Véletlenszám-generátorok

2. rész

– egy MATLAB® alapú megközelítés –

Róth Ágoston, Vas Orsolya

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

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

2. labor / 2018. október 8–11.

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 $\left(\mathcal{P}(\Omega), \cup, \cap, \overline{}, \emptyset, \Omega\right)$ struktúra Boole-algebrát alkot, azaz bármely $A, B, C \in \mathcal{P}(\Omega)$ esemény esetén teljesülnek az alábbi axiómák:
  1. asszociativitás:
  $$A \cup (B \cup C) = (A \cup B) \cup C, \ A \cap (B \cap C) = (A \cap B) \cap C;$$
  2. kommutativitás:
  $$A \cup B = B \cup A, \ A \cap B = B \cap A;$$
  3. elnyelési tulajdonság:
  $$A \cup (A \cap B) = A, \ A \cap (A \cup B) = A;$$
  4. disztributivitás:
  $$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \ A \cap (B \cup C) = (A \cap B) \cup (A \cap C);$$
  5. komplementer képzés:
  $$A \cup \overline{A} = \Omega, \ A \cap \overline{A} = \emptyset.$$

- Amennyiben az Ω eseménytér véges (végtelen), a
$$\left(\mathcal{P}(\Omega), \cup, \cap, \overline{}, \emptyset, \Omega\right)$$
eseményalgebrát végesnek (végtelennek) nevezzük.

Algebra

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

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

feltételek.

σ-algebra

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

$$\begin{aligned} A \in \mathcal{A} \quad &\Rightarrow \quad \overline{A} \in \mathcal{A}, \\ A_j \in \mathcal{A}, j \in J \quad &\Rightarrow \quad \cup_{j \in J} A_j \in \mathcal{A} \end{aligned}$$

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 $(\Omega, \mathcal{A})$ párost *véges (végtelen) eseménymezőnek* nevezzük.

Valószínűség axiomatikus értelmezése

Az $(\Omega, \mathcal{A})$ eseménymezőn értelmezett $P : \mathcal{A} \to [0,1]$ 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
$$P\left(\cup_{j \in J} A_j\right) = \sum_{j \in J} P\left(A_j\right)$$
összefüggés.

Valószínűségi mező

A $P : \mathcal{A} \to [0,1]$ valószínűségi függvénnyel felruházott véges (végtelen) $(\Omega, \mathcal{A})$ eseménymezőt véges (végtelen) *valószínűségi mezőnek* nevezzük és az $(\Omega, \mathcal{A}, P)$ 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 kívűl 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 függnek, ezért ezeket a függvényeket valószínűségi változóknak is nevezzük. Pontos értelmezésükhöz viszont szükségünk van az előző oldalakon ismertetett eseményalgebra, eseménymező és valószínűség-számítás alapvető fogalmaira.

Valószínűségi változó

Az $(\Omega, \mathcal{A})$ eseménymezőn értelmezett $X : \Omega \to \mathbb{R}$ 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 $(\Omega, \mathcal{A})$ eseménymezőn értelmezett

$$X\left(X_1, X_2, \ldots, X_n\right) : \Omega \to \mathbb{R}^n$$

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

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

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

Valószínűségi változó eloszlásfüggvénye

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

$$\left\{ \begin{array}{l} F_X : \mathbb{R} \to \mathbb{R}, \\ F_X(x) = P(X < x) \end{array} \right. \tag{1}$$

leképezést az X valószínűségi változó *eloszlásfüggvényének* nevezzük.

Valószínűségi vektor együttes eloszlásfüggvénye

Ha

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

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

$$\left\{ \begin{array}{l} F_X : \mathbb{R}^n \to \mathbb{R}, \\ F_X(x_1, x_2, \ldots, x_n) = P(X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n) \end{array} \right. \tag{2}$$

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

Diszkrét valószínűségi változó

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

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

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

$$\left\{ \begin{array}{l} f_X : \{x_i : i \in I\} \to \{p_i : i \in I\}, \\ f_X(x_i) = p_i \end{array} \right. \tag{4}$$

leképezést az X változó *relatív gyakoriság függvényének* nevezzük.

