

Segédanyag a Valószínűségszámítás tantárgyhoz

2020. január 29.

Definíció. Véges valószínűségi mező: (Ω, \mathcal{A}, P) hármas, ahol:

- $\{\omega_1, \dots, \omega_n\} = \Omega$: eseménytér; elemei: elemi események
- $\{A, B, \dots\} = \mathcal{A} \subset 2^\Omega$, ahol A, B, \dots : események
- $P : \mathcal{A} \rightarrow [0; 1]$ valószínűség
 - $P(\Omega) = 1$
 - $P(A) \geq 0$ minden A eseményre
 - $P(A \cup B) = P(A) + P(B)$ minden A és B egymást kizáró eseményre ($A \cap B = \emptyset$).

Definíció. Kolgomorov-féle vsz.-i mező: (Ω, \mathcal{A}, P) hármas, ahol:

- Ω : nemüres halmaz
- $\mathcal{A} \subset 2^\Omega$ σ -algebra ($\emptyset \in \mathcal{A}$, zárt a komplementerképzésre és a megszámlálható unióra)
- $P : \mathcal{A} \rightarrow [0; 1]$ halmazfüggvény, amelyre
 - $P(\Omega) = 1$
 - $P(A) \geq 0 \quad \forall A \in \mathcal{A}$ -ra
 - páronként kizáró $A_1, A_2, \dots \in \mathcal{A}$ eseményekre

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(\bar{A}) = 1 - P(A)$$

Permutáció: n elem összes lehetséges sorrendje: $n!$

Ismétléses permutáció: n elem összes lehetséges sorrendje, ha ezek közül k_1, \dots, k_m darab megegyezik: $\frac{n!}{k_1! \dots k_m!}$

Adott n elem, ebből k darabot kivesszünk

	Kombináció: a kihúzás sorrendje NEM számít (nem számozottak az elemek)	Variáció: a kihúzás sorrendje számít (számozottak az elemek)
Visszatevés nélkül (ismétlés nélkül)	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$	$\frac{n!}{(n-k)!}$
Visszatevéssel (ismétléssel)	$\binom{n+k-1}{k}$	n^k

Mintavétel: Adott N termék, ezek közül M selejtes. Az összes termékből kivesszünk n darabot. Mi a valószínűsége, hogy ezek között k selejtes lesz? ($k = 0, 1, \dots, n$)

- Visszatevés nélkül: $\frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$
- Visszatevéssel: $\binom{n}{k} p^k (1-p)^{n-k}$ ahol $p = \frac{M}{N}$ a selejtarány

Feltételes valószínűség: Ha B bekövetkezett, mi a valószínűsége, hogy A bekövetkezik?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (P(B) \neq 0)$$

Definíció. Teljes eseményrendszer (TER) B_1, B_2, \dots események teljes eseményrendszert alkotnak, ha

- $B_i \cap B_j = \emptyset$ minden $i \neq j$ -re
- $\bigcup_{i=1}^{\infty} B_i = \Omega$

Tétel. Teljes valószínűség tétele: Legyen B_1, B_2, \dots teljes eseményrendszer, A tetszőleges esemény, $P(B_j) > 0$ minden j -re

$$\text{Ekkor } P(A) = \sum_{j=1}^{\infty} P(A|B_j)P(B_j).$$

Tétel. Bayes-tétel: Legyen B_1, \dots, B_n teljes eseményrendszer, A tetszőleges esemény, $P(B_j) > 0$ minden j -re

$$\text{Ekkor } P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum_{j=1}^{\infty} P(A|B_j)P(B_j)}.$$

Definíció. Események függetlensége: A és B események függetlenek, ha $P(A \cap B) = P(A) \cdot P(B)$ (A esemény bekövetkezése nem befolyásolja B esemény bekövetkezését, és fordítva).

Definíció. Valószínűségi változó: $X : \Omega \rightarrow \mathbb{R}$ mérhető függvény, azaz amire $\{\omega : X(\omega) \in B\} \in \mathcal{A}$ minden $B \subseteq \mathbb{R}$ nyílt halmazra.

Definíció. Valószínűségi változó eloszlása: $Q_X(B) = P(X \in B) = P(\omega : X(\omega) \in B)$

Definíció. Diszkrét valószínűségi változó: értékészlete legfeljebb meg-

számlálhatóan végtelen, azaz $\{x_1, \dots, x_n, \dots\}$ elemekből áll.

