Valószínűségszámítás

2021. november 24.

Mészáros Szabolcs

Tárgyhonlap:
cs.bme.hu/valszam

A prezentáció anyagát és az abból készült videofelvételt a tárgy hallgatói jogosultak használni, kizárólag saját célra. A felvétel másolása, videómegosztókra való feltöltése részben vagy egészben tilos, illetve csak a tantárgyfelelős előzetes engedélyével történhet.

Feltételes várható érték, motiváció

Problémák korábbról:

1. Hogyan jellemezzünk valószínűségi változók közötti összefüggést?
2. A lineáris regresszió csak akkor működik jól, ha lineáris a kapcsolat a két val. változó között; mit csináljunk egyébként?
3. Alkalmazásokban előfordul, amikor egy val. változó eloszlása “függ” egy másik eloszlástól. Ilyenkor nehezebb várható értéket számolni.

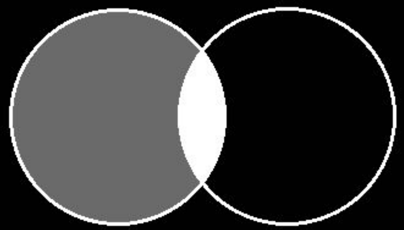
Megoldás: feltételes várható érték,

a) eseményre
b) másik valószínűségi változóra

Ismétlés: feltételes valószínűség

Feltételes valószínűség: $\mathbb{P}(B \mid A) = \dfrac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$ feltéve, hogy $\mathbb{P}(A) > 0$

Jelentés: Az A-ra fókuszálva, mi az esélye, hogy B is bekövetkezik?



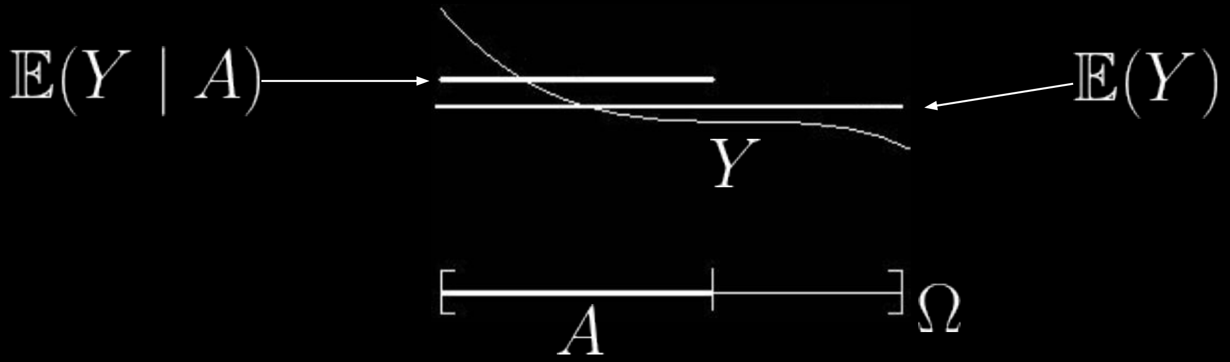
Állítás: (lásd 2. ea) Rögzített A esetén $B \mapsto \mathbb{P}(B \mid A)$ valószínűségi mérték.

Feltételes várható érték, eseményre

Rögzített jelölés: Legyen Y val. változó, és A olyan esemény, amire $\mathbb{P}(A) > 0$.

Definíció: Az Y-nak az A-ra vett *feltételes várható értéke* az Y változónak a $\mathbb{P}(\,.\mid A)$ szerinti várható értéke. Jelölés: $\mathbb{E}(Y \mid A)$

Jelentés: Y átlagos értéke, ha csak az A eseményre fókuszálunk.



Feltételes várható érték, eseményre

Lemma: Ha Y egyszerű valószínűségi változó, akkor

$$\mathbb{E}(Y \mid A) = \sum_{k \in \mathrm{Ran}(Y)} k \cdot \mathbb{P}(Y = k \mid A)$$

Megjegyzés: a várható érték definíciójához képest csak kicseréltük a valószínűséget feltételes valószínűségre.

