

Véletlenszám-generátorok

2. rész

– egy MATLAB[®] alapú megközelítés –

Róth Ágoston, Vas Orsolya

Matematika és Informatika Intézet, Babeş-Bolyai Tudományegyetem, Kolozsvár, Románia

(agoston.roth@gmail.com, vas.orsolya@yahoo.com)

2. labor / 2020. október 5–8.



Elemi események és azok tere

- Egy kísérlet lehetséges kimeneteleit *elemi eseményeknek* nevezzük.
- Az elemi események együttesen az Ω *elemi események terét* határozzák meg. Ezért az adott kísérlethez tartozó bármely eseményt felfoghatjuk úgy, mint az elemi események Ω terének olyan részhalmazát, amelyet a kérdéses eseményt előidéző elemi események alkotnak. Amennyiben az adott kísérlet lehetséges eredményei legfeljebb megszámlálhatóak az eseményteret $\Omega = \{\omega_i\}_{i \in I}$ rendszerként is jelölhetjük, ahol I legfeljebb megszámlálhatóan végtelen indexhalmaz, ω_i pedig az adott kísérlet egy lehetséges kimenetelének felel meg.
- Az Ω tér összes részhalmazának halmazát $\mathcal{P}(\Omega)$ -val jelöljük.



Eseményalgebra

- $A \left(\mathcal{P}(\Omega), \cup, \cap, \bar{}, \emptyset, \Omega \right)$ struktúra Boole-algebrát alkot, azaz bármely $A, B, C \in \mathcal{P}(\Omega)$ esemény esetén teljesülnek az alábbi axiómák:

- 1 asszociativitás:

$$A \cup (B \cup C) = (A \cup B) \cup C, \quad A \cap (B \cap C) = (A \cap B) \cap C;$$

- 2 kommutativitás:

$$A \cup B = B \cup A, \quad A \cap B = B \cap A;$$

- 3 elnyelési tulajdonság:

$$A \cup (A \cap B) = A, \quad A \cap (A \cup B) = A;$$

- 4 disztributivitás:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C);$$

- 5 komplementer képzés:

$$A \cup \bar{A} = \Omega, \quad A \cap \bar{A} = \emptyset.$$

- Amennyiben az Ω eseménytér véges (végtelen), a

$$\left(\mathcal{P}(\Omega), \cup, \cap, \bar{}, \emptyset, \Omega \right)$$

eseményalgebrát végesnek (végtelennek) nevezzük.



Algebra

Az $\mathcal{A} \subset \mathcal{P}(\Omega)$ nem üres halmazt *algebrának* nevezzük, ha teljesülnek az

$$\begin{aligned} A \in \mathcal{A} &\Rightarrow \bar{A} \in \mathcal{A}, \\ A, B \in \mathcal{A} &\Rightarrow A \cup B \in \mathcal{A} \end{aligned}$$

feltételek.

σ -algebra

Az $\mathcal{A} \subset \mathcal{P}(\Omega)$ nem üres halmazt *σ -algebrának* nevezzük, ha teljesülnek az

$$\begin{aligned} A \in \mathcal{A} &\Rightarrow \bar{A} \in \mathcal{A}, \\ A_j \in \mathcal{A}, j \in J &\Rightarrow \bigcup_{j \in J} A_j \in \mathcal{A} \end{aligned}$$

feltételek, ahol J egy megszámlálhatóan végtelen indexhalmaz.

Véges és végtelen eseménymező

Ha az $\mathcal{A} \subset \mathcal{P}(\Omega)$ halmaz alegbrát (σ -algebrát) alkot, akkor az (Ω, \mathcal{A}) párost *véges (végtelen) eseménymezőnek* nevezzük.

Valószínűség axiomatikus értelmezése

Az (Ω, \mathcal{A}) eseménymezőn értelmezett $P : \mathcal{A} \rightarrow [0, 1]$ leképezést *valószínűségi függvénynek* (röviden *valószínűségnek*) nevezzük, ha teljesíti az alábbi axiómákat:

- 1 bármely $A \in \mathcal{A}$ esemény esetén $P(A) \geq 0$;
- 2 $P(\Omega) = 1$;
- 3 bármely J legfeljebb megszámlálható indexhalmaz és bármely $\{A_j\}_{j \in J}$ páronként kölcsönösen kizáró eseményekből álló rendszer esetén teljesül a

$$P\left(\bigcup_{j \in J} A_j\right) = \sum_{j \in J} P(A_j)$$

összefüggés.

Valószínűségi mező

A $P : \mathcal{A} \rightarrow [0, 1]$ valószínűségi függvénnyel felruházott véges (végtelen) (Ω, \mathcal{A}) eseménymezőt véges (végtelen) *valószínűségi mezőnek* nevezzük és az (Ω, \mathcal{A}, P) hármassal jelöljük.



- A továbbiakban olyan kísérleteket tanulmányozunk, amelyek elemi eseményeit (lehetséges kimeneteleit) számszerűen jellemezhetjük.

Példák

- Szabályos dobókockával játszva például az elemi események az 1, 2, 3, 4, 5, 6 számoknak felelnek meg.
- Ha céltáblára lövünk, a lehetséges becsapódási pontok olyan elemi eseményeket reprezentálnak, amelyeket szintén számokkal írhatunk le: az egyes találatokat értékelhetjük aszerint, hogy a becsapódási pont a céltábla közepén levő kis körbe, vagy rendre a körülötte levő egyes körgyűrűkbe, illetve a legkülső körön kívül esik.
- Továbbá, ha egy bizonyos terméket valamely automatizált gép sorozatban gyárt, a termék szabvány szerint előírt mérete az előírt tűréshatárok között (azaz egy adott intervallumban) tetszőlegesen változhat, ezért egy többtételes gyártási folyamatot kísérletnek foghatunk fel, amely lehetséges eredményeit számszerűen jellemezhetünk a gyártási folyamat során létrejött termékek méretével.
- Általában elmondhatjuk tehát, hogy lehetőségünk van bármely véletlen kísérlet elemi eseményeihez bizonyos számértékeket rendelni, azaz bármely véletlen kísérlet eseményterén értelmezhetünk egy vagy több valós függvényt. Az így értelmezett függvények értékei az őket meghatározó elemi eseményeken keresztül a véletlentől függnnek, ezért ezeket a függvényeket valószínűségi változóknak is nevezzük. Pontos értelmezésükhöz viszont szükségünk van az előző oldalakon ismertetett eseményalgebra, eseménymező és valószínűség-számítás alapvető fogalmainra.

Valószínűségi változó

Az (Ω, \mathcal{A}) eseménymezőn értelmezett $X : \Omega \rightarrow \mathbb{R}$ leképezést *valószínűségi változónak* nevezzük, ha $\forall x \in \mathbb{R}$ esetén teljesül az

$$X^{-1}(x) = \{\omega \in \Omega : X(\omega) = x\} \in \mathcal{A}$$

tulajdonság.

Valószínűségi vektor

Az (Ω, \mathcal{A}) eseménymezőn értelmezett

$$X(X_1, X_2, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$$

leképezést *valószínűségi vektornak* nevezzük, ha $\forall (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ esetén teljesül az

$$X^{-1}(x_1, x_2, \dots, x_n) \in \mathcal{A}$$

tulajdonság. Az n természetes számot pedig a valószínűségi vektor *dimenziójának* nevezzük.

