Segédanyag a másodéves földtudomány szakos Valószínűségszámítás előadáshoz (2013/2014 ősz)*

Boros Balázs

Az 1., illetve a 2. szakaszokban összefoglaltuk a félév során tanult nevezetes diszkrét, illetve abszolút folytonos eloszlásokat, valamint azok néhány alapvető tulajdonságát. A 3. szakaszban felidézzük a Centrális Határeloszlás Tételt, valamint annak két alkalmazását. Az 1., 2. és 3. szakaszokban bemutatottakhoz ábrákat is készítettünk, melyeket az 5. szakaszban helyeztünk el. A 4. szakaszban emlékeztetünk néhány korábban tanult (főképpen analízisbeli) tudnivalóra. Az alábbiakban $(\Omega, \mathcal{A}, P)$ mindig alkalmas Kolmogorov-féle valószínűségi mezőt jelöl.

1. Nevezetes diszkrét eloszlások

1.1. Indikátor (jelölés: $\mathrm{Ind}(p)$)

Legyen $A \in \mathcal{A}$ és $p = P(A)$. Ekkor az A esemény *indikátora* a

$$\chi_A(\omega) = \begin{cases} 1, & \text{ha } \omega \in A, \\ 0, & \text{ha } \omega \in \Omega \setminus A \end{cases}$$

hozzárendeléssel értelmezett $\chi_A : \Omega \to \mathbb{R}$ valószínűségi változó. Világos, hogy χ_A lehetséges értékei a 0, illetve az 1, melyeket p, illetve $1-p$ valószínűséggel vesz fel. Könnyű számolás mutatja, hogy $EX = p$ és $D^2X = p(1-p)$.

1.2. Binomiális eloszlás (jelölés: $\mathrm{Bin}(n,p)$)

Legyen $A \in \mathcal{A}$ és $p = P(A)$. Végezünk el n kísérletet az egymástól függetlenül, és jelölje X azt, hogy ennek során az A esemény hányszor következett be. Világos, hogy X lehetséges értékei a $\{0, 1, \ldots, n\}$ halmaz elemei. Könnyen látható, hogy minden $k \in \{0, 1, \ldots, n\}$ esetén

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Az X valószínűségi változó eloszlását n rendű és p paraméterű *binomiális* eloszlásnak nevezzük. Világos, hogy $n = 1$ esetén visszakapjuk az indikátorváltozót. Kiszámolható, hogy $EX = np$ és $D^2X = np(1-p)$. Természetes úgy gondolni X-re, mint független indikátorok összege, és így a várható értékre és a szórásnégyzetre vonatkozó formula azonnal adódik.

*A jegyzetben sajtóhibák nagyon könnyen előfordulhatnak

1.3. Geometriai, más szóval Pascal eloszlás (jelölés: $\mathrm{Geo}(p)$)

Legyen $A \in \mathcal{A}$ és $p = P(A)$. Egymástól független kísérleteket végzünk, és jelölje X annak a kísérletnek a sorszámát, amikor először következett be A. Világos, hogy X lehetséges értékei az $\{1, 2, \dots\}$ halmazból valók. Könnyen látható, hogy minden $k \in \{1, 2, \dots\}$ esetén

$$P(X = k) = (1 - p)^{k-1} p.$$

Kiszámolható, hogy $EX = \frac{1}{p}$ és $D^2 X = \frac{1-p}{p^2}$.

1.4. Negatív binomiális eloszlás (jelölés: $\mathrm{NegBin}(r, p)$)

Legyen $A \in \mathcal{A}$ és $p = P(A)$. Egymástól független kísérleteket végzünk, és jelölje X annak a kísérletnek a sorszámát, amikor r-edjére következett be A. Világos, hogy X lehetséges értékei az $\{r, r + 1, \dots\}$ halmazból valók. Könnyen látható, hogy minden $k \in \{r, r + 1, \dots\}$ esetén

$$P(X = k) = \binom{k-1}{r-1} (1 - p)^{k-r} p^r.$$

Kiszámolható, hogy $EX = \frac{r}{p}$ és $D^2 X = \frac{r(1-p)}{p^2}$. Természetes úgy gondolni X-re, mint r darab független p paraméterű geometriai változó összegére, és így a várható értékre és a szórásnégyzetre vonatkozó formula azonnal adódik. Az is nyilvánvaló, hogy az $r = 1$ esetben visszakapjuk a geometriai eloszlást.

1.5. Hipergeometrikus eloszlás (jelölés: $\mathrm{HipGeo}(N, M, n)$)

Tekintsünk egy N termékből álló sokaságot, melyben M darab selejtes van ($M \le N$). Húzzunk ebből a sokaságból n-szer visszatevés nélkül ($n \le N$), és jelölje a kihúzottak közt található selejtesek számát. Világos, hogy X lehetséges értékei a

$$\max(0, n - (N - M)) \text{ és } \min(n, M)$$

közti egészek. Könnyen látható, hogy $k \in \{\max(0, n - (N - M)), \dots, \min(n, M)\}$ esetén

$$P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}.$$

Kiszámolható, hogy $EX = n\frac{M}{N}$ és $D^2 X = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}$. Könnyen látható, hogy amennyiben $X_{N,M} \sim \mathrm{HipGeo}(N, M, n)$ és $N, M \to \infty$ úgy, hogy $\frac{M}{N} \to p$ valamely $0 < p < 1$ számmal, akkor minden $0 \le k \le n$ számra

$$P(X_{N,M} = k) \to \binom{n}{k} p^k (1 - p)^{n-k}.$$

Vagyis, a visszatevés nélküli mintavételezés határértékeként megkapható a visszatevéses mintavételezés.

1.6. Poisson eloszlás (jelölés: $\mathrm{Poisson}(\lambda)$)

Legyen X olyan valószínűségi változó, melynek lehetséges értékei a $\{0, 1, \ldots\}$ halmazból valók és minden $k \in \{0, 1, \ldots\}$ esetén

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

alkalmas $\lambda > 0$ számmal. Ilyenkor X-et λ paraméterű *Poisson* eloszlásúnak mondjuk. Könnyen látható, hogy $EX = \lambda$ és $D^2X = \lambda$. Kiszámolható továbbá, hogy ha X_n binomiális eloszlású n renddel és p_n paraméterrel úgy, hogy $n \to \infty$ esetén $np_n \to \lambda$ alkalmas $\lambda > 0$ számmal, akkor minden $k \in \{0, 1, \ldots\}$ esetén

$$\lim_{n \to \infty} P(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Vagyis, a Poisson eloszlás megkapható alkalmas binomiális eloszlású valószínűségi változók eloszlásbeli limeszeként.

2. Nevezetes abszolút folytonos eloszlások

2.1. Egyenletes eloszlás (jelölés: $\mathrm{U}[a, b]$)

Azt mondjuk, hogy az X valószínűségi változó egyenletes eloszlású az $[a, b]$ intervallumon ($a < b$), ha abszolút folytonos eloszlású és sűrűségfüggvénye az

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{ha } x \in [a, b], \\ 0, & \text{ha } x \notin [a, b] \end{cases}$$

formulával értelmezett. Ekkor eloszlásfüggvényét az

$$F(x) = \begin{cases} 0, & \text{ha } x \leq a, \\ \frac{x-a}{b-a}, & \text{ha } a < x \leq b, \\ 1, & \text{ha } b < x \end{cases}$$

képlet adja. Az X valószínűségi változónak az $[a, b]$ intervallum valamely részintervallumába való esésének az esélye egyenesen arányos a kérdéses részintervallum hosszával. Kiszámolható, hogy $EX = \frac{a+b}{2}$ és $D^2X = \frac{(b-a)^2}{12}$.

2.2. Exponenciális eloszlás (jelölés: $\mathrm{Exp}(\lambda)$)

Azt mondjuk, hogy az X valószínűségi változó *exponenciális* eloszlású $\lambda > 0$ paraméterrel, ha abszolút folytonos eloszlású és sűrűségfüggvénye az

$$f(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ \lambda e^{-\lambda x}, & \text{ha } 0 < x \end{cases}$$

formulával értelmezett. Ekkor eloszlásfüggvényét az

$$F(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ 1 - e^{-\lambda x}, & \text{ha } 0 < x \end{cases}$$

képlet adja. Az exponenciális eloszlás nevezetes tulajdnosága az örökifjúság, miszerint minden $s, t > 0$ számra

$$P(X > s + t | X > s) = P(X > t).$$

Kiszámolható, hogy $EX = \frac{1}{\lambda}$ és $D^2X = \frac{1}{\lambda^2}$.

2.3. Gamma eloszlás (jelölés: $\Gamma(n, \lambda)$)

Azt mondjuk, hogy az X valószínűségi változó *Gamma* eloszlású $n \geq 1$ és $\lambda > 0$ paraméterekkel, ha abszolút folytonos eloszlású és sűrűségfüggvénye az

$$f(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}, & \text{ha } 0 < x \end{cases}$$

formulával értelmezett. Ekkor eloszlásfüggvényét az

$$F(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ 1 - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!}, & \text{ha } 0 < x \end{cases}$$

képlet adja. Az $n \geq 1$ és $\lambda > 0$ paraméterekkel rendelkező Gamma eloszlás éppen n darab független, λ paraméterű exponenciális valószínűségi változó összegének eloszlása. Ez alapján nyilvánvaló, hogy $EX = \frac{n}{\lambda}$ és $D^2X = \frac{n}{\lambda^2}$.

