

Def: Folytonos valószínűségi feltételes várható értéke: $E[X|Y=y] = \int_{\mathbb{R}} x f_{X|Y=y}(x) dx$
 Példa: megérjük a csavar Y hibáját $\rightarrow$ következtetünk arról X élettartamára

Tétel:

- 1) Ha X és Y függetlenek, akkor $E[X|Y=y] = EX$ (mert Y -nak nincs X -re vonatkozó információtartalma)
- 2) Ha $X = g(Y)$ valamely $g: \mathbb{R} \rightarrow \mathbb{R}$ függvény, akkor általában is definiálható:
 $E[X|Y=y] = g(y)$

Példa:

Legyen (X, Y) 2D normális, Σ kovarianciamatrix, μ pedig várható érték
 $E(X|Y) = ?$

$$\Sigma = \begin{bmatrix} \sigma^2_X & \text{cov}(X, Y) \\ \text{cov}(X, Y) & \sigma^2_Y \end{bmatrix}$$

1. megoldás: $f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} (x-\mu)^T \Sigma^{-1} (x-\mu)}$ $\cdot x \in \mathbb{R}^2$ Ez így nehéz!
 $E(X|Y=y) = \int_{\mathbb{R}} x \frac{f(x, y)}{f_Y(y)} dx$

2. megoldás

$$\begin{aligned} \text{cov}\left(X - \frac{\text{cov}(X, Y)}{\sigma^2_Y} Y, Y\right) &= \text{cov}(X, Y) - \text{cov}\left(\frac{\text{cov}(X, Y)}{\sigma^2_Y} Y, Y\right) = \\ &= \text{cov}(X, Y) - \frac{\text{cov}(X, Y)}{\sigma^2_Y} \text{cov}(Y, Y) = 0 \end{aligned}$$

Téj: Ha (X, Y) 2D normális $\Rightarrow (X - \beta Y, Y)$ is 2D normális, ahol $\beta \in \mathbb{R}$

$$\Rightarrow \text{cov}\left(X - \frac{\text{cov}(X, Y)}{\sigma^2_Y} Y, Y\right) = 0 \Rightarrow X - \frac{\text{cov}(X, Y)}{\sigma^2_Y} Y \text{ és } Y \text{ is független}$$

$$E[X|Y=y] = E\left[\underbrace{X - \frac{\text{cov}(X, Y)}{\sigma^2_Y} Y}_{\text{független } Y\text{-tól}} + \underbrace{\frac{\text{cov}(X, Y)}{\sigma^2_Y} Y}_{Y\text{-nak lineáris függvénye}} \mid Y=y\right] = E\left[X - \frac{\text{cov}(X, Y)}{\sigma^2_Y} Y\right] + \frac{\text{cov}(X, Y)}{\sigma^2_Y} y = m_1 - \frac{\text{cov}(X, Y)}{\sigma^2_Y} m_2 + \frac{\text{cov}(X, Y)}{\sigma^2_Y} y$$

Következésképpen: Ha $m = 0 \Rightarrow E[X|Y=y] = \frac{\text{cov}(X, Y)}{\sigma^2_Y} y$

Összefoglalás: ha (X, Y) 2D normális, akkor $E[X|Y=y]$ -nek affín függvénye

Bemutató: konstans (máshová) megjelöléssel leírhatóan ismertlen (nem máshová) megjelöléssel.

(X, Y) 2D véletlen vagy vektoros (lehet dimenzió is)

$\Rightarrow$ A leírható feladatnak tekintem: keressük azt a $g: \mathbb{R} \rightarrow \mathbb{R}$ függvényt,

mellyre $E[X - g(Y)]^2$ a lehető legkisebb! $\Rightarrow$ ez a regressziós feladat

$\hookrightarrow$ Algoritmus: $\forall h: \mathbb{R} \rightarrow \mathbb{R}$ -re $E[X - h(Y)]^2 \geq E[X - g(Y)]^2$
 $\uparrow$ kritérium hiánya

Def: Azt a g függvényt, melyre $E[X - g(Y)]^2$ a legkisebb, regressziós függvénynek nevezzük.

