A sztochasztika alapjai
MBNXK262

5. előadás: Véletlen változók függetlensége, várható érték

Kevei Péter

2022/23 tavasz

Véletlen vektorváltozók

Definíció

Az $\xi = (\xi_1, \ldots, \xi_n) : \Omega \to \mathbb{R}^n$ függvény véletlen vektorváltozó, ha minden komponense véletlen változó. Az ξ eloszlásfüggvénye

$$F(x_1, \ldots, x_n) = \mathbf{P}(\xi_1 < x_1, \ldots, \xi_n < x_n).$$

Az $(\xi_1, \ldots, \xi_n)$ véletlen vektorváltozó diszkrét, ha értékkészlete megszámlálható, és folytonos, ha van olyan f nemnegatív n-változós függvény, melyre

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \ldots \int_{-\infty}^{x_n} f(y_1, \ldots, y_n) \mathrm{d}y_n \ldots \mathrm{d}y_1$$

teljesül minden $(x_1, \ldots, x_n) \in \mathbb{R}^n$ esetén. Ilyenkor az f függvényt az ξ vektorváltozó sűrűségfüggvényének nevezzük. Az ξ_i, $i = 1, 2, \ldots, n$, változók eloszlását, peremeloszlásnak, vagy marginális eloszlásnak nevezzük.

Vvv

Állítás

Legyen $f(u, v)$ *az* (ξ, η) *véletlen vektor sűrűségfüggvénye. Ekkor* ξ *és* η *is folytonos véletlen változók* $\int_{\mathbb{R}} f(u, v)\mathrm{d}v$ *ill.* $\int_{\mathbb{R}} f(u, v)\mathrm{d}v$ *sűrűségfüggvénnyel.*

Véletlen változók függetlensége

Definíció

Az $\xi_1, \ldots, \xi_n$ függetlenek, ha minden $x_1, \ldots, x_n \in \mathbb{R}$ esetén

$$\mathbf{P}(\xi_1 < x_1, \ldots, \xi_n < x_n) = \mathbf{P}(\xi_1 < x_1) \ldots \mathbf{P}(\xi_n < x_n)$$

teljesül. Vagyis az együttes eloszlásfüggvény az egyes eloszlásfüggvények szorzata.

Függetlenség

Állítás

Legyenek ξ, η *diszkrét véletlen változók, lehetséges értékeik* $0, 1, 2, \ldots$. *Ekkor* ξ *és* η *pontosan akkor függetlenek, ha tetszőleges* k, ℓ *esetén*

$$\mathbf{P}(\xi = k, \eta = \ell) = \mathbf{P}(\xi = k)\mathbf{P}(\eta = \ell).$$

Ha ξ, η *együttesen folytonos véletlen változók, akkor ők pontosan akkkor függetlenek, ha az együttes sűrűségfüggvényük az egyes sűrűségfüggvények szorzata, azaz*

$$f(x, y) = f_\xi(x) f_\eta(y).$$

Függetlenség

A $\xi_1, \xi_2, \ldots, \xi_n$ véletlen változók pontosan akkor függetlenek, ha tetszőleges $B_1, \ldots, B_n$ szép halmazokra (mondjuk intervallumokra) a $\{\xi_1 \in B_1\}$, ..., $\{\xi_n \in B_n\}$ események függetlenek.

Függetlenség

Dobókockával 2-szer dobunk, ξ az első, η a második dobott szám. Ekkor ξ és η függetlenek.

Várható érték

Várható érték

Definíció

Ha ξ diszkrét véletlen változó $x_1, x_2, \ldots$ lehetséges értékekkel, akkor az ξ *várható értéke*

$$\mathbf{E}(\xi) = \sum_i x_i \mathbf{P}(\xi = x_i),$$

ha $\sum_i |x_i| \mathbf{P}(\xi = x_i) < \infty$.
Ha ξ folytonos véletlen változó $f(x)$ sűrűségfüggvénnyel, akkor ξ várható értéke

$$\mathbf{E}(\xi) = \int_{-\infty}^{\infty} y f(y) \mathrm{d}y,$$

ha $\int_{-\infty}^{\infty} |y| f(y) \mathrm{d}y < \infty$.