2.4. Normális más szóval Gauss eloszlás (jelölés: $\mathrm{N}(m, \sigma^2)$)

Azt mondjuk, hogy az X valószínűségi változó *standard normális* eloszlású (jelölésben $X \sim \mathrm{N}(0,1)$), ha abszolút folytonos eloszlású és sűrűségfüggvénye a

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \ (x \in \mathbb{R})$$

formulával értelmezett. Ekkor eloszlásfüggvényét a

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \mathrm{d}t \ (x \in \mathbb{R})$$

képlet adja. A Φ függvény pozitív helyeken felvett értékeinek közelítését táblázatból olvashajuk ki, míg negatív helyeken felvett értékeinek közelítését a $\Phi(-x) = 1 - \Phi(x)$ formula alapján vezethetjük vissza a pozitív helyeken felvett értékeire (a vázolt formula minden $x \in \mathbb{R}$ esetén fennáll). A táblázat Φ inverzének értékeinek megadására is használható. Látható, hogy $EX = 0$ és $D^2X = 1$.

Azt mondjuk, hogy az Y valószínűségi változó normális eloszlású $m \in \mathbb{R}$ és σ^2 paraméterekkel (ahol $\sigma > 0$), ha létezik olyan X standard normális eloszlású valószínűségi változó, mellyel $Y = \sigma X + m$. Ilyenkor Y sűrűségfüggvényét könnyen látható módon az

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-m)^2}{2\sigma^2}} \quad (y \in \mathbb{R})$$

formula adja, míg eloszlásfüggvénye az

$$F(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-m)^2}{2\sigma^2}} \mathrm{d}t \ (y \in \mathbb{R})$$

képlettel értelmezett. A várható érték és a szórásnégyzet alapvető tulajdonságai alapján nyilvánvaló, hogy $EX = m$ és $D^2X = \sigma^2$. Könnyen látható, hogy minden $y \in \mathbb{R}$ esetén fennáll $F(y) = \Phi\left(\frac{y-m}{\sigma}\right)$, vagyis a (nem feltétlenül standard) normális eloszlású változó eloszlásfüggvényének értékeinek kiszámítását visszavezethetjük Φ értékeinek megadására (utóbbi a fent említett táblázat segítségével oldható meg). Fontos, hogy amennyiben $Y_1 \sim \mathrm{N}(m_1, \sigma_1^2)$ és $Y_2 \sim \mathrm{N}(m_2, \sigma_2^2)$ függetlenek, úgy $Y_1 + Y_2 \sim \mathrm{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$ (ez nem nyilvánvaló). Triviális továbbá, hogy ha $a \neq 0$, $b \in \mathbb{R}$ és $Y \sim \mathrm{N}(m, \sigma^2)$, akkor $aY + b \sim \mathrm{N}(am + b, a^2\sigma^2)$.

3. A Centrális Határeloszlás Tétel (CHT) és néhány alkalmazása

3.1. CHT

Ha X_1, $X_2, \ldots$ független, azonos eloszlású valtozók, melyekre $0 < D^2X_1 < \infty$, akkor

$$\frac{\sum_{i=1}^{n} X_i - nEX_1}{\sqrt{n}DX_1}$$

eloszlásban standard normális eloszláshoz tart, midőn $n \to \infty$, azaz

$$P\left(\frac{\sum_{i=1}^{n} X_i - nEX_1}{\sqrt{n}DX_1} < x\right) \to \Phi(x)$$

minden $x \in \mathbb{R}$ esetén, midőn $n \to \infty$.

3.2. Fejek száma

Dobjunk fel egy szabályos érmét 10000-szer, és jelölje X a kapott fejek számát. Ekkor $X \sim \mathrm{Bin}\left(10000, \frac{1}{2}\right)$. Feladat, hogy adjunk egy, az 5000-re szimmetrikus intervallumot, melybe X beleesik 75% eséllyel. Nyilván $X = \sum_{i=1}^{10000} X_i$, ahol az összeadandók függetlenek és

$$X_i = \begin{cases} 1, & \text{ha az } i\text{-edik dobás során fejet kaptunk,} \\ 0, & \text{különben} \end{cases} \qquad (1 \leq i \leq 10000).$$

Célunk olyan $K > 0$ megadása, mellyel

$$0,75 = P(|X - 5000| \leq K).$$

Mivel $EX_1 = \frac{1}{2}$ és $DX_1 = \sqrt{\frac{1}{2}\left(1-\frac{1}{2}\right)} = \frac{1}{2}$, így a CHT miatt

$$P(|X-5000| \leq K) = P\left(\left|\sum_{i=1}^{10000} X_i - 10000 \cdot EX_1\right| \leq K\right) =$$

$$= P\left(\left|\frac{\sum_{i=1}^{10000} X_i - 10000 \cdot EX_1}{\sqrt{10000}\frac{1}{2}}\right| \leq \underbrace{\frac{K}{\sqrt{10000}\frac{1}{2}}}_{\frac{K}{50}}\right) \approx$$

$$\approx \Phi\left(\frac{K}{50}\right) - \underbrace{\Phi\left(-\frac{K}{50}\right)}_{1-\Phi\left(\frac{K}{50}\right)} =$$

$$= 2 \cdot \Phi\left(\frac{K}{50}\right) - 1.$$

Következésképpen, azt a $K > 0$ számot kell megtalálnunk, melyre

$$0,75 = 2 \cdot \Phi\left(\frac{K}{50}\right) - 1.$$

Innen $\Phi\left(\frac{K}{50}\right) = 0,875$ adódik. Mivel a standard normális eloszlás táblázata alapján $\Phi(1,15) = 0,875$, így $\frac{K}{50} = 1,15$ választással kell élnünk, ahonnan $K = 57,5$ adódik. Tehát, a 10000 érmedobásból kapott fejek száma 75% eséllyel a $[4943, 5057]$ intervallumba esik. A 15. és a 16. ábrákon szimulációs eredményeket mutatunk be.

3.3. Közvéleménykutatás

Adott N ember, akik közül M kedveli a baglyokat. Jelölje p az $\frac{M}{N}$ hányadost. Célunk p-t úgy közelíteni, hogy legalább 95%-os valószínűséggel legfeljebb $0,01$-ot tévedjünk. Megkérdezünk n embert visszatevéssel arról, hogy kedvelik-e a baglyokat. Legyen

$$X_i = \begin{cases} 1, & \text{ha az } i\text{-edik megkérdezett kedveli a baglyokat,} \\ 0, & \text{különben} \end{cases} \qquad (1 \leq i \leq n).$$

Ekkor $X_1, X_2, \ldots, X_n$ független indikátorok p paraméterrel. A p paraméterre vonatkozó becslésünk $\frac{\sum_{i=1}^n X_i}{n}$. Mivel $EX_1 = p$ és $DX_1 = \sqrt{p(1-p)}$, így a CHT miatt

$$P\left(\left|\frac{\sum_{i=1}^n X_i}{n} - p\right| \leq 0,01\right) = P\left(\left|\frac{\sum_{i=1}^n X_i - np}{n}\right| \leq 0,01\right) =$$

$$= P\left(\left|\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}}\right| \leq \frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right) \approx$$

$$\approx \Phi\left(\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right) - \underbrace{\Phi\left(-\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right)}_{1-\Phi\left(\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right)} =$$

$$= 2 \cdot \Phi\left(\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right) - 1.$$

Következésképpen, n-et olyanra kell választanunk, hogy minden $0 < p < 1$ esetén fennálljon

$$0,95 \leq 2 \cdot \Phi\left(\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right) - 1.$$

Innen $\Phi\left(\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right) \geq 0,975$ adódik. Mivel a standard normális eloszlás táblázata alapján $\Phi(1,96) = 0,975$, így fenn kell álljon $\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}} \geq 1,96$ minden $0 < p < 1$ esetén, azaz $n \geq 196^2 p(1-p)$. Mivel minden $0 < p < 1$ esetén fennáll $p(1-p) \leq \frac{1}{4}$, így $n \geq 98^2 = 9604$ választás megfelelő. Tehát, ha legalább 9604 embert kérdezünk meg, akkor legalább 95%-os eséllyel legfeljebb 1 százalékpontot tévedünk. A 17. és a 18. ábrákon szimulációs eredményeket mutatunk be.

4. Emlékeztető kombinatorikából és (főleg) analízisből

4.1. Az exponenciális függvény néhány tulajdonsága

Minden $x \in \mathbb{R}$ esetén

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots \text{ és}$$
$$e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n.$$

A l'Hôpital szabály segítségével látható, hogy

$$\lim_{x\to 0} \frac{e^x - 1}{x} = 1.$$

Ismeretes továbbá $\mathbb{R} \ni x \mapsto e^x \in \mathbb{R}$ szigorúan monoton növő és

$$\lim_{x\to-\infty} e^x = 0,$$
$$\lim_{x\to\infty} e^x = \infty.$$

4.2. A binomiális tétel

Tetszőleges x és y valós számokra, valamint n nemnegatív egészre fennáll

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k},$$

ahol $\binom{n}{k}$ definíció szerint $\frac{n!}{k!(n-k)!}$. A $0!$ értéke definíció szerint 1.

