Matematika A4
Feladatok a Tematika 9. és 11. pontjához

Szabados Tamás kurzusa / Balázs Márton gyakorlata

Feltételes várható érték

Adott két valószínűségi változó, X és Y. Az X-nek $Y = y$ feltételre vett feltételes várható értéke

$$\mathbb{E}(X \,|\, Y = y) = \begin{cases} \displaystyle\sum_i x_i \cdot \mathbb{P}\{X = x_i \,|\, Y = y\}, & \text{ha } X \text{ diszkrét,} \\ \displaystyle\int_{-\infty}^{\infty} x \cdot f(x \,|\, Y = y)\, \mathrm{d}x, & \text{ha } X \text{ folytonos.} \end{cases}$$

Mint ahogy a feltételes valószínűség egy szabályos valószínűség (tudja az axiómákat), a feltételes várható érték is egy szabályos várható érték, tehát minden tulajdonság, ami a várható értékre igaz volt, a feltételes várható értékre is igaz lesz (pl. linearitás). A feltételes súlyfüggvény illetve sűrűségfüggvény definíciója alapján könnyen ellenőrizhető, hogy a diszkrét illetve folytonos esetben

$$\sum_j \mathbb{E}(X \,|\, Y = y_j) \cdot \mathbb{P}\{Y = y_j\} = \sum_{i,\,j} x_i \cdot \mathbb{P}\{X = x_i \,|\, Y = y_j\} \cdot \mathbb{P}\{Y = y_j\} = \sum_{i,\,j} x_i \cdot \mathbb{P}\{X = x_i,\, Y = y_j\} = \mathbb{E}(X),$$

illetve

$$\int_{-\infty}^{\infty} \mathbb{E}(X \,|\, Y = y) \cdot g(y)\, \mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \cdot f(x \,|\, Y = y) \cdot g(y)\, \mathrm{d}x\, \mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \cdot h(x,\, y)\, \mathrm{d}x\, \mathrm{d}y = \mathbb{E}(X).$$

Ezt elegánsabban és rövidebben úgy is írhatjuk, hogy $\mathbb{E}[\mathbb{E}(X \,|\, Y)] = \mathbb{E}(X)$, amit *toronyszabály*nak nevezünk. Itt $\mathbb{E}(X \,|\, Y)$ az Y függvénye, és ezért maga is egy valószínűségi változó. Y-tól ugyanúgy függ, mint ahogy $\mathbb{E}(X \,|\, Y = y)$ függ y-tól.

1. X és Y közös sűrűségfüggvénye

$$h(x,\, y) = \frac{1}{y} \cdot \mathrm{e}^{-(y + x/y)}, \qquad \text{ha } x > 0,\ y > 0.$$

   (a) Mutassuk meg, hogy Y peremeloszlása exponenciális(1).

   (b) Mutassuk meg, hogy az $(X \,|\, Y = y)$ eloszlás, azaz X feltételes eloszlása az $Y = y$ feltétel mellett exponenciális$(1/y)$.

   (c) Használjuk az előbbi két pontot $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ meghatározására. Bátrabbak ellenőrizhetik, hogy $\mathbf{Cov}(X,\, Y) = 1$, miután belátták a fentiekhez hasonlóan, hogy $\mathbb{E}(X \cdot Y \,|\, Y) = Y \cdot \mathbb{E}(X \,|\, Y)$, és $\mathbb{E}(X \cdot Y) = \mathbb{E}[\mathbb{E}(X \cdot Y \,|\, Y)] = \mathbb{E}[Y \cdot \mathbb{E}(X \,|\, Y)]$.

