Valószínűségszámítás

5. előadás

Arató Miklós

2024.03.11.

Tartalomjegyzék

1. Feltételes várható érték
2. Valószínűségi vektorváltozók
   - Független valószínűségi változók
3. Abszolút folytonos eloszlású változók várható értéke

Meghatározás

Definíció: $P(A) > 0$, ξ diszkrét és véges várható értékű, ekkor ξ **feltételes várható értéke** az A feltételre nézve:
$E(\xi|A) = \sum_k x_k \cdot P(\xi = x_k|A)$.

$$\sum_k |x_k| \cdot P(\xi = x_k|A) = \sum_k |x_k| \cdot \frac{P(\xi = x_k, A)}{P(A)} \leq \frac{1}{P(A)} \sum_k |x_k| \cdot P(\xi = x_k)$$

Teljes várható érték tétel

Tétel: ξ diszkrét valószínűségi változó véges várható értékkel és $A_1, A_2, \ldots$ teljes eseményrendszer, ahol $0 < P(A_i)$. Ekkor $E\xi = \sum_n E(\xi|A_n) \cdot P(A_n)$.

Bizonyítás:

$$\sum_n E(\xi|A_n) \cdot P(A_n) = \sum_n \sum_k x_k \cdot P(\xi = x_k|A_n) \cdot P(A_n) =$$
$$\sum_k x_k \cdot \sum_n P(\xi = x_k|A_n) \cdot P(A_n) = \sum_k x_k \cdot P(\xi = x_k) = E\xi.$$

Wald-azonosság

$X_1, X_2, \dots$ független, azonos eloszlású valószínűségi változók, amelyekre létezik EX_i, és N egy tőlük független pozitív, egészértékű valószínűségi változó.

Ekkor $E(X_1 + \dots X_N) = EX_1 \cdot EN$.

$P(X_1 + \dots + X_N = y | N = n) = \frac{P(X_1+\dots+X_n=y, N=n)}{P(N=n)} =$
$P(X_1 + \dots + X_n = y) \Rightarrow E(X_1 + \dots + X_N = y | N = n) = n \cdot EX_1$
$\Rightarrow$

$$E(X_1 + \dots + X_N) = \sum_{n=1}^{\infty} E(X_1 + \dots + X_N | N = n) \cdot P(N = n) =$$

$$\sum_{n=1}^{\infty} n \cdot EX_1 \cdot P(N = n) = EX_1 \cdot EN$$

Geometriai eloszlás

$P(\eta = k) = (1-p)^{k-1} \cdot p$, ahol $k \geq 1$.

$E\eta = \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} \cdot p$.

A : az első kísérlet sikeres

$E\eta = E(\eta|A) \cdot P(A) + E(\eta|\overline{A}) \cdot P(\overline{A}) \Rightarrow$

$E\eta = 1 \cdot p + (1 + E\eta) \cdot (1-p) \Rightarrow E\eta = \frac{1}{p}$

Várhatóan mikor lesz meg mind a 24 törpünk?

$\frac{24}{24} + \frac{24}{23} + \ldots + \frac{24}{2} + \frac{24}{1} = 90,623$

szimmetrikus bolyongás

ξ : lépések száma n-ből 0-ba ($n \geq 0$), $v(n) := E\xi$.
A : az első lépésben jobbra lépünk $\Rightarrow$
$E\xi = E(\xi|A) \cdot \frac{1}{2} + E(\xi|\overline{A}) \cdot \frac{1}{2}$
$E(\xi|A) = 1 + v(n+1)$ és $E(\xi|\overline{A}) = 1 + v(n-1) \Rightarrow$
$2v(n) = 2 + v(n+1) + v(n-1)$
$v(0) = 0$
$v(n+1) - v(n) = v(n) - v(n-1) - 2 = \ldots = v(1) - v(0) - 2n.$
$\Rightarrow$
$v(n) = v(n) - v(n-1) + v(n-1) - v(n-2) + \ldots + v(1) - v(0) + v(0)$
$\Rightarrow n \cdot v(1) - n(n-1) = n(v(1) - (n-1)) \geq 0 \Rightarrow v(1) \geq n - 1$
minden n-re $\Rightarrow v(1) = +\infty \Rightarrow$
szimmetrikus bolyongásnál várhatóan végtelen sok lépésben térünk vissza a kiindulási pontba.