4.3. A mértani sor

Legyen q olyan valós szám, melyre $|q| < 1$. Ekkor

$$\sum_{k=0}^{\infty} q^n = \frac{1}{1-q}.$$

4.4. Deriválás

4.4.1. Szorzat deriválása

Legyenek $f : I \to \mathbb{R}$ és $g : I \to \mathbb{R}$ differenciálható függvények, ahol $I \subseteq \mathbb{R}$ nyílt halmaz. Ekkor $f \cdot g : I \to \mathbb{R}$ szintén differenciálható és

$$(f \cdot g)' = f' \cdot g + f \cdot g'.$$

4.4.2. Kompozíciófüggvény deriválása

Legyenek $f, g : \mathbb{R} \to \mathbb{R}$ differenciálható függvények. Ekkor $f \circ g$ differenciálható és

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x) \quad (\forall\, x \in \mathbb{R}).$$

4.4.3. Nevezetes deriváltak

Az alábbiakban a log jelölés az e alapú logaritmust jelöli.

$$\begin{aligned}
(e^x)' &= e^x \\
(a^x)' &= a^x \log a \quad (a > 0, a \neq 1) \\
(\log x)' &= \frac{1}{x} \\
(\log_a x)' &= \frac{1}{x \log a} \quad (a > 0, a \neq 1) \\
(x^\alpha)' &= \alpha x^{\alpha - 1} \quad (\alpha \neq 0)
\end{aligned}$$

4.5. Integrál

4.5.1. Nevezetes függvények integrálja

$$\begin{aligned}
\int e^x \mathrm{d}x &= e^x + C \\
\int a^x \mathrm{d}x &= \frac{a^x}{\log a} + C \\
\int x^n \mathrm{d}x &= \frac{x^{n+1}}{n+1} + C \quad (x \in \mathbb{R}, n \in \mathbb{Z}_{\geq 0}) \\
\int x^\alpha \mathrm{d}x &= \frac{x^{\alpha+1}}{\alpha+1} + C \quad (x > 0, \alpha \neq -1) \\
\int \frac{1}{x} \mathrm{d}x &= \log x + C
\end{aligned}$$

4.5.2. Newton-Leibniz formula

Legyen $F : \mathbb{R} \to \mathbb{R}$ olyan függvény, melyhez létezik olyan $f : \mathbb{R} \to \mathbb{R}$ integrálható függvény ($\int_{-\infty}^{\infty} |f(t)| \mathrm{d}t < \infty$), amellyel minden $x \in \mathbb{R}$ esetén $F(x) = \int_{-\infty}^{x} f(t) \mathrm{d}t$. Ekkor minden $a < b$

számpárra

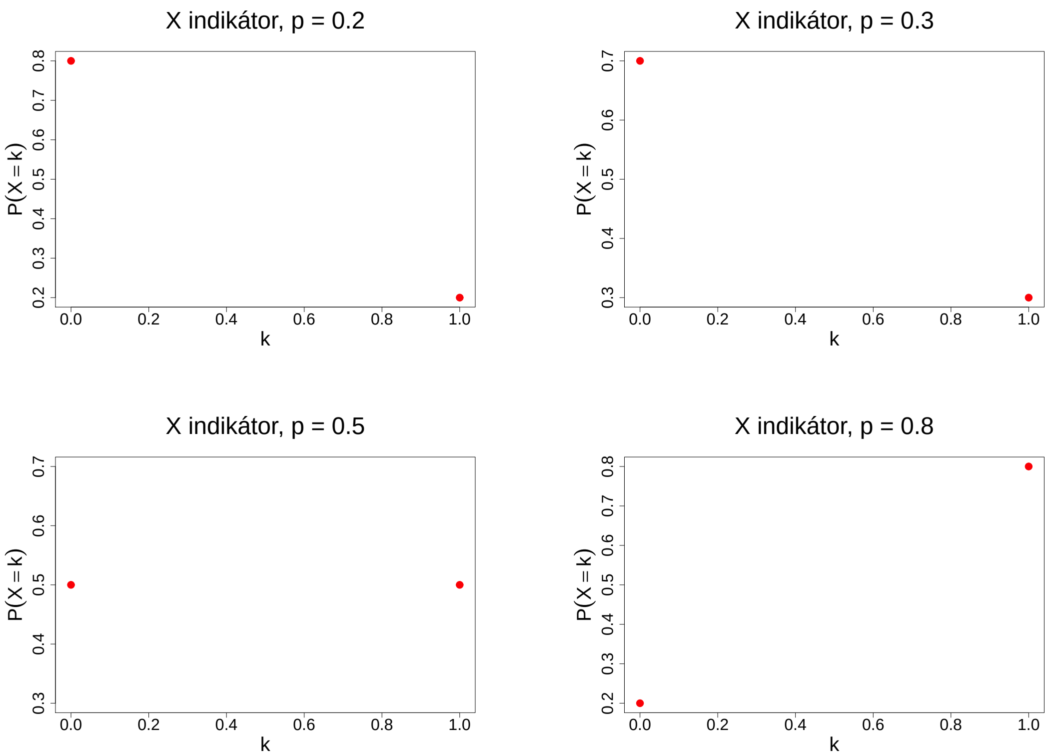
$$F(b) - F(a) = \int_a^b f(t)\mathrm{d}t,$$
$$\lim_{b\to\infty} F(b) - F(a) = \int_a^\infty f(t)\mathrm{d}t,$$
$$F(b) - \lim_{a\to-\infty} F(a) = \int_{-\infty}^b f(t)\mathrm{d}t \text{ és}$$
$$\lim_{b\to\infty} F(b) - \lim_{a\to-\infty} F(a) = \int_{-\infty}^\infty f(t)\mathrm{d}t.$$

4.5.3. Parciális integrálás

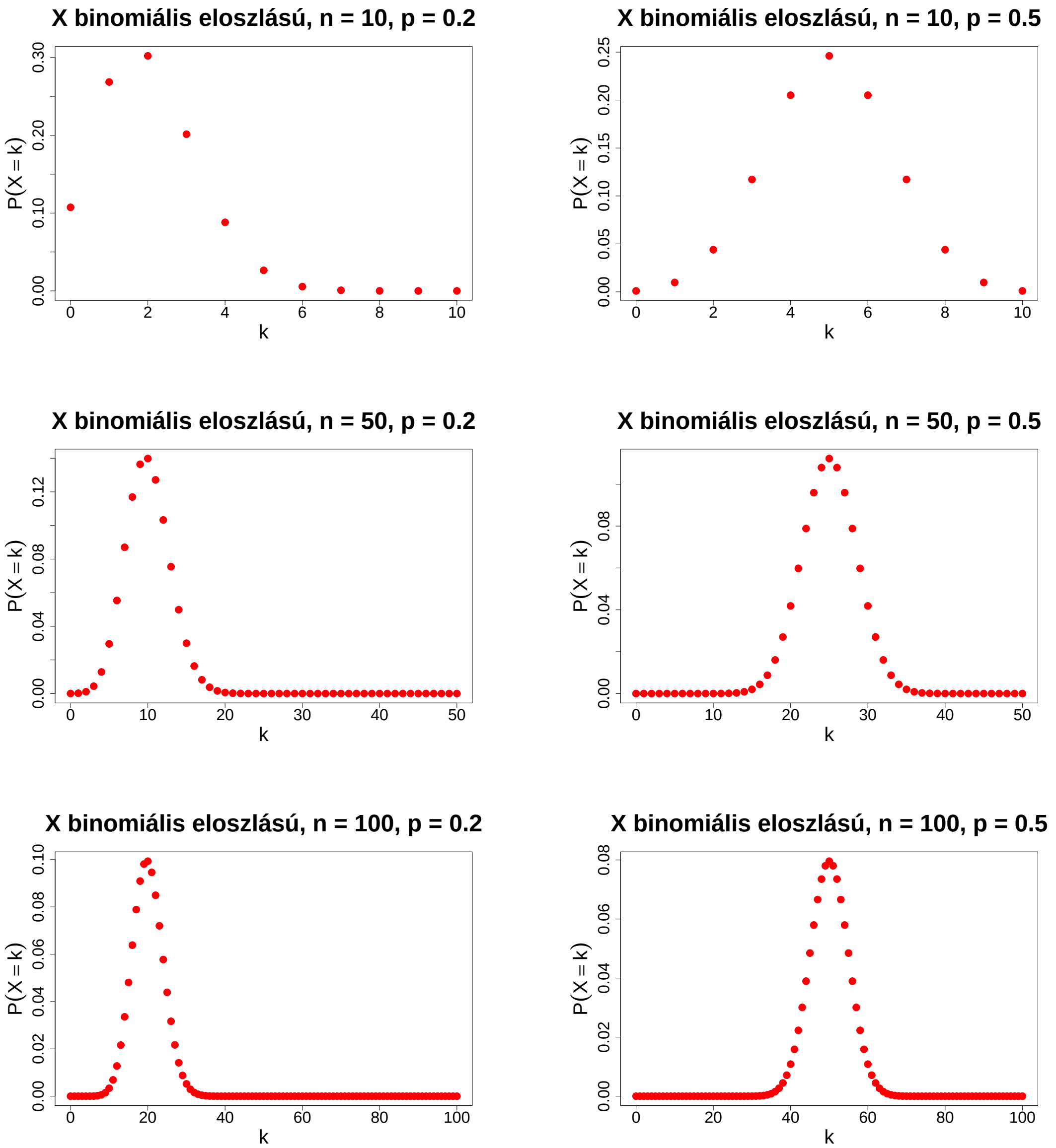
Ha $f, g : \mathbb{R} \to \mathbb{R}$ folytonosan differenciálhatóak, akkor minden $a < b$ számpárra

$$\int_a^b f'(x)g(x)\mathrm{d}x = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x)\mathrm{d}x.$$