Kockadobás

Véletlen pont

Várható érték tulajdonságai

Állítás

Legyenek ξ, η *véletlen változók,* $g: \mathbb{R} \to \mathbb{R}$, $h: \mathbb{R}^2 \to \mathbb{R}$ *olyan függvények, melyekre az állításokban szereplő várható értékek léteznek. Ekkor*

$$\mathbf{E}(g(\xi)) = \sum_{i=1}^{\infty} g(x_i)\mathbf{P}(\xi = x_i), \text{ ill. } \mathbf{E}(g(\xi)) = \int_{-\infty}^{\infty} g(y)f(y)\mathrm{d}y,$$

ahol $f(x)$ *az* ξ *sűrűségfüggvénye, és*

$$\mathbf{E}(h(\xi,\eta)) = \sum_i \sum_j h(x_i, y_j)\mathbf{P}(\xi = x_i, \eta = y_j), \text{ ill.}$$

$$\mathbf{E}(h(\xi,\eta)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f(x,y)\mathrm{d}x\mathrm{d}y,$$

ahol $f(x,y)$ *az* (ξ,η) *folytonos véletlen vektorváltozó sűrűségfüggvénye.*

Várható érték tulajdonságai

Állítás

A várható érték lineáris, azaz tetszőleges $a, b \in \mathbb{R}$ *állandókra*

$$\mathbf{E}(a\xi + b) = a\mathbf{E}(\xi) + b.$$

Várható érték tulajdonságai

Állítás

Ha $a \leq \xi \leq b$*, akkor* $a \leq \mathbf{E}(\xi) \leq b$ *tetszőleges* $a, b \in \mathbb{R}$ *számok esetén.*

Várható érték tulajdonságai

Állítás

$\mathbf{E}(\xi + \eta) = \mathbf{E}(\xi) + \mathbf{E}(\eta)$.

Várható érték tulajdonságai

Állítás

Ha $\xi_1, \xi_2, \ldots, \xi_n$ *véletlen változók, akkor*

$$\mathbf{E}\left(\sum_{i=1}^{n} \xi_i\right) = \sum_{i=1}^{n} \mathbf{E}(\xi_i).$$

Várható érték tulajdonságai

Állítás

Ha ξ és η függetlenek, akkor $\mathbf{E}\left(g_1(\xi)g_2(\eta)\right) = \mathbf{E}\left(g_1(\xi)\right)\mathbf{E}\left(g_2(\eta)\right)$*. Speciálisan, ha ξ és η függetlenek, akkor* $\mathbf{E}(\xi\eta) = \mathbf{E}(\xi)\mathbf{E}(\eta)$*.*

Példa

Csodaország munka törvénykönyve szerint egy cég minden munkása fizetett szabadságot kap azokon a napokon, amikor legalább az egyiküknek születésnapja van. Ezen napok kivételével azonban az év minden napján mindenkinek dolgoznia kell. Minden munkás 1 TV-készüléket készít egy nap alatt. Hány alkalmazottat vegyen fel a cégtulajdonos, ha azt akarja, hogy a gyártott TV-készülékek számának a várható értéke maximális legyen?

Csodaország

Csodaország

Huffman-kód

$\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ véges halmaz, a *forrásábécé*.
Kód $f : \mathcal{X} \to \{$véges 0-1 sorozatok $\}$
Az f-hez tartozó lehetséges kódszavak $f(x_1), f(x_2), \ldots, f(x_n)$.

Huffman-kód

$\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ véges halmaz, a *forrásábécé*.
Kód $f : \mathcal{X} \to \{$véges 0-1 sorozatok $\}$
Az f-hez tartozó lehetséges kódszavak $f(x_1), f(x_2), \ldots, f(x_n)$.
Az olyan kódolások érdekelnek, melyek egyértelműen dekódolhatók. Az f kód *prefix*, ha a lehetséges kódszavak közül egyik sem folytatása a másiknak. Világos, hogy egy prefix kód egyértelműen dekódolható. Jelölje $x \in \mathcal{X}$ esetén $|f(x)|$ a kódszó hosszát.

