

- η -ról kiszámolható, hogy $\frac{1}{n-1} \eta^2 \sim \text{Béta}(\frac{1}{2}, \frac{n-2}{2})$ és $\eta \stackrel{p}{=} -\eta$, azaz η szimmetrikus eloszlású, amiből F_η megkapható.

4. 2021.03.01.

4.1 Becslés, torzítatlanság, hatásosság

Definíció 13. $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$. $g : \Theta \rightarrow \mathbb{R}$.

T statisztika a $g(\theta)$ torzítatlan becslése, ha $\mathbb{E}_\theta(T) = g(\theta)$.

T a $g(\theta)$ hatásos becslése, ha torzítatlan $\mathbb{D}_\theta^2 T$ véges minden $\theta \in \Theta$ -ra és tetszőleges T' torzítatlan becslésre $\mathbb{D}_\theta^2 T \leq \mathbb{D}_\theta^2 T'$.

Tétel 14 (Rao–Blackwell–Kolmogorov). S elégséges a \mathcal{P} eloszláscsaládra és T torzítatlan $g : \Theta \rightarrow \mathbb{R}$ -re. Ekkor $\mathbb{E}_\theta(T | S) = h(S)$ közös változata torzítatlan $g(\theta)$ -ra és $\mathbb{D}_\theta^2 h(S) \leq \mathbb{D}_\theta^2 T$.

Ha az S statisztika teljes is, akkor $\mathbb{E}_\theta(T | S)$ hatásos becslés.

4.1.1 Alsó korlát torzítatlan becslés szórásnégyzetére

- $\mathbb{P}_0, \mathbb{P}_1$ eloszlások, f_0, f_1 sűrűséggel az \mathfrak{X} mintatéren, $\mathbb{P}_1 \ll \mathbb{P}_0$.
- T statisztika, $\mathbb{E}_i(T) = g(i)$ $i = 0, 1$.

$$g(1) - g(0) = \mathbb{E}_1(T) - \mathbb{E}_0(T) = \mathbb{E}_0\left(T\left(\frac{f_1}{f_0} - 1\right)\right)$$

- $\mathbb{E}_0(f_1/f_0 - 1) = 0$ és $U = f_1/f_0 - 1$ jelöléssel

$$|g(1) - g(0)| = |\text{cov}_0(T, U)| \leq \mathbb{D}_0(T)\mathbb{D}_0(U) \implies \mathbb{D}_0^2 T \geq \frac{(g(1) - g(0))^2}{\mathbb{D}_0^2(U)}.$$

Példa

- X n elemű minta $U(0, \theta)$ eloszlásból, $\theta > 0$.
- T torzítatlan becslés θ -ra.
- $0 < \theta_1 = \alpha\theta_0 < \theta_0$, mellett $\mathbb{P}_1 \ll \mathbb{P}_0$ és $U(x) = \frac{f_1(x)}{f_0(x)} - 1$, $\frac{f_1(x)}{f_0(x)} = \alpha^{-n} \mathbf{1}_{(\max x_i < \theta_1)}$.

$$\mathbb{D}_0^2 U = \mathbb{D}_0^2 \frac{f_1}{f_0} = \alpha^{-n} - 1 \implies \mathbb{D}_0^2 T \geq \theta_0^2 \sup_{\alpha \in (0,1)} \frac{(1-\alpha)^2}{\alpha^{-n} - 1} \geq C \frac{\theta_0^2}{n^2}, \quad (\alpha = 1 - \frac{1}{n}).$$

4.1.2 Példa. Indikátor minta

- X n elemű minta θ paraméterű indikátor mintából, $\theta \in (0, 1)$.
- T torzítatlan becslés $p(\theta)$ -ra, ahol p legfeljebb n -edfokú polinom.
- \mathbb{P}_θ ekvivalens a számláló mértékkel minden θ mellett és $\theta_1, \theta_0 \in (0, 1)$ esetén

$$\lim_{\theta_1 \rightarrow \theta_0} \frac{1}{\theta_1 - \theta_0} \left(\frac{f_{\theta_1}(x)}{f_{\theta_0}(x)} - 1 \right) = \partial_\theta \ln f_{\theta_0}(x) = \frac{s}{\theta_0} - \frac{n-s}{1-\theta_0}, \quad \text{ahol } s = \sum x_i.$$

•

$$p'(\theta_0) = \lim_{\theta_1 \rightarrow \theta_0} \frac{\mathbb{E}_{\theta_0}\left(T\left(\frac{f_{\theta_1}}{f_{\theta_0}} - 1\right)\right)}{\theta_1 - \theta_0} = \text{cov}_{\theta_0}(T, \partial_\theta \ln f_{\theta_0}) \leq \mathbb{D}_{\theta_0}(T)\mathbb{D}_{\theta_0}(\partial_\theta \ln f_{\theta_0})$$

•

$$\mathbb{D}_\theta^2 T \geq \frac{(p'(\theta))^2}{\mathbb{D}_\theta^2 \partial_\theta \ln f_\theta} = (p'(\theta))^2 \frac{1}{n} \theta(1-\theta), \quad \text{hiszen} \quad \mathbb{D}_\theta^2 \partial_\theta \ln f_\theta = \mathbb{D}_\theta^2 \frac{\sum X_i}{\theta(1-\theta)} = \frac{n}{\theta(1-\theta)}.$$

Ez a Cramer-Rao egyenlőtlenség speciális esete.

- $p(\theta) = \theta$ esetén az alsó becslés a mintaátlag szórásnégyzete, azaz a mintaátlag hatásos becslése a valószínűségnek, vagyis θ -nak.

4.2 Cramer–Rao-egyenlőtlenség, információs határ

Tétel 15. • $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ dominált család f_θ sűrűségekkel, $\Theta \subset \mathbb{R}^p$.

- $\ell(\theta) = \ell(\theta, x) = \ln f_\theta(x)$ a **loglikelihood** függvény.
- $\theta \in \text{int } \Theta$, $\ell' = \ell'(\theta) = \partial_\theta \ell(\theta)$ (p -dimenziós sorvektor) létezik és $\mathbb{E}_\theta(\ell'(\theta)) = 0$.
- $I(\theta) = \Sigma_\theta((\ell'(\theta))^T) = \mathbb{E}_\theta(\ell'(\theta))^T \ell'(\theta)$ invertálható. $I(\theta)$ neve **Fisher információ**.
- $T : \mathfrak{X} \rightarrow \mathbb{R}^q$ statisztikára $g(\theta) = \mathbb{E}_\theta(T)$ deriválható és $g'(\theta) = \int_{\mathfrak{X}} T(x) \partial_\theta f_\theta(x) \lambda(dx) = \mathbb{E}_\theta(T \ell'(\theta))$.

Ekkor

$$\Sigma_\theta(T) \geq G I^{-1}(\theta) G^T, \quad \text{ahol } G = g'(\theta) \in \mathbb{R}^{q \times p} \text{ és}$$

- $A \in \mathbb{R}^{q \times p}$ mátrixszal $T - A \partial_\theta \ell(\theta)$ q -dimenziós vektor változó.

$$\begin{aligned} 0 \leq \Sigma_\theta(T - A(\ell')^T) &= \Sigma_\theta(T) - A \mathbb{E}_\theta((\ell')^T T^T) - \mathbb{E}_\theta(T \ell') A^T + A \mathbb{E}_\theta((\ell')^T \ell') A^T \\ &= \Sigma_\theta(T) - A G^T - G A^T + A I(\theta) A^T \end{aligned}$$

- Ha itt $A = G I^{-1}(\theta)$, akkor

$$0 \leq \Sigma_\theta(T) - G^T I^{-1}(\theta) G \implies \Sigma_\theta(T) \geq G^T I^{-1}(\theta) G.$$

Ha $\mathbb{E}_\theta(\ell'(\theta)) = 0$, akkor az n elemű minta Fisher információjára $I_n(\theta) = I_1(\theta)$ az alsó becslés $1/n G^T I_1^{-1} G$.

4.2.1 Cramer-Rao egyenlőtlenség, speciális esetek

Tétel 16. • $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ dominált család f_θ sűrűségekkel, $\Theta \subset \mathbb{R}^p$. $\ell(\theta)$ loglikelihooddal.

- $\theta \in \text{int } \Theta$, ℓ' létezik, $\mathbb{E}_\theta(\ell'(\theta)) = 0$, $I(\theta) = \mathbb{E}_\theta(\ell'(\theta))^T \ell'(\theta)$ invertálható.
- $T : \mathfrak{X} \rightarrow \mathbb{R}^q$ statisztikára $g(\theta) = \mathbb{E}_\theta(T)$ deriválható és $g'(\theta) = \mathbb{E}_\theta(T \ell'(\theta))$

$$\Sigma_\theta(T) \geq G I^{-1}(\theta) G^T, \quad \text{ahol } G = g'(\theta) \in \mathbb{R}^{q \times p} \text{ és}$$

Következmény 1. • $\Theta \subset \mathbb{R}$, $g : \Theta \rightarrow \mathbb{R}$. Ha a fenti feltételek teljesülnek, akkor

$$\mathbb{D}_\theta^2(T) \geq \frac{(g'(\theta))^2}{I(\theta)}$$

- $g(\theta) = \theta$. $T = \hat{\theta}$ a paraméter torzítatlan becslése. Ha a fenti feltételek teljesülnek, akkor

$$\Sigma_\theta(\hat{\theta}) \geq I^{-1}(\theta).$$

4.3 Fisher információ

Definíció 14. • $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ dominált eloszláscsalád az \mathfrak{X} mintatéren, ahol $\Theta \subset \mathbb{R}^p$ nyílt.

- $f_\theta = \frac{d\mathbb{P}_\theta}{d\lambda}$ ahol λ a domináló mérték. $\ell(\theta) = \ell(\theta, x) = \ln f_\theta(x)$ a **loglikelihood** függvény.
- $\ell'(\theta) = \partial_\theta \ell(\theta, x)$ p -dimenziós sorvektor létezik tetszőleges θ esetén \mathbb{P}_θ majdnem minden x mellett.

$$I(\theta) = \mathbb{E}_\theta(\ell'(\theta)^T \ell'(\theta))$$

$I : \Theta \rightarrow \mathbb{R}^{p \times p}$ a minta **Fisher információja**, ha a várható érték létezik.

- $I(\theta)$ pozitív szemidefinit mátrix. Ha $a \in \mathbb{R}^p$, akkor $a^T I(\theta) a = \mathbb{E}_\theta((\ell'(\theta)a)^2) \geq 0$.
- A definícióban a sűrűségek „szép” változatát használjuk, pl. $(\theta, x) \mapsto f_\theta(x)$ mérhető, sőt általában θ -ban folytonosan deriválható.
- A Fisher információ nem függ a domináló mérték megválasztásától. Ha $\mathbb{P}_0 \in \mathcal{P}'$, akkor tetszőleges λ domináló mértékre

$$f_\theta = \frac{d\mathbb{P}_\theta}{d\mathbb{P}_0} \frac{d\mathbb{P}_0}{d\lambda} \implies \ell(\theta) = \ln \frac{d\mathbb{P}_\theta}{d\mathbb{P}_0} + \ln \frac{d\mathbb{P}_0}{d\lambda} \implies \partial_\theta \ell(\theta) = \partial_\theta \ln \frac{d\mathbb{P}_\theta}{d\mathbb{P}_0}.$$

- Erősebb feltételek mellett érdemes nézni, pl. gyenge regularitási feltétel.

4.3.1 Formulák

- $\ell(\theta) = \ell(\theta, x) = \ln f_\theta(x)$ a **loglikelihood** függvény. $f_\theta(x) \geq 0$, azaz $\ell(\theta, x) = -\infty$ előfordulhat. Megállapodás $\partial_\theta f_\theta(x) = 0$, ha nem létezik, vagy $f_\theta(x) = 0$.

$$\partial_\theta \ell(\theta, x) = \frac{\partial_\theta f_\theta(x)}{f_\theta(x)} = \begin{cases} 0 & f_\theta(x) = 0 \\ \frac{\partial_\theta f_\theta(x)}{f_\theta(x)} & f_\theta(x) > 0 \end{cases}$$

- Ha $\Theta \subset \mathbb{R}^p$ nyílt, akkor $\ell'(\theta)$ p -dimenziós sorvektor értékű statisztika

$$\begin{aligned} I(\theta) &= \mathbb{E}((\ell'(\theta))^T \ell'(\theta)) \\ &= \int_{\mathfrak{X}} \frac{(\partial_\theta f_\theta(x))^T \partial_\theta f_\theta(x)}{f_\theta^2(x)} f_\theta(x) \lambda(dx) \\ &= \int_{\mathfrak{X}} \frac{(\partial_\theta f_\theta(x))^T \partial_\theta f_\theta(x)}{f_\theta(x)} \lambda(dx) \\ &= 4 \int_{\mathfrak{X}} (\partial_\theta \sqrt{f_\theta(x)})^T \partial_\theta \sqrt{f_\theta(x)} \lambda(dx) \end{aligned}$$

4.3.2 Fisher információ exponenciális családban

Emlékeztető. \mathcal{P} exponenciális családot alkot, ha alkalmas domináló mérték és paraméterezés mellett $\ln f_\theta(x) = \theta T(x) - b(\theta)$ alakú, $\theta \in \Theta$.

- Ha Θ nyílt, de legalább $\theta \in \text{int } \Theta$, akkor $\theta + \alpha \in \Theta$ -ra

$$e^{b(\theta+\alpha)} = \int_{\mathfrak{X}} e^{(\theta+\alpha)T(x)} \lambda(dx) = e^{b(\theta)} \int_{\mathfrak{X}} e^{\alpha T(x)} e^{\theta T(x) - b(\theta)} \lambda(dx) = e^{b(\theta)} \mathbb{E}_\theta(e^{\alpha T})$$

- $M(\alpha) = e^{b(\theta+\alpha)-b(\theta)}$ a T statisztika momentum generáló függvénye \mathbb{P}_θ alatt.
- Kiszámoljuk, hogy ha a momentum generáló függvény véges az origó egy környezetében, akkor ott sima (C^∞) és a nullabeli deriváltak T momentumait adják. $\ln M(\alpha) = b(\theta + \alpha) - b(\theta)$ a kumuláns generáló függvény, ez is sima és a nullabeli deriváltakra

$$\mathbb{E}_\theta(T) = \frac{M'(0)}{M(0)} = (\ln M)'(0) = b'(\theta), \quad \Sigma_\theta T = \frac{M''(0)M(0) - M'(0)^T M'(0)}{M^2(0)} = (\ln M)''(0) = b''(\theta).$$

- $\ell'(\theta) = T - b'(\theta) = T - \mathbb{E}_\theta(T)$.

$$\mathbb{E}_\theta(\ell'(\theta)) = 0, \quad I(\theta) = \mathbb{E}_\theta(\ell'(\theta)^T \ell'(\theta)) = \Sigma_\theta(T) = b''(\theta) = -\mathbb{E}_\theta(\ell''(\theta))$$

4.3.3 Momentum generáló függvény, deriválhatóság

Állítás 13. Ha $M(\alpha) = \mathbb{E}(e^{\alpha \cdot X}) < \infty$, az origó egy környezetében, akkor ott M sima és $M'(0) = \mathbb{E}(X)^T$, $M''(0) = \mathbb{E}(XX^T)$.

- Létezik $\varepsilon > 0$, hogy $\mathbb{E}(e^{h\|X\|}) < \infty$, ha $h < \varepsilon$.

\mathbb{R}^p a normák ekvivalensek, elég $\|x\|_1 = \frac{1}{p} \sum |x_i|$ -re számolni. Mivel \exp konvex $e^{\|x\|_1} \leq \frac{1}{p} \sum e^{|x_i|}$ és

$$e^{h|X_i|} \leq e^{hX_i} + e^{-hX_i}, \quad \mathbb{E}(e^{h\|X\|_1}) \leq \frac{1}{p} \sum_i (M(he_i) + M(-he_i)) < \infty$$

ahol e_1, \dots, e_p a természetes bázis \mathbb{R}^p -ben.

- ha $\|\alpha\|$ kicsi, akkor $|\alpha \cdot X| < \varepsilon \|X\|$ és a dominált konvergencia tétel miatt

$$M(\alpha) = \mathbb{E}\left(\sum_k \frac{(\alpha \cdot X)^n}{n!}\right) = \sum_n \frac{\mathbb{E}((\alpha \cdot X)^n)}{n!} \implies \partial_{i_1, \dots, i_k} M(0) = \mathbb{E}\left(\prod_j X_{i_j}\right)$$

4.3.4 Példa

- X_1, \dots, X_n exponenciális eloszlású minta, ismeretlen $\lambda > 0$ paraméterrel.
- A minta eloszlásainak a családja exponenciális családot alkot, a domináló mérték a Lebesgue mérték, $\mathfrak{X} = (0, \infty)^n$

$$f_\lambda(x) = \lambda^n e^{-\lambda \sum x_i} = \exp\{\lambda T(x) - b(\lambda)\}, \quad T(x) = -\sum x_i, \quad b(\lambda) = -n \ln(\lambda)$$

- $I(\lambda) = b''(\lambda) = n\lambda^{-2}$
- Definíció alapján számolva:

$$\ell(\lambda) = n \ln \lambda - \lambda \sum X_i, \quad \ell'(\lambda) = \frac{n}{\lambda} - \sum X_i$$

$$\text{Itt } -\ell'(\lambda) = \sum X_i - \mathbb{E}(\sum X_i) \text{ és } I(\lambda) = \mathbb{D}_\lambda^2(\sum X_i) = n\mathbb{D}_\lambda^2(X_1) = n\lambda^{-2}$$

4.3.5 Átparaméterezés és Fisher információ

- $\Xi \subset \mathbb{R}^q$, nyílt. $h : \Xi \rightarrow \Theta$ injektív, differenciálható. ($q \leq p$)
- $H = h'$
- $\mathcal{P}' = \{\mathbb{P}_{h(\xi)} : \xi \in \Xi\}$.
- $\mathcal{J}(\xi)$ a Fisher információ \mathcal{P}' -ben.
- Láncszabály

$$\partial_\xi \log f_{h(\xi)}(x) = \partial_\theta \log f_\theta(x)|_{\theta=h(\xi)} H(\xi)$$

•

$$\mathcal{J}(\xi) = H(\xi)^T I(h(\xi)) H(\xi)$$

- Ha $q = 1$, akkor $\mathcal{J}(\xi) = (\theta'(\xi))^2 I(\theta(\xi))$.

4.3.6 Átparaméterezés és információs határ

- $\Xi, \Theta \subset \mathbb{R}^p$, nyílt, $h : \Xi \rightarrow \Theta$ diffeomorfizmus.
- $H = h'$ invertálható
- $\mathcal{P}' = \{\mathbb{P}_{h(\xi)} : \xi \in \Xi\}$.
- $\mathcal{J}(\xi)$ a Fisher információ \mathcal{P}' -ben, $\mathcal{J}(\xi) = H(\xi)^T I(h(\xi)) H(\xi)$.

$$I(h(\xi))^{-1} = H(\xi) \mathcal{J}^{-1}(\xi) H^T(\xi)$$

- $g(\theta) = \mathbb{E}_\theta(T)$, $G = g'$
- $\mathbb{E}_{h(\xi)}(T) = g(h(\xi))$, $(g \circ h)' = (G \circ h)H$
- Információs határ. Ha a Cramer-Rao egyenlőtlenség feltételei teljesülnek

$$\begin{aligned} \Sigma_{h(\xi)}(T) &\geq G(h(\xi)) H \mathcal{J}(\xi)^{-1} H^T G(h(\xi))^T \\ &= G(h(\xi)) I(h(\xi))^{-1} G(h(\xi))^T \end{aligned}$$

Az információs határ nem változik.

4.3.7 Példák

Példa, indikátor minta Fisher információja

- X_1, \dots, X_n indikátor minta $p \in (0, 1)$ ismeretlen paraméterrel.
- X eloszlásainak a családja exponenciális családot alkot a $\theta = \ln \frac{p}{1-p}$ paraméterezéssel:

$$f_p(x) = p^{\sum x_i} (1-p)^{n-\sum x_i} = \exp \left\{ \ln \frac{p}{1-p} \sum x_i + n \ln(1-p) \right\} = \exp \{ \theta S(x) - n \ln(1 + e^\theta) \}$$

ahol S a mintaösszeg, $p = 1/(1 + e^{-\theta})$ és $b(\theta) = -n \ln(1-p) = -n \ln(1 - \frac{1}{1+e^{-\theta}}) = n \ln(1 + e^\theta)$.

Ellenőrzés: $e^{b(0)} = |\mathfrak{X}| = 2^n$

$$e^{b(\theta)} = e^{b(0)} \mathbb{E}_{\theta=0}(e^{\theta S}) = 2^n \mathbb{E}_{\theta=0}^n e^{\theta X_1} = 2^n \left(\frac{1}{2} + \frac{1}{2} e^\theta \right)^n = (1 + e^\theta)^n$$

•

$$I(\theta) = b''(\theta) = n \left(\frac{e^\theta}{1 + e^\theta} \right)' = n \left(\frac{1}{1 + e^{-\theta}} \right)' = n \frac{e^{-\theta}}{(1 + e^{-\theta})^2} = np(1 - p) = \mathbb{D}^2 S$$

- Eredeti paraméterezés mellett $\mathcal{I}(p) = (\theta'(p))^2 I(\theta(p))$

$$\theta(p) = \ln p - \ln(1 - p), \quad \theta'(p) = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{p(1 - p)}, \quad \mathcal{I}(p) = \frac{n}{p(1 - p)}$$

Példa, normális minta Fisher információj

- $X = (X_1, \dots, X_n) \sim N(\mu, \sigma^2)$ eloszlású minta ismeretlen $\mu \in \mathbb{R}$, $\sigma > 0$ paraméterrel.

$$f_{\mu, \sigma^2}(x) = \left(\frac{1}{\sigma \sqrt{2\pi}} \right)^n \exp \left\{ -\frac{1}{2\sigma^2} \sum (x_i - \mu)^2 \right\} = \exp \left\{ -\frac{1}{2} \tau \sum x_i^2 + \rho \sum x_i - b(\tau, \rho) \right\}$$

•

$$\ell(\mu, \sigma) = C - n \ln \sigma - \frac{1}{2} \cdot \sum_i \frac{(X_i - \mu)^2}{\sigma^2}$$

$$\ell'(\mu, \sigma) = \left(\sum \frac{X_i - \mu}{\sigma^2}, \sum \frac{(X_i - \mu)^2}{\sigma^3} - \frac{n}{\sigma} \right) = \frac{1}{\sigma} \left(\sum Z_i, \sum (Z_i^2 - 1) \right), \quad \text{ahol } Z_i = \frac{X_i - \mu}{\sigma} \sim N(0, 1)$$

- $\mathbb{E}_{\mu, \sigma}(\ell'(\mu, \sigma)) = 0$,

$$I(\mu, \sigma) = \Sigma_{\mu, \sigma}(\ell'(\mu, \sigma)) = \frac{n}{\sigma^2} \Sigma((Z_1, Z_1^2 - 1)) = \frac{1}{\sigma^2} \begin{pmatrix} \mathbb{D}^2 Z & \mathbb{E}(Z^3) \\ \mathbb{E}(Z^3) & \mathbb{D}^2 Z^2 \end{pmatrix} = \frac{1}{\sigma^2} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

- A minta eloszlásai exponenciális családot alkotnak

HF. $b = ?$ Fisher információ a természetes paraméterezés mellett.

- A minta eloszlása eltolás és skála paraméteres család. HF. Hogyan számolható a Fisher információ eltolás és skála paraméteres családban.

Egy mintaelem sűrűsége $f_{\mu, \sigma}(x) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right)$, $\mu \in \mathbb{R}$, $\sigma > 0$. f „szép” sűrűségfüggvény, $I(\mu, \sigma) = ?$, pl. $f(x) = \frac{1}{\pi}(1 + x^2)^{-1}$ Cauchy eloszlás.

4.4 Gyenge regularitási feltétel

Definíció 15 ((R) , azaz gyenge regularitási feltétel). $\mathcal{P} = \{f_\theta d\lambda : \theta \in \Theta\}$ dominált mértékcsalád, $\Theta \subset \mathbb{R}^p$ nyílt.

1. λ majdnem minden x -re $\theta \mapsto \sqrt{f_\theta(x)}$ folytonosan differenciálható.

2. $I(\theta)$ véges θ -ban folytonos, nem szinguláris ($\det I(\theta) > 0$).

- (R) nem függ a domináló mértéktől.

- Ha \mathcal{P} exponenciális családot alkot, akkor (R) lényegében mindig teljesül.

- Ha $\mathcal{P} U(0, \theta)$ -ből származó n -elemű minta eloszlásainak a családja, akkor $\theta \mapsto \sqrt{f_\theta(x)} = \theta^{-n/2} \mathbf{1}_{(\max x_i < \theta)}$ nem deriválható folytonosan majdnem minden $x \in (0, \infty)^n$ -re. Nem teljesíti (R) -et.

- Ha $\Theta_0 \subset \Theta$ nyílt, akkor $\mathcal{P}_0 = \{\mathbb{P}_\theta : \theta \in \Theta_0\}$ is teljesíti (R) -et.

4.4.1 Bederiválhatóság gyenge regularitási feltétel mellett

Tétel 17. \mathcal{P} teljesíti (R) -et. T r -dimenziós statisztika, $\mathbb{E}_\theta(|T|^2) < \infty$ és lokálisan korlátos θ -ban. Ekkor $\theta \mapsto \mathbb{E}_\theta(T)$ folytonosan differenciálható és $\partial_\theta \mathbb{E}_\theta(T) = \mathbb{E}_\theta(T\ell'(\theta))$.

- Elég $r = 1$ -re igazolni. Legyen $g(\theta) = \mathbb{E}_\theta(T\ell'(\theta)) = \int_{\mathfrak{X}} T(x) \partial_\theta f_\theta(x) \lambda(dx)$.

$$\begin{aligned} \mathbb{E}_{\theta_1}(T) - \mathbb{E}_{\theta_0}(T) &= \int_{\mathfrak{X}} T(x) (f_{\theta_1}(x) - f_{\theta_0}(x)) \lambda(dx) \\ &= \int_{\mathfrak{X}} T(x) \int_0^1 \partial_\theta f_{\theta_0 + u(\theta_1 - \theta_0)}(x) (\theta_1 - \theta_0) du \lambda(dx) = \int_0^1 g(\theta_0 + u(\theta_1 - \theta_0)) du (\theta_1 - \theta_0) \end{aligned}$$

- Elég g folytonossága, ugyanis

$$\begin{aligned} \mathbb{E}_{\theta_1}(T) - \mathbb{E}_{\theta_0}(T) &= \int_0^1 g(\theta_0 + u(\theta_1 - \theta_0)) du (\theta_1 - \theta_0) \\ &= g(\theta_0)(\theta_1 - \theta_0) + \int_0^1 g(\theta_0 + u(\theta_1 - \theta_0)) - g(\theta_0) du (\theta_1 - \theta_0) \\ &= g(\theta_0)(\theta_1 - \theta_0) + o(|\theta_1 - \theta_0|), \quad \text{ha } \theta_1 \rightarrow \theta_0. \end{aligned}$$

4.4.2 $\theta \mapsto \mathbb{E}_\theta(T\ell'(\theta))$ folytonos

Lemma 3. Ha $\theta \mapsto \sqrt{f_\theta(x)}$ folytonosan differenciálható λ majdnem minden x -re, $\theta \mapsto (I(\theta))_{i,i}$ $i = 1, \dots, p$ véges értékű, folytonos és $\sup_{\theta \in \Theta} \mathbb{E}_\theta(T^2) < \infty$.

Ekkor $\theta \mapsto \mathbb{E}_\theta(T\ell'(\theta))$ folytonos.

•

$$g(\theta) = \mathbb{E}_\theta(T\ell'(\theta)) = \int_{\mathfrak{X}} T(x) \frac{\partial_\theta f_\theta(x)}{f_\theta(x)} f_\theta(x) \lambda(dx) = \int_{\mathfrak{X}} T(x) \sqrt{f_\theta(x)} 2\partial_\theta \sqrt{f_\theta(x)} \lambda(dx)$$

- Kell: $g(\theta_n) \rightarrow g(\theta_0)$, ha $\theta_n \rightarrow \theta_0$.

Legyen $h_n(x) = T(x) \sqrt{f_{\theta_n}(x)}$, $k_n(x) = 2\partial_\theta \sqrt{f_{\theta_n}(x)}$. Ezzel a jelöléssel:

$$g(\theta_n) = \int h_n k_n d\lambda, \quad g(\theta_0) = \int h_0 k_0 d\lambda, \quad \text{Cél: } \int h_n k_{n,i} d\lambda \rightarrow \int h_0 k_{0,i} d\lambda, \quad i = 1, \dots, p$$

- Itt

$$\sup_n \int h_n^2 d\lambda = \sup_n \mathbb{E}_{\theta_n}(T^2) < \infty, \quad \int k_{n,i}^2 d\lambda = (I(\theta_n))_{i,i} \rightarrow (I(\theta_0))_{i,i} = \int k_{0,i}^2 d\lambda, \quad i = 1, \dots, p$$

és $\theta \mapsto \sqrt{f_\theta(x)} \in C^1$ λ m.m. x miatt $h_n \rightarrow h_0$ és $k_n \rightarrow k_0$ λ -m.m. és feltehető, hogy $\lambda(\mathfrak{X}) = 1$.

4.4.3 Technikai lemma

Lemma 4. $h_n, h_0, k_n, k_0 : \mathfrak{X} \rightarrow \mathbb{R}$, λ valószínűségi mérték $\mathcal{B}(\mathfrak{X})$ -en.

$h_n \rightarrow h_0$, $k_n \rightarrow k_0$ λ m.m., $\sup_n \int h_n^2 d\lambda < \infty$, $\int k_n^2 d\lambda \rightarrow \int k_0^2 d\lambda < \infty$.

Ekkor $\int h_n k_n d\lambda \rightarrow \int h_0 k_0 d\lambda$.

Jelölés $\int h = \int_{\mathfrak{X}} h d\lambda$. Vázlat:

- Csonkolás. ($\varepsilon > 0$, \approx_ε jelentése a két oldal eltérése legfeljebb ε)

Ha elég nagy K -ra és elég nagy n_0 -tól kezdve

$$\int h_n k_n \approx_{2\varepsilon} \int h_n k_n \mathbf{1}_{(|k_n| < 2|k_0| < K)} \approx_\varepsilon \int h_0 k_0 \mathbf{1}_{(2|k_0| < K)} \approx_\varepsilon \int h_0 k_0 \implies \int h_n k_n \rightarrow \int h_0 k_0$$

- Cél: ha K olyan nagy, hogy $R \int k_0^2 \mathbf{1}_{(2|k_0| \geq K)} < \varepsilon^2$, ahol $R = \sup_n \int h_n^2$, akkor

$$\begin{aligned} \limsup_{n \rightarrow \infty} \int h_n^2 \int k_n^2 (1 - \mathbf{1}_{(|k_n| < 2|k_0| < K)}) &\leq 4R \int k_0^2 \mathbf{1}_{(2|k_0| \geq K)} < 4\varepsilon^2 \quad (*) \\ \lim_{n \rightarrow \infty} \left| \int h_n k_n \mathbf{1}_{(|k_n| < 2|k_0| < K)} - \int h_0 k_0 \mathbf{1}_{(2|k_0| < K)} \right| &= 0 \end{aligned}$$

- Cauchy-Schwarz: K választása \implies utolsó lépés hibája $\leq \varepsilon$,

$(*) \implies$ első lépés hibája $\leq 2\varepsilon$, elég nagy n_0 -tól kezdve.

Részletesen:

Állítás 14. $k_n, k_0 : \mathfrak{X} \rightarrow \mathbb{R}$, $k_n \rightarrow k_0$ λ m.m., $\int k_n^2 \rightarrow \int k_0^2 < \infty$.

Ekkor $\int k_n^2 \mathbf{1}_{(|k_n| \geq 2|k_0|)} \rightarrow 0$.

$k_n^2 \mathbf{1}_{(|k_n| < 2|k_0|)} \rightarrow k_0^2$. Dominált konvergencia tétel ($4k_0^2$ integrálható majoráns)

$$\int k_n^2 \mathbf{1}_{(|k_n| \leq 2|k_0|)} \rightarrow \int k_0^2 = \lim \int k_n^2 \implies \int k_n^2 (1 - \mathbf{1}_{(|k_n| < 2|k_0|)}) \rightarrow 0$$

Következmény 2.

$$\limsup_{n \rightarrow \infty} \int k_n^2 (1 - \mathbf{1}_{(|k_n| < 2|k_0| < K)}) \leq 4 \int k_0^2 \mathbf{1}_{(2|k_0| \geq K)}$$

$$\begin{aligned} \int k_n^2 (1 - \mathbf{1}_{(|k_n| < 2|k_0| < K)}) &= \int k_n^2 \mathbf{1}_{(|k_n| \geq 2|k_0|)} + \int k_n^2 \mathbf{1}_{(|k_n| < 2|k_0| \geq K)} \\ &\leq \int k_n^2 \mathbf{1}_{(|k_n| \geq 2|k_0|)} + \int 4k_0^2 \mathbf{1}_{(2|k_0| \geq K)} \rightarrow \int 4k_0^2 \mathbf{1}_{(2|k_0| \geq K)} \end{aligned}$$

Állítás 15. $h_n, h_0, k_n, k_0 : \mathfrak{X} \rightarrow \mathbb{R}$, $\lambda(\mathfrak{X}) = 1$, $h_n \rightarrow h_0$, $k_n \rightarrow k_0$ λ m.m., $R = \sup_n \int h_n^2 d\lambda < \infty$.