Valószínűségi változó eloszlásfüggvénye

Tekintsük az (Ω, \mathcal{A}, P) valószínűségi mezőn értelmezett $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változót! Ekkor az

$$\begin{cases} F_X : \mathbb{R} \rightarrow \mathbb{R}, \\ F_X(x) = P(X \leq x) \end{cases} \quad (1)$$

leképezést az X valószínűségi változó *eloszlásfüggvényének* nevezzük.

Valószínűségi vektor együttes eloszlásfüggvénye

Ha

$$X(X_1, X_2, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$$

az (Ω, \mathcal{A}, P) valószínűségi mezőn értelmezett valószínűségi vektor, akkor az

$$\begin{cases} F_X : \mathbb{R}^n \rightarrow \mathbb{R}, \\ F_X(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) \end{cases} \quad (2)$$

többszörös leképezést az X valószínűségi vektor *együttes eloszlásfüggvényének* nevezzük.



Diszkrét valószínűségi változó

Az (Ω, \mathcal{A}) eseménymezőn értelmezett olyan X valószínűségi változót, amely legfeljebb megszámlálhatóan végtelen értéket vesz fel, *diszkrét valószínűségi változónak* nevezünk.

Diszkrét valószínűségi változó eloszlása és relatív gyakoriság függvénye

Az (Ω, \mathcal{A}, P) valószínűségi mezőn értelmezett X diszkrét valószínűségi változó eloszlásán az

$$X \left(\begin{array}{c} x_i \\ p_i \end{array} \right)_{i \in I} \quad (3)$$

táblázatot értjük, ahol:

- I vagy véges, vagy megszámlálhatóan végtelen indexhalmaz (utóbbi esetben $I \cong \mathbb{N}$);
- az $\{x_i\}_{i \in I}$ halmaz az X valószínűségi változó értékkészletének felel meg;
- a $p_i = P(X = x_i)$ valószínűség az x_i érték *relatív gyakoriságát* (megjelenési valószínűségét) jelöli és az

$$\begin{cases} f_X : \{x_i : i \in I\} \rightarrow \{p_i : i \in I\}, \\ f_X(x_i) = p_i \end{cases} \quad (4)$$

leképezést az X változó *relatív gyakoriság függvényének* nevezzük.

Diszkrét valószínűségi változó eloszlásfüggvénye

A (3)-as diszkrét valószínűségi változóhoz társított

$$F_X(x) = P(X \leq x) \in [0, 1], x \in \{x_i\}_{i \in I}$$

leképezést *eloszlásfüggvénynek*, vagy *relatív kumulatív (összegzett) gyakoriság függvénynek* nevezzük.

Következmény

Figyelembe véve, hogy a (3)-as diszkrét valószínűségi változó esetén az

$$\{(X = x_i)\}_{i \in I}$$

események egy teljes eseményrendszert határoznak meg – azaz az $(X = x_i)$ események $(i \in I)$ páronként kizáróak és egyesítésük a biztos eseményt eredményezi –, következik, hogy

$$F_X(x_i) = P(X \leq x_i) = P\left(\bigcup_{j \in J_i} (X = x_j)\right) = \sum_{j \in J_i} P(X = x_j) = \sum_{j \in J_i} f_X(x_j),$$

ahol

$$J_i = \{j \in I : x_j \leq x_i\}.$$



Következmény – folytatás

Amennyiben az $\{x_i\}_{i \in I}$ halmaz elemei növekvően rendezettek is, az eloszlásfüggvényre az

$$F_X(x_i) = \sum_{j \in I: j < i} P(X = x_j) = \sum_{j \in I: j < i} f_X(x_j)$$

egyszerűbb kifejezést kapjuk. A továbbiakban bármely diszkrét valószínűségi változó értékkészletét növekvően rendezettnek tekintjük.

Megjegyzés

A MATLAB[®] számos beépített diszkrét eloszlásfüggvénnyel rendelkezik, viszont ezek implementációja az

$$F_X(x) = P(X \leq x), x \in \{x_i\}_{i \in I}$$

képletre épül: azaz a $<$ összehasonlító műveletet a \leq logikai operátorra cserélték le. Mivel az a célunk, hogy a saját függvényeink által adott eredményeket összehasonlíthassuk a beépített függvények által generáltakkal, a saját diszkrét eloszlásfüggvényeink implementációja során az utóbbi értelmezést használjuk majd mi is.



Diszkrét valószínűségi vektorok együttes relatív gyakoriság függvénye

Az

$$X_i \left(\begin{array}{c} x_{i,j_i} \\ p_{i,j_i} \end{array} \right)_{j_i \in J_i}, \quad i = 1, 2, \dots, n, \quad n \geq 1$$

diszkrét valószínűségi változókból képezett n -dimenziós $X(X_1, X_2, \dots, X_n)$ valószínűségi vektorhoz társított

$$\begin{cases} f_X : \{x_{1,j_1}\}_{j_1 \in J_1} \times \{x_{2,j_2}\}_{j_2 \in J_2} \times \dots \times \{x_{n,j_n}\}_{j_n \in J_n} \rightarrow \mathbb{R}, \\ f_X(\xi_1, \xi_2, \dots, \xi_n) = P(X_1 = \xi_1, X_2 = \xi_2, \dots, X_n = \xi_n) \end{cases}$$

leképezést az X vektor *együttes relatív gyakoriság függvényének* nevezzük.



Folytonos valószínűségi változó

Tekintsük az (Ω, \mathcal{A}, P) valószínűségi mezőn értelmezett $F_X : \mathbb{R} \rightarrow \mathbb{R}$ eloszlásfüggvényű $X : \Omega \rightarrow \mathbb{R}$ valószínűségi változót! Azt mondjuk, hogy az X valószínűségi változó *folytonos*, ha az F_X eloszlásfüggvény abszolút folytonos, azaz ha létezik egy olyan $f_X : \mathbb{R} \rightarrow \mathbb{R}$ valós függvény, amelyre fennáll az

$$F_X(x) = \int_{-\infty}^x f_X(t) dt \quad (5)$$

egyenlőség minden $x \in \mathbb{R}$ valós számra.

Folytonos valószínűségi változó sűrűségfüggvénye

Ha az X valószínűségi változó folytonos, akkor az előző értelmezésbeli f_X leképezést az X valószínűségi változó *sűrűségfüggvényének* nevezzük.



Folytonos valószínűségi változó sűrűségfüggvényének tulajdonságai

Adott f_X sűrűségfüggvényű és F_X eloszlásfüggvényű folytonos X valószínűségi változóra teljesülnek az alábbi kijelentések:

- 1 minden $x \in \mathbb{R}$ valós számra fennáll a $\frac{d}{dx} F_X(x) = f_X(x)$ azonosság;
- 2 bármely $x \in \mathbb{R}$ értékre teljesül az $f_X(x) \geq 0$ egyenlőtlenség;
- 3 érvényes az $\int_{\mathbb{R}} f_X(t) dt \equiv 1$ összefüggés;
- 4 minden $x \in \mathbb{R}$ esetén $P(X = x) = 0$, másrészt tetszőleges $a < b$ valós számokra pedig igaz a

$$\begin{aligned}
 P(a < X < b) &= P(a \leq X < b) \\
 &= P(a \leq X \leq b) \\
 &= P(a < X \leq b) \\
 &= \int_a^b f_X(t) dt
 \end{aligned}$$

egyenlőséglánc.