Meghatározás

Definíció: $\underline{\xi} : \Omega \to \mathbb{R}^p$ p-dimenziós **valószínűségi vektorváltozó**, ha minden $1 \leq i \leq p$-re ξ_i valószínűségi változó.

Meghatározás

Definíció: $\underline{\xi} : \Omega \to \mathbb{R}^p$ p-dimenziós **valószínűségi vektorváltozó**, ha minden $1 \leq i \leq p$-re ξ_i valószínűségi változó.

Definíció: A $\underline{\xi}$ valószínűségi vektorváltozó **diszkrét**, ha értékkészlete véges vagy megszámlálható, azaz léteznek olyan $\underline{x_k}$ p-dimenziós vektorok és A_k teljes eseményrendszer, hogy $\underline{\xi} = \sum_k \underline{x_k} \cdot \chi_{A_k}$.

Meghatározás

Definíció: $\underline{\xi}:\Omega\to\mathbb{R}^p$ p-dimenziós **valószínűségi vektorváltozó**, ha minden $1\le i\le p$-re ξ_i valószínűségi változó.

Definíció: A $\underline{\xi}$ valószínűségi vektorváltozó **diszkrét**, ha értékkészlete véges vagy megszámlálható, azaz léteznek olyan $\underline{x_k}$ p-dimenziós vektorok és A_k teljes eseményrendszer, hogy $\underline{\xi}=\sum_k \underline{x_k}\cdot\chi_{A_k}$.

Definíció: $\underline{\xi}$ **eloszlásfüggvénye**:
$F_{\underline{\xi}}(\underline{x}):=P(\xi_1<x_1,\dots,\xi_p<x_p)$.

Meghatározás

Definíció: $\underline{\xi} : \Omega \to \mathbb{R}^p$ p-dimenziós **valószínűségi vektorváltozó**, ha minden $1 \le i \le p$-re ξ_i valószínűségi változó.

Definíció: A $\underline{\xi}$ valószínűségi vektorváltozó **diszkrét**, ha értékkészlete véges vagy megszámlálható, azaz léteznek olyan $\underline{x_k}$ p-dimenziós vektorok és A_k teljes eseményrendszer, hogy $\underline{\xi} = \sum\limits_k \underline{x_k} \cdot \chi_{A_k}$.

Definíció: $\underline{\xi}$ **eloszlásfüggvénye**:
$F_{\underline{\xi}}(\underline{x}) := P(\xi_1 < x_1, \ldots, \xi_p < x_p)$.

Definíció: $\underline{\xi}$ abszolút folytonos eloszlású, ha $\exists f \ge 0$, hogy

$$F_{\underline{\xi}}(\underline{x}) = \int\limits_{-\infty}^{x_1} \cdots \int\limits_{-\infty}^{x_p} f(s_1, \ldots, s_p) ds_p \ldots ds_1.$$

f : **sűrűségfüggvény**.

Példák

Az $\{1, 2, 3, 4\}$ számokból húzunk visszatevés nélkül. Jelöljük X-el az első húzott számot és Y-al a másodikat.

Példák

Az $\{1, 2, 3, 4\}$ számokból húzunk visszatevés nélkül. Jelöljük X-el az első húzott számot és Y-al a másodikat.
Ekkor $P(X = i, Y = j) = \frac{1}{12}$, ha $1 \leq i, j \leq 4$, $i \neq j$ és 0 különben.

Példák

Az $\{1, 2, 3, 4\}$ számokból húzunk visszatevés nélkül. Jelöljük X-el az első húzott számot és Y-al a másodikat.
Ekkor $P(X = i, Y = j) = \frac{1}{12}$, ha $1 \leq i, j \leq 4, i \neq j$ és 0 különben.