Diszkrét valószínűségi változó eloszlásfüggvénye

A (3)-as diszkrét valószínűségi változóhoz társított

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

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

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

ahol

$$J_i = \left\{j \in I : x_j < x_i\right\}.$$

Következmény – folytatás

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

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

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

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

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

$$X_i \begin{pmatrix} x_{i,j_i} \\ p_{i,j_i} \end{pmatrix}_{j_i \in J_i}, \; i = 1, 2, \ldots, n, \; n \geq 1$$

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

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

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

Folytonos valószínűségi változó

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

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

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

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

egyenlőséglánc.

Folytonos valószínűségi vektorok

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

$$F_X\left(x_1, x_2, \ldots, x_n\right) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f_X\left(t_1, t_2, \ldots, t_n\right) \mathrm{d}t_1 \mathrm{d}t_2 \ldots \mathrm{d}t_n \tag{6}$$

egyenlőséget minden $\left(x_1, x_2, \ldots, x_n\right) \in \mathbb{R}^n$ 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 $X(X_1, X_2, \ldots, X_n)$ valószínűségi vektorra teljesülnek az alábbi kijelentések:

1. minden $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ pontra fennáll a
$$\frac{\partial^n}{\partial x_1 \partial x_2 \ldots \partial x_n} F_X(x_1, x_2, \ldots, x_n) = f_X(x_1, x_2, \ldots, x_n)$$
összefüggés;
2. bármely $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ pontra teljesül az
$$f_X(x_1, x_2, \ldots, x_n) \geq 0$$
egyenlőtlenség;
3. érvényes az
$$\int_{\mathbb{R}} \int_{\mathbb{R}} \ldots \int_{\mathbb{R}} f_X(t_1, t_2, \ldots, t_n) \, \mathrm{d}t_1 \mathrm{d}t_2 \ldots \mathrm{d}t_n \equiv 1$$
azonosság;

Együttes sűrűségfüggvény tulajdonságai – folytatás

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

$$P(X \in T) = \int\int_T \ldots \int f_X(t_1, t_2, \ldots, t_n)\, \mathrm{d}t_1 \mathrm{d}t_2 \ldots \mathrm{d}t_n$$

képlettel határozhatjuk meg;

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

$$f_{X_i}(t) = \underbrace{\int_{\mathbb{R}}\int_{\mathbb{R}} \ldots \int_{\mathbb{R}}}_{n-1} f_X(t_1, t_2, \ldots, t_{i-1}, t, t_{i+1}, t_n)\, \mathrm{d}t_1 \mathrm{d}t_2 \ldots \mathrm{d}t_{i-1} \mathrm{d}t_{i+1} \ldots \mathrm{d}t_n,$$

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

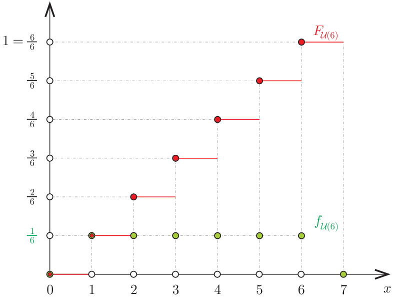
alakban írhatjuk minden $i = 1, 2, \ldots, n$ index esetén.

Diszkrét egyenletes eloszlás

Az $n \geq 1$ természetes szám által meghatározott

$$X \begin{pmatrix} 1 & 2 & \cdots & n \\ \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \end{pmatrix}$$

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.

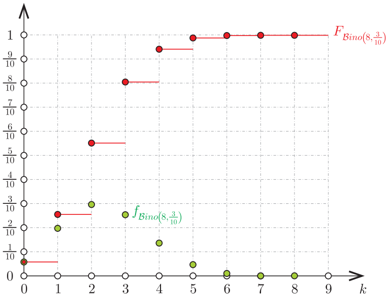
- Ez a legegyszerűbb diszkrét valószínűség-eloszlás, amelyet döntési alternatívaként használhatunk egy adott esemény bekövetkezésére, vagy meghiúsulására.

Binomiális eloszlás

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

$$X \left(\begin{array}{c} k \\ \binom{n}{k} p^k (1-p)^{n-k} \end{array} \right)_{k \in \{0,1,\ldots,n\}}$$

valószínűségi változót (n,p)-*paraméterezésű binomiális eloszlásúnak* nevezzük és $\mathcal{B}ino\,(n,p)$ módon jelöljük.