Folytonos valószínűségi vektorok

Tekintsük az (Ω, \mathcal{A}, P) valószínűségi mezőn értelmezett $F_X : \mathbb{R}^n \rightarrow \mathbb{R}$ együttes eloszlásfüggvényű $X(X_1, X_2, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ valószínűségi vektort! Azt mondjuk, hogy az X valószínűségi vektor *folytonos*, ha az együttes eloszlásfüggvénye abszolút folytonos, vagyis ha létezik olyan $f_X : \mathbb{R}^n \rightarrow \mathbb{R}$ többváltozós valós értékű függvény, amely teljesíti az

$$F_X(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f_X(t_1, t_2, \dots, t_n) dt_1 dt_2 \dots dt_n \quad (6)$$

egyenlőséget minden $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ pontra. Ugyanekkor az f_X leképezést az X vektor *együttes sűrűségfüggvényének* nevezzük.



Együttes sűrűségfüggvény tulajdonságai

Adott f_X együttes sűrűségfüggvényű és F_X együttes eloszlásfüggvényű $X(X_1, X_2, \dots, X_n)$ valószínűségi vektorra teljesülnek az alábbi kijelentések:

- ❶ minden $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ pontra fennáll a

$$\frac{\partial^n}{\partial x_1 \partial x_2 \dots \partial x_n} F_X(x_1, x_2, \dots, x_n) = f_X(x_1, x_2, \dots, x_n)$$

összefüggés;

- ❷ bármely $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ pontra teljesül az

$$f_X(x_1, x_2, \dots, x_n) \geq 0$$

egyenlőtlenség;

- ❸ érvényes az

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f_X(t_1, t_2, \dots, t_n) dt_1 dt_2 \dots dt_n \equiv 1$$

azonosság;



Együttes sűrűségfüggvény tulajdonságai – folytatás

- 4 tetszőleges $T \subseteq \mathbb{R}^n$ tartományra az $(X = (X_1, X_2, \dots, X_n) \in T)$ esemény valószínűségét a

$$P(X \in T) = \int \int \dots \int_T f_X(t_1, t_2, \dots, t_n) dt_1 dt_2 \dots dt_n$$

képlettel határozhatjuk meg;

- 5 az X_i komponens (valószínűségi változó) perem-sűrűségfüggvényét az

$$f_{X_i}(t) = \underbrace{\int_{\mathbb{R}} \int_{\mathbb{R}} \dots \int_{\mathbb{R}}}_{n-1} f_X(t_1, t_2, \dots, t_{i-1}, t, t_{i+1}, t_n) dt_1 dt_2 \dots dt_{i-1} dt_{i+1} \dots dt_n,$$

$$t \in \mathbb{R}$$

alakban írhatjuk minden $i = 1, 2, \dots, n$ index esetén.

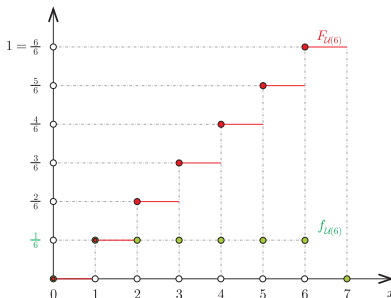


Diszkrét egyenletes eloszlás

Az $n \geq 1$ természetes szám által meghatározott

$$X \left(\begin{array}{cccc} 1 & 2 & \cdots & n \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{array} \right)$$

véges diszkrét valószínűségi változót *n*-edrendű *egyenletes eloszlásúnak* nevezzük és az $\mathcal{U}(n)$ kifejezéssel jelöljük.



1. ábra. A dobókocka oldalait jellemző diszkrét egyenletes eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Bernoulli-eloszlás

A $p \in (0, 1)$ valószínűséggel leírt

$$X \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$$

valószínűségi változóvalt p -paraméterű *Bernoulli-eloszlásúnak* nevezzük és a $\mathcal{B}ern(p)$ módon jelöljük.

- Ez a legegyszerűbb diszkrét valószínűség-eloszlás, amelyet döntési alternatívaként használhatunk egy adott esemény bekövetkezésére, vagy megghiúsulására.

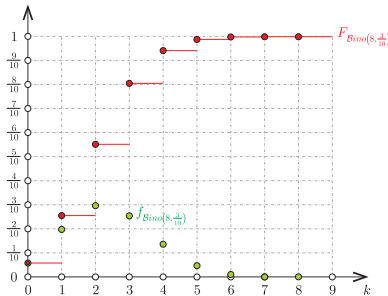


Binomiális eloszlás

Az $n \geq 1$ természetes számtól és $p \in (0, 1)$ valószínűségtől függő

$$X \left(\binom{n}{k} p^k (1-p)^{n-k} \right)_{k \in \{0, 1, \dots, n\}}$$

valószínűségi változót (n, p) -paraméterezésű *binomiális eloszlásúnak* nevezzük és $Bino(n, p)$ módon jelöljük.



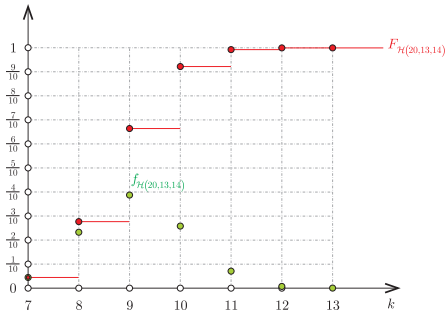
2. ábra. $Bino(8, \frac{3}{10})$ -eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Hipergeometrikus eloszlás

Az $N \geq 1$, $0 \leq M \leq N$, $0 \leq n \leq N$ természetes számokkal értelmezett

$$X \left(\frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \right)_{\max\{0, n-N+M\} \leq k \leq \min\{n, M\}}$$

diszkrét valószínűségi változót (N, M, n) -paraméterezésű hipergeometrikus eloszlásúnak nevezzük és $\mathcal{H}(N, M, n)$ módon jelöljük.



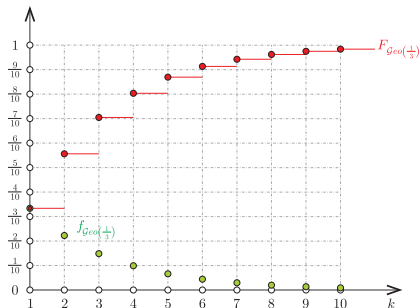
3. ábra. $\mathcal{H}(20, 13, 14)$ -eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Geometriai eloszlás

A $p \in (0, 1)$ valószínűség által meghatározott

$$X \left(\binom{k}{(1-p)^{k-1} p} \right)_{k \geq 1}$$

megszámlálhatóan végtelen diszkrét valószínűségi változót p -paraméterű geometriai eloszlásúnak nevezzük és a $\mathcal{Geo}(p)$ kifejezéssel jelöljük.