Ekkor

$$\int h_n k_n \mathbf{1}_{(|k_n| < 2|k_0| < K)} \rightarrow \int h_0 k_0 \mathbf{1}_{(2|k_0| < K)}$$

- $h_0 k_n \mathbf{1}_{(|k_n| < 2|k_0| < K)} \rightarrow h_0 k_0 \mathbf{1}_{(2|k_0| < K)}$ és $|h_0|K$ integrálható majoráns.

$$\lim \int h_0 k_n \mathbf{1}_{(|k_n| < 2|k_0| < K)} = \int h_0 k_0 \mathbf{1}_{(2|k_0| < K)}$$

- h_n L^2 -ben korlátos, ezért egyenletesen integrálható. Emlékeztető:

$$\sup_n \int |h_n| \mathbf{1}_{(|h_n| \geq L)} \leq \sup_n \frac{1}{L} \int h_n^2 = \frac{R}{L} \rightarrow 0 \quad \text{ha } L \rightarrow \infty.$$

- $\{h_n : n \geq 1\}$ egyenletesen integrálható, $h_n \rightarrow h_0$ m.m. $\implies \int |h_n - h_0| \rightarrow 0 \implies$

$$\int (h_n - h_0) k_n \mathbf{1}_{(|k_n| < 2|k_0| < K)} \rightarrow 0$$

•

$$\int h_0 k_0 \mathbf{1}_{(2|k_0| < K)} = \lim \int h_0 k_n \mathbf{1}_{(|k_n| < 2|k_0| < K)} = \lim \int h_n k_n \mathbf{1}_{(|k_n| < 2|k_0| < K)}.$$

4.4.4 (R) következményei

Definíció 16 ((R), azaz gyenge regularitási feltétel). $\mathcal{P} = \{f_\theta d\lambda : \theta \in \Theta\}$ dominált mértékcsalád, $\Theta \subset \mathbb{R}^p$ nyílt.

1. λ majdnem minden x -re $\theta \mapsto \sqrt{f_\theta(x)}$ folytonosan differenciálható.

2. $I(\theta)$ véges θ -ban folytonos, nem szinguláris ($\det I(\theta) > 0$).

Lemma 5 (Már igazoltuk). $(R) + \sup_{\theta \in \Theta} \mathbb{E}_\theta(|T|^2) < \infty \implies g(\theta) = \mathbb{E}_\theta(T)$ deriválható, $g'(\theta) = \mathbb{E}_\theta(T\ell'(\theta))$ folytonos, ahol $\ell(\theta, x) = \log f_\theta(x)$.

- (R) teljesül, $T \equiv 1$. Ekkor $0 = \mathbb{E}_\theta(\ell'(\theta)) = \int_{\mathfrak{X}} \partial_\theta f_\theta(x) \lambda(dx) \implies n$ elemű minta Fisher információja az egy elemű minta $I_1(\theta)$ Fisher információjának n -szerese.
- (R) teljesül, T torzítatlan $g(\theta)$ -ra, $\mathbb{E}_\theta(|T|^2)$ lokálisan korlátos. Cramer-Rao egyenlőtlenség feltételei teljesülnek:
 $(\ell'$ létezik, $\mathbb{E}_\theta(\ell'(\theta)) = 0$, $I(\theta) = \mathbb{E}_\theta(\ell'(\theta))^T \ell'(\theta))$ invertálható. $g(\theta) = \mathbb{E}_\theta(T)$ deriválható és $g'(\theta) = \mathbb{E}_\theta(T\ell'(\theta))$)

$$\Sigma_\theta(T) \geq g'(\theta)I(\theta)^{-1}g'(\theta)^T = \frac{1}{n}g'(\theta)I_1(\theta)^{-1}g'(\theta)^T$$

5. 2021.03.08.

5.1 Fisher információ folyt.

5.1.1 Gyenge regularitási feltétel exponenciális családban

Állítás 16 (Emlékeztető). Ha Y p -dimenziós vektorváltozó és $Y \in L^2$, akkor $\mathbb{P}(Y \in \mathbb{E}(Y) + \text{im } \Sigma(Y)) = 1$

$a \in \mathbb{R}^p$ -re $\mathbb{D}^2 a^T Y = a^T \Sigma(Y) a \implies \forall a \perp \text{im } \Sigma(Y)$ vektorra $\mathbb{D}^2 a^T Y = 0$ és $\mathbb{P}(a^T(Y - \mathbb{E}(Y)) = 0) = 1$.
 a_1, \dots, a_r ($\text{im } \Sigma(Y)$)[⊥] bázisa, $\mathbb{P}(a_i^T(Y - \mathbb{E}(Y)) = 0, i = 1, \dots, r) = 1$ és így $\mathbb{P}(Y - \mathbb{E}(Y) \in \text{im } \Sigma(Y)) = 1$.

- $\mathcal{P} = \{f_\theta d\lambda : \theta \in \Theta\}$ exponenciális családot alkot és $\Theta \subset \mathbb{R}^p$ nyílt.

$$f_\theta(x) = \exp\{\theta T(x) - b(\theta)\}, \quad \sqrt{f_\theta(x)} = \exp\{\frac{1}{2}\theta T(x) - \frac{1}{2}b(\theta)\}$$

Itt $b : \Theta \rightarrow \mathbb{R}$ sima, azaz $\theta \mapsto \sqrt{f_\theta(x)}$ folytonosan differenciálható.

- $I(\theta) = b''(\theta) = \Sigma_\theta(T)$, azaz $I(\theta)$ véges értékű és folytonos.
- A \mathbb{P}_θ eloszlások ekvivalensek, ugyanez igaz T eloszlásaira is. Tetszőleges V affin altérre $\mathbb{P}_{\theta_0}(T \in V) = 1 \iff \mathbb{P}_\theta(T \in V) = 1$ minden $\theta \in \Theta$ -ra.
- (R) teljesül, ha T a \mathbb{R}^p nem egy valódi affin altérből veszi fel az értékeit, pl. ha a paraméterezés egyértelmű, akkor (R) teljesül.
- Ha $T \in V \subset \mathbb{R}^p$ valódi V affin altérrel, akkor alacsonyabb dimenziós paraméterezést is választhatunk és (R) teljesül.

Azaz, ha exponenciális családuk van nyílt paraméterterrel, akkor (R) lényegében mindig teljesül.

5.1.2 További példák

- X Cauchy eloszlás eltolás paraméteres családjából származó n elemű minta. X eloszlásai nem alkotnak exponenciális családot, de

$$\sqrt{f_\theta(x)} = \prod \pi^{-1/2} (1 + (x_i - \theta)^2)^{-1/2}$$

folytonosan deriválható, és

$$\ell(\theta, x) = c + \sum_i \frac{-1}{2} \ln(1 + (x_i - \theta)^2), \quad \ell'(\theta, x) = \sum_i \frac{x_i - \theta}{1 + (x_i - \theta)^2}$$

$$\mathbb{E}_\theta(\ell'(\theta)) = n \int_{\mathbb{R}} \frac{x}{\pi(1+x^2)^2} dx = 0 \quad I(\theta) = \mathbb{D}_\theta^2 \ell'(\theta) = n \int_{\mathbb{R}} \frac{x^2}{(1+x^2)^2} \frac{1}{\pi(1+x^2)} dx < \infty$$

$I(\theta)$ véges, folytonos és nem nulla. X eloszlásainak családja teljesíti a gyenge regularitási feltételt.

- X a $(0, \theta)$ intervallumon egyenletes eloszlású, $\theta > 0$. Ekkor $\mathbb{P}_\theta(X < y) = 1 \wedge \frac{y}{\theta}$

$$\partial_\theta \mathbb{P}_\theta(X < y) = \frac{-y}{\theta^2} \mathbf{1}_{(y < \theta)}, \quad \mathbb{E}_\theta(\mathbf{1}_{(X < y)} \ell'(\theta)) = \int_0^\infty \mathbf{1}_{(x < y)} \partial_\theta f_\theta(x) dx = \int_0^\theta \mathbf{1}_{(x < y)} \frac{-1}{\theta^2} dx = \frac{-y \wedge \theta}{\theta^2}$$

X eloszlásainak a családja nem teljesíti (R) -et.

5.1.3 Független minták Fisher információjának összeadódik

Definíció 17 ((R) , azaz gyenge regularitási feltétel). $\mathcal{P} = \{f_\theta d\lambda : \theta \in \Theta\}$ dominált mértékcsalád, $\Theta \subset \mathbb{R}^p$ nyílt.

λ m.m. x -re $\theta \mapsto \sqrt{f_\theta(x)} \in C^1$. $I(\theta)$ véges, θ -ban folytonos, és $I(\theta)$ nem szinguláris ($\det I(\theta) > 0$).

Állítás 17. X, Y minden $\theta \in \Theta$ mellett függetlenek és eloszlásaikra teljesül (R) .

Ekkor (X, Y) együttes eloszlásaira is teljesül (R) és $I_{X,Y}(\theta) = I_X(\theta) + I_Y(\theta)$.

- X eloszlásainak domináló mértéke λ_X , Y -ra λ_Y . Ekkor (X, Y) eloszlásait $\lambda_X \otimes \lambda_Y$ dominálja.
- $f_{(X,Y),\theta}(x, y) = f_{X,\theta}(x)f_{Y,\theta}(y)$. Ha $\sqrt{f_{X,\theta}(x)} \sqrt{f_{Y,\theta}(y)}$ folytonosan deriválható θ -ban λ_X majdnem minden x -re ill. λ_Y majdnem minden y -ra, akkor $\sqrt{f_{X,Y,\theta}(x, y)}$ is folytonosan differenciálható $\lambda_X \otimes \lambda_Y$ m.m. (x, y) -ra.
- $\ln f_{X,Y,\theta}(x, y) = \ln f_{X,\theta}(x) + \ln f_{Y,\theta}(y)$. Ebből $I_{X,Y}(\theta) = I_X(\theta) + I_Y(\theta)$ következik.
- $I_{X,Y}(\theta)$ folytonos, véges értékű.
- $I_{X,Y} \geq I_X$, amiből $\ker I_{X+Y} \subset \ker I_X = \{0\}$, $I_{X,Y}$ is invertálható, minden θ -ra.

5.1.4 $S(X)$ Fisher információjának nem lehet X információ tartalmánál nagyobb

Állítás 18. X eloszlásainak a családjára teljesül (R) . S statisztika. Ekkor S eloszlásainak a családjára is dominált.

Ha S eloszlásainak családjára is teljesül (R) , akkor $I_{S(X)}(\theta) \leq I_X(\theta)$.

- $\lambda \circ S^{-1}$ domináló mérték S eloszlásaihoz:

$$\mathbb{P}_\theta(S \in H) = \int_{\mathcal{X}} f_\theta(x) \mathbf{1}_{(S(x) \in H)} \lambda(dx) = \int_H g_\theta(s) \lambda \circ S^{-1}(ds), \quad g_\theta(S) = \frac{d\mathbb{P}_\theta|_{\sigma(S)}}{d\lambda|_{\sigma(S)}}.$$

- Ha S eloszlásai is teljesítik (R) -et, $\mathbb{P}_\theta(S \in H)$ -re adott mindkét formulában az integrálon belül lehet deriválni θ szerint. $\ell_X(\theta, x) = \ln f_\theta(x)$ és $\ell_S(\theta, s) = \ln g_\theta(s)$ jelöléssel

$$\int_{\mathcal{X}} \partial_\theta f_\theta(x) \mathbf{1}_{(S(x) \in H)} \lambda(dx) = \mathbb{E}_\theta(\mathbf{1}_{(S \in H)} \ell'_X(\theta, X)) = \int_H g'_\theta(s) \lambda \circ S^{-1}(ds) = \mathbb{E}_\theta(\mathbf{1}_{(S \in H)} \ell'_S(\theta, S))$$

- $\mathbb{E}_\theta(\ell'_X(\theta, X)|S) = \ell'_S(\theta, S)$ és a teljes szórásnégyzet tétel szerint $I_S(\theta) \leq I_X(\theta)$.

5.1.5 Elégséges statisztika Fisher információjának

Állítás 19. X minta eloszlásainak családjára teljesül (R) , S elégséges statisztika. Ekkor S eloszlásainak a családjára is teljesül (R) és $I_{S(X)}(\theta) = I_X(\theta)$.

- (R) nem függ a domináló mértéktől, $\mathbb{P}_0 \in \mathcal{P}'$ választható. ($\mathcal{P}' = \{\sum c_i \mathbb{P}_{\theta_i} : c_i \geq 0, \sum c_i = 1\}$).
- S elégséges, ezért \mathbb{P}_0 -ra vonatkozó sűrűségek $f_\theta(x) = g_\theta(S(x))$ alakúak.
- g_θ S eloszlásának a sűrűsége a $\mathbb{P}_0 \circ S^{-1}$ domináló mértékre nézve.
- $\sqrt{f_\theta(x)} = \sqrt{g_\theta(S(x))}$, azaz ha (R) teljesül, akkor $\sqrt{g_\theta(s)}$ folytonosan deriválható $\mathbb{P}_0 \circ S^{-1}$ m.m. s -re.
- $\ell'_X(\theta, X) = \ell'_S(\theta, S)$, amiből $I_X(\theta) = I_S(\theta)$ és (R) I -re vonatkozó előírásai I_S -re is igazak.

5.1.6 Fisher információ és elégséges statisztika

Állítás 20. (R) teljesül X és az S statisztika eloszlásainak családjára is. $f_\theta > 0$ a mintatéren minden $\theta \in \Theta$ -ra és Θ összefüggő nyílt.

Ha $I_X(\theta) = I_{S(X)}(\theta)$, akkor S elégséges.

- $\mathbb{P}_0 = \mathbb{P}_{\theta_0}$ $\theta_0 \in \Theta$ választható domináló mértéknek, $f_{\theta_0} \equiv 1$.
- Láttuk, hogy $\mathbb{E}_\theta((\partial_\theta \ln f_\theta)(X)|S) = (\partial_\theta \ln g_\theta)(S)$, ahol g_θ S eloszlásainak sűrűsége $\mathbb{P}_{\theta_0} \circ S^{-1}$ -re.
- A feltétel szerint $I_X = I_S$, ami csak akkor lehet, ha $\partial_\theta \ln f_\theta(x) = \partial_\theta \ln g_\theta(S(x))$ minden θ -ra \mathbb{P}_0 m.m..
- $\partial_\theta \ln f_\theta(x), \partial_\theta \ln g_\theta(x)$ \mathbb{P}_0 m.m. x -re folytonos θ -ban (R) miatt.
- $\mathbb{P}_0(\forall \theta \in \Theta, \partial_\theta \ln f_\theta(x) = \partial_\theta \ln g_\theta(S(x))) = 1$.
- $\ln f_\theta(x) - \ln g_\theta(S(x))$ majdnem minden x -re θ -ban konstans, azaz $f_\theta(x) = g_\theta(S(x))h(x)$ alakú, ahol $h(x) = f_{\theta_0}(x)/g_{\theta_0}(S(x))$. S elégséges a faktorizációs tétel miatt.

5.1.7 Összefoglalás

Ha az X, Y minta és $S(X)$ eloszlásainak családja teljesíti (R) -et, akkor

- X, Y függetlenek $I_{X,Y}(\theta) = I_X(\theta) + I_Y(\theta)$, pl. n elemű mintára $I_n(\theta) = nI_1(\theta)$.
- $I_{S(X)}(\theta) \leq I_X(\theta)$
- $I_{S(X)}(\theta) = I_X(\theta)$, ha S elégséges.
- Ha $I_{S(X)}(\theta) = I_X(\theta)$, és $f_\theta > 0$ \mathfrak{X} -en, akkor S elégséges.

5.2 Becslési módszerek

- Momentum módszer: válasszuk azt a paramétert, ami az első néhány tapasztalati momentumra illeszkedik.
- Maximum likelihood becslés: válasszuk azt a paramétert, ami a minta likelihoodját maximalizálja.
- Bayes becslés: válasszuk azt a paramétert ami az „a priori” rizikót minimalizálja (részletek később).

5.2.1 Momentum módszer indikátor mintára

„Válasszuk azt a paramétert, ami az első néhány tapasztalati momentumra illeszkedik”

- X_1, \dots, X_n indikátor minta ismeretlen $\theta \in [0, 1]$ paraméterrel.
- $\theta = \mathbb{E}_\theta(X_1)$
- $S = \sum X_i$ a mintaösszeg, akkor $\frac{1}{n}S$ a mintaátlag a várható érték „tapasztalati” becslése.
- $\hat{\theta}_{MM} = \frac{1}{n}S$.

Itt a $\phi(\theta) = \mathbb{E}_\theta(X_1)$ jelöléssel a $\phi(\hat{\theta}_{MM}) = \frac{1}{n} \sum X_i$ összefüggést oldottuk meg.

5.2.2 Maximum likelihood elv indikátor mintára

„Válasszuk az a paramétert, ami a megfigyelés likelihoodját $f_\theta(x)$ -et maximalizálja.

- X_1, \dots, X_n indikátor minta ismeretlen $\theta \in [0, 1]$ paraméterrel.
 $\hat{\theta}_{ML} = \arg \max_\theta f_\theta(x)$.
- $f_\theta(x) = \theta^{S(x)}(1 - \theta)^{n-S(x)}$, ahol $S(x) = \sum x_i$.
- Ha $S = 0$, vagy $S = n$, akkor a $\hat{\theta} = \frac{1}{n}S$.
- Ha $0 < S < n$ akkor áttérhetünk a loglikelihoodra: Likelihood egyenlet:

$$0 = \ell'(\theta, x) = \frac{S}{\theta} - \frac{n-S}{1-\theta} \iff \theta = \frac{1}{n}S$$

$$\hat{\theta}_{ML} = \frac{1}{n}S \text{ a mintaátlag.}$$

Megjegyzések:

- A paraméter teret $[0, 1]$ -nek választottuk, hogy akkor is legyen maximumhely, ha $S(X) = 0$ vagy $S(X) = n$.
- A maximum helyet gyakran a likelihood egyenlet megoldásával keressük.

5.2.3 Bayes módszer indikátor mintára

- Adatok nélkül minden paraméter érték „egyformán valószínű”. Az **a priori eloszlás** a $\Theta = (0, 1)$ paraméterterén egyenletes, $q \equiv 1$ a sűrűségfüggvénye. (Más eloszlást is választhatunk).
- θ is valószínűségi változó, $\Omega = \Theta \times \mathfrak{X}$ a valószínűségi mező alaphalmaza. $\theta : \Omega \rightarrow \Theta$ az első, $X : \Omega \rightarrow \mathfrak{X}$ a második koordináta leképezés.

$$\mathbb{P}(H) = \int_{\Theta} \int_{\mathfrak{X}} \mathbf{1}_{((t,x) \in H)} \mathbb{P}_t(dx) q(t) dt = \int_H f_t(x) q(t) dt \otimes \lambda(dx).$$

ahol λ a számláló mérték \mathfrak{X} -en, dt a „Lebesgue mérték”.

- Négyzetes veszteség esetén $R_T(\theta) = \mathbb{E}_\theta((\theta - T)^2) = \mathbb{E}((\theta - T)^2 | \theta)$ a rizikó függvény, $R_T = \mathbb{E}(R_T(\theta)) = \mathbb{E}((\theta - T)^2)$ az **a priori rizikó**. Ez az átlagos veszteség, ha a T statisztikát használjuk becslésként.
- T csak a megfigyeléstől függhet (azaz $\sigma(X)$ mérhető), ezért az a priori rizikót $T(X) = \mathbb{E}(\theta | X)$ minimalizálja. (θ, X) együttes sűrűségfüggvénye ($dt \times \lambda(dx)$ -re nézve) $f_t(x)q(t)$, a feltételes sűrűségfüggvény:

$$f_{\theta|X}(t|x) = \frac{f_{\theta,X}(t,x)}{f_X(x)} = C(x)t^{\sum x_i}(1-t)^{n-\sum x_i}$$

Azaz θ X -re vonatkozó feltételes eloszlása Béta($S+1, n+1-S$), ahol S a mintaösszeg.

- $\mathbb{E}(\theta|X) = \frac{S+1}{n+2}$, S A mintaösszeg. (Laplace simítás).
- Ha $g(\theta)$ -t szeretnénk becsülni, akkor $\mathbb{E}(g(\theta)|X)$ lenne a Bayes becslés.

5.2.4 Tapasztalati eloszlás

Definíció 18 (Tapasztalati eloszlás, tapasztalati eloszlásfüggvény). X_1, \dots, X_n minta, $P_n^* = \frac{1}{n} \sum_i \delta_{X_i}$ a tapasztalati eloszlás, $F_n^*(t) = P_n^*((-\infty, t)) = \frac{1}{n} \sum_i \mathbf{1}_{(X_i < t)}$ a tapasztalati eloszlásfüggvény.

- P_n^* véletlen valószínűségi mérték F_n^* véletlen függvény (folyamat).
- μ_n, μ valószínűségi mértékek $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ -en. $\mu_n \xrightarrow{w} \mu$, ha minden h folytonos korlátos függvényre $\int h d\mu_n \rightarrow \int h d\mu$.
 $\mu_n \xrightarrow{w} \mu$ pontosan akkor, ha $\mu_n((-\infty, t)) \rightarrow \mu((-\infty, t))$ minden olyan t -re, amire $\mu(\{t\}) = 0$.
- $F_n^*(t) \rightarrow F(t)$ egy valószínűséggel, ahol F a mintaelemek közös eloszlásfüggvénye. Ez a nagy számok törvénye a $\mathbf{1}_{(X_1 < t)}, \mathbf{1}_{(X_2 < t)}, \dots$ iid sorozatra.

Több is igaz.

Tétel 18 (Glivenko-Cantelli tétel (statisztika alaptétele)). $\sup_{t \in \mathbb{R}} |F_n^*(t) - F(t)| \rightarrow 0$ egy valószínűséggel.

Ebből $P_n^* \xrightarrow{w} P$ egy valószínűséggel következik.

5.2.5 Tapasztalati becslések

Definíció 19. \mathcal{P} eloszlások egy családja, amely tartalmazza a véges sok pontra koncentrált eloszlásokat. ϕ a \mathcal{P} -n értelmezett funkcionál. $\phi(P)$ **tapasztalati becslése** $\phi(P_n^*)$, ahol P_n^* a tapasztalati eloszlás.

Definíció 20. X_1, X_2, \dots iid. minta a P eloszlásból, $\phi(P)$ becslése n elemű mintából $T_n = T_n(X_1, \dots, X_n)$.

A T_n konzisztens, ha $T_n \xrightarrow{P} \phi(P)$, és erősen konzisztens, ha $T_n \rightarrow \phi(P)$ egy valószínűséggel.

- Példa. $N(\mu, 1)$ eloszlásból származó n elemű minta esetén kézenfekvő a paramétert a mintaátlaggal becsülni. Itt a paraméter a várható érték, ami véges tartójú eloszlásokra is létezik.

$$\mathcal{P} = \{P : \mu \text{ valószínűségi Borel mérték } \mathbb{R}\text{-en, } \int |x| P(dx) < \infty\}, \quad \phi(P) = \int x P(dx).$$

$$P_n^* \in \mathcal{P} \text{ és } \phi(P_n^*) = \int x \left(\frac{1}{n} \sum \delta_{X_i}(dx) \right) = \frac{1}{n} \sum_i X_i.$$

Ebben a példában $\phi(P_n^*) \rightarrow \phi(P)$, de ϕ nem gyengén folytonos \mathcal{P} -n, sőt $N(\mu, 1)$ -ben sem.

- A minta a P eloszlásból származik, ϕ gyengén folytonos a P pontban. Ekkor ϕ tapasztalati becslése erősen konzisztens, hiszen $P_n^* \xrightarrow{w} P$ a Glivenko-Cantelli tétel szerint.

Pl. Ha mintaelemek közös eloszlásának mediánja egyértelmű, akkor a tapasztalati medián erősen konzisztens a mediánra.

5.2.6 Momentum módszer

- $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^p$.
- Legyen $\phi(\theta) = \mathbb{E}_\theta(g(X_1))$. Tegyük fel, hogy $\phi : \Theta \rightarrow T \subset \mathbb{R}^p$ kölcsönösen egyértelmű és oda-vissza folytonos.

- $\psi = \phi(\theta)$ tapasztalati becslése

$$\hat{\psi}_n = \frac{1}{n} \sum_{i=1}^n g(X_i)$$

- $\hat{\psi}_n \rightarrow \psi$ a nagy számok törvénye szerint. Ha n elég nagy és $\psi \in \text{int } T$, akkor $\hat{\psi} \in T$ és $\hat{\theta} = \phi^{-1}(\hat{\psi})$ értelmes.
- $\hat{\psi}_n \rightarrow \psi$ egy valószínűséggel és ϕ^{-1} folytonossága miatt $\hat{\theta}_n$ erősen konzisztens, de általában nem torzítatlan.
 $\hat{\psi}_n$ torzítatlan ψ -re, de ϕ^{-1} általában nem lineáris.
- Klasszikus momentum módszernél $g(x) = (x, x^2, \dots, x^p)$.

Példák

- X_1, \dots, X_n a $(-\theta, 3\theta)$ intervallumon egyenletes eloszlásból származó minta, $\theta > 0$.
 $\mathbb{E}_\theta(X_1) = \theta$.
 $\hat{\theta}_{MM}$ a mintaátlag. Minden n -re pozitív valószínűséggel $\hat{\theta}_{MM} < 0$.
- X_1, \dots, X_n Béta(α, β), $\alpha, \beta > 0$, eloszlásból származó n elemű minta.

$$(\mathbb{E}(X_1), \mathbb{E}(X_1^2)) = \phi(\alpha, \beta), \quad \text{ahol} \quad \phi(\alpha, \beta) = \left(\frac{\alpha}{\alpha + \beta}, \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} \right)$$

$\hat{\alpha}_{MM}, \hat{\beta}_{MM}$ a $\phi(\hat{\alpha}, \hat{\beta}) = (\bar{X}, \bar{X}^2)$ egyenlet megoldása. HF.

5.3 Glivenko–Cantelli tétel

Tétel 19 (Statisztika alaptétele). X_1, X_2, \dots független azonos eloszlású változók, a közös eloszlásfüggvény F . Ekkor

$$\lim_{n \rightarrow \infty} \sup_t |F_n(t) - F(t)| = 0 \quad \text{egy valószínűséggel, ahol } F_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(X_i \leq t)}$$

Megjegyzések.

- Vektor értékű megfigyelésekre is igaz. $X_i \in \mathbb{R}^d$ értékű. Ilyenkor $t \in \mathbb{R}^d$ és $X_i \leq t$ ha minden koordinátára igaz az egyenlőtlenség.
- Balról folytonosság miatt $\sup_t |F_n(t) - F(t)| = \sup_{t \in \mathbb{Q}} |F_n(t) - F(t)|$, azaz a szuprénum valószínűségi változót definiál.
- Rögzített t -re $F_n(t)$ független $F(t)$ várható értékű indikátorok átlaga. A nagy számok törvényéből $F_n(t) \rightarrow F(t)$ egy valószínűséggel.
- CHT: rögzített t -re $\sqrt{n}(F_n(t) - F(t)) \xrightarrow{d} N(0, F(t)(1 - F(t)))$.

Bizonyítás. • $k \geq 1$. $t_i = \inf\{t \in \mathbb{R} : F(t) \geq \frac{i}{k}\}$, $i = 1, \dots, k-1$, $t_0 = -\infty$, $t_k = \infty$ és $\mathcal{S}(k) = \{(-\infty, t_i), (-\infty, t_i] \cap \mathbb{R} : i = 0, \dots, k\}$ véges halmaz rendszer.

- Ha $t \in \mathbb{R}$ létezik $A, B \in \mathcal{S}$, hogy $A \subset (-\infty, t) \subset B$ és $P(B) \leq P(A) + \frac{1}{k}$.

Ha $t \in \{t_1, \dots, t_k\}$, akkor $A = B = (-\infty, t)$ jó, különben létezik i hogy $t_i < t < t_{i+1}$ és $A = (-\infty, t_i]$, $B = (-\infty, t_{i+1})$ jó ($P(B) = F(t_{i+1}) \leq \frac{i+1}{k}$ és $P(A) = F(t_i) \geq \frac{i}{k}$).

$$F_n(t) - F(t) \leq P_n^*(B) - P(B) + P(B) - P(A), \quad F(t) - F_n(t) \leq P(B) - P(A) + P(A) - P_n^*(A)$$

- A nagy számok törvénye alapján $P_n^*(H) \rightarrow P(H)$ minden $H \in \mathcal{C}(k)$ -ra

$$\limsup_{n \rightarrow \infty} \sup_{t \in \mathbb{R}} |F_n(t) - F(t)| \leq \inf_k \limsup_{n \rightarrow \infty} \left(\max_{H \in \mathcal{S}(k)} |P_n^*(H) - P(H)| + \frac{1}{k} \right) = \inf_k \frac{1}{k} = 0$$

Definíció 21. P valószínűségi mérték $(\mathfrak{X}, \mathcal{B})$ -n. $\mathcal{C} \subset \mathcal{B}$ halmazrendszer végesen approximálható, ha minde $\varepsilon > 0$ -ra létezik $\mathcal{S}(\varepsilon) \subset \mathcal{B}$ véges rendszer úgy, hogy minden $C \in \mathcal{C}$ -re, létezik $A, B \in \mathcal{S}(\varepsilon)$, amire $A \subset C \subset B$ és $P(B) \leq P(A) + \varepsilon$.

A skaláreset azon múlt, hogy $\mathcal{C} = \{(-\infty, t) : t \in \mathbb{R}\}$ végesen approximálható, továbbá

Állítás 21. Ha \mathcal{C} végesen approximálható, akkor $\sup_{H \in \mathcal{C}} |P(H) - P_n^*(H)| \leq \varepsilon + \max_{A \in \mathcal{S}(\varepsilon)} |P(A) - P_n^*(A)|$

és a tapasztalati eloszlásra $P_n^*(A) \rightarrow P(A)$ minden A -ra egy valószínűséggel.

A vektor esethez csak annyit kell hozzátenni, hogy

Állítás 22. Ha $\mathcal{C}', \mathcal{C}''$ végesen approximálható, akkor $\mathcal{C}'(\cap) \mathcal{C}'' = \{C' \cap C'' : C' \in \mathcal{C}', C'' \in \mathcal{C}''\}$ is végesen approximálható.

□

5.3.1 Glivenko-Cantelli másképp, vázlat

X_1, X_2, \dots iid változók, közös F eloszlásfüggvénnyel. $(X'_n)_{n \geq 1} \stackrel{p}{=} (X_n)_{n \geq 1}$ és $(X'_n)_{n \geq 1}$ független $(X_n)_{n \geq 1}$ -től. $F_n(t) = \frac{1}{n} \sum_{i \leq n} \mathbf{1}_{(X_i < t)}$, $F'_n(t) = \frac{1}{n} \sum_{i \leq n} \mathbf{1}_{(X'_i < t)}$

Állítás 23 (Első randomizálás). $\mathbb{P}(\sup_t |F_n(t) - F(t)| > 2\varepsilon) \leq 2\mathbb{P}(\sup_t |F_n(t) - F'_n(t)| > \varepsilon)$, ha $n > 2\varepsilon^{-2}$.

Állítás 24 (Második randomizálás).

$$\mathbb{P}(\sup |F_n - F'_n| > \varepsilon) = 2\mathbb{P}\left(\sup_t F_n - F'_n > \varepsilon\right) = 2\mathbb{P}\left(\sup_t \frac{1}{n} \sum \xi_i (\mathbf{1}_{(X_i < t)} - \mathbf{1}_{(X'_i < t)}) > \varepsilon\right)$$

ahol ξ_1, \dots, ξ_n független véletlen előjelek, egymástól és az $(X_k), (X'_k)$ változóktól is ($\mathbb{P}(\xi = \pm 1) = 1/2$).

Állítás 25 (Hoeffding egyenlőtlenség alkalmazása). $\mathbb{P}\left(\sup_t \sum \xi_i (\mathbf{1}_{(X_i < t)} - \mathbf{1}_{(X'_i < t)}) > n\varepsilon \mid X, X'\right) \leq (2n+1) \exp\{-\frac{1}{2}n\varepsilon^2\}$

Összefoglalva: Létezik $c_1, c_2 > 0$, hogy $\mathbb{P}(\sup |F_n - F| > \varepsilon) \leq c_1 n e^{-c_2 n \varepsilon^2}$ ha $n > 8\varepsilon^{-2}$.