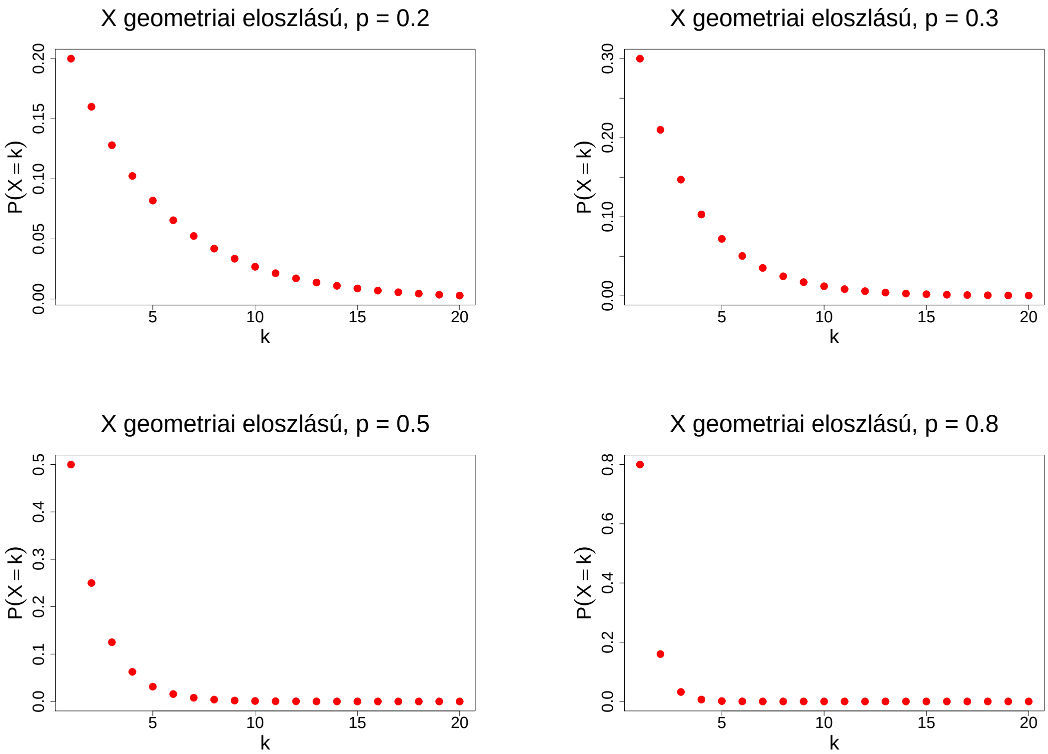
5. Ábrák



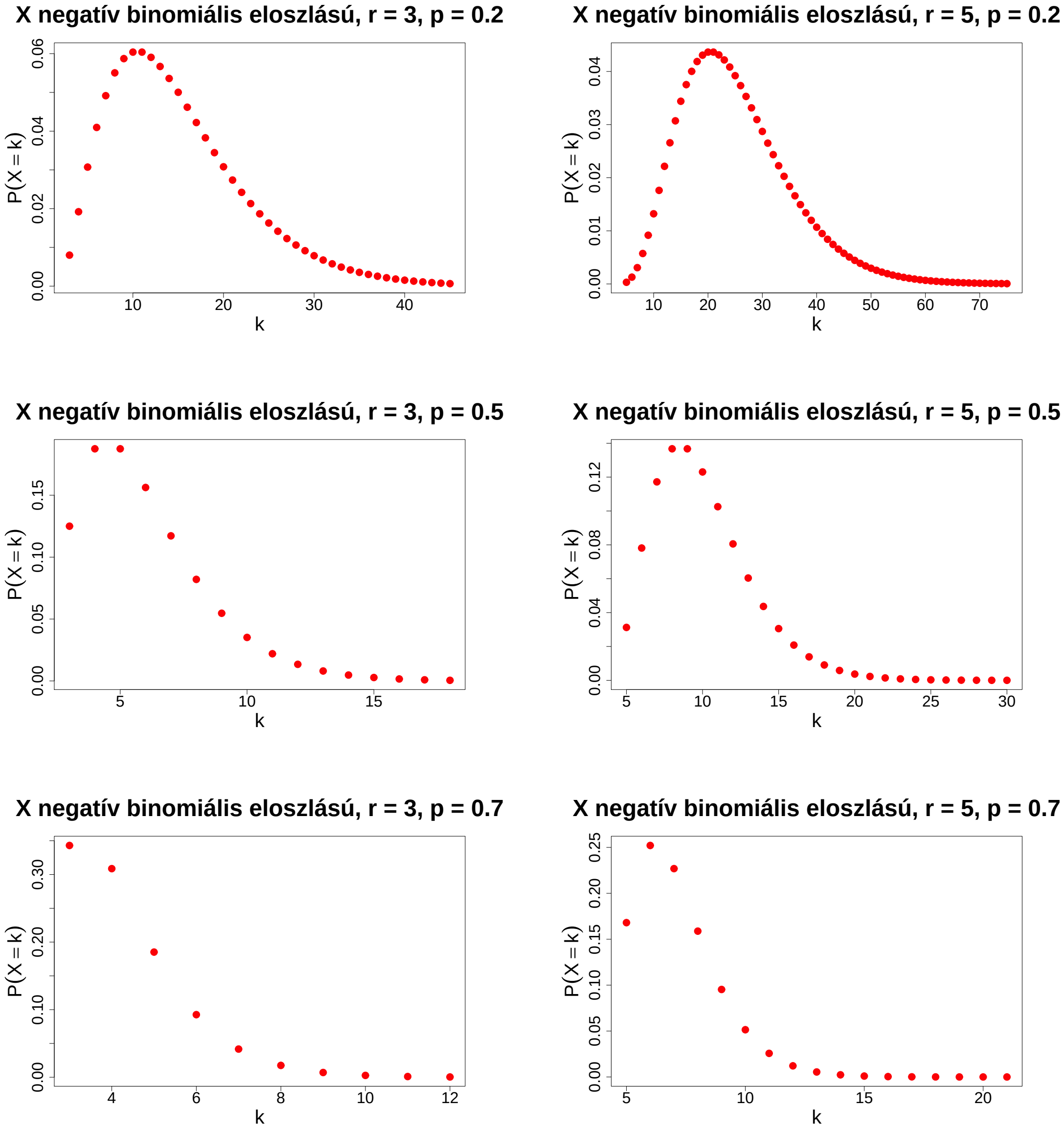
1. ábra. Az indikátorváltozó eloszlása a $p = 0.2, 0.3, 0.5$, illetve 0.8 esetekben.



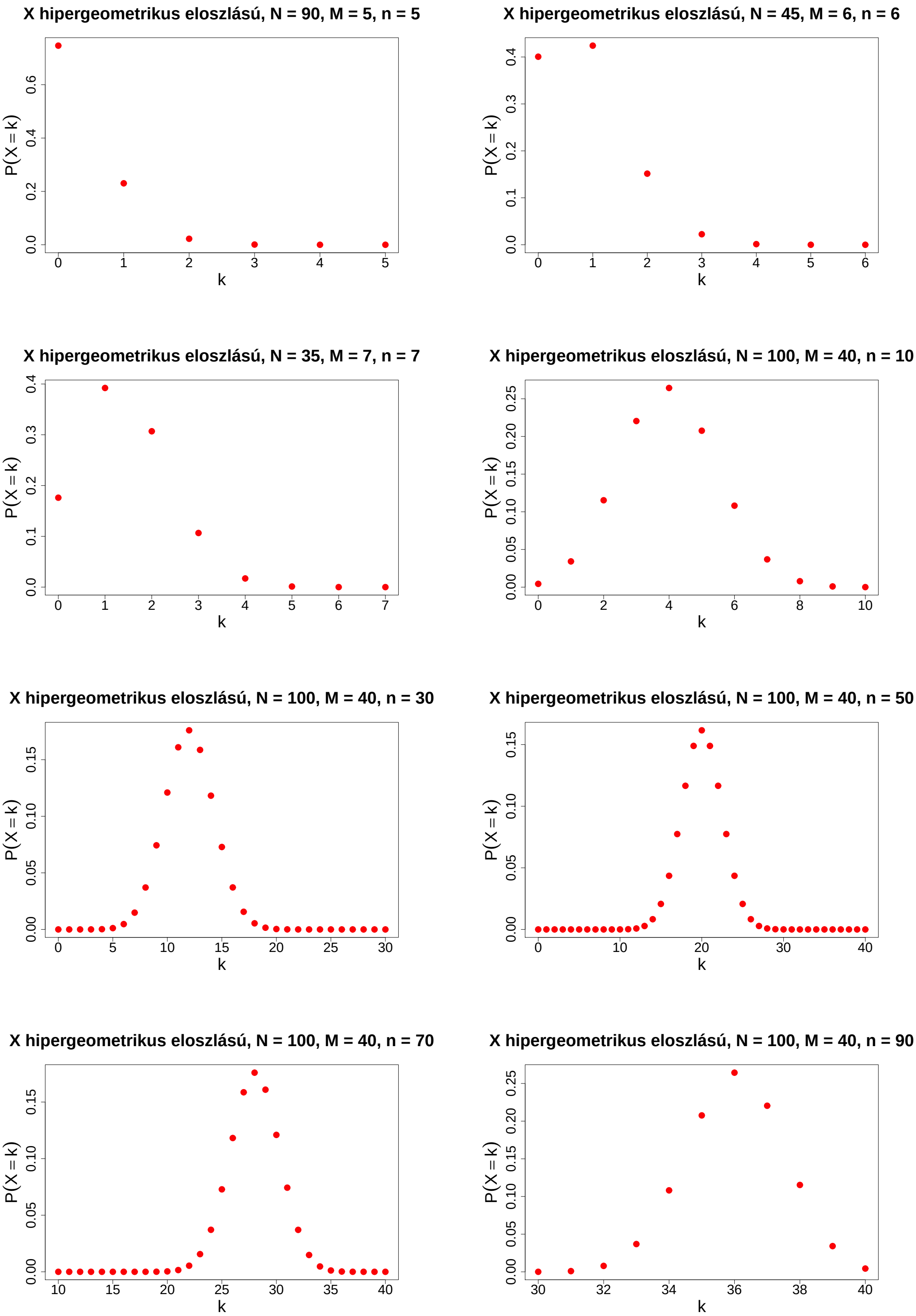
2. ábra. A binomiális eloszlás $n = 10, 50, 100$ renddel és $p = 0.2, 0.5$ paraméterrel ($3 \cdot 2 = 6$ ábra).