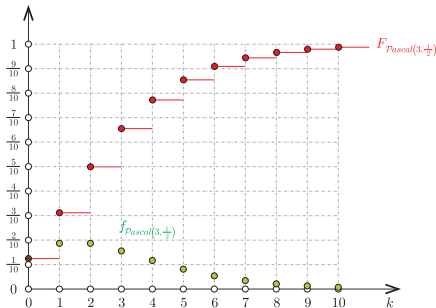
4. ábra. $\mathcal{Geo}(\frac{1}{3})$ -eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Pascal-féle vagy negatív binomiális eloszlás

Az $n \geq 1$ természetes számmal és a $p \in (0, 1)$ valószínűséggel leírt

$$X \left(\binom{n+k-1}{k} p^n (1-p)^k \right)_{k \geq 0}$$

megszámlálhatóan végtelen valószínűségi változót n -edrendű p -paraméterű Pascal-féle vagy negatív binomiális eloszlásúnak nevezzük és a $\mathcal{Pascal}(n, p)$ szimbólummal jelöljük.



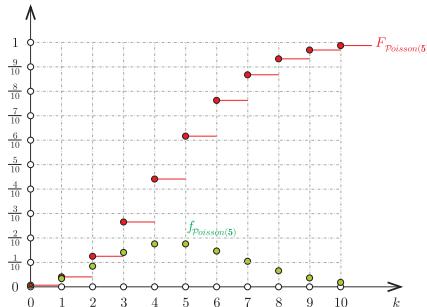
5. ábra. $\mathcal{Pascal}(3, \frac{1}{2})$ -eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Poisson-eloszlás

A $\lambda > 0$ állandó segítségével megszerkesztett

$$X \left(\frac{\lambda^k}{k!} e^{-\lambda} \right)_{k \geq 0}$$

diszkrét valószínűségi változót λ -paraméterű Poisson-eloszlásúnak nevezzük és $Poisson(\lambda)$ alakban hivatkozunk rá.



6. ábra. $Poisson(5)$ -eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Folytonos egyenletes eloszlás

Az $[a, b]$ intervallumon ($a < b$) értelmezett,

$$f_{\mathcal{U}([a,b])}(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b], \\ 0, & x \notin [a, b] \end{cases}$$

sűrűségfüggvényű folytonos valószínűségi változót *egyenletes eloszlásúnak* nevezzük és az $\mathcal{U}([a, b])$ kifejezéssel jelöljük.

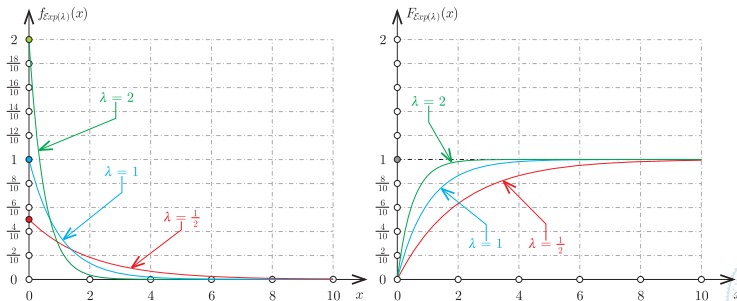


Exponenciális eloszlás

A $\lambda > 0$ valós számtól függő

$$f_{\mathcal{Exp}(\lambda)}(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0 \end{cases} \quad (7)$$

sűrűségfüggvényű folytonos valószínűségi változót λ -paraméterű exponenciális eloszlásúnak nevezzük és az $\mathcal{Exp}(\lambda)$ szimbólummal jelöljük.



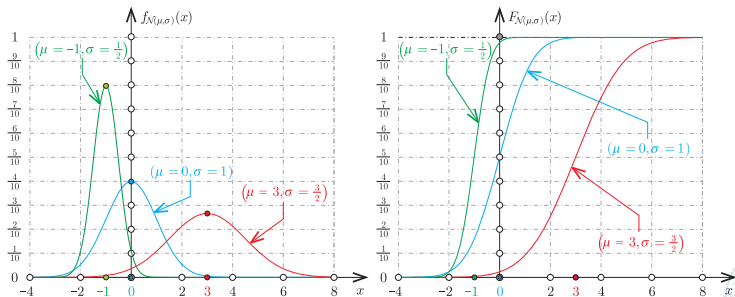
7. ábra. Az $\mathcal{Exp}(\lambda)$ -eloszlás sűrűség- és eloszlásfüggvénye a $\lambda = \frac{1}{2}, 1, 2$ paraméterek esetén

Normális eloszlás

A $\mu \in \mathbb{R}$ és $\sigma > 0$ paraméterekkel értelmezett

$$f_{\mathcal{N}(\mu, \sigma)}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R} \quad (8)$$

sűrűségfüggvényű folytonos valószínűségi változót *normális eloszlásúnak* hívjuk és az $\mathcal{N}(\mu, \sigma)$ kifejezéssel jelöljük. A $\mu = 0$ és $\sigma = 1$ sajátos esetben *standard normális eloszlású* változóról beszélünk.



8. ábra. Az $\mathcal{N}(\mu, \sigma)$ -eloszlás sűrűség- és eloszlásfüggvénye a $(\mu = 0, \sigma = 1)$, $(\mu = -1, \sigma = \frac{1}{2})$ és $(\mu = 3, \sigma = \frac{3}{2})$ beállítások esetén

Többváltozós normális eloszlású valószínűségi vektor

Az általános d -dimenziós ($d \geq 2$) normális eloszlású $X = {}^t[X_i]_{i=1}^d$ valószínűségi vektort az

$$f_{\mathcal{N}_d(\mu, \Sigma)}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det \Sigma}} e^{-\frac{1}{2}(\mathbf{x}-\mu)^t \Sigma^{-1}(\mathbf{x}-\mu)}, \mathbf{x} = {}^t[x_i] \in \mathbb{R}^d \quad (9)$$

sűrűségfüggvénnyel jellemezzük és az $\mathcal{N}_d(\mu, \Sigma)$ szimbólummal jelöljük, ahol a nem feltétlenül független X_i komponensek $\mathcal{N}(\mu_i, \sigma_i)$ paraméterű ($\mu_i \in \mathbb{R}, \sigma_i > 0$) normális eloszlású valószínűségi változók,

$$\mu = {}^t[E(X_i)]_{i=1}^d = {}^t[\mu_i]_{i=1}^d \in \mathbb{R}^d$$

az egyes komponensek várható értékeit tartalmazó vektor,

$$\begin{aligned} \Sigma &= [\text{cov}(X_i, X_j)]_{i=1, j=1}^{d, d} = [E((X_i - \mu_i)(X_j - \mu_j))]_{i=1, j=1}^{d, d} \\ &= [\rho_{ij} \sigma_i \sigma_j]_{i=1, j=1}^{d, d} \in \mathcal{M}_{d, d}(\mathbb{R}) \end{aligned}$$

a komponensek közti kovariancia együtthatókból képezett pozitív definit szimmetrikus mátrix, és

$$[\rho_{ij}]_{i=1, j=1}^{d, d} \in \mathcal{M}_{d, d}([-1, 1])$$

a komponensek közti korrelációs együtthatók szimmetrikus mátrixa.