Állítás: $\mathbb{E}(Y \mid A) = \dfrac{1}{\mathbb{P}(A)} \mathbb{E}(Y \mathbf{1}_A)$

Feltételes várható érték, eseményre

Példa: Legyen Y és Z két független, szabályos kockadobás, és legyen $X = Y + Z$. Kérdés: $\mathbb{E}(Y \mid X < 7) =?$ azaz $A = \{X < 7\}$

$$\mathbb{P}(Y = k \mid X < 7) = \frac{\mathbb{P}(Y = k, X < 7)}{\mathbb{P}(X < 7)} = \frac{6-k}{36} \Big/ \frac{15}{36} = \frac{6-k}{15}$$

$$\mathbb{E}(Y \mid X < 7) = \sum_{k=1}^{6} k \cdot \mathbb{P}(Y = k \mid X < 7) =$$

$$= \sum_{k=1}^{6} k \cdot \frac{6-k}{15} = \frac{6 \cdot 21 - 91}{15} = \frac{7}{3}$$

Feltételes várható érték, eseményre

Példa: (folyt.) Mivel $\mathbb{P}(X = x) > 0$, ezért értelmes kérdés lenne:

$\mathbb{E}(Y \mid X = x) = ?$

$$\mathbb{E}(Y \mid X = 5) = \sum_{k=1}^{6} k \cdot \mathbb{P}(Y = k \mid X = 5) =$$

$$= \sum_{k=1}^{4} k \cdot \frac{1}{4} = 2{,}5$$

Jelentés: Ha tudjuk, hogy két kockadobás összege 5, akkor az egyik kockadobás értékét átlagosan 2,5-re tippeljük.

Regresszió, diszkrét eset

Definíció: Legyenek X és Y val. változók, ahol X diszkrét. Jelölje

$$S_X = \{x \in \mathbb{R} \mid \mathbb{P}(X = x) > 0\}$$

az X *lényeges értékei*nek halmazát.

Definíció: Az Y-nak az X-re vett (diszkrét) *regresszió*ja:

$$g : S_X \to \mathbb{R} \qquad g(x) = \mathbb{E}(Y \mid X = x)$$

Megjegyzés: A regresszió az Y "legjobb" közelítése X alapján.

Regresszió, diszkrét eset

Példa: Dobunk egy szabályos kockával, majd az eredménynek megfelelő számú szabályos érmével. Legyen X a kockadobás értéke, Y pedig a fejek száma az érmedobások között. Mi Y regressziója X-re?

Tetszőleges $x \in \{1, 2, 3, 4, 5, 6\}$ esetén az Y eloszlása az $\{X = x\}$ eseményen binomiális, x és $0{,}5$ paraméterekkel, azaz

$$\mathbb{P}(Y = k \mid X = x) = \binom{x}{k} 0{,}5^x \qquad (\forall 0 \leq k \leq x)$$

Megj.: röviden, de pontatlanul írva: Y "feltételes eloszlása $B(X; 0{,}5)$"

Regresszió, diszkrét eset

Állítás: Legyen X diszkrét valószínűségi változó, $\mathbb{P}(X < x) > 0$ és $\mathbb{E}(Y) < \infty$. Ekkor

$$\mathbb{E}\big(g(X) \mid X < x\big) = \mathbb{E}\big(Y \mid X < x\big)$$

ahol $g(x) = \mathbb{E}(Y \mid X = x)$.

Értelmezés:

A regresszió “mindent tud” az Y-ról, amit tudni lehet Y-ról az X alapján, várható érték értelemben.

Regresszió, diszkrét eset

Bizonyítás:

$$\mathbb{E}\big(g(X) \mid X < x\big) = \sum_{k \in \mathrm{Ran}(X)} g(k) \cdot \mathbb{P}(X = k \mid X < x) =$$

$$= \sum_{k<x} g(k) \frac{\mathbb{P}(X = k)}{\mathbb{P}(X < x)} = \sum_{k<x} \mathbb{E}(Y \mid X = k) \frac{\mathbb{P}(X = k)}{\mathbb{P}(X < x)}$$

$$= \sum_{k<x} \frac{1}{\mathbb{P}(X = k)} \mathbb{E}(Y \mathbf{1}_{\{X=k\}}) \frac{\mathbb{P}(X = k)}{\mathbb{P}(X < x)}$$

$$= \frac{1}{\mathbb{P}(X < x)} \mathbb{E}\Big(Y \sum_{k<x} \mathbf{1}_{\{X=k\}}\Big) = \mathbb{E}(Y \mid X < x)$$

Regresszió, folytonos eset

Kérdés: Mit csináljunk folytonos esetben?
Itt ugye $\mathbb{E}(Y \mid X = x)$ értelmetlen, mert $\mathbb{P}(X = x) = 0$.