Példa

Legyen $\mathcal{X} = \{a, b, c\}$, és legyen $f_1(a) = 0$, $f_1(b) = 01$, $f_1(c) = 011$. Ekkor f_1 nem prefix kód, de könnyen látható, hogy egyértelműen dekódolható. Az $f_2(a) = 01$, $f_2(b) = 00$, $f_2(c) = 1$, kód prefix.

Legyen X egy véletlen betű, és eloszlása $\mathbf{P}(X = x_k) = p_k$, $k = 1, 2, \ldots, n$. (A forrásábécét helyettesíthetjük számokkal, és akkor a definíció szerinti véletlen változót kapunk.) Tehát p_k a x_k betű gyakorisága az adott nyelvben.

Adott f kód esetén egy hosszú szövegben az egy karakterre eső átlagos kódszóhossz:

Feltehető, hogy $p_1 \geq p_2 \geq \ldots \geq p_n$. Ha az f prefix kód optimális, akkor feltehető, hogy teljesülnek a következők:

(i) Hosszabb kódhoz ritkább betűk tartoznak, azaz

$$|f(x_1)| \leq |f(x_2)| \leq \ldots \leq |f(x_n)|.$$

Feltehető, hogy $p_1 \geq p_2 \geq \ldots \geq p_n$. Ha az f prefix kód optimális, akkor feltehető, hogy teljesülnek a következők:

(i) Hosszabb kódhoz ritkább betűk tartoznak, azaz

$$|f(x_1)| \leq |f(x_2)| \leq \ldots \leq |f(x_n)|.$$

(ii) A két legkisebb valószínűséghez tartozó kód hossza egyenlő.

Feltehető, hogy $p_1 \geq p_2 \geq \ldots \geq p_n$. Ha az f prefix kód optimális, akkor feltehető, hogy teljesülnek a következők:

(i) Hosszabb kódhoz ritkább betűk tartoznak, azaz

$$|f(x_1)| \leq |f(x_2)| \leq \ldots \leq |f(x_n)|.$$

(ii) A két legkisebb valószínűséghez tartozó kód hossza egyenlő.

(iii) $f(x_{n-1})$ és $f(x_n)$ csak az utolsó bitben térnek el.

Tétel

Tegyük fel, hogy az

$$\mathcal{X}' = \{x_1, \ldots, x_{n-2}, y_{n-1}\}$$

$(n-1)$ *elemű forrásábécé és* $p_1, \ldots, p_{n-2}, p_{n-1} + p_n$ *eloszlás esetén g egy optimális prefix kód. Ekkor az eredeti problémához tartozó optimális prefix kódot kapunk, ha az* x_{n-1}*, ill.* x_n *kódját úgy választjuk, hogy a* $g(y_{n-1})$ *kódszót kiegészítjük 0-val, ill. 1-gyel, a többi kódszót változatlanul hagyjuk.*

Példa

Legyen $n = 6$, $\mathcal{X} = \{x_1, \dots, x_6\}$, és $p_1 = 0.132$, $p_2 = 0.329$, $p_3 = 0.329$, $p_4 = 0.165$, $p_5 = 0.041$, $p_6 = 0.004$.

(i) $x_5, x_6 \to x_{56}$, $p_{56} = 0.045$;

Példa

Legyen $n = 6$, $\mathcal{X} = \{x_1, \ldots, x_6\}$, és $p_1 = 0.132$, $p_2 = 0.329$, $p_3 = 0.329$, $p_4 = 0.165$, $p_5 = 0.041$, $p_6 = 0.004$.