- Kétdimenziós esetben, a $\mu = {}^t[\mu_1, \mu_2]$ és a

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

megválasztással, a $\rho \in [-1, +1]$ korrelációs együtthatójú

$${}^t[X_1 \sim \mathcal{N}(\mu_1, \sigma_1), X_2 \sim \mathcal{N}(\mu_2, \sigma_2)]$$

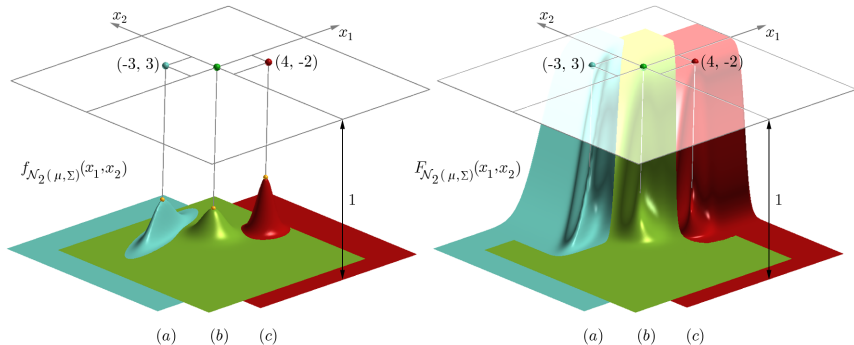
binormális eloszlású valószínűségi vektor sűrűségfüggvénye az

$$f_{\mathcal{N}_2(\mu, \Sigma)}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1-\mu_1}{\sigma_1} \right)^2 - 2\rho \cdot \frac{x_1-\mu_1}{\sigma_1} \cdot \frac{x_2-\mu_2}{\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2} \right)^2 \right]},$$

$$(x_1, x_2) \in \mathbb{R}^2$$

alakot ölti, amely valójában egy kétváltozós valós értékű függvény, azaz egy domborzatjellegű felület.





9. ábra. Különböző paraméterezésű kétdimenziós normális eloszlások sűrűség- és eloszlásfüggvényeit láthatjuk: (a) $\mu = [\mu_1, \mu_2] = [-3, 3]$, $\sigma_1 = \frac{7}{4}$, $\sigma_2 = \frac{3}{4}$, $\rho = 0$, 8 (a komponensek különböző szórásúak és korreláltak); (b) $\mu = [\mu_1, \mu_2] = [0, 0]$, $\sigma_1 = \sigma_2 = 1$, $\rho = 0$ (standard normális eloszlás); (c) $\mu = [\mu_1, \mu_2] = [4, -2]$, $\sigma_1 = \frac{3}{4}$, $\sigma_2 = \frac{2}{3}$, $\rho = 0$ (a komponensek függetlenek, de különböző szórásúak).

- A továbbiakban gyakran fogunk használni két segédfüggvényt, ezért célszerűnek tartjuk már most bevezetni ezeket.

Euler-féle gamma-függvény

Minden szigorúan pozitív valós részű $z \in \mathbb{C}$ komplex számra a

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt \quad (10)$$

integrál abszolút konvergens. Az így értelmezett leképezést *Euler-féle gamma-függvénynek* nevezzük. Parciális integrálással kimutatható a

$$\Gamma(z+1) = z\Gamma(z), \quad \forall z \in \mathbb{C} : \operatorname{Re}(z) > 0 \quad (11)$$

rekurzió is. Mivel minden $n \in \mathbb{N}$ természetes számra $\Gamma(n) = (n-1)!$, az Euler-féle gamma-függvényt a természetes számok faktoriális műveletének kiterjesztéseként foghatjuk fel. Érdekességgéppen megemlíthjük, hogy $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ (ez a speciális érték időnként előfordul majd a számításaink során).



Euler-féle béta-függvény

Bármely szigorúan pozitív valós részű $x, y \in \mathbb{C}$ komplex számpárra a

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt \quad (12)$$

integrál abszolút konvergens és *Euler-féle béta-függvénynek* nevezzük. A most értelmezett függvény kapcsolatban áll az Euler-féle gamma-függvénnyel a

$$B(x, y) = \frac{\Gamma(x) \Gamma(y)}{\Gamma(x+y)}, \quad \forall x, y \in \mathbb{C} : \operatorname{Re}(x) > 0, \operatorname{Re}(y) > 0 \quad (13)$$

összefüggésen keresztül.



Pearson-féle $\chi^2(n, \sigma)$ -eloszlás

Az

$$f_{\chi^2(n, \sigma)}(x) = \begin{cases} \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2\sigma^2}}}{2^{\frac{n}{2}} \sigma^n \Gamma\left(\frac{n}{2}\right)}, & x > 0, \\ 0, & x \leq 0 \end{cases} \quad (14)$$

sűrűségfüggvénnyel adott $\sigma > 0$ *skálázási tényezőjű* és $n \geq 1$ *szabadságfokú/alakparaméterű* ($n \in \mathbb{N}$) folytonos valószínűségi változót $\chi^2(n, \sigma)$ -eloszlásúnak nevezzük.

Megjegyzés

Amint hamarosan látni fogjuk, a $\lambda > 0$ paraméterű exponenciális, illetve az $n \geq 1$ szabadságfokú és $\sigma > 0$ skálázási tényezőjű $\chi^2(n, \sigma)$ -eloszlású folytonos valószínűségi változókat egy speciális, úgynevezett gamma-eloszlás sajátosságos paraméterezésű eseteiként kaphatjuk vissza.

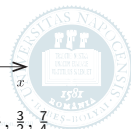
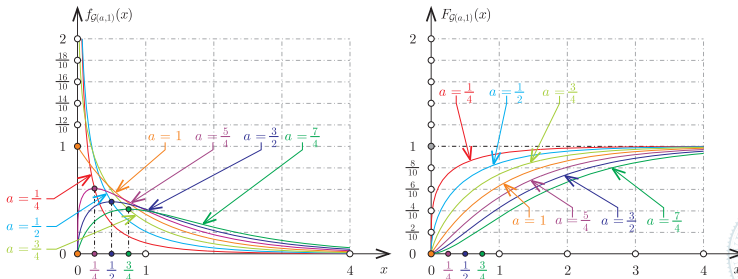


Gamma-eloszlás

Az $a > 0$ és $b > 0$ paraméterektől függő

$$f_{\mathcal{G}(a,b)}(x) = \begin{cases} \frac{1}{b^a \Gamma(a)} x^{a-1} e^{-\frac{x}{b}}, & x > 0, \\ 0, & x \leq 0 \end{cases} \quad (15)$$

sűrűségfüggvénnyel leírt folytonos valószínűségi változót *gamma-eloszlásúnak* nevezzük és a $\mathcal{G}(a,b)$ módon jelöljük. Az a és b számokra, mint *alakparaméterre*, illetve *skálázási tényezőre* hivatkozunk. Az $(a,1)$ -paraméterezésű gamma-eloszlású változót röviden *a-paraméterűnek* is nevezzük.