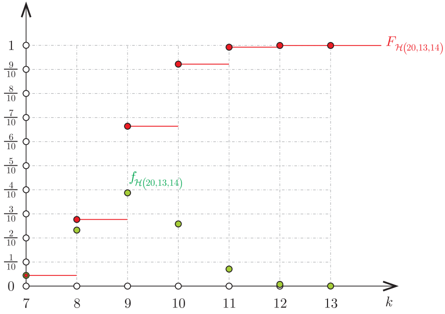
2. ábra. $\mathcal{B}ino(8, \frac{3}{10})$-eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Hipergeometrikus eloszlás

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

$$X \left(\begin{array}{c} k \\ \dfrac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \end{array} \right)_{\max\{0, n-N+M\} \leq k \leq \min\{n, M\}}$$

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



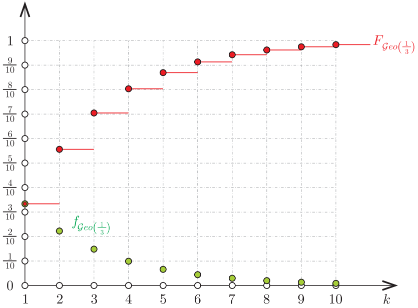
3. ábra. $\mathcal{H}(20, 13, 14)$-eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Geometriai eloszlás

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

$$X\begin{pmatrix} k \\ (1-p)^{k-1}\,p \end{pmatrix}_{k\geq 1}$$

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



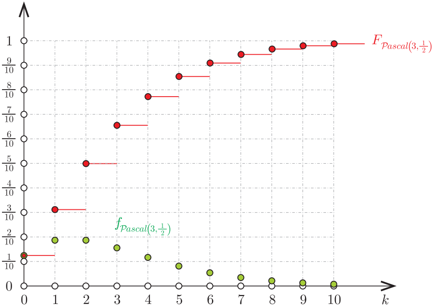
4. ábra. $\mathcal{G}eo\left(\frac{1}{3}\right)$-eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Pascal-féle vagy negatív binomiális eloszlás

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

$$X \begin{pmatrix} n+k \\ \binom{n+k-1}{k} p^n (1-p)^k \end{pmatrix}_{k \geq 0}$$

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



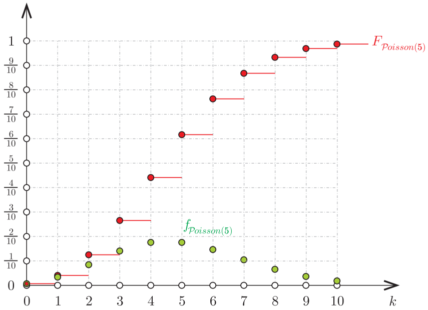
5. ábra. $\mathcal{P}ascal\left(3, \frac{1}{2}\right)$-eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Poisson-eloszlás

A $\lambda > 0$ állandó segítségével megszerkesztett

$$X \begin{pmatrix} k \\ \frac{\lambda^k}{k!} e^{-\lambda} \end{pmatrix}_{k \geq 0}$$

diszkrét valószínűségi változót *λ-paraméterű Poisson-eloszlásúnak* nevezzük és $\mathcal{P}oisson\,(\lambda)$ alakban hivatkozunk rá.



6. ábra. $\mathcal{P}oisson\,(5)$-eloszlású valószínűségi változó relatív gyakoriság és eloszlásfüggvénye

Folytonos egyenletes eloszlás

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

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

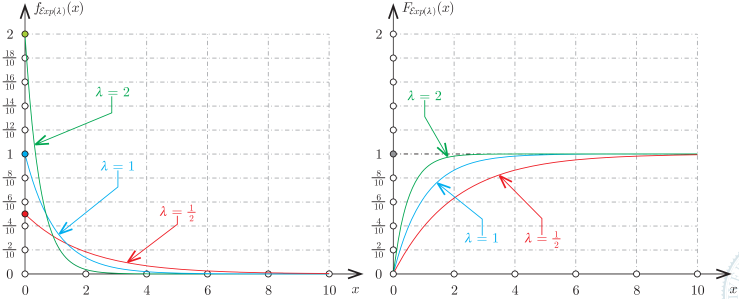
sűrűségfüggvényű folytonos valószínűségi változót *egyenletes eloszlásúnak* nevezzük és az $\mathcal{U}([a, b])$ kifejezéssel jelöljük.