Innen $\sup |F_n - F| \rightarrow 0$ egy valószínűséggel a Borel–Cantelli lemma alkalmazásával.

Első randomizálás részletei

X_1, X_2, \dots iid változók, közös F eloszlásfüggvénnyel. $(X'_n)_{n \geq 1} \stackrel{p}{=} (X_n)_{n \geq 1}$ és $(X'_n)_{n \geq 1}$ független $(X_n)_{n \geq 1}$ -től. $F_n(t) = \frac{1}{n} \sum_{i \leq n} \mathbf{1}_{(X_i < t)}$, $F'_n(t) = \frac{1}{n} \sum_{i \leq n} \mathbf{1}_{(X'_i < t)}$

Állítás 26 (Első randomizálás). $\mathbb{P}(\sup_t |F_n(t) - F(t)| > 2\varepsilon) \leq 2\mathbb{P}(\sup_t |F_n(t) - F'_n(t)| > \varepsilon)$ ha $n > 2\varepsilon^{-2}$.

- Legyen $T = T(X)$ olyan változó, hogy $|F_n - F|(T) > 2\varepsilon$, a $\{\sup |F_n - F| > 2\varepsilon\}$ eseményen.

$$\begin{aligned} \mathbb{P}(\sup |F_n - F| > 2\varepsilon) &= \mathbb{P}(|F_n - F|(T) > 2\varepsilon) \\ &= \mathbb{P}(|F_n - F|(T) > 2\varepsilon, |F'_n - F|(T) > \varepsilon) + \mathbb{P}(|F_n - F|(T) > 2\varepsilon, |F'_n - F|(T) \leq \varepsilon) \end{aligned}$$

•

$$\begin{aligned}\mathbb{P}(|F_n - F|(T) > 2\varepsilon, |F'_n - F|(T) > \varepsilon \mid (X_n)_{n \geq 1}) &= \mathbf{1}_{(|F_n - F|(T) > 2\varepsilon)} \mathbb{P}(|F'_n(t) - F(t)| > \varepsilon) \big|_{t=T} \\ &\leq \mathbf{1}_{(|F_n(T) - F(T)| > 2\varepsilon)} \sup_t \mathbb{P}(|F'_n(t) - F(t)| > \varepsilon)\end{aligned}$$

$$\text{és } \mathbb{P}(|F'_n(t) - F(t)| > \varepsilon) \leq \mathbb{D}^2 F'_n(t) \varepsilon^{-2} \leq \frac{1}{4} n^{-1} \varepsilon^{-2} < \frac{1}{2}, \text{ ha } n > 2\varepsilon^{-2}.$$

$$\bullet \quad \mathbb{P}(\sup |F_n - F| > 2\varepsilon) \leq \frac{1}{2} \mathbb{P}(\sup |F_n - F| > 2\varepsilon) + \mathbb{P}(\sup |F_n - F'_n| > \varepsilon).$$

Második randomizálás részletei

Állítás 27 (Második randomizálás).

$$\mathbb{P}(\sup |F_n - F'_n| > \varepsilon) = 2\mathbb{P}(\sup_t F_n - F'_n > \varepsilon) = 2\mathbb{P}\left(\sup_t \frac{1}{n} \sum \xi_i (\mathbf{1}_{(X_i < t)} - \mathbf{1}_{(X'_i < t)}) > \varepsilon\right)$$

ahol ξ_1, \dots, ξ_n független előjelek, egymástól és az $(X_k), (X'_k)$ változóktól is.

- $\sup F_n - F'_n \stackrel{p}{=} \sup F'_n - F_n \implies \mathbb{P}(F_n - F'_n > \varepsilon) = \mathbb{P}(F_n - F'_n < -\varepsilon).$
- Legyen $S = \sup_t \frac{1}{n} \sum \xi_i (\mathbf{1}_{(X_i < t)} - \mathbf{1}_{(X'_i < t)})$.

$$\mathbb{P}(S > \varepsilon \mid \xi_1, \dots, \xi_n) = \mathbb{P}(\sup F_n - F'_n > \varepsilon),$$

mert ξ_1, \dots, ξ_n hatása csak annyi, hogy ha $\xi_i = -1$, akkor X_i és X'_i szerepet cserél.

Formálisan. Ha $\sup F_n - F'_n = g(X, X')$, akkor $S = g(T(\xi, X, X'))$, ahol

$$T(\xi, X, X') = (Y, Y') \quad \text{és} \quad (Y_i, Y'_i) = \mathbf{1}_{(\xi_i=1)}(X_i, X'_i) + \mathbf{1}_{(\xi_i=-1)}(X'_i, X_i)$$

$T(x, X, X') \stackrel{p}{=} (X, X')$ minden $x \in \{-1, 1\}^n$ előjelsorozatra, vagyis

$$\mathbb{P}(S > \varepsilon \mid \xi) = \mathbb{P}(g(T(\xi, X, X')) > \varepsilon \mid \xi) = \mathbb{P}(g(T(x, X, X')) > \varepsilon) \big|_{x=\xi} = \mathbb{P}(g(X, X') > \varepsilon)$$

Hoeffding egyenlőtlenség

Állítás 28. Z_i független nulla várható értékű változók, $|Z_i| \leq 1$, $\varepsilon > 0$. Ekkor $\mathbb{P}(\sum Z_i > n\varepsilon) \leq e^{-\frac{1}{2}n\varepsilon^2}$

•

$$\mathbb{P}\left(\sum Z_i > n\varepsilon\right) \leq \inf_{t>0} \mathbb{E}\left(e^{t \sum Z_i}\right) e^{-tn\varepsilon}$$

- Függetlenség miatt $\mathbb{E}(e^{t \sum Z_i}) = \prod \mathbb{E}(e^{t Z_i})$
- $Z = Z_i$ -vel: $g(t) = \mathbb{E}(e^{tZ})$, g a Z momentum generáló függvénye. $\ln g(0) = 0$, $(\ln g)'(0) = \frac{g'(0)}{g(0)} = \mathbb{E}(Z) = 0$ és

$$(\ln g)''(s) = \frac{g''(s)}{g'(s)} - \left(\frac{g'(s)}{g(s)}\right)^2 = \mathbb{E}\left(Z^2 \frac{e^{sZ}}{g(s)}\right) - \mathbb{E}^2 Z \frac{e^{sZ}}{g(s)} = \mathbb{D}_s^2 Z, \quad \text{ahol } \frac{d\mathbb{P}_s}{d\mathbb{P}} = \frac{e^{sZ}}{g(s)}.$$

- $\mathbb{D}_s^2 Z \leq \mathbb{E}_s(Z^2) \leq 1$, így

$$\ln g(t) = \int_0^t (\ln g)'(s) ds = \int_0^t \int_0^s (\ln g)''(u) du ds \leq \frac{1}{2} t^2.$$

•

$$\mathbb{P}\left(\sum Z_i > n\varepsilon\right) \leq \inf_{t>0} e^{\frac{n}{2}(t^2 - 2t\varepsilon)} = \inf_{t>0} e^{\frac{n}{2}((t-\varepsilon)^2 - \varepsilon^2)} = e^{-\frac{1}{2}n\varepsilon^2}.$$

6. 2021.03.22.

6.1 Maximum likelihood becslés

\mathcal{P} dominált mértékcsalád, f_θ a sűrűségfüggvény, $\ell(\theta)$ a loglikelihood.

Definíció 22. A θ paraméter maximum likelihood becslése $\hat{\theta}(X)$, ha $f_{\hat{\theta}}(X) = \sup_{\theta} f_{\theta}(X)$.

Nem biztos, hogy létezik, nem biztos, hogy egyértelmű

- $U(\theta, \theta+1)$ eloszlásból származó minta, mintaelemszám legalább kettő. $f_\theta(x) = \mathbf{1}_{(x_n^* - 1 \leq \theta \leq x_1^*)}$, azaz ha $T(X) \in [X_n^* - 1, X_1^*]$, akkor T maximum likelihood becslése θ -nak. Pl. $T_\alpha = \alpha X_1^* + (1 - \alpha)X_n^*$, $\alpha \in [0, 1]$.

Ebben az esetben a maximum likelihood becslés nem egyértelmű.

- $f(x)$ sűrűségfüggvény \mathbb{R} -en.

$$f_{\mu, \sigma}(x) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right)$$

$\Theta = \mathbb{R} \times (0, \infty)$. $c \in \mathbb{R}$, $f(c) > 0$, x a megfigyelt érték

$$\sup_{\mu, \sigma} \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}\right) \geq \lim_{\sigma \rightarrow 0} \frac{1}{\sigma} f\left(\frac{x - (x + c\sigma)}{\sigma}\right) = \infty$$

A maximum likelihood becslés nem létezik.

6.1.1 Maximum likelihood becslés keverék eloszlásra

- X_1, \dots, X_n n elemű minta. A közös sűrűségfüggvény:

$$f(x) = \frac{1}{2} f_{\mu, \sigma}(x) + \frac{1}{2} f_{0,1}(x)$$

ahol $f_{\mu, \sigma}$ az $N(\mu, \sigma)$ eloszlás sűrűségfüggvénye.

- A loglikelihood függvény:

$$\ell(\mu, \sigma) \geq \log \frac{1}{2} f_{\mu, \sigma}(x_1) + \sum_{i=2}^n \log \frac{1}{2} f_{0,1}(x_i) \rightarrow \infty, \quad \text{ha } \mu = x_1, \sigma \rightarrow 0$$

- $\mu \in \{x_1, \dots, x_n\}$, $\sigma \rightarrow 0$ mellett a limesz ∞ .
- Nem létezik maximum likelihood becslés.
- Ez nem jelenti azt, hogy nem érdemes a loglikelihood függvény lokális maximum helyét megkeresni. EM algoritmus.

6.1.2 Maximum likelihood becslés a paraméter egy függvényére

$\psi = g(\theta)$ becslése maximum likelihood elvvel. Ha g kölcsönösen egyértelmű, akkor csak átparamétereztük a családot, $\hat{\psi} = g(\hat{\theta})$.

Ha g nem injektív, akkor az indukált likelihoodot szokták maximalizálni

$$f_\psi^*(x) = \sup\{f_\theta(x) : g(\theta) = \psi\}$$

Állítás 29. .

- Az ML becslés nem függ a domináló mérték megválasztásától.
 - Ha T elégséges statisztika és létezik ML becslés, akkor van olyan is, amelyik T függvénye.
 - Ha $\hat{\theta}$ a θ ML becslése, akkor $g(\hat{\theta})$ a $g(\theta)$ ML becslése.
 - Ha a maximum hely egyértelmű, akkor a faktorizációs tétel miatt $f_{\theta}(x) = g_{\theta}(T(x))h(x)$ csak $T(x)$ -en keresztül függ a megfigyeléstől. Ha nem egyértelmű, akkor arra kell figyelni, hogy $T(x)$ függvényében válasszunk.
- pl. $U(\theta, \theta + 1)$ esetén elégséges statisztika $T(X) = (X_1^*, X_n^*)$. $\hat{\theta} = (X_n^* - 1 + X_1^*)/2$ ML becslés T függvényei között,

$$\hat{\theta} + \sin(X_1) \frac{X_1^* - (X_n^* - 1)}{2}$$

pedig nem (feltéve, hogy a minta elemszáma legalább 2).

6.1.3 Maximum likelihood becslés exponenciális családban

- X_1, \dots, X_n n elemű minta, Θ nyílt és

$$f_{\theta}(x) = \exp\{\theta T_n(x) - b_n(\theta)\}$$

pl. Binomiális, Poisson, exponenciális, normális stb. minta.

- A loglikelihood: $\ell(\theta, x) = \theta T_n(x) - b_n(\theta)$.
- $b_n''(\theta) = \Sigma_{\theta}(T_n) \geq 0$. Láttuk, hogy $\Sigma(T_n)$ pozitív definit, azaz $\theta \mapsto \ell$ szigorúan konkáv. Ha $\ell'(\theta) = 0$, akkor θ maximum hely, azaz a maximum likelihood becslés.
- Megmutatható, hogy $T_n(n) = \sum_{i=1}^n T(x_i)$ alakú és $b_n(\theta) = nb_1(\theta)$.
- A likelihood egyenlet:

$$\ell'(\theta) = 0 \quad \Longleftrightarrow \quad \frac{1}{n} \sum_{i=1}^n T_1(x_i) = b_1'(\theta).$$

Itt $b_1'(\theta) = \mathbb{E}_{\theta}(T_1(X))$, azaz a maximum likelihood becslés „általánosított momentum módszeres” becslésnek tekinthető: azt a paramétert választjuk, amelyik $\mathbb{E}_{\theta}(T_1(X)) = \frac{1}{n} \sum_i T_1(X_i)$.

Maximum likelihood becslés aszimptotikus viselkedése exponenciális eloszlás családban

$\mathcal{P} = \{\mathbb{P}_{\theta} : \theta \in \Theta\}$, exponenciális család, Θ konvex, nyílt. f_{θ} jelöli egy mintaelem sűrűségfüggvényét.

Tétel 20. X_1, X_2, \dots független azonos eloszlású megfigyelések f_{θ_0} ($\theta_0 \in \Theta$) közös sűrűségfüggvényű eloszlásból. $\hat{\theta}_n$ a θ ML becslése az első n megfigyelés alapján. Ekkor $\hat{\theta}_n \rightarrow \theta_0$ és $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, I(\theta_0)^{-1})$.

Az ML becslés (szép eloszláscsaládban) konzisztens és aszimptotikusan normális.

- Egy mintaelem sűrűségfüggvénye:

$$f_\theta(x) = \exp\{\theta T(x) - b(\theta)\}, \quad \theta \in \Theta \subset \mathbb{R}^p, \quad \Theta \text{ nyílt.}$$

- Θ nyílt, a paraméterezés egyértelmű, azaz T értékkészlete nem lehet egy valódi affin altér része. Így $b''(\theta) = \Sigma_\theta(T(X_1))$ pozitív definit és b szigorúan konvex.
- $\hat{\theta}_n$, θ maximum likelihood becslése X_1, \dots, X_n alapján. Feltesszük, hogy létezik maximum hely, és mivel $-\ell(\theta)$ szigorúan konvex, a maximum hely a likelihood egyenlet egyetlen megoldása.

$$\hat{\theta}_n = \arg \max_{\theta} \frac{1}{n} \ell_n(\theta), \quad \Longleftrightarrow \quad \frac{1}{n} \ell'_n(\hat{\theta}_n) = 0 \quad \Longleftrightarrow \quad \frac{1}{n} \sum_{i=1}^n T(X_i) = b'(\hat{\theta}_n)$$

Maximum likelihood becslés konzisztenciája exponenciális eloszlás családban

$\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$, exponenciális család, Θ konvex, nyílt. f_θ jelöli **egy mintaelem** sűrűségfüggvényét.

Tétel 21. X_1, X_2, \dots független azonos eloszlású megfigyelések f_{θ_0} ($\theta_0 \in \Theta$) közös sűrűségfüggvényű eloszlásból. $\hat{\theta}_n$ a θ ML becslése az első n megfigyelés alapján. Ekkor $\hat{\theta}_n \rightarrow \theta_0$.

- Ha $n \rightarrow \infty$, akkor $\frac{1}{n} \sum_{i=1}^n T(X_i) \rightarrow \mathbb{E}_{\theta_0}(T(X_1)) = b'(\theta_0)$ a nagy számok törvénye szerint.
 - b'' folytonos, pozitív definit mátrix értékű $\implies b'$ invertálható és az inverz folytonos.
- Valóban, ha $\theta \in \Theta$, $e \in \mathbb{R}^p$ egységvektor:

$$e^T \cdot (b'(\theta + te) - b'(\theta)) = \int_0^t e^T b''(\theta + ue) e du \geq c \cdot (t \wedge \delta)$$

ahol $b''(\theta) \geq 2c\mathbf{I}_p$ és δ olyan kicsi, hogy $b''(\theta_1) \geq c\mathbf{I}_p$, ha $|\theta_1 - \theta| < \delta$

- b' injektív
- b' inverz folytonos: ha $b'(\theta_n) \rightarrow b'(\theta)$ akkor $|\theta_n - \theta| > \eta$ végtelen sok n -re nem teljesülhet és $\theta_n \rightarrow \theta$.
- $\hat{\theta}_n = (b')^{-1}(\frac{1}{n} \sum_{i=1}^n T(X_i)) \rightarrow (b')^{-1}(b'(\theta_0)) = \theta_0$.

6.1.4 Többdimenziós normális eloszlás

- X egy dimenziós normális változó, ha $X \stackrel{p}{=} \sigma Z + \mu$, ahol $Z \sim N(0, 1)$ és $\sigma \geq 0$, $\mu \in \mathbb{R}$.
- Elfajult eloszlás ($\sigma = 0$) is normális.
- Karakterisztikus függvény

$$\phi_Z(t) = \mathbb{E}(e^{itZ}) = e^{-\frac{1}{2}t^2}, \quad \phi_{\sigma Z + \mu}(t) = \mathbb{E}(e^{it(\sigma Z + \mu)}) = e^{it\mu} \phi_Z(t\sigma) = e^{it\mu - \frac{1}{2}t^2\sigma^2}.$$

- X d -dimenziós vektorváltozó karakterisztikus függvénye

$$\phi_X : \mathbb{R}^d \rightarrow \mathbb{C}, \quad \phi_X(t) = \mathbb{E}(\exp\{i \sum_j t_j X_j\}) = \mathbb{E}(e^{it \cdot X})$$

Vektorváltozóra is van inverziós formula, azaz X karakterisztikus függvénye X eloszlását meghatározza.

- Következmény. X eloszlását a lineáris kombinációk $\sum_j a_j X_j$ ($a \in \mathbb{R}^d$) eloszlásai meghatározzák.

Definíció 23. X d -dimenziós normális vektorváltozó, ha minden $a \in \mathbb{R}^d$ -re $\sum_j a_j X_j$ egy dimenziós normális.

- Példa. Ha Z_1, \dots, Z_d független standard normálisok, akkor $Z = (Z_1, \dots, Z_d)$. Ekkor Z normális vektorváltozó.

$$\phi_Z(t) = \mathbb{E}(e^{it \cdot Z}) = \mathbb{E}\left(\prod e^{it_j Z_j}\right) = \prod e^{-\frac{1}{2}t_j^2} = e^{-\frac{1}{2}\|t\|^2}$$

Ha $a \in \mathbb{R}^d$, akkor

$$\phi_{a \cdot Z}(t) = \mathbb{E}(e^{it a \cdot Z}) = \phi_Z(ta) = e^{-\frac{1}{2}t^2 \|a\|^2}$$

Azaz Z d dimenziós normális. Z -t **d -dimenziós standard normális** változónak hívjuk.

- Ha Z d dimenziós standard normális változó és $A \in \mathbb{R}^{k \times d}$, $m \in \mathbb{R}^k$, akkor $X = AZ + m$ k -dimenziós normális, hiszen $a \in \mathbb{R}^k$ -ra

$$a \cdot X = (A^T a) \cdot Z + a \cdot m \quad \text{egy dimenziós normális}$$

- Z d -dimenziós normális $X = AZ + m$ várható érték vektora m , kovariancia mátrixa AA^T .
- Tetszőleges pozitív szemidefinit Σ mátrix faktorizálható AA^T alakban, sőt A alsó háromszög mátrixnak is választható (Cholesky felbontás).
- Ha $X = (X_1, \dots, X_d)$ d -dimenziós normális vektor, akkor minden j -re X_j egy dimenziós normális, vagyis $X_j \in L^2$.
- Jelölje m , Σ az X várható érték vektorát, és kovariancia mátrixát. Ekkor $t \in \mathbb{R}^d$ esetén $Y = \sum_j t_j X_j \sim N(t \cdot m, t^T \Sigma t)$

$$\phi_X(t) = \mathbb{E}(e^{it \cdot X}) = \phi_{t \cdot X}(1) = e^{it \cdot m - \frac{1}{2}t^T \Sigma t}.$$

- Többdimenziós normális eloszlást a várható érték vektor és a kovariancia mátrix meghatározza.

Jelölés $N(\mu, \Sigma)$.

Emlékeztető.

Tétel 22 (Centrális határeloszlás tétel). ξ_1, ξ_2, \dots iid vektor változók, $\mathbb{E}(\xi_1) = 0$, $\Sigma(\xi_1) = \Sigma$. Ekkor $\eta_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n \xi_k \xrightarrow{d} N(0, \Sigma)$.

Ellenőrzés karakterisztikus függvénnyel. $\phi_{\eta_n}(\alpha) = \phi_{\alpha \cdot \eta_n}(1)$. $\alpha \cdot \eta_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n \alpha \cdot \xi_k \xrightarrow{d} N(0, \alpha^T \Sigma \alpha)$ és így $\phi_{\eta_n}(\alpha) \rightarrow e^{-\frac{1}{2}\alpha^T \Sigma \alpha}$, ami az $N(0, \Sigma)$ eloszlás karakterisztikus függvénye.

Példa (HF). Z standard normális, $\xi \perp Z$ és $\mathbb{P}(\xi = \pm 1) = \frac{1}{2}$. Legyen $Z_1 = Z$, $Z_2 = \xi Z$. Ekkor Z_1, Z_2 standard normálisok, de a (Z_1, Z_2) vektor változó eloszlása nem normális.

Maximum likelihood becslés aszimptotikus normalitása exponenciális eloszlás családban

$\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$, exponenciális család, Θ konvex, nyílt. f_θ jelöli **egy mintaelem** sűrűségfüggvényét.

Tétel 23. X_1, X_2, \dots független azonos eloszlású megfigyelések f_{θ_0} ($\theta_0 \in \Theta$) közös sűrűségfüggvényű eloszlásból. $\hat{\theta}_n$ a θ ML becslése az első n megfigyelés alapján. Ekkor $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, I(\theta_0)^{-1})$.

- $\hat{\theta}_n$ a likelihood egyenlet egyetlen megoldása: $\ell'_n(\hat{\theta}_n) = 0$.
- Legyen $h(t) = \frac{1}{n}(\ell'_n(t + \theta_0) - \ell'_n(\theta_0))$. Ekkor h determinisztikus $h(t) = b'(\theta_0) - b'(\theta_0 + t)$, $h'(0) = -I(\theta_0)$ nem szinguláris és $h(\hat{\theta}_n - \theta_0) = -\frac{1}{n}\ell'_n(\theta_0)$.
- A $\xi_n = \hat{\theta}_n - \theta_0$ jelöléssel $\xi_n \xrightarrow{p} 0$ a konzisztencia miatt és $\sqrt{n}h(\xi_n) \xrightarrow{d} N(0, I(\theta_0))$, a CHT szerint, hiszen

$$\sqrt{n}h(\xi_n) = -\sqrt{n}\frac{1}{n}\ell'_n(\theta_0) = n^{-1/2} \sum_{i=1}^n -\ell'_1(\theta_0, X_i)$$

Állítás 30. $\xi_n \xrightarrow{p} 0$, $\sqrt{n}h(\xi_n) \xrightarrow{d} N(0, \Sigma)$, $h(0) = 0$, létezik $h'(0)$ és $(h'(0))^{-1}$. Ekkor $h'(0)\sqrt{n}\xi_n \xrightarrow{d} N(0, \Sigma)$.

$$\sqrt{n}I(\theta_0)(\hat{\theta}_n - \theta_0) \xrightarrow{d} \xi \sim N(0, I(\theta_0)) \implies \sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} I^{-1}(\theta_0)\xi \sim N(0, I^{-1}(\theta_0)).$$

6.1.5 Lemma ellenőrzése

Emlékeztető. Ha $\zeta_n \xrightarrow{d} Z$ és $Y_n \xrightarrow{p} 0$, akkor $\zeta_n + Y_n \xrightarrow{d} Z$. Ok: elegendő egy korlátú Lipschitz-1 folytonos h -ra ellenőrizni, hogy $\mathbb{E}(h(\zeta_n + Y_n)) \rightarrow \mathbb{E}(h(Z))$ és $|\mathbb{E}(h(\zeta_n + Y_n) - h(\zeta_n))| \leq \mathbb{E}(1 \wedge |Y_n|) \rightarrow 0$.

Állítás 31. $\xi_n \xrightarrow{p} 0$, $\sqrt{n}h(\xi_n) \xrightarrow{d} N(0, \Sigma)$, $h(0) = 0$, létezik $h'(0)$ és $(h'(0))^{-1}$. Ekkor $h'(0)\sqrt{n}\xi_n \xrightarrow{d} N(0, \Sigma)$.

Elég, hogy $\sqrt{n}(h(\xi_n) - h'(0)\xi_n) \xrightarrow{p} 0$.

- h deriválható a 0-ban, és $h'(0)$ invertálható mátrix: $\varepsilon > 0$ -hoz létezik $\delta = \delta(\varepsilon) > 0$ úgy, hogy $|h(t) - h'(0)t| \leq \varepsilon|t| \leq \varepsilon\|(h'(0))^{-1}\||h'(0)t|$ ha $|t| < \delta$. Ebből

$$(1 - \varepsilon')|h'(0)t| \leq |h(t)| \leq (1 + \varepsilon')|h'(0)t|, \quad \text{ha } |t| < \delta, \text{ ahol } \varepsilon' = \varepsilon\|(h'(0))^{-1}\|$$

•

$$|h(t) - h'(0)t| < \varepsilon'|h'(0)t| \leq \frac{\varepsilon'}{1 - \varepsilon'}|h(t)| \quad \text{ha } \varepsilon' = \|(h'(0))^{-1}\|\varepsilon < 1 \text{ és } |t| < \delta(\varepsilon).$$

- $\eta > 0$

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P}(\sqrt{n}|h(\xi_n) - h'(0)\xi_n| > \eta) &\leq \inf_{\varepsilon > 0, \varepsilon' < 1} \limsup_{n \rightarrow \infty} \mathbb{P}(|\xi_n| > \delta(\varepsilon)) + \mathbb{P}\left(\sqrt{n}\frac{\varepsilon'}{1 - \varepsilon'}|h(\xi_n)| > \eta\right) \\ &= \inf_{\varepsilon > 0, \varepsilon' < 1} \mathbb{P}\left(|\xi| > \frac{\eta(1 - \varepsilon')}{\varepsilon'}\right) = 0, \quad \text{ahol } \xi \sim N(0, \Sigma). \implies \sqrt{n}(h(\xi_n) - h'(0)\xi_n) \xrightarrow{p} 0 \end{aligned}$$

6.1.6 Példa, normális eloszlás

- $N(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$. Exponenciális eloszláscsalád, de nem ez a természetes paraméterezés.
- $\theta_1 = \frac{1}{\sigma^2}$, $\theta_2 = \frac{\mu}{\sigma^2}$, $(\mu, \sigma) = h(\theta)$.
- $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, I^{-1}(\theta_0))$.
- $\sqrt{n}(h(\hat{\theta}_n) - h(\theta_0)) \xrightarrow{d} N(0, h'(\theta_0)I^{-1}(\theta_0)h'(\theta_0)^T)$ a lemma alapján.
- $I(\theta) = \Sigma_\theta((\ell \circ h)'(\theta_0))^T = h'(\theta)^T I(\mu, \sigma) h'(\theta)$, amiből

$$I^{-1}(\theta) = (h'(\theta)^{-1}I^{-1}(\mu, \sigma)(h'(\theta)^T)^{-1} \implies I^{-1}(\mu, \sigma) = h'(\theta_0)I^{-1}(\theta_0)h'(\theta_0)^T$$

- $\sqrt{n}((\hat{\mu}_n, \hat{\sigma}_n) - (\mu_0, \sigma_0)) \xrightarrow{d} N(0, I^{-1}(\mu, \sigma))$. Az ML becslés aszimptotikusan normális.
 - $I(\mu, \sigma) = \frac{1}{\sigma^2} \text{diag}(1, 2)$ és $\hat{\mu}_n = \bar{X}$, $\hat{\sigma}_n = s_n$.
- Később kiszámoljuk, hogy normális minta esetén \bar{X} és s_n^2 függetlenek, nem csak aszimptotikusan korrelálatlanok továbbá $\sqrt{n}(\hat{\mu}_n - \mu) \sim N(0, \sigma^2)$ és $\hat{\sigma}_n^2 \sim \frac{\sigma^2}{n} \chi_{n-1}^2$

6.1.7 Példa. $\sqrt{n}(\hat{\theta}_n - \theta_0)$ aszimptotikus kovariancia mátrixa lehet szinguláris

- X_1, X_2, \dots független $U(0, 1)$ eloszlású változók. Ha $U(0, \theta)$, $\theta > 0$ eloszláscsaládban becsüljük θ -t, akkor $\hat{\theta}_n = X_n^{(n)}$. Az első n megfigyelésből számított rendezett mintát $X_1^{(n)}, \dots, X_n^{(n)}$ -nel jelöljük.
- $\theta_0 = 1$, $X_n^{(n)} - 1 \stackrel{p}{=} -X_1^{(n)}$, hiszen $X_n^{(n)} - 1 = \max_{k \leq n}(X_k - 1) = -\min_{k \leq n}(1 - X_k)$ és $1 - X_1, 1 - X_2, \dots$ független $U(0, 1)$ eloszlású változók.
- $\sqrt{n}(X_n^{(n)} - 1) \stackrel{p}{=} -\sqrt{n}X_1^{(n)}$

$$\mathbb{P}(\sqrt{n}X_1^{(n)} > \varepsilon) = \left(1 - \frac{\varepsilon}{\sqrt{n}}\right)^n \rightarrow 0 \quad \text{minden } \varepsilon > 0\text{-ra} \implies \sqrt{n}(X_n^{(n)} - 1) \stackrel{p}{\rightarrow} 0.$$

- Ha \sqrt{n} helyett n -nel normálunk, akkor

$$\mathbb{P}(nX_1^{(n)} > t) = \left(1 - \frac{t}{n}\right)^n \rightarrow e^{-t}, \quad \text{azaz } n(X_n^{(n)} - 1) \xrightarrow{d} -\exp(1)$$

a határeloszlás egységnyi paraméterű exponenciális ellentettje.

7. 2021.03.29.

7.1 Általánosítás

7.1.1 konzisztencia

X_1, X_2, \dots független azonos eloszlású megfigyelések f_{θ_0} közös sűrűségfüggvénnyel. $\Theta \subset \mathbb{R}^p$ nyílt. $f_{\theta}, \theta \in \Theta$ teljesíti a gyenge regularitási feltételt.

Alkalmas feltételek mellett a következő lépések igazak.

- Elég kicsi $\delta > 0$ -ra

$$S_n = \sup_{|t| \leq \delta} \frac{1}{n}(\ell_n(\theta_0 + t) - \ell_n(\theta_0)) + \frac{1}{4}t^T I(\theta_0)t, \quad \text{jelöléssel} \quad \lim_{n \rightarrow \infty} S_n \leq 0 \quad \text{egy valószínűséggel.}$$

- $I(\theta_0) > 0$, ezért $\frac{1}{4}t^T I(\theta_0)t \geq c|t|^2$ alkalmas $c > 0$ -val.
- Jelölje $\hat{\theta}_n$ a loglikelihood ℓ_n maximum helyét a $\{\theta : |\theta - \theta_0| \leq \delta\}$ halmazon.

$$0 \leq \frac{1}{n}(\ell_n(\hat{\theta}_n) - \ell_n(\theta_0)) \leq S_n - c|\hat{\theta}_n - \theta_0|^2 \quad \text{amiből} \quad |\hat{\theta}_n - \theta_0|^2 \leq S_n/c \rightarrow 0.$$

- Elég nagy n -től kezdve $|\hat{\theta}_n - \theta_0| < \delta$ és $\ell'_n(\hat{\theta}_n) = 0$.

Ez adja az ML becslés konzisztenciáját.

7.1.2 aszimptotikus normalitás

X_1, X_2, \dots független azonos eloszlású megfigyelések f_{θ_0} közös sűrűségfüggvénnyel. $\Theta \subset \mathbb{R}^p$ nyílt. $f_{\theta}, \theta \in \Theta$ teljesíti a gyenge regularitási feltételt.

Alkalmas feltételek mellett a következő lépések igazak.

- CHT alapján:

$$n^{-1/2}\ell'_n(\theta_0) = n^{-1/2} \sum_{i=1}^n \ell'(\theta_0, X_i) \xrightarrow{d} N(0, I(\theta_0))$$

- $\hat{\theta}_n = \arg \max_{|\theta - \theta_0| \leq \delta} \ell_n(\theta)$. Ekkor $n^{-1/2}\ell'_n(\hat{\theta}_n) \rightarrow 0$

$$n^{-1/2}(\ell'(\hat{\theta}_n) - \ell'(\theta_0)) = \left(\frac{1}{n}\ell''_n(\theta_0) + R_n \right) \sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, I(\theta_0))$$

Itt $R_n \xrightarrow{p} 0$ és $\frac{1}{n}\ell''_n(\theta_0) \rightarrow \mathbb{E}_{\theta_0}(\ell''(\theta_0, X_1)) = -I(\theta_0)$.