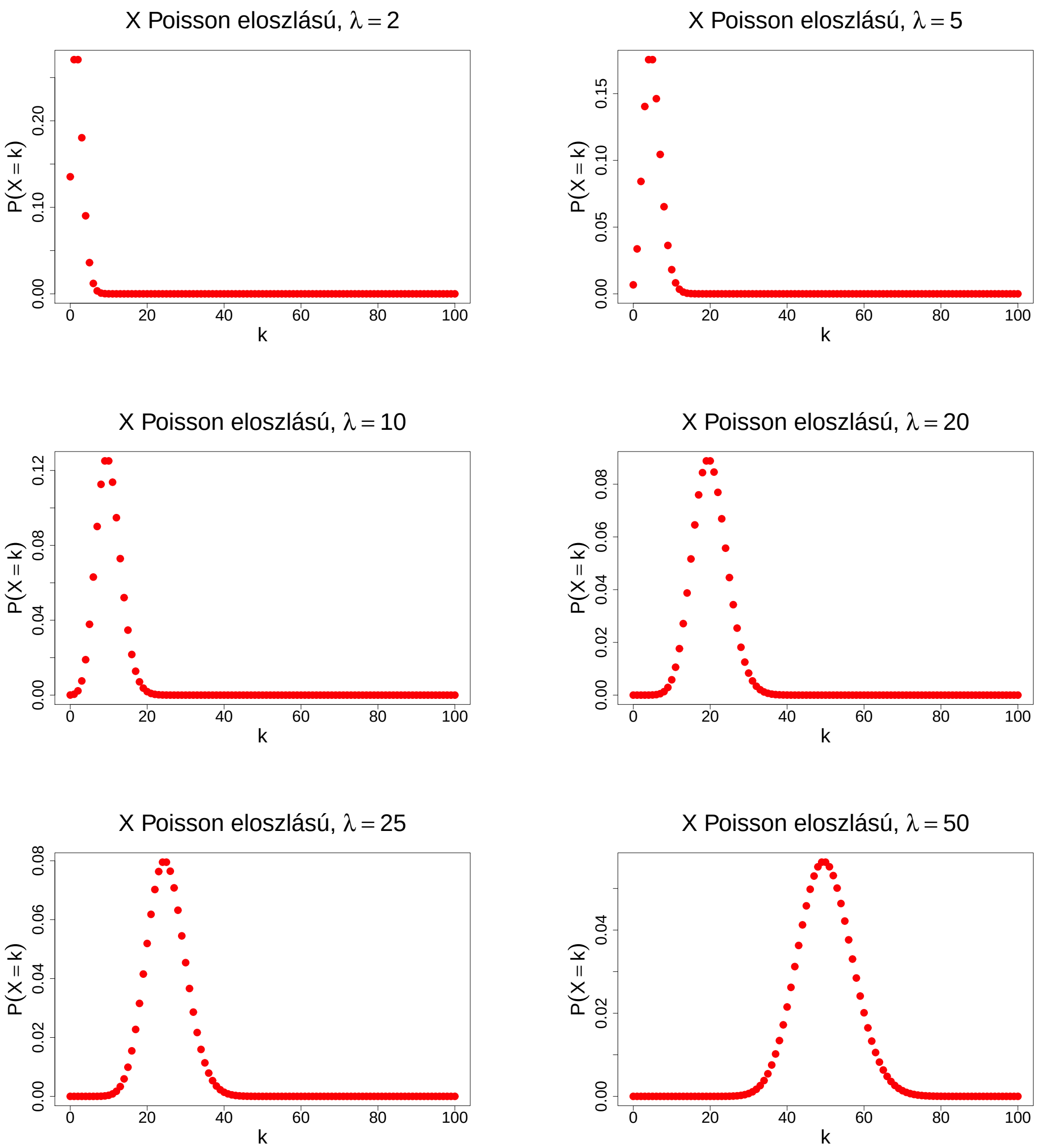
3. ábra. Az geometriai eloszlás a $p = 0.2, 0.3, 0.5$, illetve 0.8 esetekben.



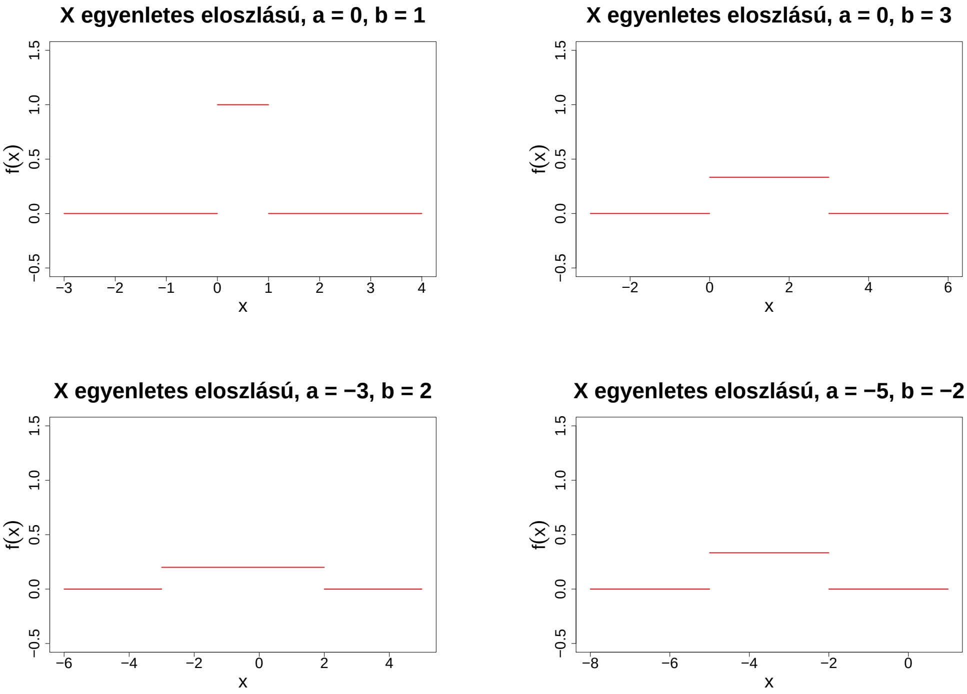
4. ábra. A negatív binomiális eloszlás $r = 3, 5$ renddel és $p = 0.2, 0.5, 0.7$ paraméterrel $(2 \cdot 3 = 6$ ábra).



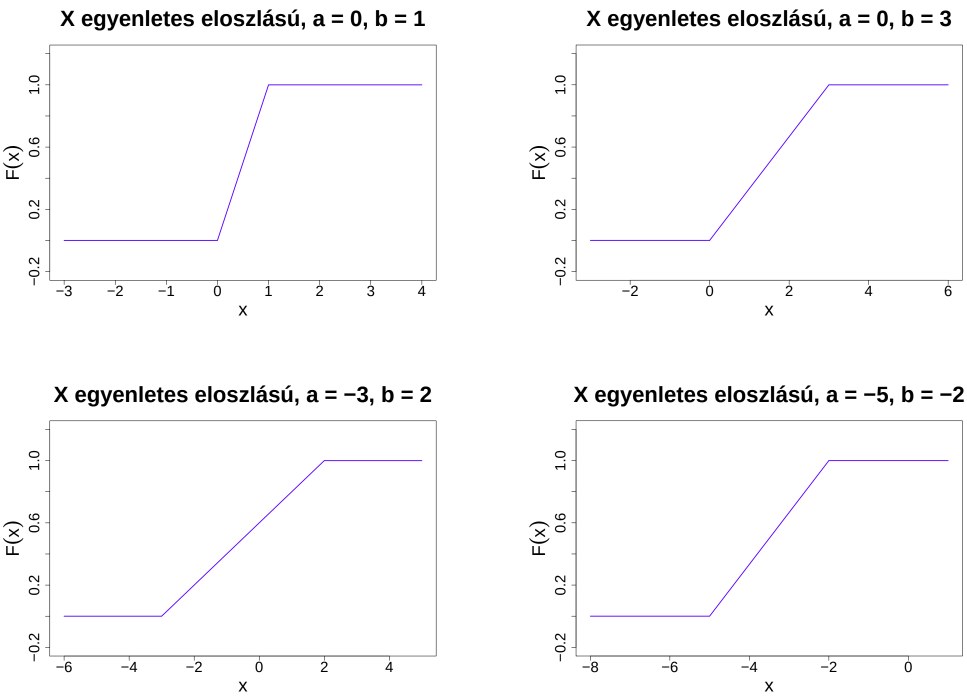
5. ábra. A hipergeometrikus eloszlás az ötöslottón, a hatoslottón, valamint a Skandináv lottón szemléltetve. Továbbá a $(100, 40, 10)$, $(100, 40, 30)$, $(100, 40, 50)$, $(100, 40, 70)$ és $(100, 40, 90)$ paraméterű hipergeometrikus eloszlás.



