

A sztochasztika alapjai

MBNXK262

6. előadás: Folytonos véletlen változók; nevezetes folytonos eloszlások

Kevei Péter

2025/26 tavasz

Ismétlés

$(\Omega, \mathcal{A}, \mathbf{P})$ valószínűségi mező

$\xi : \Omega \mapsto \mathbb{R}$ függvény *véletlen változó eloszlásfüggvénye*

$$F(x) = \mathbf{P}(\xi < x) = \mathbf{P}(\{\omega : \xi(\omega) < x\}), \quad x \in \mathbb{R}$$

Egy ξ véletlen változó *folytonos eloszlású*, ha létezik egy nemnegatív f függvény, melyre

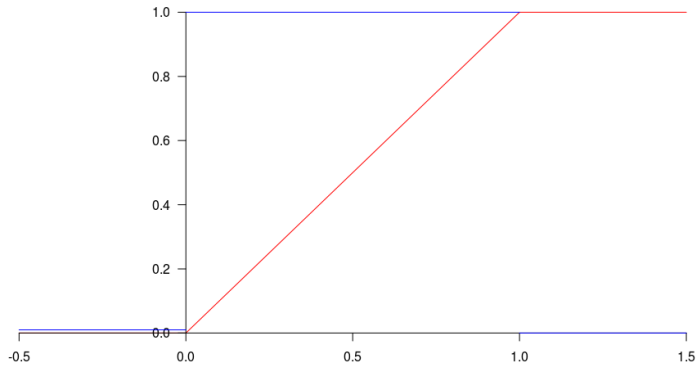
$$F(x) = \mathbf{P}(\xi < x) = \int_{-\infty}^x f(y) dy, \quad x \in \mathbb{R}.$$

Az $f(x)$ függvény a ξ véletlen változó sűrűségfüggvénye.

Egyenletes eloszlás

ξ egyenletes eloszlású az (a, b) intervallumon, $\xi \sim \text{Egy}(a, b)$,
 $-\infty < a < b < \infty$, ha sűrűségfüggvénye

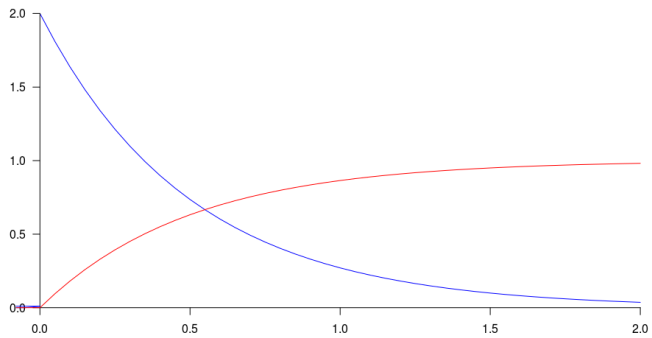
$$f(y) = \begin{cases} \frac{1}{b-a}, & \text{ha } y \in (a, b), \\ 0, & \text{különben.} \end{cases}$$



Exponenciális eloszlás

ξ λ -paraméterű exponenciális eloszlású, $\xi \sim \text{Exp}(\lambda)$, $\lambda > 0$, ha sűrűségfüggvénye

$$f(y) = \begin{cases} \lambda e^{-\lambda y}, & y \geq 0, \\ 0, & y < 0. \end{cases}$$



Tulajdonságok

- ▶ momentumok

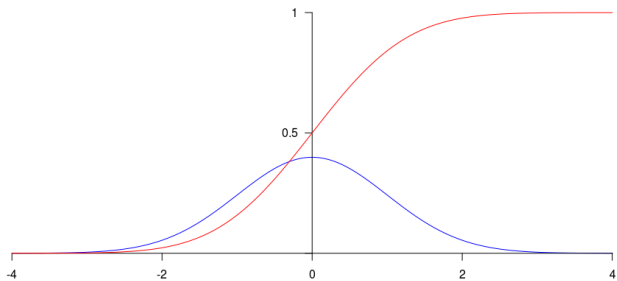
Tulajdonságok

- ▶ momentumok
- ▶ örökifjú

Normális eloszlás

ξ normális eloszlású μ és σ^2 paraméterekkel, $\xi \sim N(\mu, \sigma^2)$,
 $\mu \in \mathbb{R}, \sigma^2 > 0$, ha sűrűségfüggvénye

$$f_{\mu, \sigma}(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}.$$



Standardizálás

Állítás

Ha $\xi \sim N(\mu, \sigma^2)$, akkor $(\xi - \mu)/\sigma \sim N(0, 1)$.

[illegible]

$\xi \sim N(\mu, \sigma^2)$ és $Z \sim N(0, 1)$, akkor

$$\begin{aligned}\mathbf{P}(|\xi - \mu| \leq \lambda\sigma) &= \mathbf{P}(|Z| \leq \lambda) \\ &= \Phi(\lambda) - \Phi(-\lambda) \\ &= 2\Phi(\lambda) - 1 \\ &= \begin{cases} 0,6827, & \lambda = 1, \\ 0,9545, & \lambda = 2, \\ 0,9973, & \lambda = 3, \\ 0,9999, & \lambda = 4. \end{cases}\end{aligned}$$

Várható érték

Definíció

ξ diszkrét $x_1, x_2, \dots$ lehetséges értékekkel. 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 $f(x)$ sűrűségfüggvénnyel, akkor ξ várható értéke

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

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

Várható érték tulajdonságai

Állítás

ξ, η véletlen változók, $g : \mathbb{R} \rightarrow \mathbb{R}$, $h : \mathbb{R}^2 \rightarrow \mathbb{R}$. Ekkor

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

ahol $f(x)$ az ξ sűrűségfüggvénye, és

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

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

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

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

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

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

Állítás

Ha $\xi_1, \xi_2, \dots, \xi_n$ véletlen változók, akkor

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

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?