Definíció: Az Y-nak az X-re vett **regresszió**ja egy $g(X)$ alakú val. változó, amire

$$\mathbb{E}\big(g(X) \mid X < x\big) = \mathbb{E}\big(Y \mid X < x\big)$$

minden olyan $x \in \mathbb{R}$ esetén, amire $\mathbb{P}(X < x) > 0$. Jel.: $\mathbb{E}(Y \mid X)$

Megjegyzés:

- Diszkrét esetben: konstruktív definíció; a fenti egyenlet csak tulajdonság.
- Folytonos esetben: a fenti tulajdonsággal definiálunk.
- Az új definíció általánosítása a diszkrét definíciónak.
- A regresszió csak lényegében egyértelmű (ahogy a sűrűségfüggvény is).
- Ha $\mathbb{E}(|Y|)$ véges, akkor $\mathbb{E}(Y \mid X)$ létezik. (NB)

Regresszió, folytonos eset

Definíció: Legyen (X, Y) folytonos valószínűségi vektorváltozó, együttes sűrűségfüggvénye $f_{X,Y}$. Ekkor Y -nak az X-re vett *feltételes sűrűségfüggvénye*:

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{f_{X,Y}(x,y)}{\int_{-\infty}^{\infty} f_{X,Y}(x,u)\mathrm{d}u}$$

ha $f_X(x) \neq 0$ és 0 egyébként.

Megjegyzés: A formula a $\mathbb{P}(B \mid A) = \dfrac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$ mintát követi.

Regresszió, folytonos eset

Állítás: Az előző definíció kontextusában

$$\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y \mid x)\mathrm{d}y$$

Név: *regressziós függvény*.

Bizonyítás: a definíciót kell leellenőriznünk a jobb oldalira (jel.: $g(x)$)

$$\mathbb{E}\big(g(X) \mid X < x\big) = \frac{1}{\mathbb{P}(X < x)}\mathbb{E}\big(g(X)\mathbf{1}_{\{X<x\}}\big)$$

$$\mathbb{E}\big(g(X)\mathbf{1}_{\{X<x\}}\big) = \int_{-\infty}^{\infty} g(z)\mathbf{1}_{\{z<x\}} f_X(z)\mathrm{d}z = \int_{-\infty}^{x} g(z) f_X(z)\mathrm{d}z$$

Regresszió, folytonos eset

$$g(z)f_X(z) = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y \mid z)\mathrm{d}y \cdot f_X(z) =$$

$$= \int_{-\infty}^{\infty} y \cdot \frac{f_{X,Y}(z,y)}{f_X(z)} f_X(z)\mathrm{d}y = \int_{-\infty}^{\infty} y \cdot f_{X,Y}(z,y)\mathrm{d}y$$

Ezt visszahelyettesítve:

$$\mathbb{E}\big(g(X) \mid X < x\big) = \frac{1}{\mathbb{P}(X < x)} \int_{-\infty}^{x} \int_{-\infty}^{\infty} y \cdot f_{X,Y}(z,y)\mathrm{d}y\mathrm{d}z$$

$$= \frac{1}{\mathbb{P}(X < x)} \mathbb{E}(Y \mathbf{1}_{\{X<x\}}) = \mathbb{E}(Y \mid X < x)$$

Regresszió, folytonos eset

Példa:

$$f_{X,Y}(x,y) = \begin{cases} 15x^2y & \text{ha } 0 < x < y < 1 \\ 0 & \text{egyébként} \end{cases}$$

$$\mathbb{E}(Y \mid X) = ?$$

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\mathrm{d}y = \int_x^1 15x^2y\mathrm{d}y = \frac{15}{2}(x^2 - x^4)$$

$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{15x^2y}{\frac{15}{2}(x^2 - x^4)} = \frac{2y}{1 - x^2}$$

Regresszió, folytonos eset

Példa: (folyt.) $\mathbb{E}(Y \mid X = x) =$

$$= \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y \mid x)\mathrm{d}y$$

$$= \int_{x}^{1} y \cdot \frac{2y}{1-x^2}\mathrm{d}y$$

$$= \frac{2}{3} \cdot \frac{1-x^3}{1-x^2}$$

Tehát

$$\mathbb{E}(Y \mid X) = \frac{2}{3} \cdot \frac{1-X^3}{1-X^2}$$

Megjegyzés:

- a regresszió val. változót ad,
- a regresszió nem szimmetrikus a két paraméterében.

Regresszió tulajdonságai

Állítás: Legyenek X, Y, Z val. változók. Ekkor

1. $\mathbb{E}(aY + bZ \mid X) = a\mathbb{E}(Y \mid X) + b\mathbb{E}(Z \mid X)$
2. $\mathbb{E}\big(h(X)Y \mid X\big) = h(X)\mathbb{E}(Y \mid X)$, tetsz. h folytonos fv-re.
3. Ha X és Y függetlenek, akkor $\mathbb{E}(Y \mid X) = \mathbb{E}(Y)$

Megjegyzés: Az tipikusan nem teljesül, hogy $\mathbb{E}(h(Y) \mid X)$ és $h(\mathbb{E}(Y \mid X))$ megegyezne. Pl. ha X és Y függetlenek, akkor ez azt jelentené, hogy $h(\mathbb{E}(Y)) = \mathbb{E}(h(Y))$, ami általában nem igaz.