- $I(\theta_0)$ invertálható. Létezik K_n sorozat, hogy $K_n \left(\frac{1}{n}\ell''_n(\theta_0) + R_n \right) = \mathbf{I}_p$ ha n elég nagy. Ebből $K_n \rightarrow I^{-1}(\theta_0)$ és

$$\sqrt{n}(\hat{\theta}_n - \theta_0) - K_n n^{-1/2}(\ell'(\hat{\theta}_n) - \ell'(\theta_0)) \rightarrow 0 \implies \sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, I^{-1}(\theta_0))$$

hiszen $K_n n^{-1/2}(\ell'(\hat{\theta}_n) - \ell'(\theta_0)) \xrightarrow{d} N(0, I^{-1}(\theta_0))$.

7.2 Erős regularitási feltétel

Definíció 24 ((RR)). $(R) + \theta \mapsto \ell(\theta, x) \in C^2$ és \mathcal{P} m.m. x -re és majorálható „egyenletesen integrálható” M -mel. Azaz

- $\|\partial_\theta^2 \ell(\theta, x)\| \leq M(x)$,
- $\sup_\theta \mathbb{E}_\theta(M^2) < \infty$.

Tétel 24. X_1, X_2, \dots iid sorozat f_{θ_0} sűrűségfüggvényből. (RR) teljesül a $\{f_\theta : \theta \in \Theta\}$ eloszlás-családra.

$\ell_n(\theta) = \sum_{i=1}^n \ell_1(\theta, X_i)$ az n elemű minta loglikelihoodja.

- Az $\ell'_n(\theta) = 0$ likelihood egyenletnek létezik olyan $\hat{\theta}_n$ gyöke, amely lokális maximum és erősen konzisztens
- Erre a gyökre $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, I^{-1}(\theta_0))$

(RR)-t elég lokálisan megkövetelni.

7.2.1 Erős regularitási feltétel következményei

Definíció 25 (RR). $(R) + \theta \mapsto \ell(\theta, x) \in C^2$ \mathcal{P} m.m. x -re és majorálható „egyenletesen integrálható” M -mel. Azaz

- $\|\partial_\theta^2 \ell(\theta, x)\| \leq M(x)$,
- $\sup_\theta \mathbb{E}_\theta(M^2) < \infty$.

Lemma 6. (RR) esetén

$$\eta(\varepsilon) = \mathbb{E}_\theta \left(\sup_{|t| < \varepsilon} \|\ell''(\theta + t) - \ell''(\theta)\| \right) \rightarrow 0 \quad \text{ha } \varepsilon \rightarrow 0$$

Dominált konvergencia tétel:

$$\sup_{|t| \leq \varepsilon} \|\ell''(\theta + t, x) - \ell''(\theta, x)\| \leq 2M(x)$$

és $t \mapsto \ell''(t, x)$ folytonos.

Lemma 7. (RR) esetén $\theta \mapsto \mathbb{E}_\theta(\ell''(\theta))$ folytonos.

- $\theta_n \rightarrow \theta$

$$\mathbb{E}_{\theta_n}(\ell''(\theta_n)) - \mathbb{E}_\theta(\ell''(\theta)) = \mathbb{E}_\theta(\ell''(\theta_n) - \ell''(\theta)) + \int \ell''(\theta_n, x)(f_{\theta_n}(x) - f_\theta(x))dx$$

- $\ell''(\theta_n) \rightarrow \ell''(\theta)$ \mathcal{P} m.m. és $2M$ integrálható majoráns $\implies \lim_{n \rightarrow \infty} \mathbb{E}_\theta(\ell'(\theta_n) - \ell'(\theta)) = 0$
- Levágás + Scheffé tétel

$$\begin{aligned} \left\| \int \ell''(\theta_n, x)(f_{\theta_n}(x) - f_\theta(x))dx \right\| &\leq \int \|\ell''(\theta_n, x)\| |f_{\theta_n}(x) - f_\theta(x)| dx \leq \int M(x) |f_{\theta_n}(x) - f_\theta(x)| dx \\ &\leq K \int |f_{\theta_n}(x) - f_\theta(x)| dx + \mathbb{E}_{\theta_n}(M \mathbf{1}_{(M > K)}) + \mathbb{E}_\theta(M \mathbf{1}_{(M > K)}) \end{aligned}$$

Első tag nullához tart minden rögzített K -ra a Scheffé tétel miatt. $\sup_\theta \mathbb{E}_\theta(M \mathbf{1}_{(M > K)}) \leq \frac{1}{K} \sup_\theta \mathbb{E}_\theta(M^2) \rightarrow 0$ ha $K \rightarrow \infty$. (RR) miatt.

Következmény 3. (RR) esetén $\mathbb{E}_\theta(\ell''(\theta)) = -I(\theta)$.

- (R) miatt $\int \partial_\theta f_\theta(x) dx = 0$ minden $\theta \in \Theta$ -ra.
- $e \in \mathbb{R}^p$ egység vektor

$$0 = \int \partial_\theta f_{\theta_1+te}(x) - \partial_\theta f_{\theta_1}(x) dx = \int_0^1 \int \partial_\theta^2 f_{\theta_1+ute}(x) dx du te$$

•

$$\int \partial_\theta^2 f_\theta(x) dx = \mathbb{E}_\theta(\ell''(\theta)) + I(\theta), \quad \text{hiszen} \quad (\ln f)'' = \frac{f''}{f} - \frac{(f')^T f'}{f^2}$$

- $t \rightarrow 0$ mellett, minden $e \in \mathbb{R}^p$ egységvektorra

$$0 = \int_0^1 \mathbb{E}_{\theta+ute}(\ell''(\theta+ute)) + I(\theta+ute) du e \rightarrow (\mathbb{E}_\theta(\ell''(\theta)) + I(\theta))e \implies \mathbb{E}_\theta(\ell''(\theta)) + I(\theta) = 0.$$

7.3 Maximum likelihood becslés aszimptotikus tulajdonságai

Tétel 25. (RR) esetén a likelihood egyenletnek létezik olyan $\hat{\theta}_n$ gyöke, amely lokális maximum és erősen konzisztens

- $L_n(\theta) = \frac{1}{n} \ell_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ell(\theta, X_i)$.

$$L_n(\theta+t) - L_n(\theta) = L'_n(\theta)t + t^T \left(\frac{1}{2} L''_n(\theta) + R_n(t) \right) t$$

- Hibatag:

$$R_n(t) = \int_0^1 \int_0^v L''_n(\theta+ut) - L''_n(\theta) du dv$$

és

$$\sup_{t: \|t\| \leq \delta} \|R_n(t)\| \leq \frac{1}{2} \sup_{\|t\| \leq \delta} \|L''_n(\theta+t) - L''_n(\theta)\| \leq \frac{1}{2} \frac{1}{n} \sum_{i=1}^n \sup_{\|t\| \leq \delta} \|\partial^2 \ell(t+\theta, X_i) - \partial^2 \ell(\theta, X_i)\|$$

$$\rightarrow \frac{1}{2} \mathbb{E}_\theta \left(\sup_{\|t\| \leq \delta} \|\partial_\theta^2 \ell(t+\theta, X_1) - \partial_\theta^2 \ell(\theta, X_1)\| \right) = \frac{1}{2} \eta(\delta), \quad \text{ahol } (RR) \text{ miatt } \lim_{\delta \rightarrow 0} \eta(\delta) = 0.$$

$$L_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ell(\theta, X_i).$$

$$\sup_{|t| \leq \delta} L_n(\theta+t) - L(\theta) + \frac{1}{4} t^T I(\theta) t = \sup_{|t| \leq \delta} L'_n(\theta)t + t^T \left(\frac{1}{2} L''_n(\theta) + \frac{1}{4} I(\theta) + R_n(t) \right) t$$

$$\frac{1}{2} L''_n(\theta) + \frac{1}{4} I(\theta) \rightarrow \frac{1}{2} \mathbb{E}_\theta(\ell''(\theta)) + \frac{1}{4} I(\theta) = -\frac{1}{4} I(\theta)$$

Ha δ olyan kicsi, hogy $\frac{1}{2} \eta(\delta) \mathbf{I}_p < \frac{1}{4} I(\theta)$, akkor

$$\lim_{n \rightarrow \infty} \sup_{|t| \leq \delta} t^T \left(\frac{1}{2} L''_n(\theta) + \frac{1}{4} I(\theta) + R_n(t) \right) t \leq 0$$

míg

$$L'_n(\theta)t \rightarrow \mathbb{E}_\theta(\ell'(\theta))t = 0$$

Tétel 26. (RR) esetén

- A likelihood egyenletnek létezik olyan $\hat{\theta}_n$ gyöke, amely lokális maximum és erősen konzisztens
- Erre a gyökre $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, I^{-1}(\theta))$ (\mathbb{P}_θ alatt), vagyis $\hat{\theta}_n$ aszimptotikusan normális.

Legyen $\xi_n = \hat{\theta}_n - \theta$. Ekkor $\xi_n \rightarrow 0$ egy valószínűséggel. Mivel $\hat{\theta}_n$ a likelihood egyenlet megoldása

$$0 = L'_n(\hat{\theta}_n) = L'_n(\theta) + (L''_n(\theta) + R_n)\xi_n$$

amiből

$$-(L''_n(\theta) + R_n)\sqrt{n}\xi_n = \sqrt{n}L'_n(\theta)$$

ahol nagy n -re $\sqrt{n}L'_n(\theta)$ eloszlása közel $N(0, I(\theta))$ a CHT miatt, míg $-(L''_n(\theta) + R_n)$ nagy valószínűséggel $I(\theta)$ -hoz van közel a nagy számok erős törvénye miatt.

Ebből $\sqrt{n}\xi_n$ eloszlása közel $N(0, I^{-1}(\theta))$, ha n nagy.

•

$$R_n = \int_0^1 L''_n(\theta + u\xi_n) - L''_n(\theta) du \quad \text{amiből} \quad \limsup \|R_n\| \leq \inf_{\delta > 0} \lim_{n \rightarrow \infty} \sup_{|t| \leq |\delta|} \|L''_n(\theta + t) - L''_n(\theta)\| = \inf_{\delta > 0} \eta(\delta) = 0$$

•

$$K_n = \begin{cases} -(L''_n(\theta) + R_n)^{-1} & \text{ha } (L''_n(\theta) + R_n) \text{ invertálható} \\ \text{egységmátrix} & \text{különben} \end{cases}$$

- $I(\theta)$ invertálható

$$-(L''_n(\theta) + R_n) \rightarrow I(\theta) \implies \sqrt{n}\xi_n - K_n\sqrt{n}L'_n(\theta) \rightarrow 0 \quad \text{és} \quad K_n \rightarrow I(\theta)^{-1}$$

- Azaz $\sqrt{n}\xi_n$ és $K_n\sqrt{n}L'_n(\theta)$ eloszlásbeli limesze egyszerre létezik és megegyezik

$$\sqrt{n}L'_n(\theta) \xrightarrow{d} N(0, I(\theta)), \implies K_n\sqrt{n}L'_n(\theta) \xrightarrow{d} N(0, I^{-1}(\theta)I(\theta)I^{-1}(\theta)) = N(0, I^{-1}(\theta))$$

7.3.1 Példa, keverék eloszlás

- Egy mintaelem sűrűségfüggvénye $f_{\bar{\vartheta}}(x) = \sum_k p_k g_{\theta_k}(x)$, ahol $\bar{\vartheta} = (p, (\theta_k)_{k=1}^r)$ és $p \in \Delta_r = \{p \in (0, 1)^r : \sum p_i = 1\}$, g_θ , $\theta \in \Theta$ exponenciális családot alkot.
- Megmutatható, hogy $f_{\bar{\vartheta}}$, $\bar{\vartheta} \in \bar{\vartheta} = \Delta \times \Theta^k$ teljesíti az erős regularitási feltételt, minden olyan $\bar{\vartheta} = (p, (\theta_k)_{k=1}^r)$ pont körül, amire $\theta_1, \dots, \theta_r$ mind különbözőek.
- A likelihood egyenletnek létezik olyan megoldása, ami konzisztens becsléssorozatot ad.
- A loglikelihood függvény deriváltjai:

$$\begin{aligned} \partial_{p_k} \ln f_{\bar{\vartheta}}(x) &= \sum_{i=1}^n q_{k,i}(\bar{\vartheta}) \frac{1}{p_k} \\ \partial_{\theta_k} \ln f_{\bar{\vartheta}}(x) &= \sum_{i=1}^n q_{k,i}(\bar{\vartheta}) \partial_{\theta_k} \ln g_{\theta_k}(x_i) \end{aligned} \quad \text{ahol} \quad q_{k,i}(\bar{\vartheta}) = \frac{p_k g_{\theta_k}(x_i)}{\sum_j p_j g_{\theta_j}(x_i)}$$

- A likelihood egyenlet:

$$\begin{aligned} \partial_{p_k} \ln f_{\bar{\vartheta}}(x) &= \sum_{i=1}^n q_{k,i}(\bar{\vartheta}) \frac{1}{p_k} = \lambda \implies p_k = \frac{\sum_i q_{k,i}(\bar{\vartheta})}{\sum_{i,j} q_{j,i}(\bar{\vartheta})}, \\ \partial_{\theta_k} \ln f_{\bar{\vartheta}}(x) &= \sum_{i=1}^n q_{k,i}(\bar{\vartheta}) \partial_{\theta_k} \ln g_{\theta_k}(x_i) = 0 \end{aligned}$$

7.4 EM algoritmus

- Rekurzív eljárás: $\bar{\vartheta}^{(t)}$ a paraméter a t . lépés után.

$$q_{k,i}^{(t)} = q_{k,i}(\bar{\vartheta}^{(t)}) = \frac{p_k g_{\theta_k^{(t)}}(x_i)}{\sum_j p_j g_{\theta_j^{(t)}}(x_i)} \quad \bar{\vartheta}^{(t+1)} = \arg \max_{\bar{\vartheta}} \sum_{i,k} q_{k,i}^{(t)} \ln(p_k g_{\theta_k}(x_i))$$

Azaz

$$p_k^{(t+1)} = \frac{\sum_i q_{k,i}^{(t)}}{\sum_{i,k} q_{k,i}^{(t)}}, \quad \sum_i q_{k,i}^{(t)} \partial_{\theta_k} \ln g_{\theta_k^{(t+1)}}(x_i) = 0$$

Ha $\bar{\vartheta}^{(t)}$ konvergens, akkor $\bar{\vartheta} = \lim_{t \rightarrow \infty} \bar{\vartheta}^{(t)}$ megoldja a likelihood egyenletet.

Ezt hívják EM algoritmusnak. q kiszámítása az E-lépés (expectation) $\bar{\vartheta}$ frissítése az M-lépés (maximization).

Állítás 32. $\ell(\theta^{(t)})$ a t monoton növekvő függvénye.

•

$$q_{k,i}^{(t)} = \frac{p_k g_{\theta_k^{(t)}}(x_i)}{\sum_j p_j g_{\theta_j^{(t)}}(x_i)} \quad \bar{\vartheta}^{(t+1)} = \arg \max_{\bar{\vartheta}} \sum_{i,k} q_{k,i}^{(t)} \ln(p_k g_{\theta_k}(x_i))$$

•

$$F(q, \bar{\vartheta}) = \sum_{k,i} q_{k,i} \ln(p_k g_{\theta_k}(x_i)) = \sum_{k,i} q_{k,i} \ln q_{k,i}(\bar{\vartheta}) + \sum_i \ell(\bar{\vartheta}, x_i)$$

Így

$$\ell(\bar{\vartheta}^{(t+1)}) - \ell(\bar{\vartheta}^{(t)}) = F(q^{(t)}, \bar{\vartheta}^{(t+1)}) - F(q^{(t)}, \bar{\vartheta}^{(t)}) - \sum_{k,i} q_{k,i}^{(t)} \ln \frac{q_{k,i}^{(t+1)}}{q_{k,i}^{(t)}} \geq 0,$$

mivel az utolsó nem negatív a $-\ln$ függvény konvexitása miatt

$$-\sum_k q_{k,i}^{(t)} \ln \frac{q_{k,i}^{(t+1)}}{q_{k,i}^{(t)}} \geq -\ln \left(\sum_k q_{k,i}^{(t+1)} \right) = 0$$

7.5 Becslések tulajdonságai

- $\{\mathbb{P}_\theta : \theta \in \Theta\}$ eloszláscsalád, $g(\theta)$ értékét akarjuk becsülni, erre a T statisztikát használjuk.
- Egy T becslés „jósságát” mérhetjük az átlagos veszteséggel

$$R_T(\theta) = \mathbb{E}_\theta(W(T, \theta))$$

R_T neve: rizikó függvény.

- $W(t, \theta)$ a veszteség, amit $g(\theta)$ a t értékkel történő közelítésekor kapunk.
- Négyzetes veszteség: $W(t, \theta) = (g(\theta) - t)(g(\theta) - t)^T$ ha g vektor értékű és $(g(\theta) - t)^2$, ha g skalár értékű.
- Általánosabban W lehetne tetszőleges nem negatív, első változóban konvex függvény, amire $W(g(\theta), \theta) = 0$.
- $\{\mathbb{P}_\theta : \theta \in \Theta\}$ eloszláscsalád, $g(\theta)$ értékét akarjuk becsülni, erre a T statisztikát használjuk.

- Rizikó függvény

$$R_T(\theta) = \mathbb{E}_\theta(W(T, \theta)), \quad W(t, \theta) = (g(\theta) - t)^2 \quad \text{négyzetes rizikó és veszteség}$$

- T torzítatlan, ha $\mathbb{E}_\theta(T) = g(\theta)$, ha nem torzítatlan, akkor a torzítása (bias) $b(\theta) = \mathbb{E}_\theta(T) - g(\theta)$.
- T és S ekvivalens becslések, ha $R_T(\theta) = R_S(\theta)$, minden θ -ra.
- T egyenletesen jobb (nem rosszabb), mint S , ha $R_T(\theta) \leq R_S(\theta)$, minden θ -ra és nem ekvivalensek.
- Ha \mathcal{D} becslések (statisztikák) egy családja, pl. torzítatlan becslések, akkor $T \in \mathcal{D}$ **megengedhető** (\mathcal{D} -re nézve), ha nincs nála jobb becslés \mathcal{D} -ben, azaz tetszőleges $S \in \mathcal{D}$ -re, $R_S \leq R_T$ -ből $R_S = R_T$ következik.
- $T \in \mathcal{D}$ **optimális**, ha \mathcal{D} minden eleménél jobb vagy ekvivalens, azaz $R_T \leq R_S$ minden $S \in \mathcal{D}$ -re.

Ha \mathcal{D} a torzítatlan becslések családja, akkor az optimális becslés neve **hatásos** becslés és jobb helyett hatásosabbat mondunk.

- $T \in \mathcal{D}$ **minimax**, $\sup_{\theta \in \Theta} R_T(\theta) \leq \sup_{\theta \in \Theta} R_S(\theta)$, minden $S \in \mathcal{D}$ -re.
- $R_T(\theta) = \mathbb{E}_\theta(W(T, \theta))$, pl. $R_T(\theta) = \mathbb{E}_\theta((T - g(\theta))^2)$
- $T \in \mathcal{D}$ **megengedhető** ha tetszőleges $S \in \mathcal{D}$ -re, $R_S \leq R_T$ -ből $R_S = R_T$ következik.
- $T \in \mathcal{D}$ **optimális**, $R_T = \min_{S \in \mathcal{D}} R_S$
- $T \in \mathcal{D}$ **minimax**, ha $\sup_{\theta \in \Theta} R_T(\theta) \leq \sup_{\theta \in \Theta} R_S(\theta)$, minden $S \in \mathcal{D}$ -re.

Állítás 33. Ha T optimális, akkor T megengedhető és minimax.

Ha \mathcal{D} konvex, W első változójában szigorúan konvex, akkor az optimális becslés egyértelmű (ha létezik).

Ha T_1, T_2 optimális, akkor $R_{T_1} = R_{T_2}$. Ha \mathcal{D} konvex, akkor $T = \frac{1}{2}(T_1 + T_2) \in \mathcal{D}$ és

$$\begin{aligned} R_T(\theta) &= \mathbb{E}_\theta(W(\tfrac{1}{2}T_1 + \tfrac{1}{2}T_2, \theta)) \\ &\leq \tfrac{1}{2}\mathbb{E}_\theta(W(T_1, \theta)) + \tfrac{1}{2}\mathbb{E}_\theta(W(T_2, \theta)) = \tfrac{1}{2}R_{T_1}(\theta) + \tfrac{1}{2}R_{T_2}(\theta) = R_{T_1}(\theta) = R_{T_2}(\theta) \end{aligned}$$

Ha $\mathbb{P}_\theta(T_1 \neq T_2) > 0$, akkor $\mathbb{P}_\theta(W(T, \theta) < \frac{1}{2}(W(T_1, \theta) + W(T_2, \theta))) > 0$ és $R_T(\theta) < R_{T_1}(\theta)$.

7.6 Bayes becslés

- $\theta \mapsto \mathbb{P}_\theta(A)$ mérhető minden $A \in \mathcal{B}(\mathfrak{X})$ -re, ekkor $\theta \mapsto R_T(\theta)$ mérhető.
pl. dominált mérték család és $(\theta, x) \mapsto f_\theta(x)$ mérhető.
- Q valószínűségi mérték Θ -n.
- $R_T(Q) = \int_\Theta R_T(t)Q(t)$ a priori rizikó
- T Bayes becslés (Q -Bayes), ha $R_T(Q) \leq R_S(Q)$ minden S -re.
- $\theta \mapsto \mathbb{P}_\theta(A)$ mérhető minden $A \in \mathcal{B}(\mathfrak{X})$ -re, ekkor $\theta \mapsto R_T(\theta)$ mérhető.

•

$$\mathbb{P}(A) = \int_{\Theta} \int_{\mathfrak{X}} \mathbf{1}_{((t,x) \in A)} \mathbb{P}_t(dx) Q(dt), \quad A \in \mathcal{B}(\Theta) \otimes \mathcal{B}(\mathfrak{X})$$

valószínűségi mérték $\Theta \times \mathfrak{X}$ -en. $\theta(t, x) = t$, $X(t, x) = x$ valószínűségi változók a $(\Omega = \Theta \times \mathfrak{X}, \mathcal{B}(\Theta) \otimes \mathcal{B}(\mathfrak{X}), \mathbb{P})$ valószínűségi mezőn. És

$$\mathbb{P}(X \in H | \theta) = \mathbb{P}_{\theta}(H)$$

•

$$R_T(Q) = \int_{\Theta} R_T(t) Q(dt) = \int_{\Theta} \int_{\mathfrak{X}} W(T(x), t) \mathbb{P}_t(dx) Q(dt) = \mathbb{E}(W(T, \theta))$$

- Négyzetes veszteség esetén $(\mathbb{E}(g(\theta)^2) < \infty)$

$$\mathbb{E}((T - g(\theta))^2) = \mathbb{E}((T - \mathbb{E}(g(\theta)|X))^2) + \mathbb{E}((g(\theta) - \mathbb{E}(g(\theta)|X))^2)$$

a Bayes becslés $T(X) = \mathbb{E}(g(\theta)|X)$

- $\mathbb{P}(\theta \in H|X)$ -nek létezik reguláris változata, ez az a posteriori eloszlás. Jelölés $Q^*(dt|x)$.
- A Bayes becslés az a posteriori eloszlásból számolt várható érték.

7.6.1 Aszimptotikus normalitás

f_{θ} , $\ell(\theta, x)$ egy mintaelem sűrűségfüggvénye, ill. loglikelihoodja. X_1, X_2, \dots minta \mathbb{P}_{θ_0} eloszlásból

$$L_n(\theta) = \frac{1}{n} \sum_{i=1}^n \ell(\theta, X_i)$$

Tétel 27 (Bizonyítás nélkül). • (RR) teljesül

- Minden $\varepsilon > 0$ -ra

$$\lim_{\delta \rightarrow 0} \liminf_{n \rightarrow \infty} \mathbb{P}_{\theta_0}^n \sup_{\|h\| \geq \varepsilon} L_n(\theta_0 + h) \leq L_n(\theta_0) - \delta = 1$$

- Q abszolút folytonos és $q(\theta_0) > 0$

- $\int_{\Theta} \|t\| q(t) dt < \infty$.

θ_n a Bayes becslés n elemű mintából. Ekkor $\sqrt{n}(\theta_n - \theta_0) \xrightarrow{d} N(0, I(\theta_0)^{-1})$.

Tulajdonságok

Állítás 34. Ha T Bayes becslés, akkor triviális esettől eltekintve nem torzítatlan.

Ha $T = \mathbb{E}(g(\theta)|X) = \mathbb{E}(g(\theta)|T)$ torzítatlan, akkor

$$\mathbb{E}(T|\theta) = \mathbb{E}_{\theta}(T) = g(\theta) = \mathbb{E}(T|g(\theta))$$

Lemma 8. Ha $X, Y \in L^1$, $\mathbb{E}(X|Y) = Y$ és $\mathbb{E}(Y|X) = X$, akkor $X = Y$ egy valószínűséggel

A Lemma alapján T torzítatlan Bayes becslésre $\mathbb{P}(g(\theta) = T(x)) = 1$.

Azaz Q majdnem minden t -re $\mathbb{P}_t(g(t) = T) = 1$, vagyis létezik $\Theta' \subset \Theta$ Q teljes mértékű halmaz, $\theta_1, \theta_2 \in \Theta'$, $g(\theta_1) \neq g(\theta_2)$ esetén \mathbb{P}_{θ_1} és \mathbb{P}_{θ_2} diszjunkt halmazokra koncentrálódik.

7.6.2 Lemma bizonyítása

Lemma 9.

- Ugyanis, ha $X, Y \in L^2$, akkor

$$\mathbb{E}((X - Y)^2) = \mathbb{E}(X^2) + \mathbb{E}(Y^2) - 2\mathbb{E}(XY), \quad \mathbb{E}(XY) = \mathbb{E}(\mathbb{E}(XY|Y)) = \mathbb{E}(Y^2) = \mathbb{E}(X^2)$$

- Ha $X, Y \in L^1$, akkor legyen $h(x) = x \arctg x$ szigorúan konvex (HF). $h(X), h(Y) \in L^1$ és a Jensen egyenlőtlenség alapján

$$\mathbb{E}(h(X)) \geq \mathbb{E}(h(\mathbb{E}(X|Y))) = \mathbb{E}(h(Y)) = \mathbb{E}(h(X))$$

$$h(X) > h(Y) + h'(Y)(X - Y) \quad \text{az } \{X \neq Y\} \text{ eseményen} \quad \implies \quad \mathbb{P}(X = Y) = 1$$

- X peremeloszlása \mathbb{P} alatt a prediktív eloszlás

$$\mathbb{P}_Q(H) = \int_{\Theta} \mathbb{P}_t(X \in H) Q(dt)$$

- $R_T(\theta) = \mathbb{E}_{\theta}(W(T, \theta))$, pl. $R_T(\theta) = \mathbb{E}_{\theta}((T - g(\theta))^2)$
- $T \in \mathcal{D}$ **megengedhető** ha tetszőleges $S \in \mathcal{D}$ -re, $R_S \leq R_T$ -ből $R_S = R_T$ következik.

Állítás 35. Tegyük fel, hogy $\mathbb{P}_Q \sim \mathcal{P}$. Ekkor ha T Bayes becslés, akkor megengedhető.

Legyen S statisztika, amire $R_S \leq R_T$. Ekkor S is Bayes becslés ($R_S(Q) \leq R_T(Q)$) és ezért $S = \mathbb{E}(g(\theta)|X) = T$, vagyis $\mathbb{P}(S = T) = 1$. S és T a minta függvénye

$$\mathbb{P}_Q(T = S) = 1 \quad \implies \quad T = S \text{ } \mathcal{P} \text{ majdnem mindenütt}$$

- $R_T(\theta) = \mathbb{E}_{\theta}(W(T, \theta))$, pl. $R_T(\theta) = \mathbb{E}_{\theta}((T - g(\theta))^2)$
- $T \in \mathcal{D}$ **minimax**, ha $\sup_{\theta \in \Theta} R_T(\theta) \leq \sup_{\theta \in \Theta} R_S(\theta)$, minden $S \in \mathcal{D}$ -re.

Állítás 36. Ha a T Bayes becslés rizikója konstans, akkor T minimax becslés.

S tetszőleges becslés

$$\sup_{\theta} R_S(\theta) \geq \int_{\Theta} R_S(t) Q(dt) = R_S(Q) \geq R_T(Q) = \int_{\Theta} R_T(t) Q(dt) = \sup_{\theta} R_T(\theta)$$

Állítás 37. Dominált mértékcsalád esetén, ha S elégséges statisztika, akkor létezik Bayes becslés S függvényei között is.

$$q^*(t|x) = \frac{g_{\theta}(S(x))h(x)q(t)}{\int_{\Theta} g_t(S(x))h(x)q(t)dt} = \frac{g_{\theta}(S(x))q(t)}{\int_{\Theta} g_t(S(x))q(t)dt}$$

Azaz az a posteriori sűrűség csak $S(X)$ -en keresztül függ X , a feltételes várható érték is ilyen.

7.7 Minimax becslés indikátor mintából

- $Q = \text{Béta}(\alpha, \beta)$
- $r_n(X) = \frac{1}{n} \sum_{i=1}^n X_i,$

$$\hat{\theta} = \frac{\alpha + S}{\alpha + \beta + n} = a + br_n$$

ahol $b = \frac{n}{\alpha + \beta + n}, a = \frac{\alpha}{\alpha + \beta + n}.$

•

$$\begin{aligned} R_{\hat{\theta}}(\theta) &= \mathbb{E}_{\theta}((a + br_n - \theta)^2) = (a - (1-b)\theta)^2 + b^2 \frac{\theta(1-\theta)}{n} \\ &= \theta^2 \left((1-b)^2 - \frac{b^2}{n} \right) + \theta \left(\frac{b^2}{n} - 2a(1-b) \right) + a^2 \end{aligned}$$

- R_T konstans, ha $b = \frac{\sqrt{n}}{1+\sqrt{n}},$ és $a = \frac{1}{2} \frac{1}{1+\sqrt{n}}$ és $\alpha = n \frac{a}{b} = \frac{\sqrt{n}}{2}$ és $\beta = \frac{n}{b} - \alpha - n = \sqrt{n} - \frac{\sqrt{n}}{2} = \frac{\sqrt{n}}{2}.$
- Ha $Q = \text{Béta}(\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}),$ akkor a Bayes becslés, azaz

$$T(X) = \frac{\frac{1}{2}\sqrt{n} + \sum X_i}{\sqrt{n} + n}$$

minimax

8. 2021.04.12.

8.1 Hipotézis vizsgálati feladatok

8.1.1 Példák

- A kérdés nem az ismeretlen paraméter értéke hanem az, hogy egy állítás (hipotézis) teljesül-e rá.
- Egy gyártósor működése elfogadható, ha a selejt arány 5% alatt van. Ha a gépet beállítják, akkor ez az arány 1%, de a használat során a gépsor elállítódik és a selejt arány idővel nő.
- Minden nap végén mintát veszünk a termékekből, ez alapján szeretnénk dönteni arról, hogy időszerű-e a gépsor beállítása.
- Itt a hipotézis (nullhipotézis) az, hogy selejt arány az elfogadható szint alatt van.
- Egy döntési szabály (statisztikai próba) lehet az, hogy ha a mintában a selejt arány egy *kritikus értéket* meghalad, akkor a gép beállítása mellett döntünk, ellenkező esetben folyhat tovább a termelés.
- Mekkora legyen a kritikus érték?

Ha a kritikus érték túl nagy, akkor a selejt arány elfogadhatatlanul magas lesz, ha túl kicsi akkor feleslegesen áll a termelés a gép beállítása alatt.