Huffman-kód

$\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ véges halmaz, a *forrásábécé*.

Kód $f : \mathcal{X} \rightarrow \{\text{véges 0-1 sorozatok}\}$

Az f -hez tartozó lehetséges kódszavak $f(x_1), f(x_2), \dots, f(x_n)$.

Huffman-kód

$\mathcal{X} = \{x_1, x_2, \dots, x_n\}$ véges halmaz, a *forrásábécé*.

Kód $f : \mathcal{X} \rightarrow \{\text{véges 0-1 sorozatok}\}$

Az f -hez tartozó lehetséges kódszavak $f(x_1), f(x_2), \dots, f(x_n)$.

Az f kód *prefix*, ha a lehetséges kódszavak közül egyik sem folytatása a másiknak. Jelölje $x \in \mathcal{X}$ esetén $|f(x)|$ a kódszó hosszát.

Legyen X egy véletlen betű, és eloszlása $\mathbf{P}(X = x_k) = p_k$, $k = 1, 2, \dots, n$. Tehát p_k a x_k betű gyakorisága az adott nyelvben.

Legyen X egy véletlen betű, és eloszlása $\mathbf{P}(X = x_k) = p_k$, $k = 1, 2, \dots, n$. 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 $p_1 \geq p_2 \geq \dots \geq p_n$. Ha az f prefix kód optimális, akkor feltehető, hogy teljesülnek a következők:

(i) Hosszabb kódhoz ritkább betűk tartoznak, azaz

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

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

(i) Hosszabb kódhoz ritkább betűk tartoznak, azaz

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

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

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

(i) Hosszabb kódhoz ritkább betűk tartoznak, azaz

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

(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

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

($n - 1$) elemű forrásábécé és $p_1, \dots, p_{n-2}, p_{n-1} + p_n$ 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) $x_5, x_6 \rightarrow x_{56}$, $p_{56} = 0.045$;

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) $x_5, x_6 \rightarrow x_{56}$, $p_{56} = 0.045$;

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

Példa

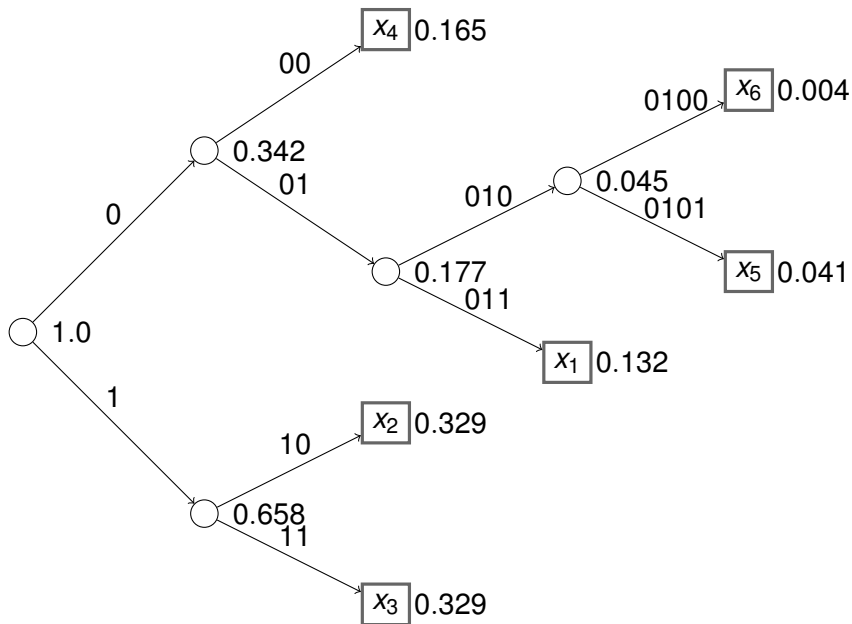
Legyen $n = 6$, $\mathcal{X} = \{x_1, \dots, x_6\}$, és $p_1 = 0.132$, $p_2 = 0.329$, $p_3 = 0.329$, $p_4 = 0.165$, $p_5 = 0.041$, $p_6 = 0.004$.

- (i) $x_5, x_6 \rightarrow x_{56}$, $p_{56} = 0.045$;
- (ii) x_1, x_{56} , $p_{156} = 0.177$;
- (iii) x_{156}, x_4 , $p_{1564} = 0.342$;

Példa

Legyen $n = 6$, $\mathcal{X} = \{x_1, \dots, x_6\}$, 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) $x_5, x_6 \rightarrow x_{56}$, $p_{56} = 0.045$;
- (ii) x_1, x_{56} , $p_{156} = 0.177$;
- (iii) x_{156}, x_4 , $p_{1564} = 0.342$;
- (iv) x_2, x_3 , $p_{23} = 0.658$.



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

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

Pl.: JPEG, MP3.