Ekkor eloszlása: $p_i := P(X = x_i) = P(\omega : X(\omega) = x_i)$

Tétel. Binomiális tétel: $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$.

Geometriai sor összege: $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$, ha $|q| < 1$.

Konvergenciatarományon belül "be lehet deriválni" egy végtelen sort, így $\sum_{n=1}^{\infty} nq^{n-1} = \frac{1}{(1-q)^2}$, ha $|q| < 1$.

Legyen (Ω, \mathcal{A}, P) valószínűségi mező, $X: \Omega \rightarrow \mathbb{R}$ valószínűségi változó.

Definíció. Várható érték X várható értéke: $EX = \int_{\Omega} X dP$, ha ez létezik.

Kiszámítását ld. később az egyes eseteknél.

Definíció. 1. momentum : $EX^l = \int_{\Omega} X^l dP$, ha ez létezik.

Definíció. X szórásnégyzete : $D^2X = E[(X-EX)^2] = EX^2 - E^2X$.

Definíció. X szórása : $DX = \sqrt{D^2X}$.

Legyen X diszkrét valószínűségi változó, ami az x_1, x_2, \dots értékeket veszi fel, p_1, p_2, \dots valószínűségekkel. Ekkor $EX = \sum_{i=1}^{\infty} x_i p_i$, ha a végtelen összeg abszolút

konvergens. Ugyanígy $EX^l = \sum_{i=1}^{\infty} (x_i)^l p_i$, ha a végtelen összeg abszolút konvergens.

Nevezetes diszkrét eloszlások:

Eloszlás neve	Jelölése	Eloszlása	EX	D ² X
Karakterisztikus (indikátorvált.)	Ind(p)	$P(X = 1) = p$ $P(X = 0) = 1 - p$	p	$p(1 - p)$
Geometriai (Pascal)	Geo(p)	$P(X = k) = p(1 - p)^{k-1}$ $k=1, 2, \dots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Hipergeometriai	Hipgeo(N, M, n)	$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$ $k=0, 1, \dots, n$	$n \frac{M}{N}$	$n \frac{M}{N} \left(1 - \frac{M}{N}\right) \left(1 - \frac{n-1}{N-1}\right)$
Binomiális	Bin(n, p)	$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$ $k=0, 1, \dots, n$	np	$np(1 - p)$
Negatív binomiális	NegBin(n, p)	$P(X = k) = \binom{k-1}{n-1} p^n (1 - p)^{k-n}$ $k=n, n+1, \dots$	$\frac{n}{p}$	$\frac{n(1-p)}{p^2}$
Poisson	Poi(λ)	$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ $k=0, 1, \dots$	λ	λ

Előfordulásuk:

- Indikátor változó: egy p valószínűségű esemény bekövetkezik-e vagy sem
- Geometriai: hányadikra következik be először egy p valószínűségű esemény
- Hipergeometriai: visszatevés nélküli mintavétel
- Binomiális: visszatevéses mintavétel
- Negatív binomiális: hányadikra következik be n . alkalommal egy p valószínűségű esemény

Állítás. Legyenek X, Y, X_1, \dots, X_n valószínűségi változók; $c, c_i, a, b \in \mathbb{R}$. Ekkor

- $E(X + Y) = EX + EY$;
- $E(cX) = cEX$;
- $E \sum_{i=1}^n c_i X_i = \sum_{i=1}^n c_i EX_i$;
- $D^2(aX + b) = a^2 D^2 X$.

Definíció. X val.változó eloszlásfüggvénye: $F_X(x) = P(X < x)$.

Amennyiben egyértelmű, melyik val.változó eloszlásfüggvényéről van szó, $F(x)$ -et írunk.

Állítás. Az eloszlásfüggvény tulajdonságai:

- $0 \leq F_X(x) \leq 1$;
- monoton növekvő;
- balról folytonos;
- $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$.

Állítás. Tetszőleges X val.változó esetén

- $P(a \leq X < b) = F(b) - F(a)$;
- $P(a < X \leq b) = F(b+) - F(a+)$.

Definíció. X val.változó abszolút folytonos, ha létezik olyan $f(x)$ függvény, amelyre $F(x) = \int_{-\infty}^x f(t) dt$. Ilyenkor $f(x)$ -et **sűrűségfüggvénynek** hívjuk.