8.1.2 Fogalmak

- **Statisztikai hipotézis.** A minta eloszlásának valamely jellemzőjére vonatkozó kijelentés. $(\Omega, \mathcal{A}, \mathcal{P})$ statisztikai mező. $\mathcal{P} = \mathcal{P}_0 \cup \mathcal{P}_1$, $(\mathcal{P}_0 \cap \mathcal{P}_1 = \emptyset)$.
 H_0 **nullhipotézis:** a valódi eloszlás $\in \mathcal{P}_0$,
 H_1 **ellenhipotézis:** a valódi eloszlás $\in \mathcal{P}_1$.
- Ha $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ paraméteres család, akkor $\mathcal{P} = \mathcal{P}_0 \cup^* \mathcal{P}_1$ megfelel egy $\Theta = \Theta_0 \cup^* \Theta_1$ felbontásnak. $H_0 : \theta \in \Theta_0$, $H_1 : \theta \in \Theta_1$.
- A bevezető példában indikátor mintánk volt $\Omega = \mathfrak{X} = \{0, 1\}^n$. $\theta \in \Theta = (0, 1)$ az ismeretlen selejt arány.
 $H_0 : \theta \leq 0.05$, $H_1 : \theta > 0.05$, $\Theta_0 = (0, 0.05]$ és $\Theta_1 = (0.05, 1)$.
- A H_0 , ill. H_1 hipotézis egyszerű, ha \mathcal{P}_0 , ill. \mathcal{P}_1 egy elemű.
- Egy hipotézis összetett, ha nem egyszerű, azaz a \mathcal{P}_0 vagy \mathcal{P}_1 több elemű.
- **Statisztikai próba.** Döntés a minta alapján H_0 -ról. Azaz $\mathfrak{X} = \mathfrak{X}_0 \cup^* \mathfrak{X}_1$. Ha a megfigyelés $X \in \mathfrak{X}_0$, akkor elfogadjuk a nullhipotézist, ellenkező esetben elutasítjuk.
 \mathfrak{X}_0 neve **elfogadási tartomány**
 \mathfrak{X}_1 neve elutasítási, vagy **kritikus tartomány**
- A bevezető példában ha c a kritikus érték, akkor $\mathfrak{X}_1 = \{x \in \mathfrak{X} : \bar{x} > c\}$.

8.2 Hiba lehetőségek

- **Elsőfajú hiba.** A nullhipotézis igaz, de elvetjük (téves elutasítás)
Másodfajú hiba. A nullhipotézis nem igaz, de mégis elfogadjuk (téves elfogadás)
- A bevezető példában:
 $H_0 : \theta \leq 0.05$, azaz elsőfajú hibát követünk el, ha a gép beállítása még nem időszerű de a mintában a selejt arány a kritikus érték felett van.
 $H_1 : \theta > 0.05$, azaz másodfajú hibát követünk el, ha a selejt arány a mintában kicsi, pedig a gép beállítása időszerű.
- **Fix mintaelemszám esetén az egyik fajta hiba valószínűségét csak a másik növelése árán lehet csökkenteni.**
- Feladat. Az elsőfajú hiba valószínűsége legyen kicsi (tipikusan nem nagyobb, mint 5%) és ezen feltétel mellett a másodfajú hiba valószínűsége legyen minél kisebb.
- Jelentés: A nullhipotézist, ha lehet elfogadjuk. Csak akkor utasítjuk el, ha az adatok ennek „ellentmondanak”.
 H_0 elfogadása csak annyit jelent, hogy az adatok alapján lehetséges hogy H_0 igaz.
 H_0 elutasítása az **informatív döntés**, ez ugyanis annyit jelent, hogy a megfigyelések H_0 ellen szólnak.

Példák.

- Gyógyszerkutatásban: H_0 : az új hatóanyag hatástalan, H_1 : az új hatóanyag hatásos.
- Növénynemesítés. H_0 : az új változat termésátlaga nem nagyobb mint a most használt fajtáké. H_1 : az új változat termésátlaga nagyobb.

8.3 Véletlenített próba, próbafüggvény, terjedelem, szint, erő

- Egy statisztikai próba leírható az elutasítási tartomány \mathfrak{X}_1 indikátorával. $\phi(x) = \mathbf{1}_{(x \in \mathfrak{X}_1)}$. Ekkor $\phi(x)$ az elutasítás valószínűsége, ha a megfigyelés x . ϕ neve **próbafüggvény** vagy **döntési függvény**.
- Véletlenített próba. $\phi : \mathfrak{X} \rightarrow [0, 1]$.
Ha $\phi : \mathfrak{X} \rightarrow \{0, 1\}$, akkor a próba nem véletlenített.
- A ϕ próba **terjedelme** az elsőfajú hiba elkövetésének maximális valószínűsége

$$\alpha(\phi) = \sup_{\theta \in \Theta_0} \mathbb{E}_{\theta}(\phi)$$

- A ϕ próba **ereje** annak a valószínűsége, hogy H_0 nem teljesül és ezt felismerjük. Ez egy Θ_1 -n értelmezett ψ függvény, az **erőfüggvény**

$$\psi(\theta) = \psi(\theta, \phi) = \mathbb{E}_{\theta}(\phi), \quad \theta \in \Theta_1$$

- Az azonos terjedelmű próbákat az erőfüggvény alapján hasonlíthatjuk össze.

Megjegyzés. Gyakran a legfeljebb α terjedelmű próbát is α terjedelműnek mondják.
Ha a próba terjedelme α , akkor a szintje $1 - \alpha$.

8.4 Statisztikai próbák jellemzői

$\phi : \mathfrak{X} \rightarrow [0, 1]$ próba függvény. $\phi(x)$ a H_0 elutasításának valószínűsége, ha a megfigyelés x .

- ϕ **torzítatlan**, ha ereje nem kisebb a terjedelménél:

$$\sup_{\theta \in \Theta_0} \mathbb{E}_\theta(\phi(X)) \leq \inf_{\theta \in \Theta_1} \mathbb{E}_\theta(\phi(X))$$

- ϕ **egyenletesen legerősebb** próba (e.l.p.), ha ereje legalább akkora, mint tetszőleges vele azonos terjedelmű próba ereje.

$$\psi(\theta, \phi) = \mathbb{E}_\theta(\phi) = \sup\{\psi(\theta, \phi') : \alpha(\phi') = \alpha(\phi)\} \quad \forall \theta \in \Theta_1$$

$$\alpha(\phi) = \sup_{\theta \in \Theta_0} \mathbb{E}_\theta(\phi).$$

Egyenletesen legerősebb próba nem feltétlenül létezik.

Ha ϕ egyenletesen legerősebb, akkor torzítatlan is, mert ereje legalább akkora, mint a $\phi'(x) = \alpha(\phi)$ próba ereje.

- ϕ **legszigorúbb**, "ha legjobban közelíti az erőfüggvény burkolóját", azaz ϕ minimalizálja

$$\sup_{\theta \in \Theta_1} (\bar{\psi}(\theta, \alpha) - \psi(\theta, \phi)) \quad \text{értékét, ahol} \quad \bar{\psi}(\theta, \alpha) = \sup\{\psi(\theta, \phi) : \alpha(\phi) = \alpha\}, \quad \theta \in \Theta_1$$

- ϕ_n n -elemű mintára épülő α terjedelmű próba. (ϕ_n) **konzisztens**, ha $\psi(\theta, \phi_n) \rightarrow 1 \quad \forall \theta \in \Theta_1$ -re.

8.5 Próbák konstruálása

- Látni fogjuk, hogy a legtöbb próba $\phi(x) = \mathbf{1}_{(T(x) > c)}$ alakú.
- T olyan statisztika, ami $\theta \in \Theta_1$ mellett tipikusan nagy értékekre vezet, míg $\theta \in \Theta_0$ esetén nem.
- T neve **próba statisztika**, c neve **kritikus érték**
- Ha S elégséges statisztika, akkor $\phi' = \mathbb{E}_\theta(\phi|S)$ nem függ θ -tól, terjedelme, ereje azonos ϕ -vel. Azaz ϕ -t elég az elégséges statisztika függvényei között keresni.
- Ha determinisztikus próbával a megengedett α terjedelem nem érhető el, akkor véletlenítéssel lehet javítani a próbán.

Azaz

$$\alpha' = \sup_{\theta \in \Theta_0} \mathbb{P}_\theta(T > c) < \alpha \quad \sup_{\theta \in \Theta_0} \mathbb{P}_\theta(T \geq c) \geq \alpha$$

akkor

$$\phi(x) = \begin{cases} 1 & T(x) > c \\ p & T(x) = c \\ 0 & T(x) < c \end{cases}$$

$$p \sup_{\theta \in \Theta_0} \mathbb{P}_\theta(T = c) = \alpha - \alpha'.$$

8.6 Neyman–Pearson lemma

Lemma 10. *Egyszerű hipotézis vizsgálati feladatban a likelihood hányados próba a legerősebb a vele azonos terjedelmű próbák között.*

Azaz, ha $c > 0$, $g : \mathfrak{X} \rightarrow [0, 1]$, és

$$\phi(x) = \begin{cases} 1 & f_1(x) > cf_0(x) \\ g(x) & f_1(x) = cf_0(x) \\ 0 & f_1(x) < cf_0(x) \end{cases}$$

akkor tetszőleges ϕ' próbafüggvényre

$$\mathbb{E}_0(\phi') \leq \mathbb{E}_0(\phi) \implies \mathbb{E}_1(\phi') \leq \mathbb{E}_1(\phi)$$

$$\mathbb{E}_1(\phi - \phi') = \int_{\mathfrak{X}} (\phi - \phi') f_1 d\lambda \geq \int_{\mathfrak{X}} (\phi - \phi') cf_0 d\lambda = c\mathbb{E}_0(\phi - \phi') \geq 0$$

Ugyanis $(\phi - \phi')(f_1 - cf_0) \geq 0$. Ez eset szétválasztással látható.

Lemma 11 (Egzisztencia). *Egyszerű hipotézis vizsgálati feladatban tetszőleges $\alpha \in (0, 1)$ -re létezik α terjedelmű likelihood hányados próba.*

Megjegyzés. Ez a legerősebb a legfeljebb α terjedelmű próbák között.

$c > 0$, $p \in [0, 1]$, megválasztásával elérhető, hogy a

$$\phi(x) = \begin{cases} 1 & f_1(x) > cf_0(x) \\ p & f_1(x) = cf_0(x) \\ 0 & f_1(x) < cf_0(x) \end{cases}$$

próba terjedelme pontosan α legyen.

Ha $c = \inf\{c > 0 : \mathbb{P}_0(f_1(X) > cf_0(X)) < \alpha\}$, akkor

$$\mathbb{P}_0(f_1(X) > cf_0(X)) \leq \alpha \leq \mathbb{P}_0(f_1(X) \geq cf_0(X)).$$

és így létezik $p \in [0, 1]$, hogy

$$\alpha = (1 - p)\mathbb{P}_0(f_1(X) > cf_0(X)) + p\mathbb{P}_0(f_1(X) \geq cf_0(X)) = \mathbb{E}_0(\phi).$$

Lemma 12 (Unicitás). *Egyszerű hipotézis vizsgálati feladatban, ha ϕ_1 legerősebb próba a legfeljebb α terjedelmű próbák között, akkor ϕ_1 likelihood hányados próba és vagy $\alpha(\phi_1) = \alpha$ vagy $\psi(\phi_1) = 1$.*

- Legyen ϕ a pontosan α terjedelmű (esetleg véletlenített) likelihood hányados próba (c a kritikus értéke). Ekkor

$$0 \geq \mathbb{E}_1(\phi - \phi_1) - c\mathbb{E}_0(\phi - \phi_1) = \int (f_1 - cf_0)(\phi - \phi_1) d\lambda \geq 0$$

- Azaz $(f_1 - cf_0)(\phi - \phi_1) = 0$ λ m.m., Vagyis ϕ_1 is likelihood hányados próba.
- Ha $\alpha(\phi_1) < \alpha$, és $\psi(\phi) < 1$, akkor véletlenítéssel növelhető lenne az ereje. Mivel ϕ_1 a legerősebb próba, ez nem lehetséges, tehát vagy $\alpha(\phi_1) = \alpha$ vagy $\psi(\phi_1) = 1$.

8.6.1 Példa

- X_1, \dots, X_n indikátor minta. $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$, ahol $\theta_0 < \theta_1$.
- Ekkor $f_i(x) = \theta_i^{\sum x_k} (1 - \theta_i)^{\sum (1-x_k)}$ és

$$\frac{f_1(x)}{f_0(x)} = \left(\frac{1 - \theta_0}{1 - \theta_1} \right)^n \cdot \left(\frac{\theta_1}{1 - \theta_1} : \frac{\theta_0}{1 - \theta_0} \right)^{\sum x_i}.$$

- A likelihood hányados a $\sum X_i$ statisztika monoton növekvő függvénye, azaz a likelihood hányados próba döntési függvénye:

$$\phi(x) = \mathbf{1}_{(\sum x_k > c)} + p \mathbf{1}_{(\sum x_k = c)},$$

ahol c és p értékét a $\mathbb{E}_0(\phi(X)) = \alpha$ összefüggés határozza meg.

- Vegyük észre, hogy a döntési függvény nem függ θ_1 konkrét értékétől, csak a $\theta_1 > \theta_0$ összefüggést használtuk. A kapott próba egyenletesen legerősebb a $H_1 : \theta > \theta_0$ ellenhipotézissel szemben is.

- X_1, \dots, X_n indikátor minta. $H_0 : \theta = \theta_0$, $H_1 : \theta = \theta_1$, ahol $\theta_0 < \theta_1$.
- A likelihood hányados a $\sum X_i$ statisztika monoton növekvő függvénye, azaz a likelihood hányados próba döntési függvénye:

$$\phi(x) = \mathbf{1}_{(\sum x_k > c)} + p \mathbf{1}_{(\sum x_k = c)}, \quad \text{ahol } c \text{ és } p \text{ értékét a } \mathbb{E}_0(\phi(X)) = \alpha \text{ összefüggés adja.}$$

- Ha n nagy, akkor $n^{-1/2} \sum_i (X_i - \theta_0)$ közelítőleg $N(0, \theta_0(1 - \theta_0))$ eloszlású és a kritikus érték ez alapján választható. $c_n = n\theta_0 + \sqrt{n\theta_0(1 - \theta_0)}\Phi^{-1}(1 - \alpha)$, nagy n -re $p = 0$ értékét vehetjük nullának.
- Emlékeztető

Állítás 38 (Hoeffding egyenlőtlenség). Z_1, \dots, Z_n független változók, $0 \leq Z_i \leq 1$, $\varepsilon > 0$. Ekkor $\mathbb{P}(\sum Z_i > \sum \mathbb{E}(Z_i) + n\varepsilon) \leq e^{-2n\varepsilon^2}$

$(1 - Z)$ -re átvérve: $\mathbb{P}(\sum Z_i < \sum \mathbb{E}(Z_i) - n\varepsilon) \leq e^{-2n\varepsilon^2}$.

$$1 - \psi(\theta_1) = \mathbb{P}_1 \left(\sum_{i=1}^n X_i < c_n \right) \leq \exp \left\{ -2n \left(\frac{n\theta_1 - c_n}{n} \right)^2 \right\} \leq e^{-2n(\theta_1 - \theta_0)^2 + a\sqrt{n} + b}$$

Azaz az erő geometriai sebességgel tart egyhez, ha minta elemszám végtelenhez tart.

8.7 Általánosítás: kritikus érték aszimptotikája

- Egyszerű hipotézis vizsgálati feladat. f_1, f_0 egy mintaelem sűrűségfüggvénye. Tegyük fel, hogy $f_1, f_0 > 0$.

•

$$\xi_i = -\log \frac{f_1(X_i)}{f_0(X_i)}, \quad \text{ekkor} \quad \mathbb{E}_0(\xi_i) = \int -\log \frac{f_1}{f_0} f_0 d\lambda = D(\mathbb{P}_0 \| \mathbb{P}_1).$$

Belátható, hogy ha $f_1 \neq f_0$, akkor $0 < \mathbb{E}_0(\xi_i)$. Tegyük fel, hogy $\mathbb{E}_0(\xi_i) < \infty$ és $\sigma^2 = \mathbb{D}_0^2 \xi < \infty$.

- A likelihood hányados próba $\sum_i \xi_i$ -vel is felírható.

$$T_n = \prod_{i=1}^n \frac{f_1(X_i)}{f_0(X_i)} > c_n \iff \sum \frac{\xi_i - D(\mathbb{P}_0 \| \mathbb{P}_1)}{\sqrt{n}\sigma} < t_n, \quad c_n = \exp\{-nD(\mathbb{P}_0 \| \mathbb{P}_1) - \sqrt{n}\sigma t_n\}.$$

- Terjedelem α . A CHT alapján

$$\sum_{i=1}^n \frac{\xi_i - D(\mathbb{P}_0 \| \mathbb{P}_1)}{\sqrt{n}\sigma} \xrightarrow{d} N(0, 1) \quad \text{ha } H_0 \text{ igaz, és } \mathbb{P}_0 \left(\sum_{i=1}^n \frac{\xi_i - D(\mathbb{P}_0 \| \mathbb{P}_1)}{\sqrt{n}\sigma} < t_n \right) \rightarrow \alpha.$$

Belátjuk, hogy ebből $t_n \rightarrow \Phi^{-1}(\alpha)$ következik.

- Ha

$$\hat{c}_n = \exp\{-nD(\mathbb{P}_0 \| \mathbb{P}_1) - \sqrt{n}\sigma\Phi^{-1}(\alpha)\}, \quad \hat{\phi}_n(x) = \mathbf{1}_{(T_n > \hat{c}_n)}, \quad \text{akkor } \alpha(\hat{\phi}_n) \rightarrow \alpha$$

8.8 Kullback-Leibler divergencia

Definíció 26. $\mathbb{P}_0, \mathbb{P}_1$ valószínűségi mértékek az (Ω, \mathcal{B}) mérhető téren.

$$D(\mathbb{P}_0 \| \mathbb{P}_1) = \begin{cases} \mathbb{E}_0\left(-\log \frac{d\mathbb{P}_1}{d\mathbb{P}_0}\right) & \text{ha } \mathbb{P}_1 \ll \mathbb{P}_0 \\ \infty & \text{különben} \end{cases}$$

- $D(\mathbb{P}_0 \| \mathbb{P}_1) = \int -\log \frac{f_1}{f_0} f_0$, ha f_1, f_0 a két sűrűségfüggvény. Ez a formula nem függ a domináló mérték választásától.
- $-\log$ konvex, ezért $\int -\log\left(\frac{f_1}{f_0}\right) f_0 \geq -\log\left(\int f_1\right) = 0$.
- $-\log$ szigorúan konvex, így egyenlőség csak akkor lehet, ha $-\log\left(\frac{f_1}{f_0}\right) = 0$ \mathbb{P}_0 mm., vagyis $f_1 = f_0$.

8.9 Két apróság

Állítás 39. Φ $N(0, 1)$ eloszlásfüggvénye, (F_n) eloszlásfüggvények sorozata és $F_n(t) \rightarrow \Phi(t)$ minden t -re. Ekkor

$$[<*>]$$

$$1. \sup |F_n - \Phi| \rightarrow 0$$

$$2. F_n(t_n) \rightarrow \Phi(t) \text{ pontosan akkor ha } t_n \rightarrow t.$$

- (1)-et láttuk a Glivenko–Cantelli tétel kapcsán. F_n, Φ monoton, ha $t_i = \Phi^{-1}\left(\frac{i}{r}\right)$, $i = 1, \dots, r-1$, akkor $F_n(t_i) \rightarrow \Phi(t_i) = \frac{i}{r}$ és $\limsup_{n \rightarrow \infty} \sup |F_n - \Phi| \leq \lim_{n \rightarrow \infty} \max_i |F_n(t_i) - \Phi(t_i)| + \frac{1}{r} \leq \frac{1}{r}$.

- (2) Ha $t_n \rightarrow t$ ($t = \pm\infty$ is lehet)

$$|F_n(t_n) - \Phi(t)| \leq |F_n(t_n) - \Phi(t_n)| + |\Phi(t_n) - \Phi(t)| \leq \sup |F_n - \Phi| + |\Phi(t_n) - \Phi(t)| \rightarrow 0$$

- (2) Ha $t_n \not\rightarrow t$, akkor létezik $(t'_n) \subset (t_n)$ részsorozat, aminek létezik limesze t' és az nem t ($t'_n \rightarrow \pm\infty$ lehetséges)

$$F_n(t'_n) \rightarrow \Phi(t') \neq \Phi(t)$$

azaz $F_n(t_n) \not\rightarrow \Phi(t)$.

Következmény. Ha $t_n = \sup\{t : F_n(t) < \alpha\}$, akkor $t_n \rightarrow \Phi^{-1}(\alpha)$.

8.10 Erő aszimptotikája

- Egyszerű hipotézis vizsgálati feladat. f_1, f_0 egy mintaelem sűrűségfüggvénye. Tegyük fel, hogy $f_1, f_0 > 0$.
- Az α terjedelmű likelihood hányados próba

$$\phi_n(x) = \mathbf{1}_{(T_n(x) > c_n)} + p_n \mathbf{1}_{(T_n(x) = c_n)}, \quad \text{ahol} \quad T_n(x) = \prod_{i=1}^n \frac{f_1(x_i)}{f_0(x_i)}, \quad c_n = e^{-nD(\mathbb{P}_0 \| \mathbb{P}_1) - \sqrt{n}\sigma t_n}, \quad t_n \rightarrow \Phi^{-1}(\alpha)$$

- ψ_n az erő. A Markov egyenlőtlenségből

$$1 - \psi_n \leq \mathbb{P}_1 \left(\prod_{i=1}^n \frac{f_1(X_i)}{f_0(X_i)} \leq c_n \right) = \mathbb{P}_1 \left(\prod_{i=1}^n \frac{f_0(X_i)}{f_1(X_i)} \geq \frac{1}{c_n} \right) \leq c_n \mathbb{E}_1 \left(\prod_{i=1}^n \frac{f_0(X_i)}{f_1(X_i)} \right) = c_n$$

•

$$1 - \psi_n \leq \exp\{-nD(\mathbb{P}_0 \| \mathbb{P}_1) + O(\sqrt{n})\}$$

Az erő geometria sebességgel tart 1-hez, ha $n \rightarrow \infty$.

9. 2021.04.19.

9.1 Többdimenziós normális eloszlás, emlékeztető

- X egy dimenziós normális változó, ha $X \stackrel{p}{=} \sigma Z + \mu$, ahol $Z \sim N(0, 1)$ és $\sigma \geq 0$, $\mu \in \mathbb{R}$.
- Elfajult eloszlás ($\sigma = 0$) is normális.
- Karakterisztikus függvény

$$\phi_Z(t) = \mathbb{E}(e^{itZ}) = e^{-\frac{1}{2}t^2}, \quad \phi_{\sigma Z + \mu}(t) = \mathbb{E}(e^{it(\sigma Z + \mu)}) = e^{it\mu} \phi_Z(t\sigma) = e^{it\mu - \frac{1}{2}t^2\sigma^2}.$$

- X d -dimenziós normális vektorváltozó, ha minden $a \in \mathbb{R}^d$ -re $\sum_j a_j X_j$ egy dimenziós normális.
- Ha Z_1, \dots, Z_d független standard normálisok, akkor $Z = (Z_1, \dots, Z_d)$. Ekkor Z d -dimenziós standard normális vektorváltozó.
- X d -dimenziós normális μ várható érték vektorral és Σ kovariancia mátrixszal. Ekkor $X \stackrel{p}{=} AZ + \mu$, ahol Z d -dimenziós standard normális és $AA^T = \Sigma$.

9.1.1 Sűrűségfüggvény

- Z d -dimenziós standard normális.

$$f_Z(t) = \prod f_{Z_j}(t_j) = \prod \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}t_j^2\right\} = (2\pi)^{-n/2} \exp\left\{-\frac{1}{2}\|t\|^2\right\}$$

- Ha A invertálható, $m \in \mathbb{R}^d$, $X = AZ + m$, akkor

$$f_X(t) = f_Z(A^{-1}(t - m)) |\det A^{-1}| = (2\pi)^{-n/2} \exp\left\{-\frac{1}{2}\|A^{-1}(t - m)\|^2\right\} |\det A^{-1}|$$

$$X \sim N(m, \Sigma = AA^T)$$

$$f_X(t) = \frac{1}{(2\pi)^{d/2} (\det \Sigma)^{1/2}} \exp\left\{-\frac{1}{2}(t - m)^T \Sigma^{-1} (t - m)\right\}$$

- Ha $U^T = U^{-1}$, vagyis U ortonormált mátrix, akkor UZ is d -dimenziós standard normális.
- Ha Σ nem invertálható, azaz nem teljes rangú, akkor az eloszlás nem abszolút folytonos a d -dimenziós Lebesgue mértékre nézve.

Ha $a \perp \text{im } \Sigma$, akkor $\mathbb{D}^2 a^T X = a^T \Sigma a = 0$ és $a^T X = a^T m$ egy valószínűséggel.

Megmutatható, hogy $\mathbb{P}(X \in \text{im } \Sigma + m) = 1$.

9.1.2 függetlenség és korrelálatlanság

Állítás 40. $X = (X_1, X_2)^T$ normális vektor változó X_1 d_1 , X_2 d_2 dimenziós.
 $\text{cov}(X_1, X_2) = 0$ pontosan akkor, ha $X_1 \perp X_2$.

- Ha $X_1 \perp X_2$, akkor $\text{cov}(X_1, X_2) = 0$. Elég a másik iránnyal foglalkozni
- X kovariancia mátrixa Σ .

$$\Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}, \quad \Sigma_{11} = A_1 A_1^T, \quad \Sigma_{22} = A_2 A_2^T, \quad A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad \Sigma = A A^T$$

- Ha $Z = (Z_1, Z_2)^T$ standard normális, akkor

$$X \stackrel{p}{=} AZ + \mu = \begin{pmatrix} A_1 Z_1 + \mu_1 \\ A_2 Z_2 + \mu_2 \end{pmatrix}$$

- $A_1 Z_1 + \mu_1 \perp A_2 Z_2 + \mu_2$ azaz $X_1 \perp X_2$.

Ha $X = (X_1, \dots, X_r)$ normális vektor változó, X_i d_i dimenziós és a keresztkovarianciák eltűnnek, akkor X_1, \dots, X_r független vektorváltozók.

9.1.3 Következmények

- $X \sim N(\mu, \Sigma)$, ekkor $AX + a \perp BX + b$ pontosan akkor ha $A\Sigma B^T = 0$.
- $X = (X_1, X_2)^T$, $X \sim N(\mu, \Sigma)$, ekkor $X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 \perp X_2$, hiszen

$$\text{cov}(X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2, X_2) = 0$$

•

$$\mathbb{P}(X_1 \in H | X_2) = \mathbb{P}(X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2 + y \in H) \Big|_{y=\Sigma_{12}\Sigma_{22}^{-1}X_2}$$

Azaz X_1 feltételes eloszlása X_2 feltétel mellett normális, $N(\mu_{1|2}, \Sigma_{11|2})$, ahol

$$\begin{aligned} \mu_{1|2} &= \mathbb{E}(X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2) + \Sigma_{12}\Sigma_{22}^{-1}X_2 = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(X_2 - \mu_2) \\ \Sigma_{11|2} &= \text{cov}(X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2, X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \end{aligned}$$

9.2 χ^2 eloszlás, vetületek eloszlása

Definíció 27. Z n -dimenziós standard normális. $\|Z\|^2 = \sum_j Z_j^2$ eloszlását n szabadságfokú khi-négyszet eloszlásnak nevezzük. Jelölés χ_n^2 .

$$\chi_n^2 = \Gamma_{\frac{n}{2}, \frac{1}{2}}.$$

Állítás 41. Z n -dimenziós standard normális, $\mathbb{R}^n = \mathcal{L}_1 \oplus \dots \oplus \mathcal{L}_r$, \mathcal{L}_i -k merőleges alterek, P_i a merőleges vetítés \mathcal{L}_i -re.

Ekkor $X^{(i)} = P_i Z \sim N(0, P_i)$, $X^{(1)}, \dots, X^{(r)}$ függetlenek és $\|X^{(i)}\|^2 \sim \chi_{\dim \mathcal{L}_i}^2$

$$\text{cov}(X^{(i)}, X^{(j)}) = P_i \text{cov}(Z, Z) P_j^T = P_i P_j = P_i \delta_{i,j}$$

Legyen $d = \dim \mathcal{L}_i$ és e_1, \dots, e_d ortonormált bázis \mathcal{L}_i -ben.

Ekkor $X^{(i)} = \sum_k (e_k \cdot Z) e_k$.

$(e_1 \cdot Z, \dots, e_d \cdot Z)$ független $N(0, 1)$ -ek.

$$\|X^{(i)}\|^2 = \sum_k (e_k \cdot Z)^2 \sim \chi_d^2.$$

9.3 Fisher-Bartlett tétel

Tétel 28. X_1, \dots, X_n független $N(\mu, \sigma^2)$ változók.

$$\bar{X} = \frac{1}{n} \sum_i X_i, \quad s_n^{*2} = \frac{1}{n-1} \sum_i (X_i - \bar{X})^2$$

\bar{X} a mintaátlag, s_n^{*2} a korrigált tapasztalati szórásnégyzet.

Ekkor $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, $s_n^{*2} \sim \frac{\sigma^2}{n-1} \chi_{n-1}^2$ és $\bar{X} \perp s_n^{*2}$.

- Feltehető $\sigma > 0$. $Z_i = \frac{X_i - \mu}{\sigma}$, Z_i -k független $N(0, 1)$ -ek, $\bar{X} = \sigma \bar{Z} + \mu$, $s_n^{*2}(X) = \sigma^2 s_n^{*2}(Z)$
- Elég $\mu = 0$, $\sigma^2 = 1$ -re ellenőrizni.
- Legyen $e = \frac{1}{\sqrt{n}}(1, \dots, 1)$, $\mathcal{L}_1 = \mathbb{R}e$, $\mathcal{L}_2 = \mathcal{L}_1^\perp$, P_1, P_2 a projekciók.
- $P_1 Z$ és $P_2 Z$ függetlenek.
- $P_1 Z = (e \cdot Z)e = (\bar{Z}, \dots, \bar{Z})^T$.
- $\|P_2 Z\|^2 = \sum_j (Z_j - \bar{Z})^2 = (n-1)s_n^{*2}(Z) \sim \chi_{n-1}^2$

A függetlenség Basu tételével is megkapható. HF. Hogyan?

9.4 Fisher-Bartlett tétel megfordítása

- Feltesszük, hogy az X -ek szórásnégyzete σ^2 véges.
- Feltehető, $\mathbb{E}(X_k) = 0$, $\sigma^2 = 1$. Legyen $\phi(t) = \mathbb{E}(e^{itX_1})$.
- Egyfelől $\mathbb{E}(\eta) = \mathbb{E}((s_n^*)^2) = \sigma^2 = 1$ miatt

$$\mathbb{E}(\eta e^{it\xi}) = \mathbb{E}(\eta) \mathbb{E}(e^{it\xi}) = \phi^n(t)$$

- Másfelől, ha $e = (1, \dots, 1)^T$,

$$\begin{aligned} \eta &= \frac{1}{n-1} \|X - \bar{X}e\|^2 = \frac{1}{n-1} (\|X\|^2 - (\bar{X})^2 \|e\|^2) \\ &= \frac{1}{n-1} \left(\sum_k X_k^2 - \frac{1}{n} \sum_k X_k^2 - \frac{2}{n} \sum_{k < \ell} X_k X_\ell \right) = \frac{1}{n} \sum_k X_k^2 - \frac{2}{n(n-1)} \sum_{k < \ell} X_k X_\ell \end{aligned}$$

Tétel 29. $n \geq 2$, X_1, \dots, X_n független, azonos eloszlásúak.

Ha $\xi = \sum_k X_k$ és $\eta = s_n^{*2}(X) = \frac{1}{n-1} \sum_k (X_k - \bar{X})^2$ független, akkor az X_i -k normális eloszlásúak.

- $\mathbb{E}(X_k) = 0$, $\mathbb{E}(X_k^2) = 1$, $\phi(t) = \mathbb{E}(e^{itX_1})$.
- Ekkor $k \neq \ell$ mellett

$$\begin{aligned} \mathbb{E}(X_k e^{it\xi}) &= \mathbb{E}(X_k e^{itX_k}) \mathbb{E}(e^{it \sum_{\ell \neq k} X_\ell}) = (-i)\phi'(t)\phi^{n-1}(t) \\ \mathbb{E}(X_k^2 e^{it\xi}) &= \mathbb{E}(X_k^2 e^{itX_k}) \mathbb{E}(e^{it \sum_{\ell \neq k} X_\ell}) = -\phi''(t)\phi^{n-1}(t) \\ \mathbb{E}(X_k X_\ell e^{it\xi}) &= \mathbb{E}(X_k e^{itX_k}) \mathbb{E}(X_\ell e^{itX_\ell}) \mathbb{E}(e^{it \sum_{r \neq k, \ell} X_r}) = -(\phi'(t))^2 \phi^{n-2}(t) \end{aligned}$$

- Láttuk, hogy

$$\phi^n(t) = \mathbb{E}(\eta e^{it\xi}) = \frac{1}{n} \sum_k \mathbb{E}(X_k^2 e^{it\xi}) - \frac{2}{n(n-1)} \sum_{k < \ell} \mathbb{E}(X_k X_\ell e^{it\xi}) = -\phi''(t) \phi^{n-1}(t) - (\phi'(t))^2 \phi^{n-2}(t)$$

- Átosztás után,

$$-1 = \frac{\phi''}{\phi} + \left(\frac{\phi'}{\phi} \right)^2 = (\ln \phi)'' \quad \Rightarrow \quad \ln \phi(t) = -\frac{1}{2} t^2$$

9.5 t eloszlás

- $X_i = \sigma Z_i + \mu$, ($i = 1, \dots, n$) $N(\mu, \sigma^2)$ eloszlású minta. $H_0 : \mu = \mu_0$ esetén

$$T = \sqrt{n} \frac{\bar{X} - \mu_0}{s_n^*} = \sqrt{n} \frac{\sigma \bar{Z}}{\sigma \sqrt{\frac{1}{n-1} \sum_i (Z_i - \bar{Z})^2}} = \frac{\sqrt{n} \bar{Z}}{\sqrt{\frac{1}{n-1} \sum_i (Z_i - \bar{Z})^2}}$$

- A Fisher-Bartlett tétel alapján

$$\sum Z_i \perp \sum (Z_i - \bar{Z})^2 \quad \text{és} \quad \sum Z_i \sim N(0, n), \quad \sum (Z_i - \bar{Z})^2 \sim \chi_{n-1}^2$$

- Azaz H_0 mellett

$$T = \sqrt{n} \frac{\bar{X} - \mu_0}{s_n^*} \stackrel{p}{=} \frac{Z_1}{\sqrt{\frac{1}{n-1} \sum_{i=2}^n Z_i^2}}$$

Ezt hívják t_{n-1} eloszlásnak ($n-1$ szabadságfokú t eloszlás).