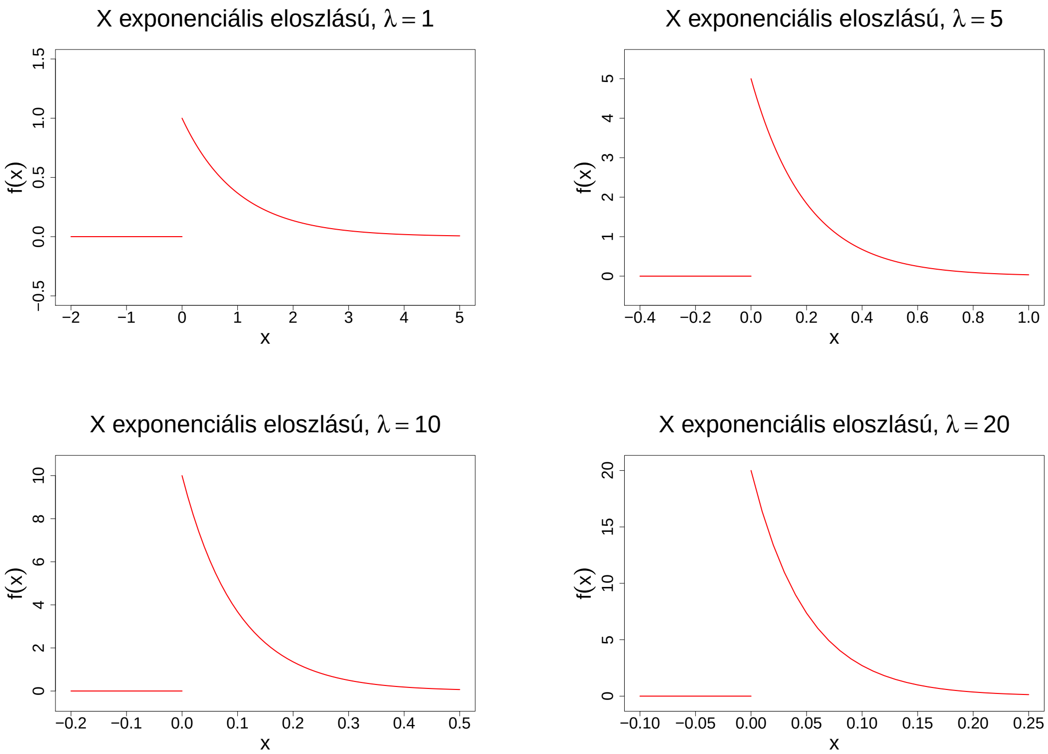
6. ábra. A Poisson eloszlás $\lambda = 2, 5, 10, 20, 25, 50$ paraméterekkel.



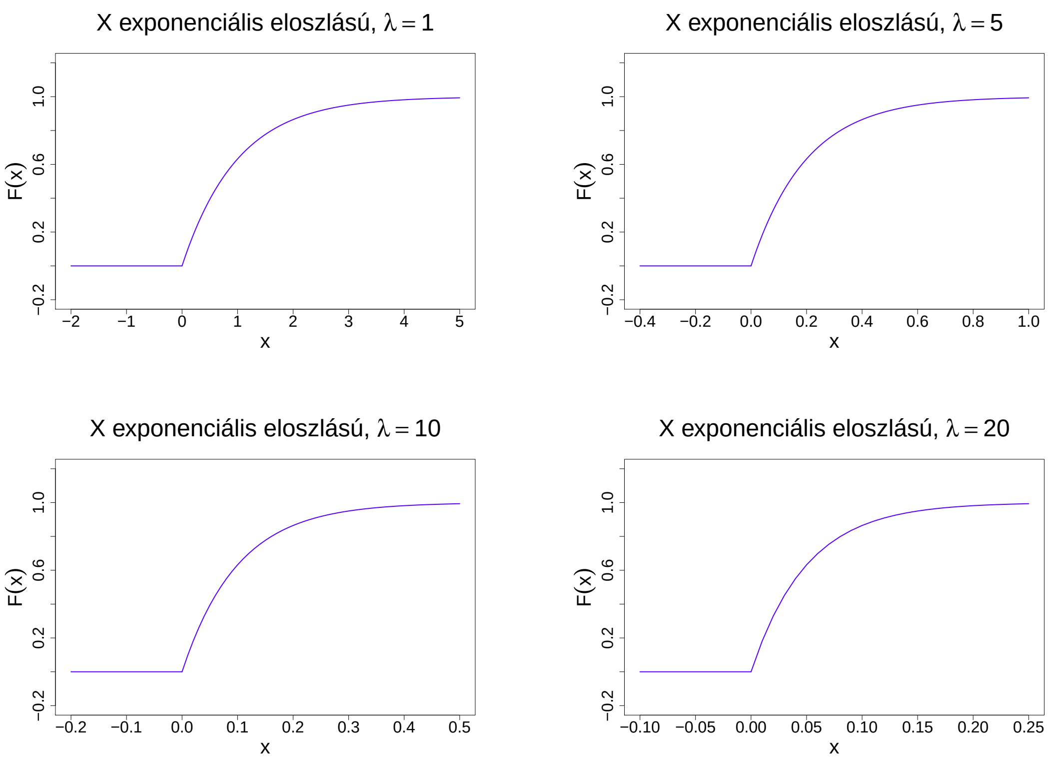
7. ábra. Az egyenletes eloszlás sűrűségfüggvénye az $[a, b] = [0, 1]$, $[0, 3]$, $[-3, 2]$, $[-5, -2]$ intervallumok esetében.



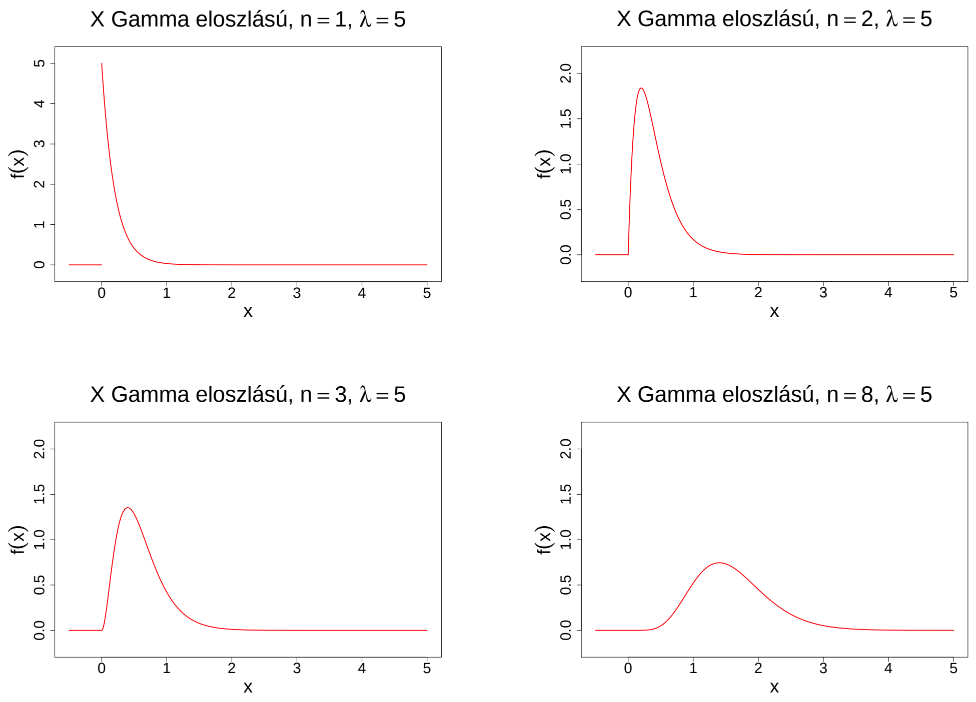
8. ábra. Az egyenletes eloszlás eloszlásfüggvénye az $[a, b] = [0, 1]$, $[0, 3]$, $[-3, 2]$, $[-5, -2]$ intervallumok esetében.



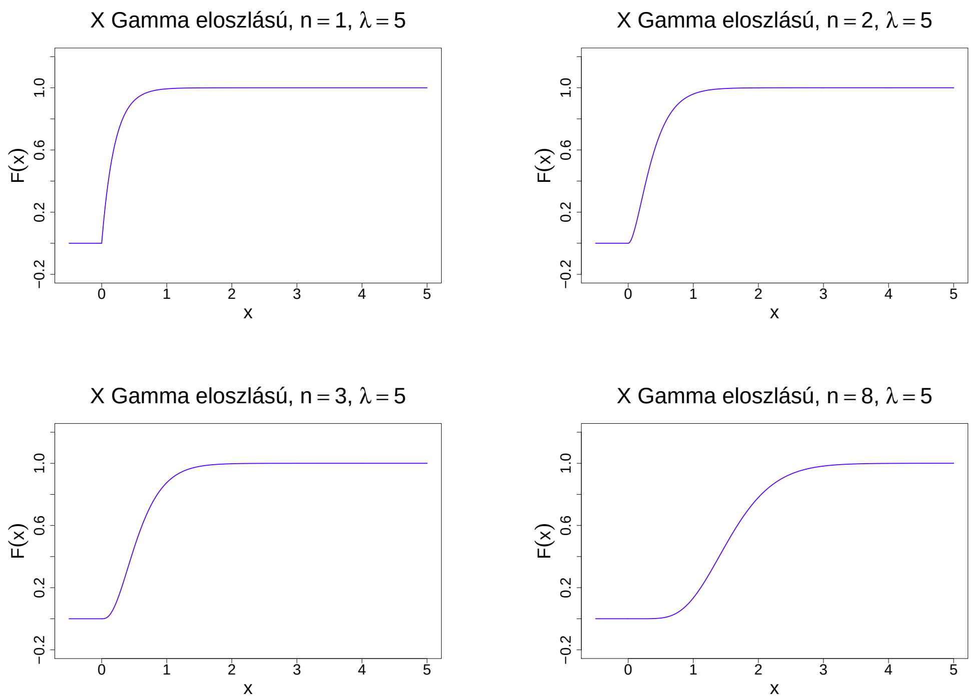
9. ábra. Az exponenciális eloszlás sűrűségfüggvénye a $\lambda = 1, 5, 10, 20$ esetekben.