Állítás. Legyen X abszolút folytonos eloszlású. Ekkor

- $f(x) = F'(x)$ (esetleg néhány pont kivételével);
- $f(x) \geq 0$;
- $\int_{-\infty}^{\infty} f(x) dx = 1$;
- $P(X = x) = 0 \quad \forall x$ -re;

• $P(a < X \leq b) = P(a \leq X < b) = F(b) - F(a)$.

Abszolút folytonos val.változó várható értéke: $EX = \int_{-\infty}^{\infty} xf(x) dx$.

Abszolút folytonos val.változó l . momentuma: $EX^l = \int_{-\infty}^{\infty} x^l f(x) dx$.

Nevezetes abszolút folytonos eloszlások:

Eloszlás neve	Jelölése	Eloszlásfüggvény	Sűrűségfüggvény	EX	D ² X
Egyenletes	E(a, b)	$\begin{cases} 0 & \text{ha } x \leq a \\ \frac{x-a}{b-a} & \text{ha } a < x \leq b \\ 1 & \text{ha } b < x \end{cases}$	$\begin{cases} \frac{1}{b-a} & \text{ha } a < x \leq b \\ 0 & \text{különben} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponenciális	Exp(λ)	$\begin{cases} 1 - e^{-\lambda x} & \text{ha } x \geq 0 \\ 0 & \text{különben} \end{cases}$	$\begin{cases} \lambda e^{-\lambda x} & \text{ha } x \geq 0 \\ 0 & \text{különben} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Standard normális	N(0, 1 ²)	$\Phi(x) = \dots$	$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad x \in \mathbb{R}$	0	1
Normális	N(m, σ^2)	...	$\frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad x \in \mathbb{R}$	m	σ^2

Állítás. Val.változó függvényének várható értéke

Legyen X val. változó; $g: \mathbb{R} \rightarrow \mathbb{R}$ függvény.

Ekkor

• $E(g(X)) = \sum_k g(x_k)p_k$, ha X diszkrét

• $E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x) dx$, ha X abszolút folytonos

Mindkét esetben a várható érték létezéséhez a szumma/integrál abszolút konvergenciájára van szükség.

Állítás. Abszolút folytonos val.változó szigorúan monoton függvényének eloszlás- és sűrűségfüggvénye

Legyen X abszolút folytonos val. változó; $g: \mathbb{R} \rightarrow \mathbb{R}$ szigorúan monoton, folytonosan differenciálható függvény, $g' \neq 0$. Ekkor

a.) $Y=g(X)$ eloszlásfüggvénye:

$$F_Y(y) = F_{g(X)}(y) = \begin{cases} F_X(g^{-1}(y)) & \text{ha } g \text{ szig. mon. növő} \\ 1 - F_X(g^{-1}(y)) & \text{ha } g \text{ szig. mon. csökkenő} \end{cases}$$

b.) $Y=g(X)$ sűrűségfüggvénye:

$$f_Y(y) = f_{g(X)}(y) = f_X(g^{-1}(y)) \cdot |[g^{-1}(y)]'| = \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}$$

Állítás. Standardizálás

Legyen $X \sim N(m, \sigma^2)$. Ekkor $\frac{X-m}{\sigma} \sim N(0, 1)$.

Állítás. $\Phi(-x) = 1 - \Phi(x)$

Állítás. $\Phi^{-1}(q) = -\Phi^{-1}(1 - q) \quad 0 < q < 1$

Definíció. Val.változók konvolúciója: Legyenek X és Y független val.változók. X és Y konvolúciójának (jel. $X*Y$) az $X+Y$ val.változót nevezük.

Állítás. A konvolúció eloszlásának meghatározása

• Diszkrét eset: $P(X + Y = k) = \sum_{l=0}^{\infty} P(X = l) \cdot P(Y = k - l)$

• Folytonos eset: $f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(u)f_Y(z - u)du$

Állítás. Legyenek X_1, \dots, X_n, X és Y független val. változók

- $X_1 \sim \text{Ind}(p), \dots, X_n \sim \text{Ind}(p) \Rightarrow X_1 + \dots + X_n \sim \text{Bin}(n, p)$
- $X \sim \text{Bin}(n, p), Y \sim \text{Bin}(m, p) \Rightarrow X + Y \sim \text{Bin}(n + m, p)$
- $X_1 \sim \text{Geo}(p), \dots, X_n \sim \text{Geo}(p) \Rightarrow X_1 + \dots + X_n \sim \text{NegBin}(n, p)$
- $X \sim \text{Poi}(\lambda_1), Y \sim \text{Poi}(\lambda_2) \Rightarrow X + Y \sim \text{Poi}(\lambda_1 + \lambda_2)$
- $X \sim N(m_1, \sigma_1^2), Y \sim N(m_2, \sigma_2^2) \Rightarrow X + Y \sim N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

Definíció. Valószínűségi vektorváltozó: $\mathbf{X}: \Omega \rightarrow \mathbb{R}^d$ mérhető függvény, azaz amire $\{\omega : \mathbf{X}(\omega) \in B\} \in \mathcal{A}$ minden $B \subseteq \mathbb{R}^d$ nyílt halmazra.