Exponenciális eloszlás

A $\lambda > 0$ valós számtól függő

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

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



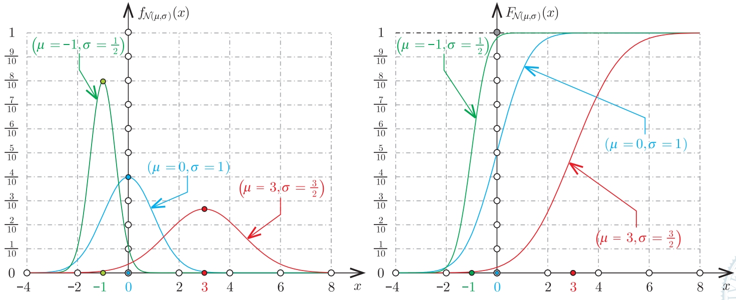
7. ábra. Az $\mathcal{E}xp(\lambda)$-eloszlás sűrűség- és eloszlásfüggvénye a $\lambda = \frac{1}{2}, 1, 2$ paraméterek esetén

Normális eloszlás

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

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

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

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

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,

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

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

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

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

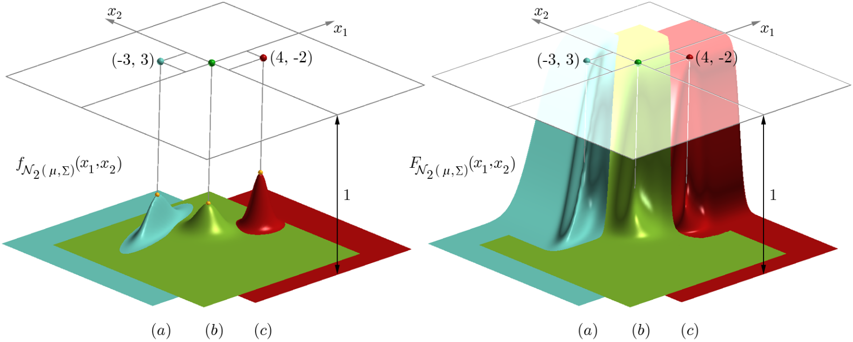
$${}^t[X_1 \sim \mathcal{N}(\mu_1, \sigma_1), X_2 \sim \mathcal{N}(\mu_2, \sigma_2)]$$

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 külünböző szórásúak és korelláltak); (*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

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

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

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

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. Érdekességképpen megemlítjük, hogy $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ (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

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

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

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

összefüggésen keresztül.

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

Az

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

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

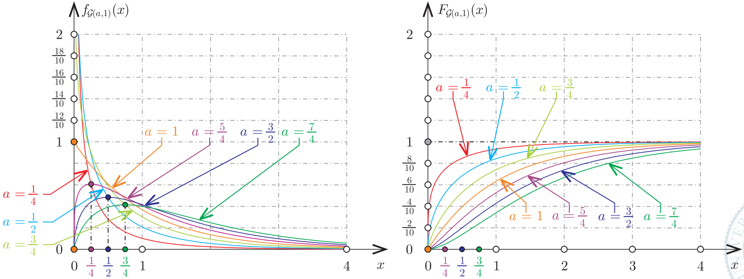
Amint hamarosan látni fogjuk, a $\lambda > 0$ paraméterű exponenciális, illetve az $n \geq 1$ szabadságfokú és $\sigma > 0$ skálázási tényezőjű $\chi^2(n,\sigma)$-eloszlású folytonos valószínűségi változókat egy speciális, úgynevezett gamma-eloszlás sajátságos paraméterezésű eseteiként kaphatjuk vissza.