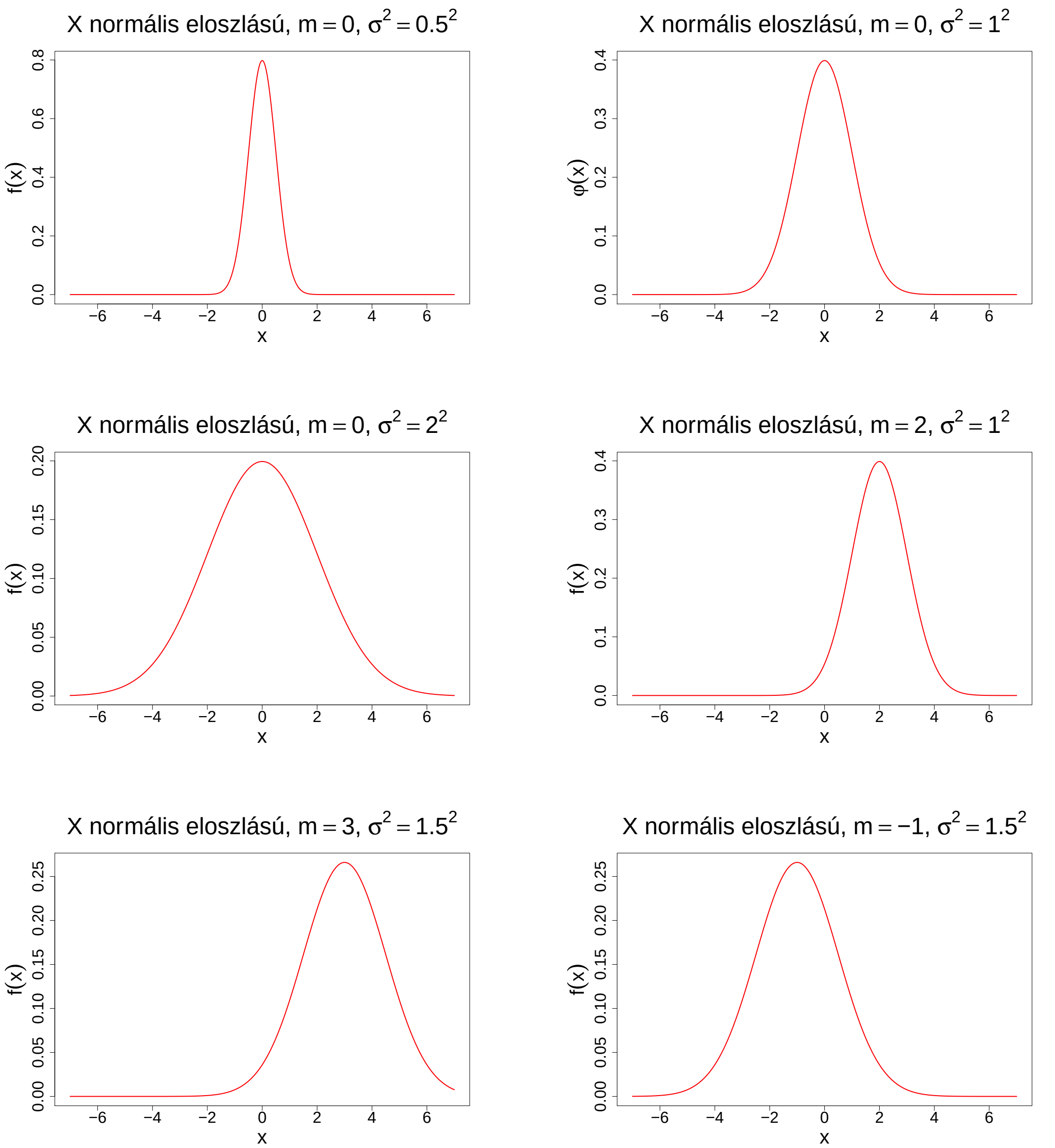
10. ábra. Az exponenciális eloszlás eloszlásfüggvénye a $\lambda = 1, 5, 10, 20$ esetekben.



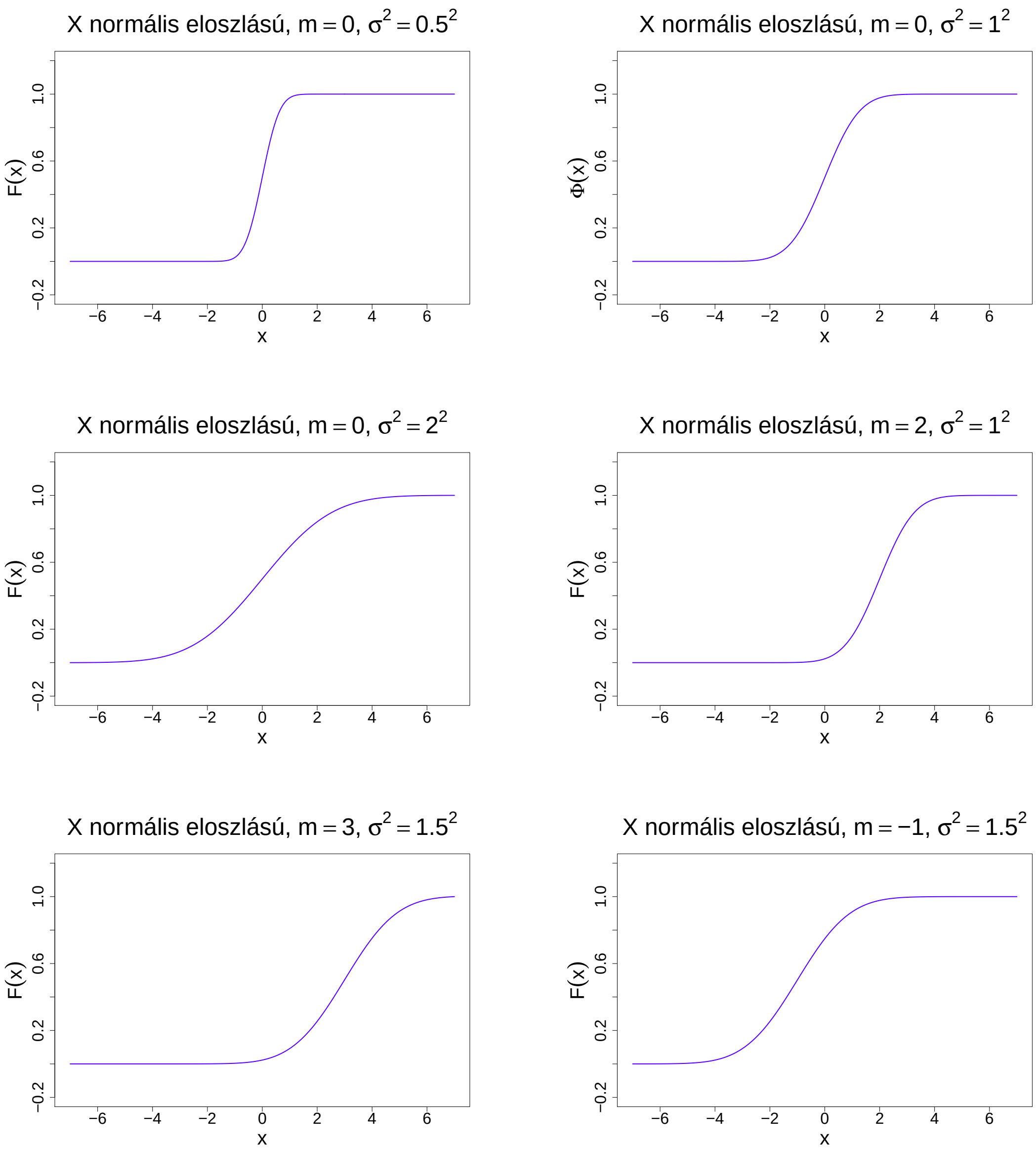
11. ábra. A gamma eloszlás sűrűségfüggvénye az $n = 1, 2, 3, 8$ paraméterek és $\lambda = 5$ mellett.



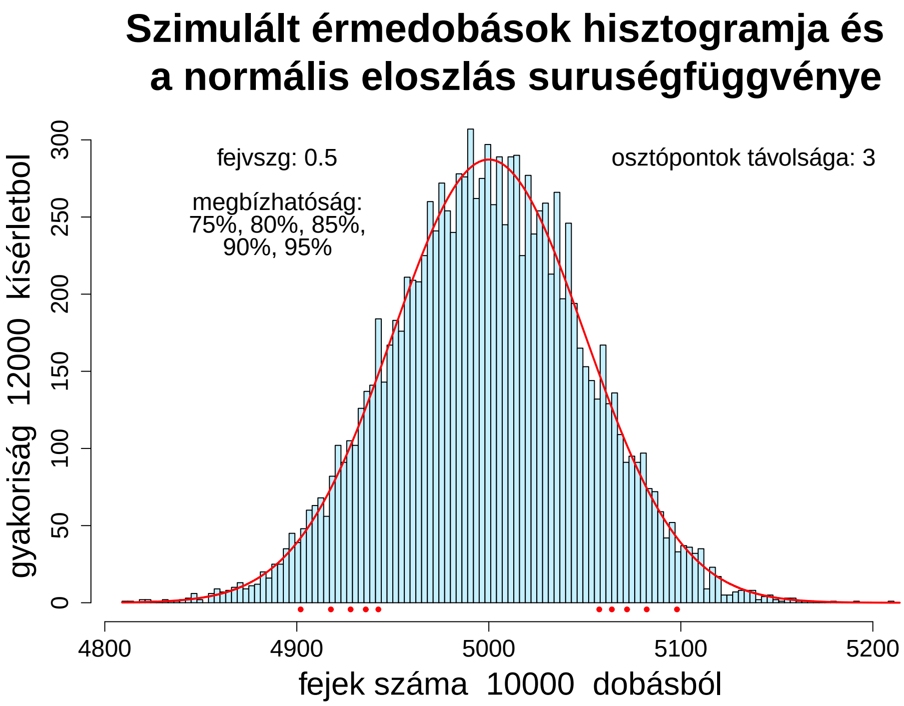
12. ábra. A gamma eloszlás eloszlásfüggvénye az $n = 1, 2, 3, 8$ paraméterek és $\lambda = 5$ mellett.