Definíció. Valószínűségi vektorváltozó eloszlása:

$$Q_{\mathbf{X}}(B) = P(\mathbf{X} \in B) = P(\omega : \mathbf{X}(\omega) \in B)$$

Definíció. X valószínűségi vektorváltozó eloszlásfüggvénye:

$$F_{\mathbf{X}}(\mathbf{x}) = P(\mathbf{X} < \mathbf{x}) = P(X_1 < x_1, \dots, X_d < x_d)$$

Állítás. Az eloszlásfüggvény tulajdonságai:

- $0 \leq F_{\mathbf{X}}(\mathbf{x}) \leq 1$;
- minden koordinátájában monoton növő;
- minden koordinátájában balról folytonos;
- $\lim_{x_1, \dots, x_d \rightarrow \infty} F_{\mathbf{X}}(x_1, \dots, x_d) = 1$;
- $\lim_{x_i \rightarrow -\infty} F_{\mathbf{X}}(x_1, \dots, x_d) = 0$ minden i -re.

Definíció. X valószínűségi vektorváltozó abszolút folytonos, ha létezik

olyan $f_{\mathbf{X}}(x_1, \dots, x_d)$ függvény, amelyre

$$F_{\mathbf{X}}(x_1, \dots, x_d) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_d} f_{\mathbf{X}}(t_1, \dots, t_d) dt_1 \dots dt_d.$$

Ilyenkor $f_{\mathbf{X}}(\mathbf{x})$ -et **sűrűségfüggvénynek** hívjuk.

Mostantól $d = 2$ lesz, és a következő jelöléseket és elnevezéseket használjuk:

- $F_{X,Y}(x, y) = P(X < x, Y < y) \rightsquigarrow$ együttes eloszlásfüggvény
- $F_X(x) = P(X < x)$
- $F_Y(y) = P(Y < y)$ \rightsquigarrow peremeloszlásfüggvények
- $f_{X,Y}(x, y) \rightsquigarrow$ együttes sűrűségfüggvény
- $f_X(x), f_Y(y) \rightsquigarrow$ peremsűrűségfüggvények

Állítás.

- $F_X(x) = \lim_{y \rightarrow \infty} F_{X,Y}(x, y)$ és $F_Y(y) = \lim_{x \rightarrow \infty} F_{X,Y}(x, y)$
- $F_{X,Y}(x, y) = \int_{-\infty}^y \int_{-\infty}^x f_{X,Y}(u, v) dudv$
- $f_{X,Y}(x, y) = \partial_y \partial_x F_{X,Y}(x, y) = \partial_x \partial_y F_{X,Y}(x, y)$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$
- $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$ és $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$

Állítás.

- X, Y függetlenek $\Leftrightarrow F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y)$
- X, Y függetlenek $\Leftrightarrow f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$
- Diszkrét esetben: X, Y függetlenek $\Leftrightarrow P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$
- X, Y függetlenek $\Rightarrow E(XY) = EX \cdot EY$

Definíció. X és Y kovarianciája: $\text{Cov}(X, Y) = E[(X - EX)(Y - EY)].$

Köv.: $\text{Cov}(X, Y) = E(XY) - EXEY.$

Elnevezés: ha $\text{Cov}(X, Y) = 0$, akkor azt mondjuk, hogy X és Y **korrelálatlanok**.

Állítás.

- Ha X és Y függetlenek egymástól, akkor korrelálatlanok is.
- Ha X és Y korrelálatlanok, akkor ebből **nem** következik, hogy függetlenek is!!!!