Gamma-eloszlás

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

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

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

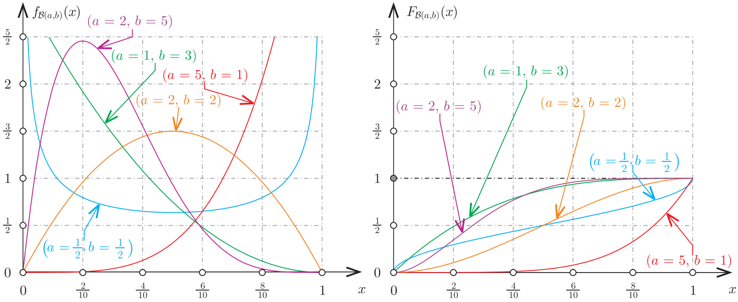
Az $a > 0$ és $b > 0$ valós paraméterektől függő

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

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

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

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



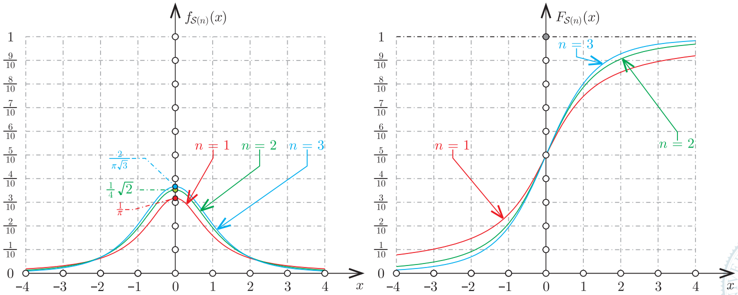
11. ábra. Az $(a, b) \in \left\{\left(\frac{1}{2}, \frac{1}{2}\right), (1, 3), (2, 2), (2, 5), (5, 1)\right\}$ paraméterpárok által meghatározott $\mathcal{B}(a, b)$-eloszlások sűrűség- és eloszlásfüggvényeinek alakja

Student-féle eloszlás

Adott $n \geq 1$ természetes szám esetén az

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

sűrűségfüggvénnyel jellemzett folytonos valószínűségi változót *n-szabadságfokú Student-eloszlásúnak* nevezzük és az $\mathcal{S}(n)$ 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

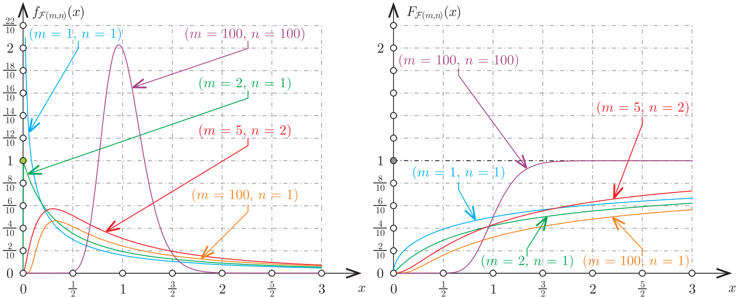
Az $m > 0$ és $n > 0$ valós számoktól függő

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

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 konstanst 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: 2018. október 15–18. (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

```
%  ---------
%  Important
%  ---------
%  All probability density[1] and cumulative distribution[2] functions have similar
%  names to *pdf and *cdf, respectively, where the asterisk must be replaced by
%  one of the listed prefixes.
%  ----------------------
%  Discrete distributions
%  ----------------------
%  'bino'  - binomial distribution
%  'nbin'  - negative binomial distribution
%  'geo'   - geometric distribution
%  'hyge'  - hypergeometric distribution
%  'poiss' - Poisson distribution
%  -----------------------
%  Continous distributions
%  -----------------------
%  'beta'  - beta distribution
%  'chi2'  - chi-square distribution
%  'exp'   - exponential distribution
%  'f'     - F or Snedecor-Fisher distribution
%  'gam'   - gamma distribution
%  'norm'  - normal distribution
%  't'     - T or Student distribution
%  'unif'  - uniform[3] distribution
```

```
% -------------------------------
% Continuous 2D distributions
% -------------------------------
% 'mvn'  - multivariate[4] normal distribution
% ----
% Help
% ----
% doc *pdf, or doc *cdf
% -------
% Example
% -------
  % define a mean value[5] and a standard deviation for $X \sim \mathcal{N}(\mu, \sigma)$
  mu    = -1;
  sigma = 2.0;
  % define an interval and create a subdivision[6]
  range           = [-6.0, 4.0];
  div_point_count = 100;
  x               = linspace(range(1), range(2), div_point_count);
  % evaluate the probability density $f_{\mathcal{N}(\mu,\sigma)}$ and the cumulative distribution $F_{\mathcal{N}(\mu,\sigma)}$
  f               = normpdf(x, mu, sigma);
  F               = normcdf(x, mu, sigma);
  % finally, plot[7] both functions
  plot(x, f, x, F);
```