Definíció 28. $Z_0, \dots, Z_n \sim N(0, 1)$ függetlenek.

$$\frac{Z_0}{\sqrt{\frac{1}{n} \sum_{i=1}^n Z_i^2}}$$

eloszlását t_n eloszlásnak (n szabadságfokú t eloszlás) hívjuk.

- t_1 azonos a Cauchy eloszlással, ugyanis nem más, mint

$$\frac{Z_1}{|Z_2|} = \frac{R \sin \phi}{R |\cos \phi|} \stackrel{p}{=} \frac{\sin \phi}{\cos \phi} = \operatorname{tg} \phi, \quad \text{ahol } R^2 = Z_1^2 + Z_2^2 \sim \exp(1/2) \text{ és } \phi = \arg(Z_2, Z_1) \sim U(0, 2\pi)$$

- Általános n -re (sűrűségfüggvény transzformációs formulával) a t_n eloszlás sűrűségfüggvénye (HF.)

$$\frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

- $n \rightarrow \infty$ esetén $t_n \xrightarrow{w} N(0, 1)$, ugyanis a nevező 1-hez tart.

9.6 Klasszikus próbák

9.6.1 u -próba

- X_1, \dots, X_n minta $N(\mu, \sigma^2)$ eloszlásból, σ ismert, μ ismeretlen.

- Esetek:

$H_0 : \mu = \mu_0, H_1 : \mu < \mu_0$ egyoldali ellenhipotézis

$H_0 : \mu = \mu_0, H_1 : \mu > \mu_0$ egyoldali ellenhipotézis

$H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$ kétoldali ellenhipotézis

- Próba statisztika:

$$T(X) = \sqrt{n} \frac{\bar{X} - \mu_0}{\sigma}$$

- T eloszlása H_0 mellett $N(0, 1)$
- $u_\alpha = \Phi^{-1}(1 - \alpha)$, azaz $\mathbb{P}(Z > u_\alpha) = \alpha$, ha $Z \sim N(0, 1)$
- α terjedelmű próba. Kritikus tartomány:

$H_1 : \mu < \mu_0$ esetén $T < -u_\alpha$

$H_1 : \mu > \mu_0$ esetén $T > u_\alpha$

$H_1 : \mu \neq \mu_0$ esetén $|T| > u_{\alpha/2}$

9.6.2 t -próba

- X_1, \dots, X_n minta $N(\mu, \sigma^2)$ eloszlásból, μ, σ ismeretlen.

- Esetek:

$H_0 : \mu = \mu_0, H_1 : \mu < \mu_0$ egyoldali ellenhipotézis

$H_0 : \mu = \mu_0, H_1 : \mu > \mu_0$ egyoldali ellenhipotézis

$H_0 : \mu = \mu_0, H_1 : \mu \neq \mu_0$ kétoldali ellenhipotézis

- Próba statisztika:

$$T(X) = \sqrt{n} \frac{\bar{X} - \mu_0}{s_n^*}, \quad \text{ahol} \quad s_n^{*2} = \frac{1}{n-1} \sum_i (X_i - \bar{X})^2 \quad \text{a korrigált tapasztalati szórásnégyzet}$$

- T eloszlása H_0 mellett t_{n-1} Student-féle t eloszlás $n-1$ szabadságfokkal.
- t_α kritikus érték, az $1 - \alpha$ szinthez tartozó kvantilis, azaz $\mathbb{P}(T < t_\alpha) = 1 - \alpha$, ha $T \sim t_{n-1}$
- α terjedelmű próba. Kritikus tartomány:

$H_1 : \mu < \mu_0$ esetén $T < -t_\alpha$

$H_1 : \mu > \mu_0$ esetén $T > t_\alpha$

$H_1 : \mu \neq \mu_0$ esetén $|T| > t_{\alpha/2}$

9.6.3 Kétmintás u -próba

- $X_1, \dots, X_n \sim N(\mu_1, \sigma_1^2)$, $Y_1, \dots, Y_m \sim N(\mu_2, \sigma_2^2)$ független minták.
- σ_1, σ_2 ismert, μ_1, μ_2 ismeretlen.
- Esetek:
 $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 < \mu_2$ (vagy $H_1 : \mu_1 > \mu_2$) egyoldali ellenhipotézis
 $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$ kétoldali ellenhipotézis
- Próba statisztika:

$$T(X, Y) = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$

- T eloszlása H_0 mellett $N(0, 1)$
- $u_\alpha = \Phi^{-1}(1 - \alpha)$, azaz $\mathbb{P}(Z > u_\alpha) = \alpha$, ha $Z \sim N(0, 1)$
- α terjedelmű próba. Kritikus tartomány:
 $H_1 : \mu_1 < \mu_2$ esetén $T < -u_\alpha$
 $H_1 : \mu_1 > \mu_2$ esetén $T > u_\alpha$
 $H_1 : \mu_1 \neq \mu_2$ esetén $|T| > u_{\alpha/2}$
- Ezek a próbák egyenletesen legerősebbek a torzítatlan próbák között.

9.6.4 Kétmintás t -próba

- $X_1, \dots, X_n \sim N(\mu_1, \sigma^2)$, $Y_1, \dots, Y_m \sim N(\mu_2, \sigma^2)$ független minták. σ, μ_1, μ_2 ismeretlen. **A két minta szórása azonos!**
- Esetek:
 $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 < \mu_2$ (vagy $H_1 : \mu_1 > \mu_2$) egyoldali ellenhipotézis
 $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$ kétoldali ellenhipotézis
- Próba statisztika:

$$T(X, Y) = \frac{1}{\sqrt{\frac{1}{n} + \frac{1}{m}}} \cdot \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{(n-1)(s_X^*)^2 + (m-1)(s_Y^*)^2}{n+m-2}}}$$

vagy más alakban

$$T(X, Y) = \frac{\bar{X} - \bar{Y}}{\hat{\sigma}}, \quad \text{ahol} \quad \hat{\sigma}^2 = \left(\frac{1}{n} + \frac{1}{m}\right) \cdot \frac{\sum (X_i - \bar{X})^2 + \sum (Y_i - \bar{Y})^2}{n + m - 2}$$

- T eloszlása H_0 mellett t_{n+m-2}
- t_α kritikus érték, az $1 - \alpha$ szinthez tartozó kvantilis, azaz $\mathbb{P}(T < t_\alpha) = 1 - \alpha$, ha $T \sim t_{n+m-2}$
- α terjedelmű próba. Kritikus tartomány:
 $H_1 : \mu_1 > \mu_2$ esetén $T > t_\alpha$ ($H_1 : \mu_1 < \mu_2$ esetén $T < -t_\alpha$)
 $H_1 : \mu_1 \neq \mu_2$ esetén $|T| > t_{\alpha/2}$
- Ezek a próbák egyenletesen legerősebbek a torzítatlan próbák között.

9.6.5 F -próba a szórások egyezésére

- $X_1, \dots, X_n \sim N(\mu_1, \sigma_1^2)$, $Y_1, \dots, Y_m \sim N(\mu_2, \sigma_2^2)$ minta. $\sigma_1, \sigma_2, \mu_1, \mu_2$ ismeretlen.

- Esetek:

$H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 > \sigma_2^2$ egyoldali ellenhipotézis

$H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \neq \sigma_2^2$ kétoldali ellenhipotézis

- Próba statisztika:

$$T = T(X, Y) = \frac{s_X^{*2}}{s_Y^{*2}}$$

- T eloszlása H_0 mellett $F_{a,b} = \frac{\chi_a^2/a}{\chi_b^2/b}$, $a = n - 1$, $b = m - 1$. $F_{a,b}$ neve a, b szabadságfokú F eloszlás.
- Az $f_\alpha(a, b)$ kritikus érték, az $1 - \alpha$ szinthez tartozó kvantilis, azaz $\mathbb{P}(T < f_\alpha) = 1 - \alpha$, ha $T \sim F_{a,b}$
- α terjedelmű próba. Kritikus tartomány:
 - $H_1 : \sigma_1^2 > \sigma_2^2$ esetén $T > f_\alpha(a, b)$,
 - $H_1 : \sigma_1^2 \neq \sigma_2^2$ esetén $T \notin (1/f_{\alpha/2}(b, a), f_{\alpha/2}(a, b))$
- $a = b$ esetén a próba egyenletesen legerősebb a torzítatlan próbák között.
Ha $a \neq b$, akkor a megadott elutasítási tartomány nem a legerősebb próbát adja, de ezt szokás használni.

F eloszlás

- Legyen X, Y független, $X \sim \chi_a^2 = \Gamma_{a/2, 1/2}$, $Y \sim \chi_b^2 = \Gamma_{b/2, 1/2}$.
- $F_{a,b}$ nem más, mint $(X/a)/(Y/b)$ eloszlása, vagyis ha $Z \sim F_{a,b}$, akkor

$$\frac{b}{b + aZ} = \frac{Y}{X + Y} \sim \beta\left(\frac{b}{2}, \frac{a}{2}\right),$$

- Ha $\beta \sim \beta\left(\frac{b}{2}, \frac{a}{2}\right)$, akkor

$$\frac{b}{a} \left(\frac{1}{\beta} - 1 \right) \sim F_{a,b}$$

- Ha $b \rightarrow \infty$, akkor $aF_{a,b} \xrightarrow{w} \chi_a^2$

9.7 Diszkrét illeszkedés vizsgálat, χ^2 -próbák

- X_1, \dots, X_n n elemű minta. $X_k \in \{1, \dots, r\}$.
- $\mathbb{P}_\pi(X_k = i) = \pi_i$, $\pi \in \Delta = \{\pi : \pi_i \geq 0, \sum \pi_i = 1\}$.
- $p \in \Delta$, $H_0 : \pi = p$, $H_1 : \pi \neq p$.
- $f_\pi(x) = \prod_{k=1}^n \mathbb{P}(X_k = x_k) = \prod_{k=1}^n \pi_{x_k} = \prod_{i=1}^r \pi_i^{\sum_k \mathbf{1}_{(X_k=i)}}$.
- Jelölje $N_i = \sum_k \mathbf{1}_{(X_k=i)}$, $i = 1, \dots, r$ az egyes értékek gyakoriságát. Maximum likelihood becslés π -re: $\ell(\pi) = \sum N_i \ln \pi_i$.
Feltételes szélsőérték (Δ -n maximalizálunk): $\ell'(\hat{\pi}) \perp \Delta$, azaz $\hat{\pi}_i = N_i/n$ a relatív gyakoriság.

- A próba statisztika

$$T_n = T_n(X) = \sum_{i=1}^r \frac{(N_i - np_i)^2}{np_i} = n \sum \frac{(\hat{\pi}_i - p_i)^2}{p_i}$$

- $n \rightarrow \infty$ esetén H_0 mellett $T_n \xrightarrow{d} \chi_{r-1}^2$, míg H_1 mellett $T_n \rightarrow \infty$. (később belátjuk.)
- Kritikus tartomány. $T_n > c_\alpha$, ahol c_α a χ_{r-1}^2 eloszlás kritikus értéke.
- Nem egzakt próba, mert T_n H_0 alatti eloszlásának közelítéséből számoljuk a kritikus értéket.

9.7.1 Speciális eset $r = 2$

- $r = 2$ esetén $p_2 = 1 - p_1$, $N_2 = n - N_1$, így

$$\begin{aligned} T_n &= \frac{(N_1 - np_1)^2}{np_1} + \frac{(N_2 - np_2)^2}{np_2} \\ &= \frac{(N_1 - np_1)^2}{n} \left(\frac{1}{p_1} + \frac{1}{1 - p_1} \right) \\ &= \left(\frac{N_1 - np_1}{\sqrt{np_1(1 - p_1)}} \right)^2, \end{aligned}$$

•

$$\frac{N_1 - np_1}{\sqrt{np_1(1 - p_1)}} \xrightarrow{d} N(0, 1) \quad \Rightarrow \quad T_n \xrightarrow{d} \chi_1^2 \quad \text{ha } n \rightarrow \infty$$

- $r = 2$ esetén lényegében az u -próbát kapjuk vissza, H_0 mellett, nagy n -re $N_1 = \sum_{k=1}^n \mathbf{1}_{(X_k=1)}$ közelítőleg normális eloszlású np_1 várható értékkel és $n(p_1(1 - p_1))$ szórásnégyzettel.

A $H_0 : \pi_1 = p_1$ nullhipotézist a $H_1 : \pi_1 \neq p_1$ ellenhipotézis ellenében elutasítjuk, ha

$$\frac{N_1 - np_1}{\sqrt{np_1(1 - p_1)}} > u_{\alpha/2}, \quad \text{másképp írva } T_n > u_{\alpha/2}^2$$

9.7.2 T_n határeloszlása

- Legyen $Y_k = (\mathbf{1}_{(X_k=1)}, \dots, \mathbf{1}_{(X_k=r)})$, ekkor Y_1, \dots, Y_n független azonos eloszlású változók.
- $\mathbb{E}_\pi(Y_k) = \pi$ és $\text{cov}_\pi(\mathbf{1}_{(X_k=i)}, \mathbf{1}_{(X_k=j)}) = \mathbb{P}_\pi(X_k = i, X_k = j) - \pi_i \pi_j = \delta_{i,j} \pi_i - \pi_i \pi_j$.
- A centrális határeloszlás tételből H_0 , azaz $\pi = p$ mellett

$$S_n = \frac{\sum_{k=1}^n Y_k - p}{\sqrt{n}} \rightarrow N(0, \Sigma(Y)), \quad \text{ahol } \sum_{k=1}^n Y_k - p = (N_1 - np_1, \dots, N_r - np_r)$$

Így $T_n = \sum_{i=1}^r \left(\frac{N_i - np_i}{\sqrt{np_i}} \right)^2$ nem más mint

$$T_n = \|DS_n\|^2, \quad \text{ahol } D = \text{diag}(p_1^{-1/2}, \dots, p_r^{-1/2}), \quad DS_n \xrightarrow{d} N(0, D\Sigma(Y)D)$$

Itt

$$(D\Sigma(Y)D)_{i,j} = \frac{\delta_{i,j} p_i - p_i p_j}{\sqrt{p_i p_j}} = \delta_{i,j} - \sqrt{p_i} \sqrt{p_j}$$

Ha $e = (p_1^{1/2}, \dots, p_r^{1/2})$, akkor $\|e\|_2 = \sum_i p_i = 1$

$$(D\Sigma(Y)D) = I - ee^T$$

- Legyen $Y_k = (\mathbf{1}_{(X_k=1)}, \dots, \mathbf{1}_{(X_k=r)})$, ekkor Y_1, \dots, Y_n független azonos eloszlású változók.
- A centrális határeloszlás tételből H_0 , azaz $\pi = p$ mellett

$$DS_n = \left(\frac{N_1 - np_1}{\sqrt{np_1}}, \dots, \frac{N_r - np_r}{\sqrt{np_r}} \right) \xrightarrow{d} N(0, I - ee^T), \quad \text{ahol } e = (p_1^{1/2}, \dots, p_r^{1/2})$$

- $I - ee^T$ merőleges vetítés \mathbb{R}^r ortokomplementerére, így $N(0, I - ee^T)$ nem más mint $(I - ee^T)Z$ eloszlása, ahol Z r dimenziós standard normális.
- Innen

$$T_n = \|DS_n\|^2 \xrightarrow{d} \|(I - ee^T)Z\|^2 \sim \chi_{r-1}^2$$

Emlékeztető. Ha $\xi_n \xrightarrow{d} \xi$, és h folytonos, akkor $h(\xi_n) \xrightarrow{d} h(\xi)$.

10. 2021.04.26.

10.1 Diszkrét illeszkedés vizsgálat, χ^2 -próbák

- X_1, \dots, X_n n elemű minta. $X_k \in \{1, \dots, r\}$.
- $\mathbb{P}_\pi(X_k = i) = \pi_i$, $\pi \in \Delta = \{\pi : \pi_i > 0, \sum \pi_i = 1\}$.
- $p \in \Delta$, $H_0 : \pi = p$, $H_1 : \pi \neq p$.
- $f_\pi(x) = \prod_{k=1}^n \mathbb{P}(X_k = x_k) = \prod_{k=1}^n \pi_{x_k} = \prod_{i=1}^r \pi_i^{\sum_k \mathbf{1}_{(X_k=i)}}$.
- Jelölje $N_i = \sum_k \mathbf{1}_{(X_k=i)}$, $i = 1, \dots, r$ az egyes értékek gyakoriságát. Maximum likelihood becslés π -re: $\ell(\pi) = \sum N_i \ln \pi_i$.
Feltételes szélsőérték (Δ -n maximalizálunk): $\ell'(\hat{\pi}) \perp \Delta$, azaz $\hat{\pi}_i = N_i/n$ a relatív gyakoriság.

- A próba statisztika

$$T_n = T_n(X) = \sum_{i=1}^r \frac{(N_i - np_i)^2}{np_i} = n \sum \frac{(\hat{\pi}_i - p_i)^2}{p_i}$$

- $n \rightarrow \infty$ esetén H_0 mellett $T_n \xrightarrow{d} \chi_{r-1}^2$, míg H_1 mellett $T_n \rightarrow \infty$. (később belátjuk.)
- Kritikus tartomány. $T_n > c_\alpha$, ahol c_α a χ_{r-1}^2 eloszlás kritikus értéke.
- Nem egzakt próba, mert T_n H_0 alatti eloszlásának közelítéséből számoljuk a kritikus értéket.

10.1.1 Speciális eset $r = 2$

- $r = 2$ esetén $p_2 = 1 - p_1$, $N_2 = n - N_1$, így

$$\begin{aligned} T_n &= \frac{(N_1 - np_1)^2}{np_1} + \frac{(N_2 - np_2)^2}{np_2} \\ &= \frac{(N_1 - np_1)^2}{n} \left(\frac{1}{p_1} + \frac{1}{1 - p_1} \right) \\ &= \left(\frac{N_1 - np_1}{\sqrt{np_1(1 - p_1)}} \right)^2, \end{aligned}$$

•

$$\frac{N_1 - np_1}{\sqrt{np_1(1 - p_1)}} \xrightarrow{d} N(0, 1) \quad \Rightarrow \quad T_n \xrightarrow{d} \chi_1^2 \quad \text{ha } n \rightarrow \infty$$

- $r = 2$ esetén lényegében az u -próbát kapjuk vissza, H_0 mellett, nagy n -re $N_1 = \sum_{k=1}^n \mathbf{1}_{(X_k=1)}$ közelítőleg normális eloszlású np_1 várható értékkel és $n(p_1(1 - p_1))$ szórásnégyzettel.

A $H_0 : \pi_1 = p_1$ nullhipotézist a $H_1 : \pi_1 \neq p_1$ ellenhipotézis ellenében elutasítjuk, ha

$$\left| \frac{N_1 - np_1}{\sqrt{np_1(1 - p_1)}} \right| > u_{\alpha/2}, \quad \text{másképp írva } T_n > u_{\alpha/2}^2$$

10.1.2 T_n határeloszlása

- Legyen $Y_k = (\mathbf{1}_{(X_k=1)}, \dots, \mathbf{1}_{(X_k=r)})$, ekkor Y_1, \dots, Y_n független azonos eloszlású változók.
- $\mathbb{E}_\pi(Y_k) = \pi$ és $\text{cov}_\pi(\mathbf{1}_{(X_k=i)}, \mathbf{1}_{(X_k=j)}) = \mathbb{P}_\pi(X_k = i, X_k = j) - \pi_i \pi_j = \delta_{i,j} \pi_i - \pi_i \pi_j$.
- A centrális határeloszlás tételből H_0 , azaz $\pi = p$ mellett

$$S_n = \frac{\sum_{k=1}^n Y_k - p}{\sqrt{n}} \rightarrow N(0, \Sigma(Y)), \quad \text{ahol} \quad \sum_{k=1}^n (Y_k - p) = (N_1 - np_1, \dots, N_r - np_r)$$

Így $T_n = \sum_{i=1}^r \left(\frac{N_i - np_i}{\sqrt{np_i}} \right)^2$ nem más mint

$$T_n = \|DS_n\|^2, \quad \text{ahol} \quad D = \text{diag}(p_1^{-1/2}, \dots, p_r^{-1/2}), \quad DS_n \xrightarrow{d} N(0, D\Sigma(Y)D)$$

Itt

$$(D\Sigma(Y)D)_{i,j} = \frac{\delta_{i,j}p_i - p_i p_j}{\sqrt{p_i p_j}} = \delta_{i,j} - \sqrt{p_i} \sqrt{p_j}$$

Ha $e = (p_1^{1/2}, \dots, p_r^{1/2})$, akkor $\|e\|_2 = \sum_i p_i = 1$

$$(D\Sigma(Y)D) = I - ee^T$$

•

$$DS_n = \left(\frac{N_1 - np_1}{\sqrt{np_1}}, \dots, \frac{N_r - np_r}{\sqrt{np_r}} \right) \xrightarrow{d} N(0, I - ee^T), \quad \text{ahol} \quad e = (p_1^{1/2}, \dots, p_r^{1/2})$$

- $I - ee^T$ merőleges vetítés \mathbb{R}^r ortokomplementerére, így $N(0, I - ee^T)$ nem más mint $(I - ee^T)Z$ eloszlása, ahol Z r dimenziós standard normális.
- Innen

$$T_n = \|DS_n\|^2 \xrightarrow{d} \|(I - ee^T)Z\|^2 \sim \chi_{r-1}^2$$

Emlékeztető. Ha $\xi_n \xrightarrow{d} \xi$, és h folytonos, akkor $h(\xi_n) \xrightarrow{d} h(\xi)$.

10.2 Emlékeztető

- Egy megfigyelés sűrűségfüggvénye f_θ , loglikelihoodja $\ell(\theta)$. $\theta \in \Theta$, ahol $\Theta \subset \mathbb{R}^p$ nyílt.
- Ha $X = (X_1, X_2, \dots)$ független megfigyelések az f_θ eloszlásból, akkor legyen

$$\ell_n(\theta) = \ell_n(\theta, X) = \sum_{i=1}^n \ell(\theta, X_i)$$

- $\hat{\theta}_n$ θ ML becslése az első n megfigyelés alapján. Likelihood egyenlet: $\ell'_n(\hat{\theta}_n) = 0$
- $\xi_n = \frac{1}{\sqrt{n}} \ell'_n(\theta) \xrightarrow{d} N(0, I(\theta))$, ha $\mathbb{E}_\theta(\ell'(\theta)) = 0$ és $I(\theta)$ létezik és véges, pl. (R) esetén.
- Ha (RR) is teljesül, pl. exponenciális családban, vagy annak részcsaládjában, akkor

$$-\xi_n = -\frac{1}{\sqrt{n}} \ell'_n(\theta) = \frac{1}{\sqrt{n}} (\ell'_n(\hat{\theta}_n) - \ell'_n(\theta)) = \left(\frac{1}{n} \ell''_n(\theta) + R_n \right) \sqrt{n}(\hat{\theta}_n - \theta)$$

ahol $\frac{1}{n} \ell''_n(\theta) \rightarrow \mathbb{E}(\ell''(\theta)) = -I(\theta)$ és $R_n \rightarrow 0$. Ebből $\sqrt{n}(\hat{\theta}_n - \theta) = I(\theta)^{-1} \xi_n + o_p(1)$

•

$$\begin{aligned}\ell_n(\hat{\theta}_n) - \ell_n(\theta) &= \frac{1}{\sqrt{n}} \ell'_n(\theta) \sqrt{n}(\hat{\theta}_n - \theta) + \frac{1}{2} \sqrt{n}(\hat{\theta}_n - \theta)^T \frac{1}{n} \ell''_n(\theta) \sqrt{n}(\hat{\theta}_n - \theta) + o_p(1) \\ &= \frac{1}{2} \xi_n^T I(\theta)^{-1} \xi_n + o_p(1) = \frac{1}{2} |I^{-1/2}(\theta) \xi_n|^2 + o_p(1) \xrightarrow{d} \frac{1}{2} \chi_p^2.\end{aligned}$$

Állítás 42. $\Theta \subset \mathbb{R}^p$ $\Xi \subset \mathbb{R}^q$ nyílt, $h : \Xi \rightarrow \Theta$ sima és h' teljes rangú. $\{\mathbb{P}_\theta : \theta \in \Theta\}$ exponenciális eloszlás család. X_1, X_2, \dots független megfigyelések $\mathbb{P}_{h(\tau)}$ eloszlásból $\theta = h(\tau)$, ℓ_n az első n megfigyelés loglikelihoodja, $\hat{\tau}_n = \arg \max_{\tau \in \Xi} \ell_n(h(\tau))$ a τ maximum likelihood becslése az első n megfigyelés alapján.

Ekkor

$$\ell_n(h(\hat{\tau}_n)) - \ell_n(\theta) = \frac{1}{2} |PI^{-1/2}(\theta) \xi_n|^2 + o(1)$$

ahol $\xi_n = \frac{1}{\sqrt{n}} \ell'_n(\theta)$ és P merőleges vetítés $I^{1/2}(\theta)h'(\tau)$ képterére.

- $\{\mathbb{P}_{h(\tau)} : \tau \in \Xi\}$ -re a korábbi számolás alapján

$$\ell_n(h(\hat{\tau}_n)) - \ell_n(h(\tau)) = \frac{1}{2} (\xi'_n)^T J(\tau)^{-1} \xi'_n, \quad \text{ahol} \quad (\xi'_n)^T = \frac{1}{\sqrt{n}} \frac{\partial}{\partial \tau} \ell_n(h(\tau)) = \frac{1}{\sqrt{n}} \ell'_n(\theta) h'(\tau)$$

- $H = h'(\tau)$ jelöléssel $J(\tau) = H^T I(\theta) H$ és

$$\ell_n(h(\hat{\tau}_n)) - \ell_n(\theta) = \frac{1}{2} \xi_n^T H (H^T I(\theta) H)^{-1} H^T \xi_n + o(1)$$

- $A = I^{1/2} H$ jelöléssel $I^{1/2}(\theta) H (H^T I(\theta) H)^{-1} H^T I^{1/2}(\theta) = A(A^T A)^{-1} A^T = P$ merőleges vetítés $A = I^{1/2} H$ képterére és

$$\ell_n(h(\hat{\tau}_n)) - \ell_n(\theta) = \frac{1}{2} \xi_n^T I^{-1/2} I^{1/2} H (H^T I(\theta) H)^{-1} H^T I^{1/2} I^{-1/2} \xi_n + o(1) = \frac{1}{2} |PI^{-1/2} \xi_n|^2 + o(1)$$

10.3 Likelihood hányados határeloszlása

Állítás 43. • $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ exponenciális család $\Theta \subset \mathbb{R}^p$ nyílt.

- $\Xi \subset \mathbb{R}^q$ nyílt. $h : \Xi \rightarrow \Theta$, h sima, $H = h'$, $H(\tau)$ rangja q .
 - X_1, X_2, \dots független megfigyelések \mathbb{P}_θ eloszlásból, $\theta = h(\tau)$
 - $\ell_n(\theta) = \sum_{k=1}^n \ell(\theta, X_k)$
 - $\hat{\theta}_n = \arg \max_{\theta \in \Theta} \ell_n(\theta)$ ML becslés θ -ra. $\hat{\tau}_n = \arg \max_{\tau \in \Xi} \ell_n(h(\tau))$ ML becslés τ -ra.
- Ekkor

$$2(\ell_n(\hat{\theta}_n) - \ell_n(h(\hat{\tau}_n))) \xrightarrow{d} \chi_{p-q}^2 \quad \text{ha } n \rightarrow \infty.$$

Legyen $\xi_n = \frac{1}{\sqrt{n}} \ell'_n(\theta)$ és P merőleges vetítés $I^{1/2} h'(\tau)$ képterére.

$$\begin{aligned}2(\ell_n(\hat{\theta}_n) - \ell_n(h(\hat{\tau}_n))) &= 2(\ell_n(\hat{\theta}_n) - \ell_n(\theta)) - 2(\ell_n(h(\hat{\tau}_n)) - \ell_n(\theta)) \\ &= |I^{-1/2}(\theta) \xi_n|^2 - |PI^{-1/2}(\theta) \xi_n|^2 + o(1) = |(\mathbf{I}_p - P)I^{-1/2} \xi_n|^2 + o(1) \xrightarrow{d} \chi_{p-q}^2,\end{aligned}$$

mert $I^{-1/2}(\theta) \xi_n \xrightarrow{d} N(0, \mathbf{I}_p)$ és $\mathbf{I}_p - P$ merőleges vetítés egy $p - q$ dimenziós altérre.

10.4 Diszkrét illeszkedés vizsgálat, T_n határeloszlása

- $\mathbb{P}_\pi(X_k = i) = \pi_i = \exp\{\ln p_i\}$, $i = 1, \dots, r$. Egy mintaelem eloszlása exponenciális családot alkot $T(x) = (\mathbf{1}_{(x=1)}, \dots, \mathbf{1}_{(x=r)})$ (one-hot encoding), a természetes paraméterezés $(\ln p_1, \dots, \ln p_r)$, azaz $\Theta' = \{t \in \mathbb{R}^r : \sum_{i=1}^r e^{t_i} = 1\}$. Ez egy $r - 1$ dimenziós sokaság \mathbb{R}^r -ben.

- Egy lehetséges paraméterezés $\Theta = \mathbb{R}^{r-1} = \mathbb{R}^r \times \{0\}$,

$$p_i = \frac{e^{\theta_i}}{1 + \sum_{j < r} e^{\theta_j}}, \quad i < r \quad \text{és} \quad p_r = \frac{1}{1 + \sum_{j < r} e^{\theta_j}}, \quad \ell(\theta, x) = \sum_{i < r} \theta_i \mathbf{1}_{(x=i)} - \ln(1 + \sum_{j < r} e^{\theta_j}).$$

azaz $b(\theta) = -\ln p_r$ és $\ln p_i = \theta_i - b(\theta)$ ($\theta_r = 0$). $b'(\theta) = (p_1(\theta), \dots, p_{r-1}(\theta))$.

- Ezzel a paraméterezéssel és a $\hat{p}_{n,i} = \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{(X_k=i)}$ jelöléssel

$$\ell_n(\theta) = \ln p(\theta) \cdot \sum_{k=1}^n T(X_k) = n \ln p(\theta) \cdot \hat{p}_n$$

$p(\hat{\theta}_n) = \hat{p}_n$, mert \hat{p}_n a maximum likelihood becslés.

•

$$\ell_n(\hat{\theta}_n) - \ell_n(\theta^*) = n(\ln p(\hat{\theta}_n) \cdot \hat{p}_n - \ln p(\theta^*) \cdot \hat{p}_n) = n \sum_{i=1}^r \hat{p}_{n,i} \ln \left(\frac{\hat{p}_{n,i}}{p_i} \right).$$

Állítás 44. p, q két eloszlás $\{1, \dots, r\}$ -en. Ekkor

$$\sum_i q_i \log \frac{q_i}{p_i} = \frac{1}{2} \sum_i \frac{(q_i - p_i)^2}{p_i} (1 + \alpha \max_i |\frac{q_i - p_i}{p_i}|) \quad \text{ahol } \alpha \in [-1, 1].$$

- $f(x) = x \log x$ -re $f(1) = 0$, $f'(1) = 1$ és $f''(x) = 1/x$, mert $(x \log x)' = 1 + \log x$, $(x \log x)'' = 1/x$.

$$\begin{aligned} x \log x &= f(x) = f(1) + f'(1)(x-1) + \frac{1}{2} f''(\xi)(x-1)^2 \\ &= x - 1 + \frac{1}{2\xi}(x-1)^2 \quad \text{ahol } \xi \text{ alkalmas közbülső pont 1 és } x \text{ között} \end{aligned}$$

- Ha a ξ pont 1 és x között van, akkor $\frac{1}{\xi} - 1 = \alpha |\frac{1}{x} - 1|$ valamely $\alpha \in [-1, 1]$ -re.