10. ábra. A $\mathcal{G}(a,1)$ -eloszlás sűrűség- és eloszlásfüggvényének alakja az $a = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, \frac{5}{4}, \frac{3}{2}, \frac{7}{4}$ esetekben

A gamma-eloszlás sajátos esetei

A gamma-eloszlású valószínűségi változó értelmezése alapján

- a $\lambda > 0$ paraméterű exponenciális eloszlás sűrűségfüggvényét az $a = 1$ és $b = \frac{1}{\lambda}$ megválasztással,
- az $n \geq 1$ szabadságfokú ($n \in \mathbb{N}$) és $\sigma > 0$ skálázási tényezőjű $\chi^2(n, \sigma)$ -eloszlás sűrűségfüggvényét pedig az $a = \frac{n}{2}$ és $b = 2\sigma^2$

beállításokkal kapjuk vissza a (15)-ös leképezésből.



Béta-eloszlás

Az $a > 0$ és $b > 0$ valós paraméterektől függő

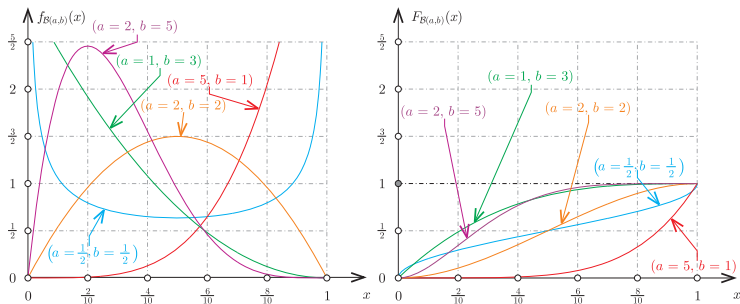
$$f_{B(a,b)}(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}, & x \in (0,1), \\ 0, & x \notin (0,1) \end{cases} \quad (16)$$

sűrűségfüggvénnyel leírt folytonos valószínűségi változót (a,b) -paraméterű béta-eloszlásúnak nevezzük, ahol az Euler-féle béta-függvénnyel értelmezett

$$B(a,b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)}$$

konstans normalizációs szerepet tölt be. Az így értelmezett valószínűségi változót a $B(a,b)$ kifejezéssel jelöljük.





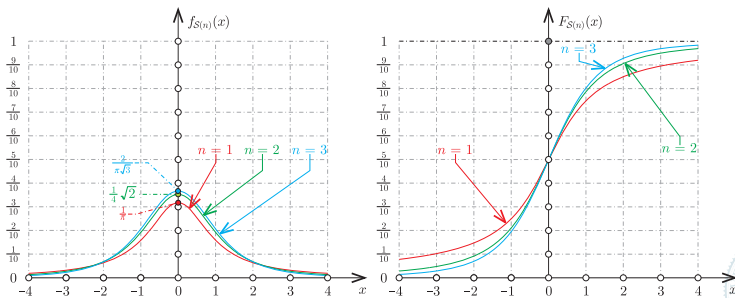
11. ábra. Az $(a, b) \in \{(\frac{1}{2}, \frac{1}{2}), (1, 3), (2, 2), (2, 5), (5, 1)\}$ paraméterpárok által meghatározott $B(a, b)$ -eloszlások sűrűség- és eloszlásfüggvényeinek alakja

Student-féle eloszlás

Adott $n \geq 1$ természetes szám esetén az

$$f_{S(n)}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad x \in \mathbb{R} \quad (17)$$

sűrűségfüggvénnyel jellemzett folytonos valószínűségi változót n -szabadságfokú Student-eloszlásúnak nevezzük és az $S(n)$ szimbólummal jelöljük. Az 1-szabadságfokú Student-eloszlást még *Cauchy-eloszlásnak* is mondjuk.



12. ábra. Az n -szabadságfokú Student-eloszlás sűrűség- és eloszlásfüggvényének alakja az $n = 1, 2, 3$ értékekre

Snedecor–Fisher-féle eloszlás

Az $m > 0$ és $n > 0$ valós számoktól függő

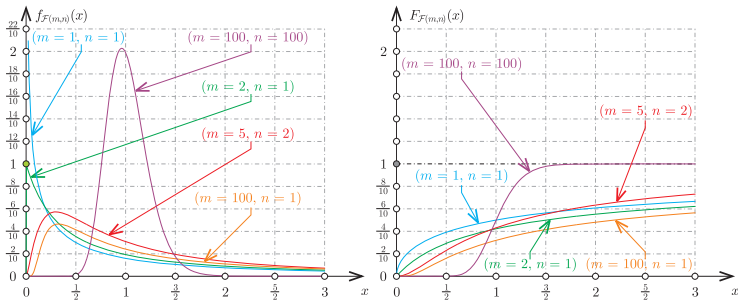
$$f_{\mathcal{F}(m,n)}(x) = \begin{cases} \frac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{x^{\frac{m}{2}-1}}{\left(1 + \frac{m}{n}x\right)^{\frac{m+n}{2}}}, & x \geq 0, \\ 0, & x < 0 \end{cases}$$

sűrűségfüggvényű folytonos valószínűségi változót (m, n) -paraméterezésű *Snedecor–Fisher-eloszlásúnak* nevezzük és $\mathcal{F}(m, n)$ módon jelöljük, ahol a

$$B\left(\frac{m}{2}, \frac{n}{2}\right) = \frac{\Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{m+n}{2}\right)}$$

normalizációs konstans az Euler-féle béta-függvénnyel értelmeztük. Nagyon sok gyakorlati alkalmazásban az m és n paraméterek csak természetes számokat vesznek fel, ilyenkor ezekre, mint *szabadságfokokra* hivatkozunk.





13. ábra. Az $(m, n) \in \{(1, 1), (2, 1), (5, 2), (100, 1), (100, 100)\}$ szabadságfokok által meghatározott Snedecor–Fisher-eloszlás sűrűség- és eloszlásfüggvényének alakjai

Leadási határidő: 2020. október 12–15. (a megfelelő laborórákon)

- A felsorolt nevezetes diszkrét, illetve folytonos eloszlások sűrűség- és eloszlásfüggvényeinek implementálása, valamint azok tetszőleges paraméterezés melletti megjelenítése kötelező! Éppen ezért a **2.**, **3.**, **4.** és az **5.** kódrészletekben egy-egy befejezendő függvényt ajánlottunk a különböző sűrűség- és eloszlásfüggvények kezelésére. Ügyeljetek arra, hogy a függvények sormátrix típusú bemenetre is helyes eredményt kell adjanak!
- Eredményeiteket hasonlítsátok össze a MATLAB[®] beépített függvényeivel generált értékekkel és ábrákkal! (A környezet által biztosított sűrűség- és eloszlásfüggvényeket az **1.** kódrészletben soroltuk fel. Már most felhívjuk a figyelmet arra, hogy időnként eltérés van a jegyzet, illetve a MATLAB[®] által feltételezett paraméterezési módban!)