Regresszió optimalitása

Állítás: Tegyük fel, hogy $\mathbb{E}(Y^2)$ véges. Ekkor a

$$\mathbb{E}\left(\left(Y - g(X)\right)^2\right)$$

pontosan akkor minimális, ha $g(X) = \mathbb{E}(Y \mid X)$ egy valséggel teljesül.

Kérdés: mikor egyezik meg a lineáris regresszió a regresszióval?

Állítás: Legyenek $Z_1, \ldots, Z_n$ független, normális eloszlású val változók,

$X = \sum_{i=1}^{n} a_i Z_i$ és $Y = \sum_{i=1}^{n} b_i Z_i$ (valamilyen együtthatókkal).

Ekkor $\mathbb{E}(Y \mid X)$ megegyezik az Y-nak az X-re vett lineáris regressziójával.

Teljes várható érték tétele, általánosan

Állítás: Legyenek X és Y val. változók, amire $\mathbb{E}(Y)$ véges. Ekkor

$$\mathbb{E}\big(\mathbb{E}(Y \mid X)\big) = \mathbb{E}(Y)$$

Megjegyzés: Elsőre nem hasonlít a teljes valószínűség tételére. Valójában itt is az gondolatmenet, hogy

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \mid A_i) \cdot \mathbb{P}(A_i)$$

- ha Y -ra komplikált kiszámolni,
- akkor számoljuk ki feltétellel (most épp $\{X = x\}$ feltételekkel), és
- vegyük a valószínűségekkel súlyozott közepét a feltételes értékeknek.

Teljes várható érték tétele, diszkrét

Állítás: Legyen $A_1, \ldots, A_n$ teljes eseményrendszer, amire $\mathbb{P}(A_i) > 0 \quad (\forall i)$

Ha $\mathbb{E}(Y)$ véges, akkor

$$\mathbb{E}(Y) = \sum_{i=1}^{n} \mathbb{E}(Y \mid A_i)\mathbb{P}(A_i)$$

Állítás: Legyen X diszkrét val. változó, és $S_X = \{x_1, x_2, \ldots\}$.

Ha $\mathbb{E}(Y)$ véges, akkor

$$\mathbb{E}(Y) = \sum_{i=1}^{\infty} \mathbb{E}(Y \mid X = x_i)\mathbb{P}(X = x_i)$$

Teljes várható érték tétele, diszkrét eset

Példa: geometriai eloszlás szórása. Legyen Y az első fejhez szükséges dobások száma, ha olyan érmét használunk, ami p eséllyel ad fej eredményt.

Legyen X pontosan akkor 1, ha az első dobás fej, különben 0.

$$\mathbb{E}(Y^2) = \mathbb{E}(Y^2 \mid X = 0)\mathbb{P}(X = 0) + \mathbb{E}(Y^2 \mid X = 1)\mathbb{P}(X = 1)$$

$$= \mathbb{E}\big((Y+1)^2\big)\mathbb{P}(X = 0) + 1 \cdot \mathbb{P}(X = 1)$$

$$= p + \mathbb{E}(Y^2 + 2Y + 1)(1-p) \quad \Rightarrow \mathbb{E}(Y^2) = \frac{2-p}{p^2}$$

$$\Rightarrow \mathbb{D}^2(Y) = \frac{1-p}{p^2}$$

Teljes várható érték tétele, folytonos eset

Állítás: Legyen X folytonos val. változó, sűrűségfüggvénye f_X.

Ha $\mathbb{E}(Y)$ véges, akkor

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} \mathbb{E}(Y \mid X = x) f_X(x) \mathrm{d}x$$

Példa: Válasszunk egyenletesen véletlenszerűen egy X számot 0 és 1 között, majd válasszunk egyenletesen véletlenszerűen egy Y számot X^2 és X között. Mennyi Y várható értéke?

Teljes várható érték tétele, folytonos eset

Mivel az Y “feltételes eloszlása” egyenletes az $[X^2, X]$ intervallumon, ezért

$$f_{Y|X}(y \mid x) = \begin{cases} \frac{1}{x-x^2} & \text{ha } x^2 < y < x \\ 0 & \text{egyébként} \end{cases}$$

$$\mathbb{E}(Y \mid X = x) = \int_{-\infty}^{\infty} y \cdot f_{Y|X}(y \mid x)\mathrm{d}y = \left[\frac{y^2}{2(x - x^2)}\right]_{x^2}^{x} = \frac{x + x^2}{2}$$

$$\mathbb{E}(Y) = \int_{-\infty}^{\infty} \mathbb{E}(Y \mid X = x) f_X(x)\mathrm{d}x = \int_0^1 \frac{x + x^2}{2}\mathrm{d}x = \frac{5}{12}$$

Köszönöm a figyelmet!