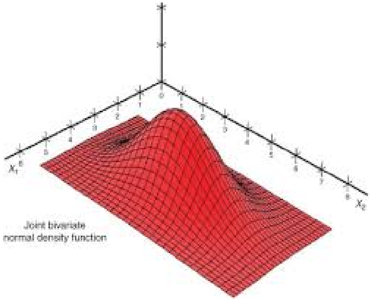
Egy kísérletnek k különböző kimenetele lehet. Az i-edik kimenetel valószínűsége p_i. Jelöljük η_i-vel az i-edik kimenetel gyakoriságát n független kísérletből. Ekkor
$P(\eta_1 = l_1, \eta_2 = l_2, \ldots, \eta_k = l_k) = \frac{n!}{l_1!l_2!\ldots l_k!} p_1^{l_1} p_2^{l_2} \ldots p_k^{l_k}, 0 \leq l_i, l_1 + \ldots, l_k = n$ (multinomiális eloszlás)

Kétdimenziós normális eloszlás

Sűrűségfüggvény:

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times$$

$$\times \exp\left\{ \frac{-1}{2(1-\rho^2)} \left(\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 \right) \right\}$$



Eloszlásfüggvény tulajdonságai

Tétel (biz. nélkül): F pontosan akkor egy p-dimenziós valószínűségi változó eloszlásfüggvénye, ha teljesülnek a következők:

1) F minden változóban monoton növő.

2) $\lim_{x_i \searrow -\infty} F_{\underline{\xi}}(\underline{x}) = 0 \ (i = 1, \ldots, p)$ és $\lim_{x_i \nearrow +\infty, i=1,\ldots,p} F_{\underline{\xi}}(\underline{x}) = 1$.

3) F minden változóban balról folytonos, és jobbról létezik a limesze.

4) Minden $a_k < b_k \ (k = 1, \ldots, p)$ esetén

$$\sum_{\varepsilon_i = 0;1} (-1)^{\sum \varepsilon_k} \cdot F\big(\varepsilon_1 a_1 + (1-\varepsilon_1) b_1 \ , \ \ldots \ , \ \varepsilon_p a_p + (1-\varepsilon_p) b_p\big) \geq 0$$

Egy $\underline{\xi}$ valószínűségi vektorváltozóra az utolsó összeg pont $P(a_k < \xi_k < b_k, k = 1, \ldots, p)$.

Meghatározás

Definíció: A $\xi_1, \ldots, \xi_n$ valószínűségi változók **függetlenek**, ha $P(\xi_1 < x_1, \ldots, \xi_n < x_n) = \prod_{i=1}^{n} F_{\xi_i}(x_i)$ minden $x_1, \ldots, x_n$-re.

Definíció: A $\xi_1, \xi_2, \ldots$ valószínűségi változók **függetlenek**, ha minden n-re $\xi_1, \ldots, \xi_n$ függetlenek.

Legyenek $\xi_1, \ldots, \xi_n$ diszkrétek. Ekkor pontosan akkor függetlenek, ha $P(\xi_1 = x_1, \ldots, \xi_n = x_n) = \prod_{i=1}^{n} P(\xi_i = x_i)$ minden x_i-re.

Legyenek $\xi_1, \ldots, \xi_n$ abszolút folytonos valószínűségi változók. Itt a függetlenség ekvivalens azzal, hogy $f_{\underline{\xi}}(\underline{x}) = \prod_{i=1}^{n} f_{\xi_i}(x_i)$.

Megjegyzések

$\xi_1, \dots, \xi_n$ függetlenekre $\{\xi_1 \in B_1\}, \dots, \{\xi_n \in B_n\}$ független események minden "szép" $B_1, \dots, B_n$ halmazra.
Konstans valószínűségi változó minden valószínűségi változótól független.
Legyenek $\xi_1, \dots, \xi_n$ függetlenek és g_i-k "szép" függvények. Ekkor $g_1(\xi_1), \dots, g_n(\xi_n)$ is függetlenek.
Legyenek $\xi_1, \dots, \xi_n$ függetlenek és h "szép", k-változós függvény. Ekkor $h(\xi_1, \dots, \xi_k), \xi_{k+1}, \dots, \xi_n$ is függetlenek.