1. Kódrészlet. A MATLAB[®] által támogatott diszkrét és folytonos eloszlások

```
1 % 

---


2 % Important
3 % 

---


4 % All probability density1 and cumulative distribution2 functions have similar
5 % names to *pdf and *cdf, respectively, where the asterisk must be replaced by
6 % one of the listed prefixes.
7 % 

---


8 % Discrete distributions
9 % 

---


10 % 'bino' – binomial distribution
11 % 'nbin' – negative binomial distribution
12 % 'geo' – geometric distribution
13 % 'hyge' – hypergeometric distribution
14 % 'poiss' – Poisson distribution
15 % 

---


16 % Continous distributions
17 % 

---


18 % 'beta' – beta distribution
19 % 'chi2' – chi-square distribution
20 % 'exp' – exponential distribution
21 % 'f' – F or Snedecor–Fisher distribution
22 % 'gam' – gamma distribution
23 % 'norm' – normal distribution
24 % 't' – T or Student distribution
25 % 'unif' – uniform3 distribution
```



```

26 % _____
27 % Continuous 2D distributions
28 % _____
29 % 'mvn' — multivariate4 normal distribution
30 % _____
31 % Help
32 % _____
33 % doc *pdf, or doc *cdf
34 % _____
35 % Example
36 % _____
37 % define a mean value5 and a standard deviation for  $X \sim \mathcal{N}(\mu, \sigma)$ 
38 mu = -1;
39 sigma = 2.0;
40 % define an interval and create a subdivision6
41 range = [-6.0, 4.0];
42 div_point_count = 100;
43 x = linspace(range(1), range(2), div_point_count);
44 % evaluate the probability density  $f_{\mathcal{N}(\mu, \sigma)}$  and the cumulative distribution  $F_{\mathcal{N}(\mu, \sigma)}$ 
45 f = normpdf(x, mu, sigma);
46 F = normcdf(x, mu, sigma);
47 % finally, plot7 both functions
48 plot(x, f, x, F);

```

¹**density** *sűrűség*; **probability** \sim **function** *relatív gyakoriság függvény (diszkrét eset), sűrűségfüggvény (folytonos eset)*

²**distribution** *eloszlás*; **discrete** \sim *diszkrét eloszlás*; **continuous** \sim *folytonos eloszlás*; **cumulative** \sim **function** *összegzett gyakoriság függvény (diszkrét eset), vagy eloszlásfüggvény (diszkrét és folytonos eset)*

³**uniform** *egyenletes*; \sim **distribution** *egyenletes eloszlás*

⁴**multivariate** *többváltozós*; \sim **normal distribution** *többváltozós normális eloszlás*

1. feladat

Az ismertetett diszkrét eloszlások figyelembevételével, fejezzétek be a 2. kódrészletbeli **DiscretePDF** függvényt!

2. Kódrészlet. Diszkrét eloszlások relatív gyakoriság függvényei

```

1 % _____
2 % Description
3 % _____
4 % The function calculates the values of different discrete probability density functions
5 % _____
6 % Input
7 % _____
8 %  $x = [x_i]_{i=1}^n$  – an increasing sequence8 of positive integers
9 % distribution_type – a string that identifies the distribution (e.g., 'Bernoulli',
10 % 'binomial', 'Poisson', 'geometric', etc.)
11 % parameters – an array of parameters which characterize the distribution
12 % specified by distribution_type
13 % _____
14 % Output
15 % _____
16 %  $f = [f(x_i)]_{i=1}^n$  – values of the given probability density function
17 function f = DiscretePDF(x, distribution_type, parameters)
18
19 % for safety reasons
20 sort(x);
    
```



```

21 x = round(x);
22
23 % get the size of the input array
24 n = length(x);
25
26 % select the corresponding distribution
27 switch (distribution_type)
28
29     case 'geometric'
30         % the  $\mathcal{Geo}(p)$ -distribution has a single parameter  $p \in [0, 1]$ 
31         p = parameters(1);
32
33         % check the validity of the distribution parameter p
34         if (p < 0 || p > 1)
35             error('Wrong_parameter!');
36         end
37
38         % allocate memory and evaluate the probability density function  $f_{\mathcal{Geo}(p)}$ 
39         % for each element of the input array  $x = [x_i]_{i=1}^n$ 
40         f = zeros(1, n);
41
42         q = 1 - p;
43         for i = 1:n
44             % check the validity of the current value xi
45             if (x(i) < 1)
46                 error('Incorrect_input_data!');
47             else
48                 f(i) = q^(x(i) - 1) * p; % i.e.,  $f_{\mathcal{Geo}(p)}(x_i) = (1 - p)^{x_i} \cdot p, i = 1, 2, \dots, n$ 
49             end
50         end
51     end

```



```
51
52     % handle another discrete distribution type
53     ...
54 end
```

2. feladat

Az ismertetett folytonos eloszlások figyelembevételével, fejezzétek be a 3. kódrészletbeli **ContinuousPDF** függvényt!

3. Kódrészlet. Folytonos eloszlások sűrűségfüggvényei

```

1 % _____
2 % Description
3 % _____
4 % The function evaluates different continuous probability density functions.
5 %
6 % _____
7 % Input
8 % _____
9 %  $x = [x_i]_{i=1}^n$       – an increasing sequence of real numbers
10 % distribution_type – a string that identifies the distribution (e.g., 'exponential',
11 %                  'normal', 'chi2', 'gamma', 'beta', 'Student', etc.)
12 % parameters       – an array of parameters which characterize the distribution
13 %                  specified by distribution_type
14 %
15 % _____
16 % Output
17 % _____
18 %  $f = [f_i]_{i=1}^n = [f(x_i)]_{i=1}^n$  – values of the given probability density function
19 %
20 function f = ContinuousPDF(x, distribution_type, parameters)

```



```

21
22 % for safety reasons
23 x = sort(x);
24
25 % get the size of the input array
26 n = length(x);
27 % select the corresponding distribution
28 switch (distribution_type)
29
30     case 'normal'
31         % the  $\mathcal{N}(\mu, \sigma)$ -distribution has two parameters, where  $\mu \in \mathbb{R}$  and  $\sigma > 0$ 
32         mu = parameters(1);
33         sigma = parameters(2);
34
35         % check the validity of the distribution parameters
36         if (sigma <= 0)
37             error('The standard deviation must be a strictly positive number!');
38         end
39
40         % Allocate memory and evaluate the probability density function  $f_{\mathcal{N}(\mu, \sigma)}$ 
41         % for each element of the input array  $x = [x_i]_{i=1}^n$ .
42         %
43         % Note that, in this special case, this can be done in a single line of code,
44         % provided that one uses the dotted arithmetical operators of MATLAB®.
45
46         % use the formula  $f_{\mathcal{N}(\mu, \sigma)}(x_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}$ ,  $x_i \in \mathbb{R}$ 
47         f = (1.0 / sqrt(2.0 * pi) / sigma) * exp(-(x - mu).^ 2 / 2.0 / sigma^2);
48

```




```
49     % handle another continuous distribution type
50     ...
51 end
```



3. feladat

- Felhasználva az eloszlásfüggvények monoton növekedési tulajdonságát, írjatok hatékony függvényeket diszkrét és folytonos valószínűségi változók eloszlásfüggvényeinek kiértékelésére.
- Az implementálás során használjátok fel a 2. és 3. kódrészletekben adott [DiscretePDF](#), illetve [ContinuousPDF](#) függvények befejezett változatait!
- Folytonos valószínűségi változók esetén, amikor csak lehetséges, használjátok fel az eloszlásfüggvény explicit alakját, minden más esetben az improprius integrálok közelítésére alkalmazzátok vagy az egyenközü felosztást feltételező

$$\int_a^b f(t) dt \approx \frac{b-a}{3n} \left(f(t_0) + 2 \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor - 1} f(t_{2j}) + 4 \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor} f(t_{2j-1}) + f(t_n) \right),$$

$$t_j = a + j \cdot \frac{b-a}{n}, j = 0, 1, \dots, n, n \geq 1$$

általános Simpson-féle kvadratúra-képletet, vagy a MATLAB[®] beépített, adaptív Simpson-féle kvadratúrára épülő **quad** függvényt!