2. Egy bányász a bánya egy termében rekedt. A teremből három ajtó nyílik: az első ajtó 3 órányi út végén a szabadba vezet. A második ajtó egy alagútba nyílik, mely 5 órányi séta után visszavezet ugyanebbe a terembe. A harmadik ajtó szintén egy alagútba nyílik, mely 7 órányi séta után vezet vissza ugyanebbe a terembe. A bányász minden alkalommal amikor ebbe a terembe ér, e három ajtó közül választ egyet egyenlő valószínűséggel, az előző választásoktól függetlenül. Legyen X a szabadba kijutáshoz szükséges idő, Y pedig a legelső alkalommal kiválasztott ajtó sorszáma. Érveljünk amellett, hogy $\mathbb{E}(X \,|\, Y = 1) = 3$, $\mathbb{E}(X \,|\, Y = 2) = 5 + \mathbb{E}(X)$, és $\mathbb{E}(X \,|\, Y = 3) = 7 + \mathbb{E}(X)$. Alkalmazzunk egy toronyszabályt ezen feltételes várható értékekkel, és így mutassuk meg, hogy $\mathbb{E}(X) = 15$.

3. Legyenek $X_1,\, X_2,\, \ldots,\, X_n$ független és azonos eloszlású valószínűségi változók. Határozzuk meg

$$\mathbb{E}(X_1 \,|\, X_1 + X_2 + \cdots + X_n = x)$$

   értékét. Segítség: tudjuk, hogy $\mathbb{E}(X_1 + X_2 + \cdots + X_n \,|\, X_1 + X_2 + \cdots + X_n = x)$ mivel egyenlő, ezután már csak a feltételes várható érték additivitását kell használnunk, és azt, hogy a változók azonos eloszlásúak.

Kétdimenziós normális eloszlás

A kétdimenziós normális eloszlás sűrűségfüggvénye

$$h(x,\,y)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\varrho^2}}\cdot\exp\Big\{-\frac{1}{2(1-\varrho^2)}\Big[\Big(\frac{x-\mu_x}{\sigma_x}\Big)^2+\Big(\frac{y-\mu_y}{\sigma_y}\Big)^2-2\varrho\,\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\Big]\Big\}.$$

Az ilyen sűrűségfüggvényt követő $(X,\,Y)$ pár esetén X peremeloszlása $\mathcal{N}(\mu_x,\,\sigma_x^2)$, Y peremeloszlása $\mathcal{N}(\mu_y,\,\sigma_y^2)$, és a kettő korrelációja ϱ.

1. Mutassuk meg, hogy ha $(X,\,Y)$ eloszlása kétdimenziós normális, akkor
   (a) X feltételes eloszlása az $Y=y$ feltétel mellett normális, $\mu_x+\varrho\frac{\sigma_x}{\sigma_y}(y-\mu_y)$ várható értékkel, és $\sigma_x^2(1-\varrho^2)$ szórásnégyzettel. (Ez egy kicsit számolós, de egy ügyes változócsere sokat segít.)
   (b) X és Y korrelációja ϱ.
   (c) X és Y független akkor és csak akkor, ha $\varrho=0$.
2. Íme hogyan lehet egy általános kétdimenziós normális eloszlást független normálisak összegére bontani. Rögzítsünk $\sigma_x,\,\sigma_y>0$, $\mu_x,\,\mu_y\in\mathbb{R}$, és $-1<\varrho<1$ paramétereket. Legyen Y egy normális eloszlású valószínűségi változó μ_y várható értékkel és σ_y szórással. Legyen Z egy másik, Y-tól független normális valószínűségi változó, $\mu_x-\varrho\cdot\mu_y\cdot\frac{\sigma_x}{\sigma_y}$ várható értékkel és $(1-\varrho^2)\cdot\sigma_x^2$ szórásnégyzettel. Definiáljuk az $X:=\varrho\cdot\frac{\sigma_x}{\sigma_y}\cdot Y+Z$ változót. Mutassuk meg, hogy az $(X,\,Y)$ pár eloszlása kétdimenziós normális, a megfelelő μ_x és μ_y várható értékekkel, σ_x és σ_y szórásokkal, és ϱ korrelációval.
3. Legyen X standard normális eloszlású, és I X-től független, $\mathbb{P}\{I=1\}=\mathbb{P}\{I=0\}=1/2$ eloszlással. Definiáljuk a következő valószínűségi változót:
   $$Y:=\begin{cases}X & \text{, ha } I=1,\\ -X & \text{, ha } I=0.\end{cases}$$
   Azaz: Y (X-től független) egyenlő eséllyel lesz X vagy $-X$.
   (a) Független-e X és Y?
   (b) Független-e I és Y?
   (c) Mutassuk meg, hogy Y standard normális eloszlású.
   (d) Mutassuk meg, hogy $\mathbf{Cov}(X,\,Y)=0$.
4. Legyenek X és Y független standard normális valószínűségi változók, és $U:=(X+Y)/\sqrt{2}$, $V:=(X-Y)/\sqrt{2}$. Mutassuk meg, hogy U és V szintén független standard normális változók.