[1] **density** *sűrűség*; **probability** ∼ **function** *relatív gyakoriság függvény (diszkrét eset), sűrűségfüggvény (folytonos eset)*

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

[3] **uniform** *egyenletes*; ∼ **distribution** *egyenletes eloszlás*

[4] **multivariate** *többváltozós*; ∼ **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

```
% ------------
% Description
% ------------
% The function calculates the values of different discrete probability density functions
% ------
% Input
% ------
% x = [x_i]_{i=1}^n          - an increasing sequence[8] of positive integers
% distribution_type          - a string that identifies the distribution (e.g., 'Bernoulli',
%                              'binomial', 'Poisson', 'geometric', etc.)
% parameters                 - an array of parameters which characterize the distribution
%                              specified by distribution_type
% -------
% Output
% -------
% f = [f(x_i)]_{i=1}^n       - values of the given probability density function
function f = DiscretePDF(x, distribution_type, parameters)

% for safety reasons
sort(x);
```

```
x = round(x);

% get the size of the input array
n = length(x);

% select the corresponding distribution
switch (distribution_type)

    case 'geometric'
        % the Geo(p)-distribution has a single parameter p ∈ [0, 1]
        p = parameters(1);

        % check the validity of the distribution parameter p
        if (p < 0 || p > 1)
            error('Wrong parameter!');
        end

        % allocate memory and evaluate the probability density function f_Geo(p)
        % for each element of the input array x = [x_i]_{i=1}^n
        f = zeros(1, n);

        q = 1 - p;
        for i = 1:n
            % check the validity of the current value x_i
            if (x(i) < 1)
                error('Incorrect input data!');
            else
                f(i) = q^(x(i) - 1) * p; % i.e., f_Geo(p)(x_i) = (1 - p)^(x_i) · p, i = 1, 2, ..., n
            end
        end
```

```

    % handle another discrete distribution type
     ...
end
```

[8]**sequence** *sorozat*; **increasing** $\sim$ *növekvő sorozat*; **random** $\sim$ *véletlen sorozat*

2. feladat

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

3. Kódrészlet. Folytonos eloszlások sűrűségfüggvényei

```
% -----------
% Description
% -----------
% The function evaluates different continuous probability density functions.
%
% -----
% Input
% -----
% x = [x_i]_{i=1}^n              - an increasing sequence of real numbers
% distribution_type             - a string that identifies the distribution (e.g., 'exponential',
%                                 'normal', 'chi2', 'gamma', 'beta', 'Student', etc.)
% parameters                    - an array of parameters which characterize the distribution
%                                 specified by distribution_type
%
% ------
% Output
% ------
% f = [f_i]_{i=1}^n = [f(x_i)]_{i=1}^n   - values of the given probability density function
%
function f = ContinuousPDF(x, distribution_type, parameters)
```

```matlab

% for safety reasons
x = sort(x);

% get the size of the input array
n = length(x);
% select the corresponding distribution
switch (distribution_type)

    case 'normal'
        % the $\mathcal{N}(\mu,\sigma)$-distribution has two parameters, where $\mu \in \mathbb{R}$ and $\sigma > 0$
        mu    = parameters(1);
        sigma = parameters(2);

        % check the validity of the distribution parameters
        if (sigma <= 0)
            error('The standard deviation must be a strictly positive number!');
        end

        % Allocate memory and evaluate the probability density function $f_{\mathcal{N}(\mu,\sigma)}$
        % for each element of the input array $\mathbf{x} = [x_i]_{i=1}^n$.
        %
        % Note that, in this special case, this can be done in a single line of code,
        % provided that one uses the dotted arithmetical operators of MATLAB®.


        % use the formula $f_{\mathcal{N}(\mu,\sigma)}(x_i) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$, $x_i \in \mathbb{R}$
        f = (1.0 / sqrt(2.0 * pi) / sigma) * exp(-(x - mu).^ 2 / 2.0 / sigma^2);