- A képletekben esetlegesen megjelenő Euler-féle béta- és gamma-függvények kiértékelésére szintén kvadratúra képletet, vagy a MATLAB[®] által biztosított **gamma** és **beta** függvényeket használhatjátok.



3. feladat – folytatás

- Vegyétek észre, hogy amennyiben az $\int_a^b f(t) dt$ integrálban valamelyik, vagy mindkét határ a plusz/mínusz végtelent jelképezi, akkor véges intervallumra vezethetjük vissza a számolásokat az

$$\int_a^b f(t) dt = \begin{cases} \int_0^1 \frac{f(\ln(u))}{u} du, & a = -\infty, b = 0 \\ \int_0^1 \frac{f(\ln(-u))}{u} du, & a = 0, b = \infty, \\ \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{f(\tan(u))}{\cos^2(u)} du, & a = -\infty, b = \infty \end{cases}$$

átalakításoknak megfelelően, ahol rendre a $t = \ln(u)$, $t = \ln(-u)$, illetve a $t = \tan(u)$ változócserét alkalmaztuk!

3. feladat – folytatás

- A 4. és 5. kódrészletekben a befejezendő függvények fejlécét adtuk meg. Ne felejtsetek el a bemeneti adatok helyességét is ellenőrizni!

4. Kódrészlet. Diszkrét eloszlásfüggvények kiértékelése

```

1 % _____
2 % Description
3 % _____
4 % The function evaluates the values of different discrete cumulative distribution
5 % functions.
6 %
7 % _____
8 % Input
9 % _____
10 %  $x = [x_i]_{i=1}^n$           – an increasing sequence of positive integers
11 %
12 % distribution_type      – a string that identifies the distribution (e.g., 'Bernoulli',
13 %                        'binomial', 'Poisson', 'geometric', etc.)
14 %
15 % parameters            – an array of parameters which characterize the distribution
16 %                        specified by distribution_type
    
```



```

17 % -----
18 % Output
19 % -----
20 %  $F = [F_i]_{i=1}^n = [F(x_i)]_{i=1}^n$  – values of the given cumulative distribution function
21 %
22 % -----
23 % Hint
24 % -----
25 % Since the input sequence  $x = [x_i]_{i=1}^n \subset \mathbb{N}$  is increasing it is sufficient to calculate
26 % the values
27 %
28 %  $F_1 = \sum_{i=i_{\min}}^{x_1} f(i), F_2 = F_1 + \sum_{i=x_1+1}^{x_2} f(i), \dots, F_n = F_{n-1} + \sum_{i=x_{n-1}+1}^{x_n} f(i),$ 
29 %
30 % where  $f$  denotes the probability density function that corresponds to  $F$  and  $i_{\min}$ 
31 % represents the minimal integer value of the distribution (e.g., in case of the
32 % geometric distribution  $i_{\min} = 1$ , while in the case of the Poisson distribution  $i_{\min} = 0$ ).
33 %
34 function F = DiscreteCDF(x, distribution_type, parameters)

```



5. Kódrészlet. Folytonos eloszlásfüggvények kiértékelése

```

1 % _____
2 % Description
3 % _____
4 % The function evaluates the values of different continuous cumulative distribution
5 % functions.
6 %
7 % _____
8 % Input
9 % _____
10 %  $\mathbf{x} = [x_i]_{i=1}^n$       – an increasing sequence of real numbers
11 %
12 % distribution_type – a string that identifies the distribution (e.g., 'exponential',
13 %                      'normal', 'chi2', 'gamma', 'beta', 'Student', etc.)
14 %
15 % parameters       – an array of parameters which characterize the distribution
16 %                      specified by distribution_type
17 %
18 % _____
19 % Output
20 % _____
21 %  $\mathbf{F} = [F_i]_{i=1}^n = [F(x_i)]_{i=1}^n$  – values of the given cumulative distribution function

```



```
22 % _____
23 % Hint
24 % _____
25 % Since the input sequence  $\mathbf{x} = [x_i]_{i=1}^n \subset \mathbb{R}$  is increasing it is sufficient to calculate
26 % the values
27 %
28 %  $F_1 = \int_{x_{\min}}^{x_1} f(t) dt$ ,  $F_2 = F_1 + \int_{x_1}^{x_2} f(t) dt$ , ...,  $F_n = F_{n-1} + \int_{x_{n-1}}^{x_n} f(t) dt$ ,
29 %
30 % where  $f$  denotes the probability density function that corresponds to  $F$  and  $x_{\min}$ 
31 % represents the minimal value of the random variable (e.g., in case of the gamma
32 % distribution  $x_{\min} = 0$ , while in the case of the normal distribution  $x_{\min} = -\infty$ ).
33 %
34 function F = ContinuousCDF(x, distribution_type, parameters)
```



4. feladat (+1 pont elméleti megoldással együtt)

- Határozzátok meg az $\alpha \in \mathbb{R}$ paraméter értékét úgy, hogy az

$$f_X(x) = \begin{cases} \alpha \cdot \frac{x^2-1}{12}, & -2 < x \leq -1, \\ \frac{x+1}{6}, & -1 < x \leq 1, \\ \frac{x+5}{18}, & 1 < x \leq 2, \\ \alpha \cdot (x-2)^2, & 2 < x \leq 3, \\ 0, & \text{különben} \end{cases}$$

leképzés egy folytonos valószínűségi változó sűrűségfüggvénye legyen, majd határozzátok meg annak eloszlásfüggvényét is!

- A kapott kifejezések alapján egészítsétek ki a 3. és 5. kódrészleteket, valamint azok segítségével jelenítsétek is meg az adott változó sűrűség- és eloszlásfüggvényét!



5. feladat (+1 pont elméleti megoldással együtt)

- Határozzátok meg a $p \in (0, 1)$ valószínűség értékét úgy, hogy az

$$X \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \frac{1}{36} & \frac{1}{18} & p & \frac{1}{9} & \frac{1}{36} & 3p & \frac{1}{12} & \frac{1}{3} \end{pmatrix}$$

diszkrét valószínűségi változó eloszlástáblázata helyessé váljon, majd írjátok fel az X valószínűségi változó eloszlásfüggvényét is!

- A kapott kifejezések alapján egészítsétek ki a 2. és 4. kódrészleteket, valamint azok segítségével jelenítsétek is meg az adott változó relatív gyakoriság és eloszlásfüggvényét is!