Diszkrét konvolúciós formula (emlékeztető)

Legyenek ξ és η függetlenek, értékkészletük pedig $\{x_k\}$ és $\{y_l\}$.

$$P(\xi+\eta=z)=P(\bigcup_{x_k+y_l=z}\{\xi=x_k,\eta=y_l\})=$$

$$\sum_{x_k+y_l=z}P(\xi=x_k,\eta=y_l)=\sum_{x_k+y_l=z}P(\xi=x_k)\cdot P(\eta=y_l).$$

Konvolúciós formula

Legyenek ξ és η független, abszolút folytonos valószínűségi változók. Ekkor $\xi+\eta$ is abszolút folytonos eloszlású, és sűrűségfüggvénye

$$f_{\xi+\eta}(x) = \int\limits_{-\infty}^{+\infty} f_\xi(x-y)\cdot f_\eta(y)\,dy = \int\limits_{-\infty}^{+\infty} f_\xi(y)\cdot f_\eta(x-y)\,dy.$$

Exponenciálisak konvolúciója

$\xi_1, \dots, \xi_n$ független λ-exponenciális valószínűségi változók
$\eta_n = \xi_1 + \ldots + \xi_n \Rightarrow$.

Állítás: η_n sűrűségfüggvénye $g_n(x) = \begin{cases} 0 & x \le 0 \\ \frac{x^{n-1} \cdot \lambda^n \cdot e^{-\lambda x}}{(n-1)!} & x > 0 \end{cases}$.

Bizonyítás: n-re vonatkozó teljes indukció
$n = 1$ rendben . Tegyük fel, hogy n-ig igaz az állítás.
$\Rightarrow (n+1)$-re:

$$g_{n+1}(x) = f_{(\xi_1+\ldots+\xi_n+\xi_{n+1})}(x) = \int_{-\infty}^{\infty} \underbrace{f_{\xi_1+\ldots+\xi_n}(x-y)}_{g_n(x-y)} \cdot \underbrace{f_{\xi_{n+1}}(y)}_{g_1(y)} \, dy =$$

$$= \int_0^{\infty} \frac{(x-y)^{n-1} \lambda^n e^{-\lambda(x-y)}}{(n-1)!} =$$

$$=\lambda e^{-\lambda y} dy = \frac{\lambda^{n+1} e^{-\lambda x}}{n!} \int_0^x n(x-y)^{n-1} dy = \frac{x^n \lambda^{n+1}}{n!} e^{-\lambda x} (x > 0).$$

"Buszok száma"

Olyan autóbuszjárat, ahol a buszok követési ideje egymástól független, azonos λ-exponenciális eloszlású.
ξ_1 : az első busz beérkezési ideje, ξ_2 : az első és a második busz érkezése közötti idő, stb. N busz érkezik $[0, t)$-ben.

$$P(N = n) = P(N \geq n) - P(N \geq n+1) = P(\eta_n < t) - P(\eta_{n+1} < t) =$$

$$\int_0^t \frac{x^{n-1}\lambda^n e^{-\lambda x}}{(n-1)!} dx - \int_0^t \frac{x^n \lambda^{n+1} e^{-\lambda x}}{n!} dx =$$

$$\frac{\lambda^n}{n!} \left\{ \int_0^t n x^{n-1} e^{-\lambda x} dx - \underbrace{\int_0^t x^n \lambda e^{-\lambda x} dx}_{[x^n e^{-\lambda x}]_0^t - \int_0^t n x^{n-1} e^{-\lambda x} dx} \right\} = \frac{(\lambda t)^n e^{-\lambda t}}{n!},$$

Normálisak konvolúciója

Legyenek X és Y független normálisak,
$X \sim N(0,1)$, $Y \sim N(0,s^2)$. Ekkor összegük sűrűségfüggvénye.