Általános regresszió

Adott az $(X,\,Y)$ kétdimenziós valószínűségi változó az együttes eloszlásával (például folytonos esetben sűrűségfüggvényével). X-et megfigyelve Y-t szeretnénk közelíteni egy $k(X)$ alakú tippelő függvénnyel. A közelítés azt jelenti, hogy az elkövetett $(Y-k(X))^2$ négyzetes hiba átlagát szeretnénk minimalizálni. Pontosabban azt a k függvényt keressük, amire $\mathbb{E}(Y-k(X))^2$ minimális. Az órán tanult tétel kimondja, hogy ebben az esetben a megoldás a feltételes várható érték: $k(x)=\mathbb{E}(Y\,|\,X=x)$. Ha pedig az elkövetett abszolút hiba átlagát, azaz $\mathbb{E}(|Y-k(X)|)$-et szeretnént minimalizálni, akkor a lehető legjobb tippelés az Y feltételes mediánja. Azaz ebben az esetben az $F_{Y|X}(y\,|\,X=x)=\int_{-\infty}^{y}g(v\,|\,X=x)\,\mathrm{d}v$ feltételes eloszlásfüggvényt kell 1/2-del egyenlővé tenni, és belőle y-t, mint x függvényét kifejezni.

1. A Duna holnaputáni budapesti vízállását akajuk becsülni a mai bécsi vízállásból. Bár a két vízállás közt szoros kapcsolat van, azért pontosan nem lehet megmondani a vízállást, mindkettőt egy-egy valószínűségi változó írja le. Tegyük fel, hogy mindkét vízállást egy 0 és 1 közti számmal tudjuk jellemezni, melynek legyen az együttes sűrűségfüggvénye $h(x,\,y)=\frac{6}{5}\cdot(x+(y-1)^2)$, ha $0<x<1$ és $0<y<1$.
   (a) Határozzuk meg a budapesti vízállás eloszlását a bécsi ismeretében, azaz a feltételes sűrűségfüggvényt.
   (b) Mi annak a valószínűsége, hogy Budapesten alacsonynak nevezhető (azaz 0 és 1/2 közé esik) a vízállás, ha Bécsben x volt? (Mennyi ez $x=1/3$-ra?)

(c) Ha már ismerjük a bécsi vízállást, mire tippelünk a budapestire, ha a lehető legkisebb átlagos négyzetes hibát akarjuk elkövetni?

(d) És ha az átlagos abszolút hibát akarjuk minimalizálni?

2. $X = \mathrm{RND1}$, $Y = \mathrm{RND1} \cdot \mathrm{RND2}$ eloszlás esetén láttuk, hogy az együttes sűrűségfüggvény $h(x, y) = 1/x$, ha a $0 < y < x < 1$ háromszögön vagyunk. Mi a regressziós görbe, ha az abszolút hibát, illetve ha a négyzetes hibát szeretnénk minimalizálni?

3. Az egységkörön választunk egyenletes eloszlás szerint egy (X, Y) pontot. Az X koordináta ismeretében hogyan közelítenénk $|Y|$-t, feltéve, hogy a hiba abszolútértéknégyzetét szeretnénk minimalizálni?

4. Többpártrendszer esetén az egyes pártokra leadott szavazatok százalékos aránya valószínűségi változó. A Zöldek az összes szavazatok X, a Demokraták az összes szavazatok Y hányadát kapják, együttes eloszlásuk a $h(x, y) = 24xy$, ha $0 < x,\ 0 < y,\ x + y < 1$ sűrűségfüggvényt követi. Ha a Demokraták az összes szavazatok 40%-át kapták, mire tippelünk, mennyit kaptak a Zöldek?