(i) $x_5, x_6 \to x_{56}$, $p_{56} = 0.045$;

(ii) x_1, x_{56}, $p_{156} = 0.177$;

Példa

Legyen $n = 6$, $\mathcal{X} = \{x_1, \dots, x_6\}$, és $p_1 = 0.132$, $p_2 = 0.329$, $p_3 = 0.329$, $p_4 = 0.165$, $p_5 = 0.041$, $p_6 = 0.004$.

(i) $x_5, x_6 \to x_{56}$, $p_{56} = 0.045$;

(ii) x_1, x_{56}, $p_{156} = 0.177$;

(iii) x_{156}, x_4, $p_{1564} = 0.342$;

Példa

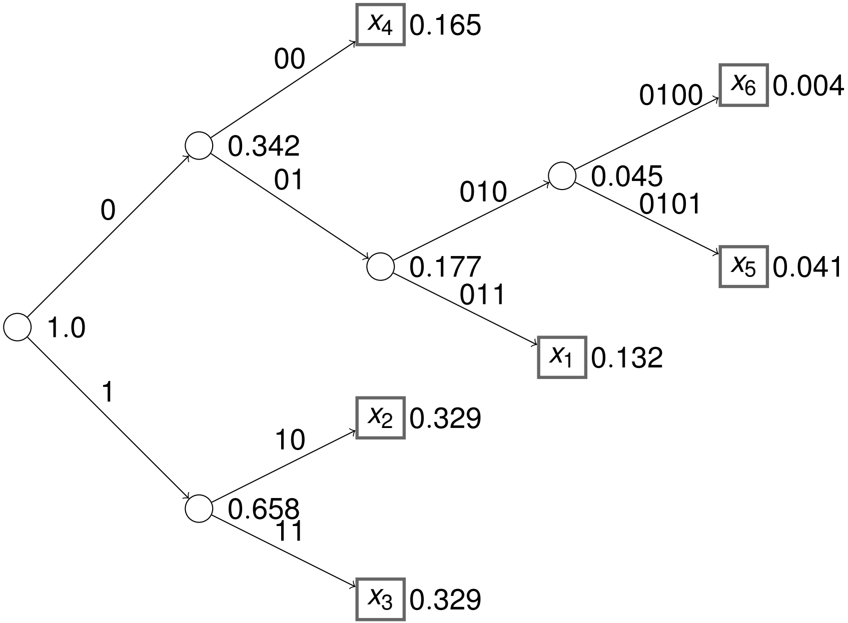
Legyen $n = 6$, $\mathcal{X} = \{x_1, \ldots, x_6\}$, és $p_1 = 0.132$, $p_2 = 0.329$, $p_3 = 0.329$, $p_4 = 0.165$, $p_5 = 0.041$, $p_6 = 0.004$.

(i) $x_5, x_6 \to x_{56}$, $p_{56} = 0.045$;

(ii) x_1, x_{56}, $p_{156} = 0.177$;

(iii) x_{156}, x_4, $p_{1564} = 0.342$;

(iv) x_2, x_3, $p_{23} = 0.658$.



x_4 0.165
00
x_6 0.004
0100
0.342
01
010
0.045
0101
0
x_5 0.041
0.177
011
1.0
x_1 0.132
1
x_2 0.329
10
0.658
11
x_3 0.329

Így az optimális kód:

x_1	x_2	x_3	x_4	x_5	x_6
011	10	11	00	0101	0100
0.132	0.329	0.329	0.165	0.041	0.004

A várható érték:

$$\begin{aligned}\mathbf{E}(f(X)) &= 0.132 \cdot 3 + 0.329 \cdot 2 + 0.329 \cdot 2 \\ &\quad + 0.165 \cdot 2 + 0.041 \cdot 4 + 0.004 \cdot 4 = 2.22\end{aligned}$$

Entrópia

Az optimális várható kódhosszra teljesül, hogy

$$\sum_{k=1}^{n} p_k \log_2 \frac{1}{p_k} \leq \mathbf{E}(|f(\xi)|) < \sum_{k=1}^{n} p_k \log_2 \frac{1}{p_k} + 1.$$