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(x-y)^2}{2}\right\} \frac{1}{\sqrt{2\pi}s} \exp\left\{-\frac{y^2}{2s^2}\right\} dy =$$

$$\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2\pi}s} \exp\left\{-\frac{x^2}{2}\right\} \int_{-\infty}^{\infty} \exp\left\{-\frac{y^2(1+s^2)-2s^2xy}{2s^2}\right\} dy =$$

$$\frac{1}{\sqrt{2\pi(1+s^2)}} \exp\left\{-\frac{x^2}{2(1+s^2)}\right\} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\frac{s^2}{1+s^2}}} \exp\left\{-\frac{(y-x)^2}{2\frac{s^2}{1+s^2}}\right\} dy =$$

$$\frac{1}{\sqrt{2\pi(1+s^2)}} \exp\left\{-\frac{x^2}{2(1+s^2)}\right\} \Rightarrow X+Y \sim N(0,1+s^2)$$

Normálisak konvolúciója (folyt.)

Legyenek ξ és η független normálisak,
$\xi \sim N(\mu_1, \sigma_1^2), \eta \sim N(\mu_2, \sigma_2^2)$. Ekkor összegük:

$$\xi + \eta = \sigma_1 \left(\frac{\xi - \mu_1}{\sigma_1} + \frac{\sigma_2}{\sigma_1} \frac{\eta - \mu_2}{\sigma_2} \right) + \mu_1 + \mu_2$$

$$\frac{\xi - \mu_1}{\sigma_1} + \frac{\sigma_2}{\sigma_1} \frac{\eta - \mu_2}{\sigma_2} \sim N(0, 1 + \frac{\sigma_2^2}{\sigma_1^2}) \Rightarrow$$

$$\sigma_1 \left(\frac{\xi - \mu_1}{\sigma_1} + \frac{\sigma_2}{\sigma_1} \frac{\eta - \mu_2}{\sigma_2} \right) \sim N(0, \sigma_1^2 + \sigma_2^2) \Rightarrow$$

$$\xi + \eta \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

Meghatározás és tulajdonságok

Definíció: $\xi \geq 0$ **abszolút folytonos eloszlású**, ekkor $E\xi = \int_{\mathbb{R}} x \cdot f_\xi(x)\, dx$.

Definíció: ξ **abszolút folytonos eloszlású**, ekkor $E\xi = E\xi^+ - E\xi^-$, ha $\min(E\xi^+, E\xi^-) < \infty$.

$Eg(\xi) = \int_{\mathbb{R}} g(x) \cdot f_\xi(x)\, dx$

$\xi \geq 0 \Rightarrow E\xi = \int_0^\infty (1 - F_\xi(y))\, dy$

ξ, η függetlenek, $E|\xi|$, $E|\eta|$ végesek. $\Rightarrow E|\xi \cdot \eta|$ is véges és $E(\xi \cdot \eta) = E(\xi) \cdot E(\eta)$.

Abszolút folytonos eloszlású példák

1) $\xi \sim E(a,b)$ esetén $E\xi = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{a+b}{2}$.

Abszolút folytonos eloszlású példák

1) $\xi \sim E(a,b)$ esetén $E\xi = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{a+b}{2}$.

2) λ-**exponenciális** eloszlás várható értéke
$E\xi = \int_0^\infty e^{-\lambda y}\, dy = \frac{1}{\lambda}$.

Abszolút folytonos eloszlású példák

1) $\xi \sim E(a,b)$ esetén $E\xi = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{a+b}{2}$.

2) λ-**exponenciális** eloszlás várható értéke $E\xi = \int_0^\infty e^{-\lambda y}\, dy = \frac{1}{\lambda}$.

3) $\xi \sim N(0,1)$ esetén $E\xi = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}\, dx = 0$, hiszen a sűrűségfüggvény szimmetrikus, így az integrálban egy páratlan függvény szerepel (továbbá az integrál konvergens, mert elég nagy x-re $x \cdot e^{-\frac{x^2}{2}}$ felülről becsülhető az e^{-x} függvénnyel).
Általánosan pedig $m + \sigma\xi \sim N(m, \sigma^2)$ esetén $E(m + \sigma\xi) = m + \sigma \cdot E\xi = m$.