Tétel: Ha $EX^2 < \infty$, akkor $g(y) := E[X | Y = y]$ függvény a regressziós függvény

$\Rightarrow$ Nem bizonyított, csak az áll: egymillió állítás

Állítás: Melyik az a $c \in \mathbb{R}$ konstans, melyre $E[X - c]^2$ a lehető legkisebb, ha $EX^2 < \infty$?

$\hookrightarrow$ Valóban: $c = EX$

Algoritmus: Egy val vektorhoz a legkisebb konstans (nagyon érdekes) a várható értéke.

Diz: $m = EX \Rightarrow E[X - c]^2 = E[X^2 - 2Xc + c^2] = E[X^2] - 2cEX + c^2$

Ez akkor a legkisebb, ha $c^2 - 2cEX = c^2 - 2cm$ a legkisebb

$$\left. \begin{array}{l} c^2 - 2cm = (c - m)^2 \geq -m^2 \\ \text{Ha } c = m, \text{ akkor } (c - m)^2 - m^2 = -m^2 \end{array} \right\} c = m \text{ a minimumhely}$$

Megj: ha (X, Y) dimenzió 2D vagy vektoros, $\mathcal{R}(Y) = \{y_1, \dots, y_n\}$, $\mathcal{R}(X) = \{x_1, \dots, x_n\}$

$$E[X | Y = y_i] = \sum_j x_j P(X = x_j | Y = y_i) = \sum_j x_j \frac{P(X = x_j, Y = y_i)}{P(Y = y_i)} \leftarrow \text{líneáris kombináció}$$

$\uparrow$ értelmezhető, mert $P(Y = y_i) > 0$

Megj: ha (X, Y) 2D dimenzió vagy vektoros, akkor is $g(y) = E[X | Y = y]$ a regressziós függvény

Statistika: jelenségek adatait összegyűjtő, rendező és értelmező tudomány.
 $\Rightarrow$ Amivel mi foglalkozunk: matematikai statistika

Adatok: $\xi_1, \xi_2, \xi_3, \dots$ független, azonos eloszlású val. v. vektorok

Def: $(\xi_1, \xi_2, \dots, \xi_n)$ -t n elemű mintának hívjuk.

Megj: $n = j^2$, ha n négyzetszám $\Rightarrow$ sokszor sajnos nem teljesül (pl. biológiai-ban)

Def: Legyen $\mathcal{I} \subseteq \mathbb{R}^2$ egy halmaz lehetséges paraméterek halmaza.
 $\Rightarrow$ Ennek elemét $\mathcal{I}$ -vel rendeljük jellel

Példák:

- 1) $E\xi_1$ -t keressük. Ekkor $\Theta = \mathbb{R}$ és $\mathcal{I} \in \mathbb{R}$ a lehetséges várható értékek
- 2) Ha pl. ξ_1 valószínűségi eloszlás, akkor tudjuk, hogy $E\xi_1 \geq 0$. Pl. $\Theta = \mathbb{R}_+$
- 3) Érdemes mintát $E\xi_1$ és $D^2\xi_1$ is. $\Theta = \mathbb{R} \times [0, \infty)$, $\mathcal{I} = (\mu, \sigma^2) \in \mathbb{R} \times [0, \infty)$
 $\Rightarrow$ Keressük μ -t és σ^2 -t

Konkrét példa 2)-re:

$\xi_1, \dots, \xi_n$: N db lámpa élettartama

Megfigyeljük a $\xi_1(\omega), \dots, \xi_n(\omega)$ mérési eredményeket

Tudni akarjuk az $E\xi_1$ -t $\Rightarrow \mathcal{I} \in \Theta = [0, \infty) \leftarrow$ paraméterek lehetséges értékei

Becsüléselmélet:

Def: Legyen $(\xi_1, \dots, \xi_n)$ egy N elemű minta, $f_N: \mathbb{R}^n \rightarrow \mathbb{R}$ függvény.
 Ekkor $f_N(\xi_1, \dots, \xi_n)$ -t statisztikai függvénynek vagy statisztikának hívjuk.