•

$$\begin{aligned} \sum_i q_i \log \frac{q_i}{p_i} &= \sum_i p_i f\left(\frac{q_i}{p_i}\right) = \sum_i p_i \left(\frac{q_i}{p_i} - 1\right) + \frac{1}{2} \sum_i p_i \left(\frac{q_i}{p_i} - 1\right)^2 (1 + \alpha_i |\frac{q_i}{p_i} - 1|) \\ &= \left(\frac{1}{2} \sum_i \frac{(q_i - p_i)^2}{p_i}\right) (1 + \alpha \max_i |\frac{q_i - p_i}{p_i}|) \quad \alpha \in [-1, 1] \end{aligned}$$

- $\hat{p} \rightarrow p$ esetén $\max_i |\frac{\hat{p}_{n,i} - p_i}{p_i}| \rightarrow 0$

- $2(\ell_n(\hat{p}) - \ell_n(p)) = n \sum_i \hat{p}_{n,i} \ln \frac{\hat{p}_{n,i}}{p_i} = n \sum_i \frac{(\hat{p}_{n,i} - p_i)^2}{p_i} + o(1) \xrightarrow{d} \chi_{r-1}^2$.

10.5 Függetlenség vizsgálat χ^2 -próbával

- Két szempont szerint osztályozunk, azaz a megfigyelések (X_k, Y_k) , $k = 1, \dots, n$.
- $X_k \in \{1, \dots, r\}$, $Y_k \in \{1, \dots, s\}$.

- Kérdés: független-e X és Y ? Azaz ha az együttes eloszlás π , $\pi_{i,\cdot} = \sum_j \pi_{i,j}$, $\pi_{\cdot,j} = \sum_i \pi_{i,j}$ akkor $H_0 : \pi_{i,j} = \pi_{i,\cdot} \pi_{\cdot,j}$, $H_1 : \exists i, j, \pi_{i,j} \neq \pi_{i,\cdot} \pi_{\cdot,j}$.

- Gyakoriságok $n_{i,j} = \sum_k \mathbf{1}_{(X_k=i, Y_k=j)}$, $n_{\cdot,j} = \sum_i n_{i,j}$, $n_{i,\cdot} = \sum_j n_{i,j}$

•

$$T_n = \sum_{i=1}^r \frac{(n_{i,j} - \frac{n_{i,\cdot} n_{\cdot,j}}{n})^2}{n \frac{n_{i,\cdot} n_{\cdot,j}}{n^2}} = n \sum_{i=1}^r \frac{(n_{i,j} - \frac{n_{i,\cdot} n_{\cdot,j}}{n})^2}{n_{i,\cdot} n_{\cdot,j}} \xrightarrow{d} \chi_{(r-1)(s-1)}^2$$

10.5.1 T_n határeloszlása

- $(\frac{n_{i,j}}{n})_{i,j}$ a $(\pi_{i,j})_{i,j}$ eloszlás maximum likelihood becslése az összes lehetséges eloszlás $rs - 1$ dimenziós családjában.
- $\frac{n_{i,\cdot} n_{\cdot,j}}{n^2}$ a maximum likelihood becslés a szorzat alakú eloszlások $(r-1) + (s-1)$ dimenziós családjában.
- A likelihood hányados logaritmus közelítőleg $rs - 1 - (r-1) - (s-1) = (r-1)(s-1)$ szabadságokú χ^2 eloszlású.
- A likelihood hányados logaritmus

$$\begin{aligned} 2 \left(\sum_{i,j} n_{i,j} \log \frac{n_{i,j}}{n} - \sum_{i,j} n_{i,j} \log \frac{n_{i,\cdot} n_{\cdot,j}}{n^2} \right) &= 2n \sum_{i,j} \hat{p}_{i,j} \log \frac{\hat{p}_{i,j}}{\hat{p}_{i,\cdot} \hat{p}_{\cdot,j}} \\ &= n \sum_{i,j} \frac{(\hat{p}_{i,j} - \hat{p}_{i,\cdot} \hat{p}_{\cdot,j})^2}{\hat{p}_{i,\cdot} \hat{p}_{\cdot,j}} + o(1) \\ &= n \sum_{i=1}^r \frac{(n_{i,j} - \frac{n_{i,\cdot} n_{\cdot,j}}{n})^2}{n_{i,\cdot} n_{\cdot,j}} = T_n + o(1) \end{aligned}$$

10.6 Homogenitás vizsgálat χ^2 -próbával

- X_1, \dots, X_n , $X_i \in \{1, \dots, r\}$, π eloszlásból.
- Y_1, \dots, Y_m , $Y_i \in \{1, \dots, r\}$, ρ eloszlásból.
- Gyakoriságok: $n_i = \sum_k \mathbf{1}_{(X_k=i)}$, $m_i = \sum_k \mathbf{1}_{(Y_k=i)}$.
- Kérdés $H_0 : \pi = \rho$, $H_1 : \pi \neq \rho$.
- Öntsük össze a megfigyeléseket és $Z_i = 1$ ha az i . megfigyelés az első mintából, $Z_i = 2$ ha a második mintából származik.
- Új kérdés független-e X, Z ?

•

$$n_{i,j} = \begin{cases} n_i & j=1 \\ m_i & j=2 \end{cases}, \quad n_{i,\cdot} = n_i + m_i, \quad n_{\cdot,1} = n, n_{\cdot,2} = m$$

•

$$\begin{aligned} T_n &= (n+m) \sum_{i,j} \frac{(n_{i,j} - \frac{n_{i,\cdot} n_{\cdot,j}}{n+m})^2}{n_{i,\cdot} n_{\cdot,j}} \\ &= (n+m) \sum_i \frac{(n_i - \frac{(n_i+m_i)n}{n+m})^2}{(n_i+m_i)n} + (n+m) \sum_i \frac{(m_i - \frac{(n_i+m_i)m}{n+m})^2}{(n_i+m_i)m} = nm \sum_i \frac{(\frac{n_i}{n} - \frac{m_i}{m})^2}{\frac{n_i}{n} + \frac{m_i}{m}} \xrightarrow{d} \chi_{r-1}^2 \end{aligned}$$

10.7 Becsléses illeszkedés vizsgálat

- Példa. X_1, X_2, \dots, X_n mintáról szeretnénk eldönteni, hogy Poisson eloszlásból származik-e.
- A mintából maximum likelihood módszerrel megbecsüljük a λ paramétert, majd a megfigyeléseket osztályokba soroljuk és χ^2 próbát végzünk.

$$T_n = \sum_i \frac{(n_i - np_i(\hat{\lambda}))^2}{np_i(\hat{\lambda})}$$

- T_n eloszlása közelítőleg χ^2 a szabadságfok $r - 1 - p$, ahol p a becsült paraméterek száma.
- A konkrét feladatban $p = 1$, T_n közelítőleg $r - 2$ szabadságfokú.

10.8 Megjegyzések χ^2 próbákhoz

- A próbastatisztika eloszlását a határeloszlással közelíti.
- Nagy mintára használható.
- Minden cellában legyen néhány megfigyelés.
- Ökölszabály: minden cella gyakoriság legalább 4-6, ha ez nem teljesül, akkor cella összevonás.

11. 2021.05.03.

11.1 Folytonos illeszkedés vizsgálat

- X_1, X_2, \dots, X_n minta F folytonos egy dimenziós eloszlásból.

- F_0 adott folytonos eloszlás függvény.

- $H_0 : F = F_0$, lehetséges alternatívák

$$H_1 : F \geq F_0, F \neq F_0$$

$$H_1^* : F \neq F_0$$

- Statisztikák:

$$D_n^+ = \sup_t F_n(t) - F_0(t), \quad D_n = \sup_t |F_n(t) - F_0(t)| \quad \text{Kolmogorov-Szmirnov}$$

$$\omega_n^2 = \int_{\mathbb{R}} (F_n(t) - F_0(t))^2 dF_0(t), \quad \text{Cramer-von Mises}$$

Állítás 45. A Kolmogorov-Szmirnov és Cramer-von Mises statisztikák eloszlásmentesek, azaz H_0 fennállása esetén eloszlásuk nem függ F_0 -tól.

11.1.1 Kolmogorov-Szmirnov statisztika eloszlásmentessége

- X_1, \dots, X_n n -elemű minta F_0 folytonos eloszlásfüggvényből.

- $X_0^* = -\infty < X_1^* < X_2^* < \dots < X_n^* < X_{n+1}^* = \infty$ egy valószínűséggel, mert $i \neq j$ -re $\mathbb{P}(X_i = X_j) = \mathbb{E}(\mathbb{P}(X_i = X_j | X_j)) = \mathbb{E}(\mathbb{P}(X_i = y) |_{y=X_j}) = 0$.

- (X_i^*, X_{i+1}^*) -n F_n konstans, F_0 monoton,

$$D_n^+ = \sup_t F_n(t) - F_0(t) = \max_{1 \leq i \leq n} (F_n(X_i) - F_0(X_i)) \vee (F_n(X_i+) - F_0(X_i))$$

- $F_n(X_i^*) = \frac{i-1}{n}$ és $F_n(X_i^*+) = \frac{i}{n}$. Azaz

$$D_n^+ = \max_i \left(\frac{i-1}{n} - F_0(X_i^*) \right) \vee \left(\frac{i}{n} - F_0(X_i^*) \right)$$

- $U_i = F_0(X_i)$ $(0, 1)$ -en egyenletes eloszlású minta, $F_0(X_i^*) = U_i^*$

- D_n^+ eloszlása tetszőleges F_0 esetén ugyanaz.

- Hasonlóan $D_n = \sup_t |F_n(t) - F(t)|$ -re.

11.1.2 Cramer-von Mises statisztika eloszlásmentes

- X_1, \dots, X_n n -elemű minta F_0 folytonos eloszlásfüggvényből.

$$\omega_n^2 = \int (F_n(t) - F_0(t))^2 dF_0(t)$$

- Ha X_0 a mintától független F_0 eloszlású

$$\omega_n^2 = \mathbb{E}((F_n(X_0) - F_0(X_0))^2 | X_1, \dots, X_n)$$

- $U_0 = F_0(X_0)$, $U_i = F_0(X_i)$ $i = 1, \dots, n$.

$$F_n^U(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(U_i < t)}$$

- Mivel $\mathbb{P}(U_0 \in \{U_1, \dots, U_n\}) = 0$, ezért $F_n(X_0) = F_n^U(U_0)$ és

$$\omega_n^2 = \mathbb{E}((F_n^U(U_0) - U_0)^2 | U_1, \dots, U_n)$$

azaz ω_n^2 eloszlása minden folytonos F_0 esetén ugyanaz.

Folytonos illeszkedés vizsgálat folyt

- X_1, X_2, \dots, X_n minta F folytonos egy dimenziós eloszlásból.
- F_0 adott folytonos eloszlás függvény.
- $H_0 : F = F_0$, lehetséges alternatívák
 $H_1 : F \geq F_0$, $F \neq F_0$
 $H_1^* : F \neq F_0$
- Statisztikák:

$$D_n^+ = \sup_t F_n(t) - F_0(t), \quad D_n = \sup_t |F_n(t) - F_0(t)| \quad \text{Kolmogorov-Szmirnov}$$

$$\omega_n^2 = \int_{\mathbb{R}} (F_n(t) - F_0(t))^2 dF_0(t), \quad \text{Cramer-von Mises}$$

- Kolmogorov-Szmirnov próba kritikus tartománya
 H_1 esetben $D_n^+ > c$ esetén elvetjük a nullhipotézist
 H_1^* esetben $D_n > c$ esetén elvetjük a nullhipotézist
- Cramer-von Mises próba: $\omega_n^2 > c$ esetén vetjük el H_0 -t.
- Nem túl nagy n -re léteznek táblázatok.
- A skálázott próba statisztikának létezik határeloszlása, általában a pontos eloszlást a határeloszlással közelítjük.

11.2 Skálázott próbastatisztikák határeloszlása

•

$$D_n^+ = \sup_t F_n(t) - F_0(t), \quad D_n = \sup_t |F_n(t) - F_0(t)|$$

•

$$\lim_{n \rightarrow \infty} \mathbb{P}(\sqrt{n} D_n^+ > x) = e^{-2x^2} = \frac{\phi(2x)}{\phi(0)},$$

ahol ϕ a standard normális sűrűség függvény.

Más szóval $n(D_n^+)^2 \xrightarrow{d} \exp\{2\}$.

•

$$\lim_{n \rightarrow \infty} \mathbb{P}(\sqrt{n}D_n < x) = \sum_{k \in \mathbb{Z}} (-1)^k e^{-2k^2 x^2} = \sum_{k \in \mathbb{Z}} (-1)^k \frac{\phi(2kx)}{\phi(0)}$$

•

$$n\omega_n^2 \xrightarrow{d} \sum_{k=1}^{\infty} \frac{Z_k^2}{k^2 \pi^2}, \quad \text{ahol } Z_1, Z_2, \dots \text{ független } N(0, 1)\text{-ek.}$$

11.2.1 Brown híd, mese

- Legyen U_1, \dots, U_n $(0, 1)$ -n egyenletes minta.

$$X^{(n)} = \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n (\mathbf{1}_{(U_i \leq t)} - t) \right)_{t \in [0, 1]} = \sqrt{n}(F_n(t) - t)_{t \in [0, 1]}$$

valószínűségi változók paraméterezett családja, azaz folyamat.

- Ha $\underline{t} = (t_1, \dots, t_r) \in [0, 1]^r$ akkor

$$X^{(n)}_{\underline{t}} = \frac{1}{\sqrt{n}} \left(\sum_{k=1}^n Y_k - \mathbb{E}(Y_k) \right), \quad \text{ahol } Y_k = (\mathbf{1}_{(U_k < t_1)}, \dots, \mathbf{1}_{(U_k < t_r)}).$$

$X^{(n)}_{\underline{t}} \xrightarrow{d} N(0, \Sigma(Y))$, ahol $\text{cov}(\mathbf{1}_{(U < s)}, \mathbf{1}_{(U < t)}) = \mathbb{P}(U < s, U < t) - \mathbb{P}(U < s)\mathbb{P}(U < t) = s \wedge t - st$.

- $t \mapsto X_t^n(\omega)$ jobbról folytonos, balról határértékkel rendelkező függvény (trajektória).

$$D[0, 1] = \{w : [0, 1] \rightarrow \mathbb{R} : w \text{ jobbról folytonos, balról határértékkel rendelkezik}\}$$

$D[0, 1]$ alkalmas metrikával teljes szeparábilis metrikus tér. $X^{(n)}$ véletlen elem ebből a térből.

- Megmutatható, hogy $X^{(n)}$ eloszlásban konvergens, az eloszlásbeli limesz neve Brown híd.

$(B_t)_{t \in [0, 1]}$ **Brown híd**, ha folytonos trajektóriájú Gauss folyamat, $\mathbb{E}(B_t) = 0$, $\text{cov}(B_t, B_s) = s \wedge t - st$.

- $\phi(w) = \max_s w(s), \max_s |w(s)|, \int_0^1 w^2(s) ds$ folytonos funkcionálok $D[0, 1]$ -en.

$$\sqrt{n}D_n^+ \xrightarrow{d} \max_{s \in [0, 1]} B_s, \quad \sqrt{n}D_n \xrightarrow{d} \max_{s \in [0, 1]} |B_s|, \quad n\omega_n^2 \xrightarrow{d} \int_0^1 B_s^2 ds, \quad \text{ahol } B \text{ Brown híd}$$

11.2.2 Skálazott Cramer–von Mises statisztika határeloszlása

- $(B_t)_{t \in [0, 1]}$ **Brown híd**, ha folytonos trajektóriájú Gauss folyamat, $\mathbb{E}(B_t) = 0$, $\text{cov}(B_t, B_s) = s \wedge t - st$.

$$X^{(n)} = \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n (\mathbf{1}_{(U_i \leq t)} - t) \right)_{t \in [0, 1]} = \sqrt{n}(F_n(t) - t)_{t \in [0, 1]}, \quad X^{(n)} \xrightarrow{d} B \quad \text{a } D[0, 1] \text{ metrikus térben.}$$

- $e_n(t) = \sqrt{2} \sin(k\pi t)$. $\{e_n : n \geq 1\}$ teljes ortonormált rendszer (bázis) az $L^2([0, 1])$ Hilbert térben.

$$n\omega_n^2 \stackrel{p}{=} \int_0^1 (X_s^{(n)})^2 ds \xrightarrow{d} \int_0^1 B_s^2 ds = \sum_{n=1}^{\infty} \left(\int_0^1 B_s e_n(s) ds \right)^2$$

Itt a $\xi_n = \int_0^1 B_s e_n(s) ds$ változók együttesen normálisak. $\mathbb{E}(\xi_n) = \int_0^1 \mathbb{E}(B_s) e_n(s) ds = 0$ és

$$\text{cov}(\xi_n, \xi_m) = \mathbb{E} \left(\int_0^1 \int_0^2 B_s e_n(s) B_t e_m(t) ds dt \right) = \int_{[0,1]^2} \text{cov}(B_s, B_t) e_n(s) e_m(t) ds dt = \mathbf{1}_{(n=m)} \frac{1}{n^2 \pi^2}$$

Azaz a ξ_n változók függetlenek és $Z_n = n\pi\xi_n$ független $N(0, 1)$ -ek sorozata.

- $n\omega_n^2 \xrightarrow{d} \sum_{n=1}^{\infty} \frac{Z_n^2}{n^2 \pi^2}$

11.2.3 Skálázott Kolmogorov–Szmirnov statisztika határeloszlása

- $(B_t)_{t \in [0,1]}$ **Brown híd**, ha folytonos trajektóriájú Gauss folyamat, $\mathbb{E}(B_t) = 0$, $\text{cov}(B_t, B_s) = s \wedge t - st$.

$$X^{(n)} = \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n (\mathbf{1}_{(U_i \leq t)} - t) \right)_{t \in [0,1]} = \sqrt{n}(F_n(t) - t)_{t \in [0,1]}, \quad X^{(n)} \xrightarrow{d} B \quad \text{a } D[0,1] \text{ metrikus térben.}$$

- $Z \sim N(0, 1)$ független a B Brown hídtól és $\beta_t = B_t + tZ$.

$$\text{cov}(\beta_s, \beta_t) = \text{cov}(B_s, B_t) + \text{cov}(sZ, tZ) = \min(s, t) = s \wedge t.$$

β **Brown mozgás**, azaz folytonos trajektóriájú Gauss folyamat nulla várható értékkel és $s \wedge t$ kovariancia függvényvel.

-

$$\mathbb{P} \left(\max_{s \in [0,1]} B_s \geq x \right) = \lim_{\varepsilon \rightarrow 0} \mathbb{P} \left(\max_{s \leq 1} B_s + sZ \geq x \mid |Z| < \varepsilon \right) = \lim_{\varepsilon \rightarrow 0} \mathbb{P} \left(\max_{s \leq 1} \beta_s \geq x \mid |\beta_1| < \varepsilon \right)$$

- $u < s < t$ esetén $\text{cov}(\beta_t - \beta_s, \beta_u) = 0$, azaz $\{\beta_u : u \in [0, s]\}$ és $\{\beta_t - \beta_s : t \in [s, 1]\}$ függetlenek. $\{\beta_t - \beta_s : t \in [s, 1]\} \stackrel{p}{=} \{-(\beta_t - \beta_s) : t \in [s, 1]\}$ és így

$$\{\beta_t : t \in [0, 1]\} \stackrel{p}{=} \{\tilde{\beta}_t = \beta_{t \wedge s} - (\beta_t - \beta_{t \wedge s}) : t \in [0, 1]\}$$

Ez a tükrözési elv determinisztikus s időpontra. Igaz marad akkor is, ha pl. $\tau = \inf\{s \in [0, 1] : \beta_s \geq x\}$.

- $Z \sim N(0, 1)$ független a B Brown hídtól és $\beta_t = B_t + tZ$. β Brown mozgás.

$$\mathbb{P} \left(\max_{s \in [0,1]} B_s \geq x \right) = \lim_{\varepsilon \rightarrow 0} \mathbb{P} \left(\max_{s \leq 1} \beta_s \geq x \mid |\beta_1| < \varepsilon \right)$$

- $x > 0$, $\tau = \inf\{s : \beta_s \geq x\}$. Ekkor $\tilde{\beta}_t = \beta_{t \wedge \tau} - (\beta_t - \beta_{t \wedge \tau}) = 2\beta_{t \wedge \tau} - \beta_t$ is Brown mozgás és

$$\mathbb{P} \left(\max_{s \leq 1} \beta_s \geq x, |\beta_1| < \varepsilon \right) = \mathbb{P} \left(\max_{s \leq 1} \tilde{\beta}_s \geq x, |\tilde{\beta}_1 - 2x| < \varepsilon \right) = \mathbb{P}(|\tilde{\beta}_1 - 2x| < \varepsilon)$$

- $\tilde{\beta}_1 \sim N(0, 1)$

$$\lim_{n \rightarrow \infty} \mathbb{P}(\sqrt{n}D_n^+ \geq x) = \mathbb{P} \left(\max_{s \leq 1} B_s \geq x \right) = \lim_{\varepsilon \rightarrow 0} \mathbb{P} \left(\max_{s \leq 1} \beta_s \geq x \mid |\beta_1| < \varepsilon \right) = \lim_{\varepsilon \rightarrow 0} \frac{\mathbb{P}(|\tilde{\beta}_1 - 2x| < \varepsilon)}{\mathbb{P}(|\tilde{\beta}_1| < \varepsilon)} = \frac{\phi(2x)}{\phi(0)}$$

11.3 Klasszikus próbák optimalitása

11.3.1 Emlékeztető. Neyman-Pearson lemma

- **Egyszerű** a hipotézis vizsgálati feladat, ha Θ_0, Θ_1 egyelemű. Ekkor $H_0 : f = f_0, H_1 : f = f_1$ alakú a feladat, ahol f_0, f_1 a két eloszlás $\mathbb{P}_0, \mathbb{P}_1$ sűrűségfüggvénye alkalmas domináló mértékre nézve (pl. $\lambda = \frac{1}{2}(\mathbb{P}_0 + \mathbb{P}_1)$).
- **Likelihood hányados** statisztika f_1/f_0

Lemma 13. *Egyszerű hipotézis vizsgálati feladat, $f_0 > 0$. Tegyük fel, hogy a likelihood hányados f_1/f_0 folytonos eloszlású H_0 és H_1 mellett is.*

Ekkor

- Minden $\alpha \in (0, 1)$ -re létezik c_α úgy, hogy $\phi = \mathbf{1}_{(f_1 > c_\alpha f_0)}$ terjedelme α . (Azaz $\mathbb{E}_0(\phi) = \alpha$)
- ϕ a legerősebb α terjedelmű próba.
- Ha $\alpha(\bar{\phi}) \leq \alpha$ és $\mathbb{E}_1(\bar{\phi}) = \mathbb{E}_1(\phi)$, akkor $\bar{\phi} = \phi$ ($\mathbb{P}_0, \mathbb{P}_1$ majdnem mindenütt).

11.4 Az u próba

11.4.1 Egyoldali ellenhipotézis mellett, az u -próba egyenletesen legerősebb

A feladat: X_1, \dots, X_n minta $N(\mu, \sigma^2)$ eloszlásból, σ ismert. $H_0 : \mu = \mu_0, H_1 : \mu > \mu_0$.

A sűrűségfüggvény

$$f_\mu(x) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2\right\} = h(x) \exp\{\mu \tilde{T}(x) - b(\mu)\}$$

ahol

$$h(x) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{1}{2\sigma^2} \sum_i x_i^2}, \quad b(\mu) = n \frac{\mu^2}{2\sigma^2}, \quad \tilde{T}(x) = \frac{\sum_i x_i}{\sigma^2}$$

- Elegendő a $\mu_0 = 0$ esetet nézni.
- Egyszerű hipotézis vizsgálati feladat ($0 < \mu_1$), $H_0 : \mu = 0, H_1 : \mu = \mu_1$.
- A likelihood hányados statisztika

$$\frac{f_1}{f_0}(x) = \exp\{(\mu_1 - 0)\tilde{T}(x) - (b(\mu_1) - b(0))\} \quad \tilde{T} \text{ szigorúan monoton növekvő függvénye}$$

- T eloszlása folytonos, azaz a legerősebb próba nem randomizált és

$$\mathfrak{X}_1 = \{\tilde{T} > c\} \quad \text{alakú, ahol} \quad \mathbb{P}_0(\tilde{T} > c) = \alpha$$

- X_1, \dots, X_n minta $N(\mu, \sigma^2)$ eloszlásból, σ ismert. Egyszerű hipotézis vizsgálati feladat ($0 < \mu_1$), $H_0 : \mu = 0, H_1 : \mu = \mu_1$.
- α terjedelmű likelihood hányados próba

$$\mathfrak{X}_1 = \{\tilde{T} > c\} \quad \text{alakú, ahol} \quad \tilde{T} = \frac{1}{\sigma^2} \sum_i X_i, \quad \mathbb{P}_0(\tilde{T} > c) = \alpha$$

- A likelihood hányados próba nem függ $\mu_1 \in (0, \infty)$ -től.

- A fenti próba egyenletesen legerősebb a $H_1 : \mu > 0$ ellenhipotézisre is.

Ugyanis, ha a ϕ döntési függvénnyel megadott próba terjedelme α és valamely $\mu_1 > 0$ pontban az ereje nem kisebb, mint a fenti próba ereje, akkor a Neyman–Pearson lemma miatt ϕ is likelihood hányados próba, vagyis azonos a fenti próbával.

- Az elutasítási tartomány

$$\left\{ \sum_i X_i > \sigma^2 c \right\} = \{T(X) > u_\alpha\}, \quad \text{ahol} \quad T(X) = \sqrt{n} \frac{\bar{X}}{\sigma}, \quad \mathbb{P}_0(T > u_\alpha) = \alpha.$$

$T(X) \sim N(0, 1)$ a $H_0 : \mu = 0$ hipotézis mellett.

11.4.2 $H_1 : \mu \neq 0$ -ra nincs egyenletesen legerősebb próba

- $H_0 : \mu = 0$, $H_1^* : \mu = \mu_1$ feladatra, a likelihood hányados próba a legerősebb.
- Mivel a likelihood hányados eloszlása H_0 mellett folytonos a likelihood hányados próba nem randomizált, így egyértelmű.
- Ha ϕ egyenletesen legerősebb lenne, akkor minden $\mu_1 \neq 0$ -ra egybeesne a H^* feladathoz tartozó likelihood hányados próbával.
- $\mu_1 > 0$ esetén $\phi(x) = \mathbf{1}_{(T(x) > u_\alpha)}$
- $\mu_1 < 0$ esetén $\phi(x) = \mathbf{1}_{(T(x) < -u_\alpha)}$
- Ezek egyszerre nem teljesülhetnek, nincs ilyen ϕ és nincs egyenletesen legerősebb próba.

11.4.3 $H_0 : \mu = 0$, $H_1 : \mu \neq 0$ -ra, az u -próba egyenletesen legerősebb torzítatlan próba

Θ konvex, nyílt és (R) teljesül, $\Theta_0 = \{\theta_0\}$, ϕ torzítatlan. Ekkor $\psi(\theta) = \mathbb{E}_\theta(\phi)$ folytonosan differenciálható, minimumhely $\theta = \theta_0$ -ban van és $\mathbb{E}_{\theta_0}(\ell'(\theta_0)) = 0$.

$$\text{cov}_{\theta_0}(\phi, \ell'(\theta_0)) = \mathbb{E}_{\theta_0}(\phi \ell'(\theta_0)) = \psi'(\theta_0) = 0.$$

Állítás 46 (Bizonyítás később). $\Theta \subset \mathbb{R}$ nyílt, $\{\mathbb{P}_\theta : \theta \in \Theta\}$ exponenciális eloszláscsalád.

$$\phi = \mathbf{1}_{(\ell'(0) \notin [a_1, a_2])}, \quad \Phi = \{\bar{\varphi} : \mathbb{E}_0(\bar{\varphi}) = \mathbb{E}_0(\phi), \quad \mathbb{E}_0(\bar{\varphi} \ell'(0)) = 0\}$$

Ha $\phi \in \Phi$, akkor ϕ egyenletesen legerősebb próba Φ -ben a $H_0 : \theta = 0$, $H_1 : \theta \neq 0$ feladatra és $a_1 < 0 < a_2$.

- $\{N(\mu, \sigma^2) : \mu \in \mathbb{R}\}$, $\sigma^2 > 0$ rögzített, esetén ($\theta = \mu$)

$$\ell(\mu, x) = \ln h(x) + \mu \tilde{T}(x) - b(\mu), \quad \text{ahol} \quad \tilde{T}(x) = \frac{\sum_i x_i}{\sigma^2}, \quad b(\theta) = \frac{n\mu^2}{2\sigma^2} \quad \ell'(\mu) = \frac{n}{\sigma^2}(\bar{X} - \mu)$$

- u próba próba statisztikája: $T = \frac{\sqrt{n}}{\sigma} \bar{X}$, döntési függvénye: $\phi = \mathbf{1}_{(|T| > u_{\alpha/2})}$. ϕ a fenti alakú.
- ϕ torzítatlan: \mathbb{P}_μ alatt $T \sim N(\sqrt{n} \frac{\mu}{\sigma}, 1)$, a $\psi(\theta) = \mathbb{E}_\theta(\phi)$ jelöléssel, ha $Z \sim N(0, 1)$

$$\psi(\theta) = \mathbb{P}\left(-u_{\alpha/2} - \frac{\sqrt{n}\mu}{\sigma} < Z < +u_{\alpha/2} - \frac{\sqrt{n}\mu}{\sigma}\right), \quad \psi'(\theta) = c\left(f_Z\left(u_{\alpha/2} - \frac{\sqrt{n}\mu}{\sigma}\right) - f_Z\left(-u_{\alpha/2} - \frac{\sqrt{n}\mu}{\sigma}\right)\right)$$

Állítás 47. $\Theta \subset \mathbb{R}$ nyílt, $\{\mathbb{P}_\theta : \theta \in \Theta\}$ exponenciális eloszláscsalád.

$$\phi = \mathbf{1}_{(\ell'(0) \notin [a_1, a_2])}, \quad \Phi = \{\bar{\varphi} : \mathbb{E}_0(\bar{\varphi}) = \mathbb{E}_0(\phi), \quad \mathbb{E}_0(\bar{\varphi} \ell'(0)) = 0\}.$$

Ha $\phi \in \Phi$, akkor ϕ egyenletesen legerősebb próba Φ -ben a $H_0 : \theta = 0$, $H_1 : \theta \neq 0$ feladatra.

Lemma 14 (Bizonyítás később). $\theta_1 < 0 < \theta_2$, $p \in (0, 1)$. $H_0 : f = f_0$, $H_1^* : f = pf_{\theta_1} + (1-p)f_{\theta_2}$.

Ekkor létezik $p = p(\mu_1, \mu_2)$ úgy, hogy H^* -hoz ϕ a legerősebb a vele azonos terjedelmű próbák között és

$$\exists \lim_{\theta_2 \rightarrow 0} \frac{p(\theta_1, \theta_2)}{\theta_2} \in (0, \infty), \quad \lim_{\theta_1 \rightarrow 0} \frac{1-p(\theta_1, \theta_2)}{\theta_1} \in (0, \infty).$$

- Ezt felhasználva, ha $\bar{\phi} \in \Phi$, akkor

$$\lim_{\theta_2 \rightarrow 0} \frac{1}{\theta_2} (\mathbb{E}_{\theta_2}(\bar{\phi}) - \mathbb{E}_0(\bar{\phi})) = \mathbb{E}_0(\bar{\phi}'(0)) = 0, \quad \lim_{\theta_2 \rightarrow 0} \frac{1}{\theta_2} (\mathbb{E}_{\mu_2}(\phi) - \mathbb{E}_0(\phi)) = \mathbb{E}_0(\phi'(0)) = 0$$

és

$$\frac{1}{p(\theta_1, \theta_2)} \mathbb{E}_{\theta_2}(\bar{\phi} - \phi) \rightarrow 0, \quad \text{ha } \theta_2 \rightarrow 0 \text{ és } \mathbb{E}_0(\bar{\phi}) = \mathbb{E}_0(\phi)$$

- Mivel ϕ legerősebb próba H^* -ra

$$p\mathbb{E}_{\theta_1}(\bar{\phi}) + (1-p)\mathbb{E}_{\theta_2}(\bar{\phi}) \leq p\mathbb{E}_{\theta_1}(\phi) + (1-p)\mathbb{E}_{\theta_2}(\phi), \quad \implies \mathbb{E}_{\theta_1}(\phi - \bar{\phi}) \geq \frac{1-p(\theta_1, \theta_2)}{p(\theta_2)} \mathbb{E}_{\theta_2}(\bar{\phi} - \phi) \rightarrow 0.$$

Lemma bizonyítása

- $\theta_1 < 0 < \theta_2$. $H_0 : f = f_0$, $H_1^* : f = pf_1 + (1-p)f_2$, ahol

$$f_i = f_{\theta_i}, \quad f_{\theta}(x) = h(x)e^{\theta \cdot T(x) - b(\theta)},$$

Feltehető, hogy $T = \ell'(0)$.