Állítás. A kovariancia tulajdonságai:

Legyenek X, Y, X_1, \dots, X_n valószínűségi változók, $a, b \in \mathbb{R}$. Ekkor

- $\text{Cov}(X, X) = D^2 X$
- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- $\text{Cov}(X, a) = 0$
- $\text{Cov}(aX, bY) = ab \text{Cov}(X, Y)$
- $D^2(X + Y) = D^2 X + D^2 Y + 2 \text{Cov}(X, Y)$
- $D^2(\sum_{i=1}^n X_i) = \sum_{i=1}^n D^2 X_i + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$
- X, Y függetlenek $\Rightarrow \text{Cov}(X, Y) = 0$

Definíció. X és Y korrelációja: $R(X, Y) = \frac{\text{Cov}(X, Y)}{D^2 X D^2 Y}.$

A korreláció két valószínűségi változó lineáris kapcsolatát méri:

- $R > 0 \Rightarrow$ pozitív a kapcsolat
- $R < 0 \Rightarrow$ negatív a kapcsolat
- $R^2 \sim 1 \Rightarrow$ erős a kapcsolat
- $R^2 \sim 0.5 \Rightarrow$ közepes a kapcsolat
- $R^2 \sim 0 \Rightarrow$ gyenge a kapcsolat

Legyenek X és Y valószínűségi változók. Y -nak X -re vonatkozó feltételes várható értéke – $E(Y|X)$ – precíz definiálására nem vállalkozok, úgy gondoljunk rá, mint egy valószínűségi változóra; és ha X egy adott értéket vesz fel – $E(Y|X = x)$ –, akkor mint konkrét számra.

$E(Y|X)$ abszolút folytonos eloszlások esetén a következő képlettel számítható:

$$E(Y|X) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy \Big|_{x=X}$$

ahol $f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$ a feltételes sűrűségfüggvény.

Állítás. Legyen g mérhető függvény.

- $E[g(X)|X] = g(X)$
- X, Y függetlenek $\Rightarrow E(Y|X) = EY$

Feladat: Y val. változót szeretnénk közelíteni X val. változó tetszőleges függvénye segítségével:

$$E|Y - f(X)|^2 \longrightarrow \min_f \rightsquigarrow \text{Megoldása: } f_{opt} = E(Y|X)$$

Feladat: Y val. változót szeretnénk közelíteni X val. változó lineáris függvénye segítségével:

$$E[Y - (aX + b)]^2 \longrightarrow \min_{a,b} \rightsquigarrow \text{Megoldása: } a_{opt} = \frac{Cov(X,Y)}{D^2(X)}$$

$$b_{opt} = EY - a_{opt}EX$$

Tétel. Markov-egyenlőtlenség: Legyen $g : \mathbb{R} \rightarrow \mathbb{R}$ monoton növekvő függvény, $X \geq 0$ val. változó, $\varepsilon > 0$ tetsz.

$$\text{Ekkor } P(X \geq \varepsilon) \leq \frac{E[g(X)]}{g(\varepsilon)}.$$

$$\text{Spec., ha } g(x) = x \Rightarrow P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}$$

$$\text{Tétel. Csebisev-egyenlőtlenség: } P(|X - EX| \geq \varepsilon) \leq \frac{D^2(X)}{\varepsilon^2}.$$

Legyen X_1, X_2, \dots valószínűségi változók sorozata.

Definíció. 1 valószínűségű konvergencia:

$$X_n \xrightarrow{n \rightarrow \infty} X \text{ 1 valószínűséggel, ha } P(\{\omega : X_n(\omega) \xrightarrow{n \rightarrow \infty} X(\omega)\}) = 1.$$

Definíció. Gyenge konvergencia: $X_n \xrightarrow{n \rightarrow \infty} X$ gyengén, ha az eloszlásfüggvényeikre $F_n(x) \xrightarrow{n \rightarrow \infty} F(x)$ F minden folytonossági pontjában.

Tétel. Nagy számok törvénye (NSZT):

Legyenek X_1, X_2, \dots i.i.d. val. változók, $EX_1 = m < \infty$.

Ekkor $\frac{X_1 + \dots + X_n}{n} \xrightarrow{n \rightarrow \infty} m$ 1 valószínűséggel.

Tétel. Centrális határeloszlás tétel (CHT):

Legyenek X_1, X_2, \dots i.i.d. val. változók, $EX_1 = m$, $D^2(X_1) = \sigma^2 < \infty$. Ekkor

$$\frac{X_1 + \dots + X_n - nm}{\sqrt{n\sigma}} \xrightarrow{n \rightarrow \infty} N(0,1) \text{ gyengén, azaz } P\left(\frac{X_1 + \dots + X_n - nm}{\sqrt{n\sigma}} < x\right) \xrightarrow{n \rightarrow \infty} \Phi(x).$$