5. Legyen (X, Y) egyenletes eloszlású a $(0, 0),\ (1, 0),\ (0, 2)$ pontok által meghatározott háromszögön. Számítsuk ki Y-nak X-re vonatkozó regressziós függvényét.

6. Magyarországon a 18 év feletti férfiak testmagasságának átlagos értéke 178 cm, szórása 10 cm. Nőknél ugyanezek az adatok 166 cm és 8cm. Focimeccseken a drukkerek 10%-a nő, a többiek férfiak. Mindkét nem testmagasságágának eloszlását normálisnak véve:

   (a) Mi annak a valószínűsége, hogy egy 170 cm-nél alacsonyabb szurkoló nő?

   (b) Adjuk meg x függvényében annak a valószínűségét, hogy egy x cm magas drukker férfi.

   (c) Hogyan tippeljünk a szurkolók testmagasságából a nemükre, ha a célunk az, hogy a lehető legnagyobb valószínűséggel helyesen tippeljünk?

Lineáris regresszió

A gyakorlatban gyakran nem tudjuk az együttes sűrűségfüggvényt meghatározni. Ilyenkor könnyebb lehet a várható értéket, a szórást, a kovarianciát kiszámolni. Ezek segítségével már meg tudjuk mondani, hogy melyik az a lineáris függvény, amivel tippelve a hiba négyzetének várható értéke a legkisebb.

$$y - \mathbb{E}(Y) = \frac{\mathbf{Cov}(X, Y)}{\mathbb{D}^2(X)} \cdot (x - \mathbb{E}(X)).$$

Ez egy kicsit másképp:

$$\frac{y - \mathbb{E}(Y)}{\mathbb{D}(Y)} = \varrho(X, Y) \cdot \frac{x - \mathbb{E}(X)}{\mathbb{D}(X)},$$

ahol $\varrho(X, Y) = \frac{\mathbf{Cov}(X, Y)}{\mathbb{D}(X)\mathbb{D}(Y)}$ az X és Y korrelációs együtthatója. Ha $\varrho(X, Y) = \pm 1$, akkor a két valószínűségi változó közt lineáris függés van. Ha $\varrho(X, Y) = 0$, akkor a változók *korrelálatlanok* (ami nem szükségképpen jelenti azt, hogy függetlenek volnának).

1. Egy kétdimenziós valószínűségi változó sűrűségfüggvénye $\frac{1}{6}xy$ $(0 < x < 2,\ x < y < 2x)$. Milyen $k(y)$ függvénnyel érdemes a második koordinátából az elsőt tippelni, ha az a célunk, hogy a tipelésnél elkövetett hiba négyzetének átlagos értéke a sok kísérlet esetén minél kisebb legyen,

   (a) ha feltesszük, hogy $k(y)$ lineáris,

   (b) ha $k(y)$ tetszőleges valós lehet?

2. X és Y együttes sűrűségfüggvénye $h(x, y) = 60xy^2$, ha $0 \le x \le 1,\ 0 \le y \le 1 - x$. Határozzuk meg a kovarianciájukat. Tegyük fel, hogy a második koordinátát tudjuk megfigyelni és az elsőt ezen megfigyelt adattól függően becsüljük az $x = \frac{2}{3}(1 - y)$ képlet alapján. Van-e ennél jobb módszer, ha a négyzetes eltérés hibáját akarjuk minimalizálni?

3. Statisztikai adatok alapján annak a valószínűsége, hogy ikerszületéskor mindkét gyerek fiú, 0.32, annak a valószínűsége, hogy mindkét gyermek lány, 0.28. Annak a valószínűsége, hogy az első iker fiú és a második lány ugyanannyi, mint fordítva. Jelölje X illetve Y az első, illetve a második gyerek nemét, legyen a felvett értékük fiú esetén 1, lány esetén 0. Számítsuk ki X és Y korrelációs együtthatóját. Hogyan tippelnénk Y ismeretében X-re lineáris függvénnyel, ha a tippelés átlagos négyzetes hibáját akarjuk minimalizálni?