- A likelihood hányados

$$\frac{pf_1 + (1-p)f_2}{f_0}(x) = pe^{\theta_1 t - b(\theta_1)} + (1-p)e^{\theta_2 t - b(\theta_2)}|_{t=T(x)} = g(T(x))$$

Ez a t -nek konvex függvénye, $\lim_{t \rightarrow \pm\infty} = \infty$

- A kritikus tartomány $\{T \notin [a_1, a_2]\}$ alakú $a_1 < 0 < a_2$, ha

$$g(a_1) = g(a_2), \quad \text{azaz } p(e^{\theta_1 a_1} - e^{\theta_1 a_2})e^{-b(\theta_1)} = (1-p)(e^{\theta_2 a_2} - e^{\theta_2 a_1})e^{-b(\theta_2)}$$

Ennek megoldása p -re $p(\theta_1, \theta_2)$. Ha θ_1 rögzített és $\theta_2 \rightarrow 0$, akkor $p(\theta_1, \theta_2) \rightarrow 0$

- Ha θ_2 -vel osztunk és $\theta_2 \rightarrow 0$, akkor

$$\lim_{\theta_2 \searrow 0} \frac{p(\theta_1, \theta_2)}{\theta_2} (e^{\theta_1 a_1} - e^{\theta_1 a_2})e^{-b(\theta_1)} = \lim_{\theta_2 \searrow 0} (1-p(\theta_1, \theta_2))e^{-b(\theta_2)} \left(\frac{e^{\theta_2 a_2} - 1}{\theta_2} - \frac{e^{\theta_2 a_1} - 1}{\theta_2} \right) = e^{-b(0)}(a_2 - a_1)$$

11.5 Egyenletesen legerősebb próbák zavaró paraméter mellett

- t -próba esetét nézzük részletesebben.
- $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ minta, csak μ érdekel minket.
- Elég a $H_0 : \mu = 0$ esetet nézni.
- Az eloszláscsalád dominált

$$f_{\mu, \sigma^2}(x) = \exp \left\{ -\frac{1}{2\sigma^2} \sum (x_i - \mu)^2 + c(\sigma) \right\} = \exp \left\{ -\frac{1}{2\sigma^2} \sum x_i^2 + \frac{\mu}{\sigma^2} \sum x_i + c(\mu, \sigma) \right\}$$

- Átparaméterezés: $\theta = \frac{n\mu}{\sigma^2}$, $T(x) = \bar{x}$, $\tau = -\frac{n}{2\sigma^2}$, $S(x) = \frac{1}{n} \sum x_i^2 = s_n^2 + T^2$.

- Átparaméterezés után a kérdés $H_0 : \theta = 0$, $H_1 : \theta > 0$ ($H_1 : \theta < 0$) vagy kétoldali ellenhipotézis esetében $H_1 : \theta \neq 0$.
- Az átparaméterezett család sűrűségfüggvénye:

$$f_{\theta,\tau}(x) = \exp\{\theta T(x) + \tau S(x) + b(\theta, \tau)\}$$

alakú.

- A paraméterter eredetileg $\mathbb{R} \times (0, \infty)$, átparaméterezés után $\mathbb{R} \times (-\infty, 0)$.
- T, S elégséges statisztika. b a kumuláns generáló függvénnyel van kapcsolatban.
- Átparaméterezés után a kérdés $H_0 : \theta = 0$, $H_1 : \theta > 0$ ($H_1 : \theta < 0$) vagy kétoldali ellenhipotézis esetében $H_1 : \theta \neq 0$.
- Az átparaméterezett család sűrűségfüggvénye:

$$f_{\theta,\tau}(x) = \exp\{\theta T(x) + \tau S(x) + b(\theta, \tau)\}$$

alakú.

- T, S elégséges statisztika. b a (T, S) kumuláns generáló függvényével van kapcsolatban.
- Tetszőleges ϕ döntési függvényre $\mathbb{E}(\phi|(T, S))$ ugyanolyan jó próbát ad, vagyis elég a (T, S) statisztika eloszlásainak a családját nézni.
- Ha \mathcal{P} dominál mértéke \mathbb{P}_0 (valószínűségi mérték), akkor (T, S) eloszlásainak a családját pl. $Q_0 = \mathbb{P}_0 \circ (T, S)^{-1}$ dominálja és a sűrűségfüggvények

$$g_{\theta,\tau}(t, s) = \exp\{\theta t + \tau s - b(\theta, \tau)\}$$

alakúak.

- Q_0 mérték a $\mathbb{R} \times \mathbb{R}^p$ téren, de nem feltétlenül szorzat mérték!
- Nézzük a feladatot az S -re vett feltétel mellett.

11.5.1 (T, S) feltételes eloszlásai

- \mathbb{P}_0 dominál mérték X eloszlásaira, $Q_0 = \mathbb{P}_0 \circ (T, S)^{-1}$.
- T, S eloszlása $\mathbb{P}_{\theta,\tau}$ mellett abszolút folytonos Q_0 -ra nézve, a sűrűségfüggvény

$$g_{\theta,\tau}(t, s) = \exp\{\theta t + \tau s + b(\theta, \tau)\}$$

- Legyen $Q(H, S)$ a $\mathbb{P}_0(T \in H|S)$ feltételes eloszlás reguláris változata.

$$Q_0(A \times B) = \mathbb{P}_0(T \in A, S \in B) = \mathbb{E}_0(\mathbb{P}_0(T \in A|S) \mathbf{1}_{(S \in B)}) = \int_B Q(A, s) \tilde{Q}(ds), \quad \text{ahol} \quad \tilde{Q} = \mathbb{P}_0 \circ S^{-1}$$

- Ezzel a jelöléssel

$$\begin{aligned} \mathbb{P}_{\theta,\tau}(T \in B, S \in A) &= \int \mathbf{1}_{(t \in B)} \mathbf{1}_{(s \in A)} e^{\theta t + \tau s + b(\theta, \tau)} Q_0(dt, ds) \\ &= \int_A \int_B e^{\theta t} Q(dt, s) e^{\tau s + b(\theta, \tau)} \tilde{Q}(ds) = \mathbb{E}_{\theta,\tau}(h(B, S)) \end{aligned}$$

ahol

$$h(B, s) = \int_B e^{\theta t + \tilde{b}(\theta, s)} Q(dt, s), \quad \text{és} \quad e^{-\tilde{b}(\theta, s)} = \int e^{\theta t} Q(dt, s)$$

- Összefoglalva, $\mathbb{P}_{\theta,\tau}(T \in H|S)$ feltételes eloszlás reguláris változata $S = s$ mellett csak θ -tól függ és egy paraméteres exponenciális családot alkot $Q(dt, s)$ domináló mértékkel

Egyenletesen legerősebb próbák zavaró paraméter mellett folyt.

- $H_0 : \theta = 0, H_1 : \theta > 0$. $\phi = \phi(S, T)$ torzítatlan α terjedelmű próba. Ekkor $\psi(\theta, \tau) = \mathbb{E}_{\theta, \tau}(\phi)$ folytonos és a torzítatlanság miatt $\psi(0, \tau) = \alpha$ minden τ -ra.
- A $\mathbb{P}_{0, \tau}$ egy paraméteres család is exponenciális S teljes és elégséges statisztikával

$$f_{0, \tau}(x) = \exp\{\tau S(x) + b(0, \tau)\}$$

$\{(0, \tau) : (0, \tau) \in \Theta\}$, a Θ_0, Θ_1 közös határa.

- Ha S teljes a $\{\mathbb{P}_{0, \tau} : \tau\}$ családra, akkor azt mondjuk, S a **határon teljes**.
- Mivel S teljes a határon, ezért

$$\mathbb{E}_{0, \tau}(\mathbb{E}_{0, \tau}(\phi(T, S)|S) - \alpha) = 0 \implies \mathbb{E}_{0, \tau}(\phi(T, S)|S) = \alpha$$

- Ha ϕ torzítatlan α terjedelmű, akkor $\mathbb{E}_{(0, \tau)}(\phi) = \alpha$ minden τ -ra.
 S teljessége miatt $\phi(T, S)$ α terjedelmű a feltételes változatra is, azaz a $\mathbb{P}_{\theta, \tau}(T \in H|S)$ eloszlás családra.
- A feltételes változatra a likelihood hányados

$$\frac{e^{\theta_1 t + \tilde{b}(\theta_1, s)}}{e^{\tilde{b}(0, s)}} \quad t = T \text{ monoton növekvő függvénye}$$

Egyoldali ellenhipotézis esetén a feltételes változatra a legerősebb próba alakja

$$\phi(T, S) = \mathbf{1}_{(T > c(S))}$$

11.5.2 A t -próba egyenletesen legerősebb a torzítatlan próbák között

- $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ minta, csak μ érdekel minket.
- A $\theta = \frac{n\mu}{\sigma^2}$, $T(x) = \bar{x}$, $\tau = -\frac{n}{2\sigma^2}$, $S(x) = \frac{1}{n} \sum x_i^2 = s_n^2 + T^2$ átparaméterezés után

$$f_{\mu, \sigma^2}(x) = f_{\theta, \tau}(x) = \exp\{\theta T(x) + \tau S(x) + b(\theta, \tau)\} \quad H_0 : \theta = 0, H_1 : \theta > 0.$$

- Áttértünk a $\mathbb{P}_{\theta, \tau}(T \in H|S)$ feltételes eloszlásokra. Ha $\phi = \phi(T, S)$ torzítatlan α terjedelmű, akkor α terjedelmű a feltételes feladatra is $\mathbb{E}_{0, \tau}(\phi(T, S)|S) = \alpha$.
- A feltételes feladatban egy paraméteres család marad $\mathbb{P}_{\theta, \tau}(T \in H|S)$ nem függ τ -tól. A likelihood hányados T monoton növekvő függvénye az egyenletesen legerősebb próba $\phi(t, s) = \mathbf{1}_{(t > c(s))}$ alakú.
- A t próba $\phi(T, S) = \mathbf{1}_{(T/\sqrt{S-T^2} > c)}$ torzítatlan, mert T/s_n eloszlása minden $(0, \tau)$ párra ugyanaz.
- Mivel $c \geq 0$ -ra

$$T > c\sqrt{S-T^2} \iff T^2 > c^2(S-T^2), T > 0 \iff T > \frac{c}{\sqrt{1+c^2}}\sqrt{S} = \bar{c}\sqrt{S}$$

$\phi = \mathbf{1}_{(T > \bar{c}\sqrt{S})}$ alakú és egyenletesen legerősebb a feltételes feladatra. Hasonlóan $c < 0$ -ra

- Ha $\bar{\phi}$ tetszőleges α terjedelmű torzítatlan. Feltehető, hogy $\bar{\phi} = \phi(T, S)$ és ekkor

$$\mathbb{E}_{\theta, \tau}(\phi - \bar{\phi}) = \mathbb{E}_{\theta, \tau}(\mathbb{E}_{\theta, \tau}(\phi - \bar{\phi}|S)) \geq 0 \quad \forall \theta, \tau$$

12. 2021.05.10

12.1 Lineáris regresszió, lineáris modell

- (x_i, y_i) , $i = 1, \dots, n$ megfigyelések. Van-e valamilyen függvénykapcsolat x és y között?
Pl. y jövedelem, x életkor. y lakásár, x a lakás jellemzői.
- y neve függő, vagy eredmény változó, x a független vagy magyarázó változó.
Becsüljük az $f \in \mathcal{F}$ függvényt az (x_i, y_i) adatok alapján, ha $y_i = f(x_i) + \varepsilon_i$, ahol ε_i zaj, azaz független $N(0, \sigma^2)$ változók és σ^2 ismeretlen.
- \mathcal{F} általában véges dimenziós vektortér, ha f_1, \dots, f_p bázis \mathcal{F} -ben, és $\underline{x}_i = (f_1(x_i), \dots, f_p(x_i))$, akkor \mathcal{F} tetszőleges eleme $\sum_i \beta_i f_i$ alakú és β meghatározása a feladat.
- Maximum likelihood becslés β -ra és σ^2 -re

$$\ell(\beta, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \underline{x}_i \cdot \beta)^2$$

Vektor alakban y az y_i értékekből alkotott (oszlop) vektor. Az X mátrix i . sora \underline{x}_i .

$$\ell(\beta, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \|y - X\beta\|^2$$

- $\ell(\beta, \sigma^2)$ maximális, ha $\|y - X\beta\|^2$ minimális és $\sigma^2 = \frac{1}{n} \|y - X\hat{\beta}\|^2$
- β maximum likelihood becslése a négyzetes veszteség $\|y - X\beta\|^2$ minimalizálásával kapható. Legkisebb négyzetes becslésnek is hívják (LNB)

12.1.1 Példa (lineáris regresszió)

- A magyarázó változó lineáris (affin) függvényei között keressük a függvénykapcsolatot:
 $f(x) = \beta_0 + x\beta_1$.

$$\|y - X\beta\|^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2 = n(\bar{y} - \beta_1 \bar{x} - \beta_0)^2 + \sum (y_i - \bar{y} - \beta_1(x_i - \bar{x}))^2$$

- $\hat{\beta}_1 = \frac{\frac{1}{n} \sum_i (y_i - \bar{y})(x_i - \bar{x})}{\frac{1}{n} \sum_i (x_i - \bar{x})^2}$, $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$.
- Geometriai kép: y és x -ből levonjuk az átlagot, feltehető, hogy $\bar{x} = 0$, $\bar{y} = 0$. Az $y \in \mathbb{R}^n$ vektort vetítjük az $x \in \mathbb{R}^n$ által kifeszített egy dimenziós altérre. A vetület $\frac{\langle x \rangle y}{\|x\|^2} x$ és $\hat{\beta}_1 = \frac{\langle x \rangle y}{\|x\|^2}$.
- $\hat{\sigma}^2 = \frac{1}{n} \|y - \hat{\beta}_0 - \hat{\beta}_1 x\|^2$ nem torzítatlan σ^2 -re.

Itt $\hat{\beta}_0 + \hat{\beta}_1 x$ az $y \in \mathbb{R}^n$ vektor vetülete az x és a konstans 1 vektor által kifeszített kétdimenziós altérre. P a merőleges vetítés erre az altérre:

$$Py = P(\beta_0 + \beta_1 x + \varepsilon) = \beta_0 + \beta_1 x + P\varepsilon, \quad \|y - \hat{\beta}_0 - \hat{\beta}_1 x\|^2 = \|(I - P)\varepsilon\|^2, \quad \|y - \hat{\beta}_0 - \hat{\beta}_1 x\|^2 \sim \sigma^2 \chi_{n-2}^2$$

$\frac{1}{n-2} \|y - \hat{\beta}_0 - \hat{\beta}_1 x\|^2$ torzítatlan σ^2 -re.

12.2 Lineáris modell

- Általánosabban. X $n \times p$ mátrix, a mátrix i . sora x_i : a magyarázó változók értéke az i . megfigyelésben.

$$y_i = x_i \cdot \beta + \varepsilon_i, \quad \text{ahol } \varepsilon_i \sim N(0, \sigma^2) \text{ függetlenek}$$

Vektor alakban $y = X\beta + \varepsilon$.

X első oszlopa lehet 1, akkor β_1 a konstans hatás (intercept).

- β maximum likelihood becslése: $\hat{\beta} = \arg \min_{\beta} \|y - X\beta\|^2$ és σ^2 maximum likelihood becslése $\hat{\sigma}^2 = \frac{1}{n} \|y - X\hat{\beta}\|^2$

- $\hat{y} = X\hat{\beta}$ az y vetülete X képterére. $\text{im } X \perp \ker X^T$, azaz $X^T(y - X\hat{\beta}) = 0$, azaz $X^T y = (X^T X)\hat{\beta}$ és $\beta = (X^T X)^{-1} X^T y$, ha az inverz létezik, azaz X rangja p .

Ha $(X^T X)$ nem invertálható, akkor $\hat{\beta}$ nem egyértelmű, de $\hat{y} = X\hat{\beta}$ igen. Ilyenkor a legkisebb normájú $\hat{\beta}$ -t szokás venni, ami $(X^T X)$ Moore-Penrose inverzével kapható.

- $\hat{\beta}$ a likelihood egyenlet megoldásából is megkapható. A likelihood egyenlet β -ra:

$$\ell(\beta, \sigma^2) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - x_i \cdot \beta)^2, \quad \partial_{\beta} -\frac{1}{2} \|y - X\beta\|^2 = \sum_{i=1}^n (y_i - x_i \beta) x_i = (y - X\beta)^T X = 0$$

12.3 $\hat{\beta}$, $\hat{\sigma}^2$ eloszlása

- n megfigyelés, p magyarázó változó. X $n \times p$ méretű mátrix, a sorai a magyarázó változókból álló vektorok.

$$y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 \mathbf{I}_n)$$

Ha $X^T X$ invertálható, azaz X rangja p , akkor $\hat{\beta} = (X^T X)^{-1} X^T y$.

- $\hat{\beta}$ normális eloszlású és

$$\mathbb{E}(\hat{\beta}) = (X^T X)^{-1} X^T \mathbb{E}(X\beta + \varepsilon) = \beta,$$

$$\Sigma(\hat{\beta}) = \Sigma((X^T X)^{-1} X^T y) = (X^T X)^{-1} X^T \Sigma(y) X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1}.$$

- σ^2 becslése a **rezidulis négyzet összeg** (RSS) $\|y - X\hat{\beta}\|^2$ alapul, ahol $X\hat{\beta} = \hat{y}$ az y vetülete X képterére.

$y - \hat{y}$ nem más mint y vetülete X képterének ortokomplementerére. $y = X\beta + \varepsilon$ miatt $y - \hat{y}$ az ε zaj vektor merőleges vetülete X képterének ortokomplementerére. $\|y - \hat{y}\|^2 \sim \sigma^2 \chi_{\dim(\text{im } X)^\perp}^2 = \sigma^2 \chi_{n-p}^2$.

$\hat{\sigma}^2 = \frac{1}{n-p} \|y - X\hat{\beta}\|^2$ torzítatlan becslés σ^2 -re.

- $\hat{\beta}$ és $\hat{\sigma}^2$ függetlenek, mert $\hat{\beta}$ -t y $\text{im } X$ -re vett vetületéből, míg $\hat{\sigma}^2$ -et y $\text{im } X^\perp$ -ra vett vetületéből számoljuk.

12.4 Hipotézis vizsgálat normális lineáris modellben

•

$$y \sim N(X\beta, \sigma^2 \mathbf{I}_n), \quad \beta \in \mathbb{R}^p, \sigma > 0$$

- Kérdés: Van-e fölösleges magyarázó változó, azaz olyan j , hogy $\beta_j = 0$, vagy általánosabb igaz-e, hogy $C\beta = 0$.

- $H_0 : C\beta = 0, H_1 : C\beta \neq 0$.
- $\mathbb{R}^n = \mathcal{L}^R \oplus \mathcal{L}^1 \oplus \mathcal{L}^0$ három merőleges komponensre bomlik, ahol $\mathcal{L}^R = \text{im } X^\perp$, $\mathcal{L}^0 = \{X\beta : C\beta = 0\}$ és $\mathcal{L}^1 = \text{im } X \cap (\mathcal{L}^0)^\perp$.
- H_0 fenállása esetén $X\beta \in \mathcal{L}^0$ és y vetülete $\mathcal{L}^R \oplus \mathcal{L}^1$ -re azonos $\varepsilon = y - X\beta \sim N(0, \sigma^2 \mathbf{I}_n)$ vetületével.
- Azaz H_0 mellett $y = y^R + y^1 + X\hat{\beta}_0$, ahol $\hat{\beta}_0 = \arg \min_{\beta: C\beta=0} \|y - X\beta\|^2$ és $y^1 \in \mathcal{L}^1$, $y^R \in \mathcal{L}^R$.
- $\mathcal{L}^1, \mathcal{L}^R$ merőlegesek, ezért y^R és y^1 függetlenek és $\|y^R\|^2 \sim \sigma^2 \chi_{\dim \mathcal{L}^R}^2 = \sigma^2 \chi_{n-p}^2$, $\|y^1\|^2 \sim \sigma^2 \chi_{\dim \mathcal{L}^1}^2 = \sigma^2 \chi_q^2$.
Ha H_0 nem teljesül, akkor $\|y^1\|^2 = \|P_{\mathcal{L}^1}(X\beta + \varepsilon)\|^2$ nem centrális χ^2 eloszlású. H_1 mellett $\|y^1\|^2$ nagy(obb) értékre vezet.
- H_0 mellett $T = \frac{\frac{1}{q}\|y^1\|^2}{\frac{1}{n-p}\|y^R\|^2} \sim F_{q,n-p}$ eloszlású, míg H_1 mellett T nagy értékeket ad. F próbát végeztünk egyoldali ellenhipotézis ellen.

12.5 Példa

$$y_i = \sin(x_i) + \varepsilon_i, \quad \varepsilon_i \sim N(0, 0.25^2)$$

$$x_i \sim U(0, 5), \quad i = 1, \dots, 30$$

fokszám	0	1	2	3	5	10	20
dim im X	1	2	3	4	6	11	16
RSS	20.618	7.048	4.913	3.317	3.249	2.671	1.209
hatsigma^2	0.711	0.252	0.182	0.128	0.135	0.141	0.086
F-hányados	NA	53.909	11.734	12.507	0.25	0.822	3.389
p-érték	NA	0	0.002	0.002	0.781	0.549	0.033

- Itt egymásba ágyazott modelleket hasonlítottunk össze. Az **F-hányados** statisztikák függetlenek.
- $\dim \text{im } X$ elméletben a polinom fonszáma plusz egy. A számolt értékek ennél kisebbek lehetnek.
Ok: a becslés QR faktorizációval történik, ami az X mátrix oszlopait ortogonalizálja. Ha egy oszlop „majdnem” a korábbiak által kifeszített altérben van, akkor azt az oszlopot és az együtthatóját kihagyjuk a becslésből.

12.6 ANOVA táblázat R-ben

R-ben részletesebb táblázat is kapható:

```
anova(model)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
x	1	13.570	13.570	96.514	0.000
I((x/5)^(2))	1	2.135	2.135	15.186	0.001
I((x/5)^(3))	1	1.596	1.596	11.349	0.003
I((x/5)^(4))	1	0.067	0.067	0.475	0.499
I((x/5)^(5))	1	0.001	0.001	0.006	0.937
I((x/5)^(6))	1	0.007	0.007	0.050	0.826
I((x/5)^(7))	1	0.068	0.068	0.487	0.494
I((x/5)^(8))	1	0.276	0.276	1.966	0.177
I((x/5)^(9))	1	0.203	0.203	1.443	0.244
I((x/5)^(10))	1	0.023	0.023	0.167	0.687
Residuals	19	2.671	0.141	NA	NA

Itt az egyes magyarázó változók hatása, ortogonalizálva van. $\text{rss}_i = \min\{\|y - X\beta\| : \beta_j = 0, j > i\}$, akkor az i . magyarázó változó hatása $\text{rss}_{i-1} - \text{rss}_i$.

A szórásnégyzet becslése $\frac{1}{n-p}\|y - X\hat{\beta}\|^2$

12.7 Wald próba

$$y \sim N(X\beta, \sigma^2 \mathbf{I}_n), \quad \beta \in \mathbb{R}^p, \sigma > 0$$

- $\hat{\beta} = (X^T X)^{-1} X^T y$. Láttuk, hogy $\mathbb{E}(\hat{\beta}) = \beta$ és $\Sigma(\hat{\beta}) = (X^T X)^{-1} \sigma^2$.
- $\hat{\sigma}^2 = \frac{1}{n-p}\|y - X\hat{\beta}\|^2$ független $\hat{\beta}$ -től.

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{((X^T X)^{-1})_{j,j} \hat{\sigma}^2}} \sim t_{n-p}$$

- A $H_0 : \beta_j = 0$, $H_1 : \beta_j \neq 0$ ellenében a fenti próba statisztikával is vizsgálható.

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.013	0.538	0.023	0.981
x	1.167	2.036	0.573	0.572
I((x/5)^(2))	-4.210	63.073	-0.067	0.947
I((x/5)^(3))	-22.374	161.538	-0.139	0.891
I((x/5)^(4))	25.664	179.836	0.143	0.888
I((x/5)^(5))	-5.875	72.113	-0.081	0.936

12.8 Lineáris becslés, becsülhető függvény

$$y_i = x_i \cdot \beta + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$$

A paramétereket az $\{(y_i, x_i) : i = 1, \dots, n\}$ adatok alapján becsüljük. Gyakran a cél y^* becslése, ha a magyarázó változó értéke x^* adott.

Definíció 29. T lineáris becslés, ha $T(y)$ az y lineáris függvénye, létezik $B \in \mathbb{R}^{q \times n}$, $T(y) = By$.
 ψ a paraméter becsülhető függvénye, ha létezik rá torzítatlan lineáris becslés.

- Ha ψ becsülhető, akkor $\psi(\beta) = \mathbb{E}(T) = BX\beta = C\beta$.
- Ha X rangja p , akkor β minden lineáris függvénye becsülhető.
 Ugyanis $\hat{\beta} = (X^T X)^{-1} X^T y$ torzítatlan becslés β -ra és ha $\psi(\beta) = C\beta$, akkor $C\hat{\beta}$ torzítatlan lineáris becslés ψ -re.

- ψ becslhető $\iff \psi(\beta) = C\beta = BX\beta$ alakú.
Ugyanis, $\mathbb{E}(X\hat{\beta}) = P_X\mathbb{E}(y) = P_X\mathbb{E}(\varepsilon + X\beta) = P_XX\beta = X\beta$, azaz $X\hat{\beta}$ torzítatlan $X\beta$ -ra, ha $\hat{\beta}$ legkisebb négyzetes becslés.

12.9 Gauss–Markov tétel

Tétel 30. Tegyük fel, hogy $\psi(\beta) = C\beta$ becslhető és $\hat{\beta}$ tetszőleges legkisebb négyzetes becslés β -ra.

- $C\hat{\beta}$ torzítatlan, lineáris becslés ψ -re,
- $C\hat{\beta}$ a torzítatlan lineáris becslések között optimális a négyzetes veszteségfüggvényre nézve, azaz kovariancia mátrixa a legkisebb.
- $C\hat{\beta}$ nem függ az $\hat{\beta}$ LNB megválasztásától.
- ψ becslhető $\iff \psi(\beta) = C\beta = BX\beta$ alakú. Feltehető, hogy $B = BP_X$, azaz $(\text{im } X)^\perp \subset \ker B$.
- $\hat{\beta}$ LNB, $X\hat{\beta} = P_Xy$. Ebből

$$\mathbb{E}(X\hat{\beta}) = P_X\mathbb{E}(y) = P_XX\beta = X\beta, \implies \mathbb{E}(C\hat{\beta}) = B\mathbb{E}(X\hat{\beta}) = BX\beta = C\beta$$

- Ha $T = Dy$ torzítatlan lineáris becslés ψ -re, akkor $\mathbb{E}(Dy) = D\mathbb{E}(y) = DX\beta$, másrészt a torzítatlanság miatt $DX\beta = BX\beta$ minden β -ra, vagyis $B = DP_X$. Innen

$$\Sigma(Dy) = D\Sigma(y)D^T = \sigma^2 DD^T = \sigma^2 D(P_X + (\mathbf{I}_n - P_X))D^T \geq \sigma^2 BB^T = B\Sigma(P_Xy)B^T = \Sigma(C\hat{\beta})$$

- Ha $\hat{\beta}$ és \hat{b} LNB β -ra, akkor $\frac{1}{2}(\hat{\beta} + \hat{b})$ is az $\Sigma(C\hat{\beta}) = \Sigma(C\hat{b}) = \Sigma(C\frac{1}{2}(\hat{\beta} + \hat{b}))$. Ugyanakkor

$$\Sigma(C\frac{1}{2}(\hat{\beta} + \hat{b})) + \Sigma(C\frac{1}{2}(\hat{\beta} - \hat{b})) = \frac{1}{2}(\Sigma(C\hat{\beta}) + \Sigma(C\hat{b})) = \Sigma(C\hat{\beta}) \implies \Sigma(C(\hat{\beta} - \hat{b})) = 0 \iff C\hat{\beta} = C\hat{b}$$

12.10 Kereszt-validáció

- (y_i, x_i) , $i = 1, \dots, n$. X az x_i -kből számolt magyarázó változókat $\phi(x_i)$, mint sorvektorokat tartalmazó mátrix. Pl. $\phi(x) = (1, x, x^2, \dots, x^{p-1})$.
- $y = X\beta + \varepsilon$ lineáris modell. (y_*, x_*) új adat, x_* ismert, y_* -ot szeretnénk megbecsülni. $\phi(x_*)$ az x_* -ból számolt magyarázó változókból álló sorvektor.
- Milyen ϕ -t válasszunk?

A korábbi példában statisztikai módszerekkel próbáltunk választani.

Válasszuk azt, ami a legjobban teljesít az illesztés során nem látott adatokon, a legkisebb a generalizációs hibája.

- Generalizációs hiba becsléséhez, az adatokat két részre osztjuk. Az egyiket a modellt illesztjük, a másikon az illesztett modell hibáját számoljuk.
- Lineáris modellnél könnyen számolható, hogy mi lenne y_i becslése, ha az i . adatot nem használnánk $\hat{\beta}$ meghatározásához.

Ok. $X^T X$ egy egyrangú mátrix-szal $(x_i \otimes x_i)$ módosul, míg $X^T y$ $x_i y_i$ -vel. Ezek hatása a residuálisra $y_i - \hat{y}_{(i)} = \frac{1}{1 - x_i(X^T X)^{-1}x_i^T}(y_i - \hat{y}_i)$.

12.10.1 Példa

- Átlagos négyzetes hiba becslése: $\frac{1}{n} \sum_i |y_i - \hat{y}_{(i)}|^2$, ahol $\hat{y}_{(i)} = \phi(x_i) \cdot \hat{\beta}_{(i)}$ és $\hat{\beta}_{(i)}$ az i . adat kihagyásával kapott becslés.

fokszám	0	1	2	3	5	10	20
átlagos hiba	0.735	0.269	0.213	0.148	0.218	0.679	102.79

- Az általánosítási hiba (becslése) a fokszámmal először csökken, majd nő.
- A hiba kétrészre bontható: $\mathbb{E}(y|x) = \sin(x)$ -et kell közelíteni x polinomjaival a $[0, 5]$ intervallumon. Minél nagyobb a fokszám, annál pontosabb a közelítés.

A fokszám növelésével a X mátrix képterének a dimenziója is nő ezzel $\hat{\beta}$ kovariancia mátrixa is és így összegében a y becslésének a szórása is nő.

- Általánosabban: $y = f(x) + \varepsilon$ és f -et az adatok alapján egy véges dimenziós függvénytérből akarjuk közelíteni, pl. polinomok a példában.

ϕ_1, \dots, ϕ_p egy bázis a függvénytérből és az illesztést a $y = \phi(x)\beta + \varepsilon$ összefüggés alapján végezzük.

$\hat{\beta}$ a β becslése az adatok alapján.

x egy új adat $y = f(x) + \varepsilon$ (független az adatoktól, amiből $\hat{\beta}$ számoltuk). Ekkor

$$\mathbb{E}((y - \hat{y})^2) = \mathbb{D}^2 y + \mathbb{D}^2 \hat{y} + (\mathbb{E}(y) - \mathbb{E}(\hat{y}))^2$$

12.11 Regularizáció, Rigde regresszió és LASSO

- Ahelyett, hogy a függvényteret változtatnánk, (polinom maximális fokszáma a példában), lehet úgy is egyszerű modellt választani, hogy a bonyolultságot büntetjük.

$$\hat{\beta} = \arg \min_{\beta} \frac{1}{2} \sum_{i=1}^n (y_i - x_i \cdot \beta)^2 + \lambda \psi(\beta),$$

ahol $\lambda > 0$ és ψ adott, pl. $\psi = \frac{1}{2} \|\beta\|^2$.

- Ha $\psi(\beta) = \frac{1}{2} \|\beta\|_2^2$, akkor az eljárást **ridge** regressziónak hívják, ha $\psi(\beta) = \|\beta\|_1$ akkor a neve **LASSO**. A kettő kombinációja elasztikus háló, (elastic net).
- A minimum helyen a β szerinti derivált eltűnik. Rigde regresszió esetén

$$-X^T(y - X\beta) + \lambda\beta = 0, \quad \implies \quad X^T y = (X^T X + \lambda)\beta \quad \implies \quad \hat{\beta} = (X^T X + \lambda)^{-1} X^T y$$

- LASSO esetében nincs zárt alakban megoldás, de pl. koordinátánkénti minimalizálással könnyen számolható.
- λ megválasztása kereszt validációval történhet.