Megj: A becsüléselméletben olyan $f_N(\xi_1, \dots, \xi_n)$ statisztikát keressük, hogy bármely

$\mathcal{I} \in \Theta$ esetén $f_N(\xi_1, \dots, \xi_n)$ közel legyen $\mathcal{I}$ -hez (jól becsülje $\mathcal{I}$ -t), ha $\xi_1, \dots, \xi_n$ paraméterre $\mathcal{I}$

Példa: $\xi_1, \xi_2, \dots$ lámpák élettartama, $\mathcal{I} \in \Theta = (0, \infty)$ lehetséges értékek $E\xi_1$ számára

Def: $\bar{\xi}_N := \frac{\xi_1 + \dots + \xi_n}{N}$: mintaközép vagy tapasztalati közép

Megj: $\bar{\xi}_N$ egy statisztika, mert $f_N(y_1, \dots, y_n) := \frac{y_1 + \dots + y_n}{N}$

Parancs: $\bar{\xi}_N = f_N(\xi_1, \dots, \xi_n)$

Példa: $\xi_1, \dots, \xi_n$: N lámpa élettartama

$E\xi_1 = ?$, $D^2\xi_1 = ?$, $\Theta = [80, 300] \times [0, 10000]$
 $\mu = E\xi_1$, $\sigma^2 = D^2\xi_1$

$\mathcal{I} = (\mu, \sigma^2)$ -t akarjuk becsülni

μ becsülése: $\bar{\xi}_N$

σ^2 becsülése?

Def: $s_N^2 := \frac{\sum_{i=1}^n (\xi_i - \bar{\xi}_N)^2}{N}$ - t tapasztalati varianciának nevezzük

Def: Legyen $f_N(\xi_1, \dots, \xi_n)$, $N \in \mathbb{N}$ statisztika egy sorozata.

Def: Legyen $t_N(\xi_1, \dots, \xi_N)$, $N \in \mathbb{N}$ statisztika egy sorozata.

Azt mondjuk, hogy ez egy konszisztens becsléssorozat,

ha $t_N(\xi_1, \dots, \xi_N) \xrightarrow{N \rightarrow \infty} \vartheta$ valószínűleg, ha ϑ az igazi paraméter.

Példa: Legyen $\xi_1, \dots, \xi_N \dots$ független - sorozat. $E|\xi_i| < \infty$.

Ekkor $t_N(\xi_1, \dots, \xi_N) = \bar{\xi}_N$ konszisztens becsléssorozat a $\vartheta = E\xi_1$ várható értékre.

$\bar{\xi}_N = \frac{\xi_1 + \dots + \xi_N}{N} \xrightarrow[N \rightarrow \infty]{\text{1 valószínűleg}} E\xi_1 \Rightarrow$ Megg: A statisztikát a Nagy Számok Törvény ilteti

Példa: Alkalmam feltétel miatt $s_N^* = \frac{\sum_{i=1}^N (\xi_i - \bar{\xi}_N)^2}{N}$ valószínűleg $D^2 \xi_1$,

tehát $s_N^* : n \in \mathbb{N}$ konszisztens becsléssorozat a $D^2 \xi_1$ -re

Def: Az $t_N(\xi_1, \dots, \xi_N)$ statisztika torzítatlan, ha $E t_N(\xi_1, \dots, \xi_N) = \vartheta$,

ha ϑ igazi paraméter.

Példa: ξ_N torzítatlan becslés $\vartheta = E\xi_1$ -re,

$$\text{hiszen } E \bar{\xi}_N = E \frac{\xi_1 + \dots + \xi_N}{N} = \frac{1}{N} [E\xi_1 + \dots + E\xi_N] = \frac{N E\xi_1}{N} = E\xi_1 = \vartheta$$

Megg: Ha N kicsi, akkor jó, ha torzítatlan a becslésünk.