```

```
    % handle another continuous distribution type
     ...
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 egyenközű felosztást feltételező

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

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

  általános Simpson-féle kvadratúra-képletet, vagy a MATLAB® beépített, adaptív Simpson-féle kvadratúrára épülő quad függvényét!
- 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 $\int_a^b f(t)\,\mathrm{d}t$ integrálban valamelyik, vagy mindkét határ a plusz/mínusz végtelent jelképezi, akkor véges intervallumra vezethetjük vissza a számolásokat az

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

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

3. feladat – folytatás

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

4. Kódrészlet. Diszkrét eloszlásfüggvények kiértékelése

```
% -------------
% Description
% -------------
% The function evaluates the values of different discrete cumulative distribution
% functions.
%
% -------
% Input
% -------
% x = [x_i]_{i=1}^n        - an increasing sequence of positive integers
%
% distribution_type  - a string that identifies the distribution (e.g., 'Bernoulli',
%                      'binomial', 'Poisson', 'geometric', etc.)
%
% parameters         - an array of parameters which characterize the distribution
%                      specified by distribution_type
```

```matlab
% ———
% Output
% ———
% $\mathbf{F} = [F_i]_{i=1}^n = [F(x_i)]_{i=1}^n$ – values of the given cumulative distribution function
%
% ——
% Hint
% ——
% Since the input sequence $\mathbf{x} = [x_i]_{i=1}^n \subset \mathbb{N}$ is increasing it is sufficient to calculate
% the values
%
% $F_1 = \sum_{i=i_{\min}}^{x_1} f(i)$, $F_2 = F_1 + \sum_{i=x_1+1}^{x_2} f(i)$, ..., $F_n = F_{n-1} + \sum_{i=x_{n-1}+1}^{x_n} f(i)$,
%
% where $f$ denotes the probability density function that corresponds to $F$ and $i_{\min}$
% represents the minimal integer value of the distribution (e.g., in case of the
% geometric distribution $i_{\min} = 1$, while in the case of the Poisson distribution $i_{\min} = 0$).
%
function F = DiscreteCDF(x, distribution_type, parameters)
```

5. Kódrészlet. Folytonos eloszlásfüggvények kiértékelése

```
% -----------
% Description
% -----------
% The function evaluates the values of different continuous cumulative distribution
% functions.
%
% -----
% Input
% -----
% x = [x_i]_{i=1}^n        - an increasing sequence of real numbers
%
% distribution_type        - a string that identifies the distribution (e.g., 'exponential',
%                            'normal', 'chi2', 'gamma', 'beta', 'Student', etc.)
%
% parameters               - an array of parameters which characterize the distribution
%                            specified by distribution_type
%
% ------
% Output
% ------
% F = [F_i]_{i=1}^n = [F(x_i)]_{i=1}^n - values of the given cumulative distribution function
```

```
% ———
% Hint
% ———
% Since the input sequence $\mathbf{x} = [x_i]_{i=1}^{n} \subset \mathbb{R}$ is increasing it is sufficient to calculate
% the values
%
% $F_1 = \int_{x_{\min}}^{x_1} f(t)\,\mathrm{d}t,\ F_2 = F_1 + \int_{x_1}^{x_2} f(t)\,\mathrm{d}t, \ldots,\ F_n = F_{n-1} + \int_{x_{n-1}}^{x_n} f(t)\,\mathrm{d}t$,
%
% where $f$ denotes the probability density function that corresponds to $F$ and $x_{\min}$
% represents the minimal value of the random variable (e.g., in case of the gamma
% distribution $x_{\min} = 0$, while in the case of the normal distribution $x_{\min} = -\infty$).
%
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

$$f(x) = \begin{cases} 0, & x < 0, \\ \alpha \cdot xe^{-(1+\lambda)x}, & x \geq 0 \end{cases}$$

  leképezés egy folytonos valószínűségi változó sűrűségfüggvénye legyen, majd számítsátok ki annak eloszlásfüggvényét is, ahol $\lambda > -1$ tetszőlegesen állítható alakparaméter!

- A kapott kifejezések alapján egészítsétek ki a 3. és 5. kódrészleteket, valamint azok segítségével jelenítsétek is meg az adott változó sűrűség- és eloszlásfüggvényét különböző $\lambda > -1$ alakparaméterek mellett!