Jeg: tapasztalati mérésiátlagot nem torzítatlan becslés $D^2 \xi_1$ -re,

$$\text{mert } E s_N^* = \frac{N-1}{N} D^2 \xi_1 < D^2 \xi_1$$

Def: $A_2(s_N^*) = \frac{\sum_{i=1}^N (\xi_i - \bar{\xi}_N)^2}{N-1}$ -et korrigált tapasztalati mérésiátlagunk nevezzük

Állítás: $(s_N^*)^2 : N \in \mathbb{N}$ konszisztens becsléssorozat $D^2 \xi_1$ -re és $(s_N^*)^2$ -re torzítatlan becslés

Megg: Természetesen $D^2 \xi_1$ mérési konszisztens becslés:

$$s_N = \sqrt{\frac{\sum_{i=1}^N (\xi_i - \bar{\xi}_N)^2}{N}}, \quad s_N^* = \sqrt{\frac{\sum_{i=1}^N (\xi_i - \bar{\xi}_N)^2}{N-1}} \leftarrow \text{nem torzítatlan}$$

Példa: N -re dolunk egy érvel.

$$P(\xi_i = 1) = P(\text{fej}) = p \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} i = 1, \dots, N$$

$$P(\xi_i = 0) = P(\text{írá}) = 1-p$$

$\Rightarrow (\xi_1, \dots, \xi_N)$ mintából becsüljük meg p valószínűsít

Ötlet: minden $\vartheta \in \Theta$ paraméter esetén megvárunk, hogy nulla-e valójában feltételezhető ez a minta,

és azt a $\hat{\vartheta}_N$ paramétert választjuk, melyre ez az előf. a legmeggyőzőbb.

$\Rightarrow$ Legyen $p \in (0, 1)$ rögzített. Milyen valószínűségi: $x_1, x_2, \dots, x_N \in \{0, 1\}$

$$P(\text{a paraméter } p \text{ és } x_1, \dots, x_N \text{-t kaptuk}) = \prod_{i=1}^N p^{x_i} (1-p)^{1-x_i}$$

$$\Rightarrow \text{Jegyzet: ha } \xi_i = x_i = 0, \text{ akkor } p^{x_i} (1-p)^{1-x_i} = 1-p$$

$$\text{ha pedig } \xi_i = x_i = 1, \text{ akkor } p^{x_i} (1-p)^{1-x_i} = p$$

$$P(\text{a paraméter } p \text{ és } x_1, \dots, x_N \text{-t kaptuk}) = p^{n(\sum_{i=1}^N x_i)} \cdot (1-p)^{n(N - \sum_{i=1}^N x_i)}$$

$\Rightarrow$ válasszuk azt a p_N^* -t, melyre ez a legmeggyőzőbb

$$g(p) = p^{\sum_{i=1}^N x_i} \cdot (1-p)^{N - \sum_{i=1}^N x_i}$$

$$\ln(g(p)) = \underbrace{\left(\sum_{i=1}^N x_i\right) \cdot \ln(p)}_{*} + \underbrace{\left(N - \sum_{i=1}^N x_i\right) \cdot \ln(1-p)}_{\#}$$

$$\frac{\partial}{\partial p} \ln(g(p)) = \frac{*}{p} - \frac{\#}{1-p} \Rightarrow \frac{*}{p} = \frac{\#}{1-p} \Rightarrow \hat{p}_N = \frac{\xi_1 + \dots + \xi_N}{N} \quad \leftarrow \text{maximum likelihood becslés}$$

Maximum likelihood becslés: (Ha nincs jobb ötlet becsülésre, ezt használjuk)

Recept: $(\xi_1, \dots, \xi_N)$: N elemű minta

Maximum likelihood kezelése (Ha nincs jólléti lecsúszás, ezt használjuk)

Recept: $(\xi_1, \dots, \xi_N)$: N elemű minta

feltételezve, hogy ξ_i független elemek is $f_{\theta}(x)$: $x \in \mathbb{R}$ a sűrűségfüggvény, θ az ismeretlen paraméter
Ekkor $(\xi_1, \dots, \xi_N)$ együttes sűrűségfüggvénye

Nagyon könnyű az azt jelenti, hogy $h_{\theta}(x_1, x_2, \dots, x_N)$ annak a "valószínűség" (likelihood, hogy $\xi_1 = x_1, \dots, \xi_N = x_N$

Def: $l(\theta; x_1, \dots, x_N) := \ln h_{\theta}(x_1, \dots, x_N)$ a log-likelihood függvény.

Legyen továbbá θ_N a $\theta \mapsto l(\theta, \xi_1, \dots, \xi_N)$ maximuma

Ekkor $\hat{\theta}_N$ a maximum likelihood becslés a θ paraméterre