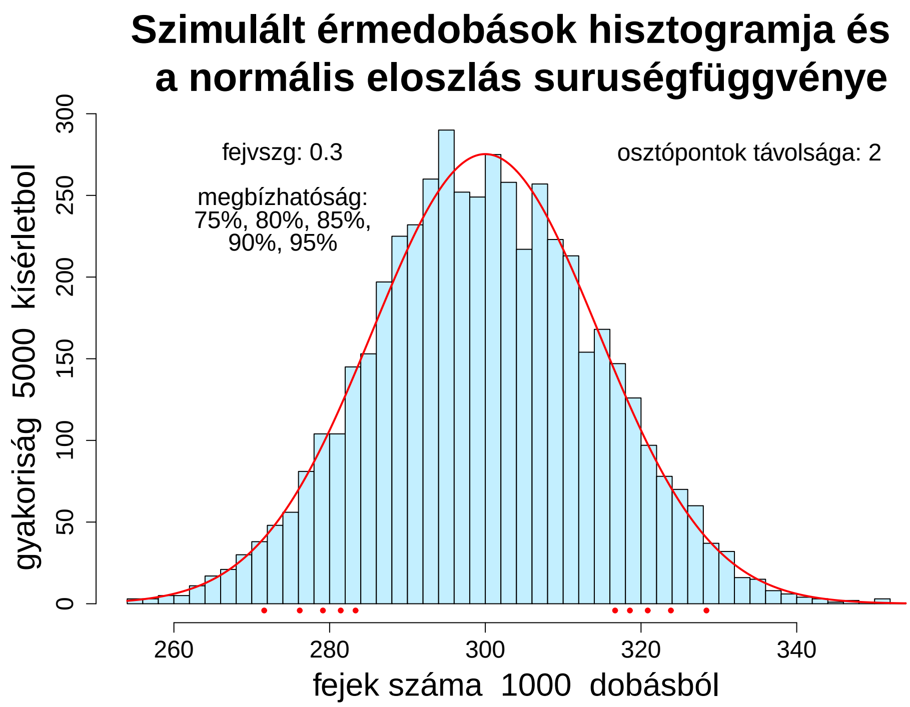
13. ábra. A normális eloszlás sűrűségfüggvénye az $(m, \sigma^2) = (0, 0.5^2)$, $(0, 1^2)$, $(0, 2^2)$, $(2, 1^2)$, $(3, 1.5^2)$, $(-1, 1.5^2)$ paraméterek esetén.



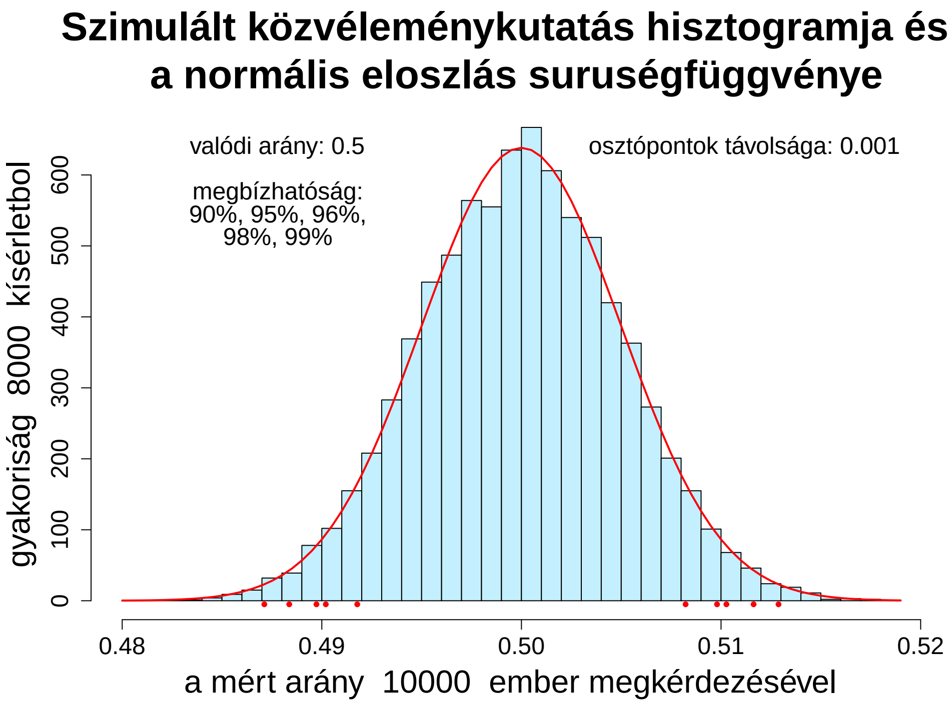
14. ábra. A normális eloszlás eloszlásfüggvénye az $(m, \sigma^2) = (0, 0.5^2)$, $(0, 1^2)$, $(0, 2^2)$, $(2, 1^2)$, $(3, 1.5^2)$, $(-1, 1.5^2)$ paraméterek esetén.



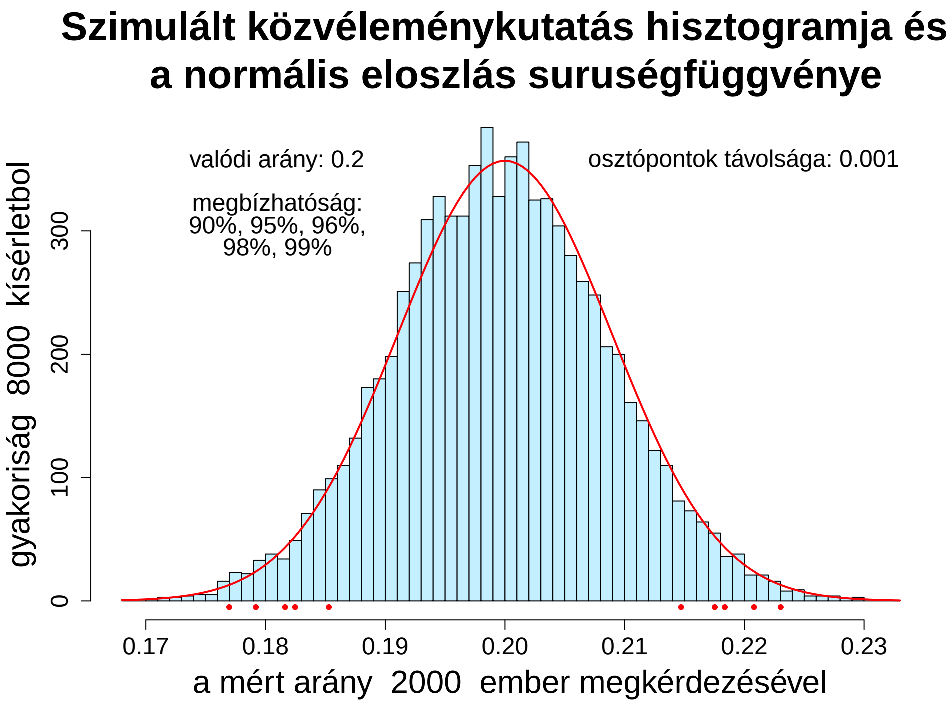
15. ábra. Szimuláltunk 10000 érmedobást 12000-szer (mind a 12000 alkalommal feljegyeztük, hogy hány fejet kaptunk) és hisztogramon árbázoltuk a kapott értékek gyakoriságát. A jobb láthatóság kedvéért a vízszintes tengelyen 3-3 értéket összevontunk, így a függőleges tengelyen mért gyakoriságok azt jelzik, hogy a 12000 kísérlet során hányszor kaptunk a megfelelő, 3 egészet tartalmazó intervallumba eső értéket. A CHT alapján $\frac{X-5000}{50}$ közel standard normális eloszlású, így X közel $\mathrm{N}(5000, 50^2)$ eloszlású, ennek sűrűségfüggvényét ábrázoltuk pirossal (36000-rel felszorozva, hiszen a hisztogram tégláinak területe éppen 36000, mivel minden tégla vízszintes éle 3 hosszú és a téglák függőleges éleinek összhossza 12000). A piros pöttyök a jelzett megbízhatóságú konfidenciaintervallum végpontjait jelölik (azaz a 75%, 80%, 85%, 90% és 95%-hoz tartozó olyan intervallumokat, amelybe az a kapott fejek száma a megfelelő eséllyel beleesik). Leolvashatjuk például, hogy a fejek száma 95% eséllyel a $[4903, 5097]$ intervallumba esik.



16. ábra. Szimuláltunk 1000 érmedobást (az érme fejvalószínűsége 0.3, míg írásvalószínűsége 0.7) 5000-szer (mind az 5000 alkalommal feljegyeztük, hogy hány fejet kaptunk) és hisztogramon árbázoltuk a kapott értékek gyakoriságát. A jobb láthatóság kedvéért a vízszintes tengelyen 2-2 értéket összevontunk, így a függőleges tengelyen mért gyakoriságok azt jelzik, hogy az 12000 kísérlet során hányszor kaptunk a megfelelő, 2 egészet tartalmazó intervallumba eső értéket. A CHT alapján $\frac{X-300}{\sqrt{1000}\sqrt{0.3\cdot 0.7}}$ közel standard normális eloszlású, így X közel $\mathrm{N}(300, (\sqrt{210})^2)$ eloszlású, ennek sűrűségfüggvényét ábrázoltuk pirossal (2000-rel felszorozva, hiszen a hisztogram tégláinak területe éppen 10000, mivel minden tégla vízszintes éle 2 hosszú és a téglák függőleges éleinek összhossza 5000). A piros pöttyök a jelzett megbízhatóságú konfidenciaintervallum végpontjait jelölik (azaz a 75%, 80%, 85%, 90% és 95%-hoz tartozó olyan intervallumokat, amelybe az a kapott fejek száma a megfelelő eséllyel beleesik). Leolvashatjuk például, hogy a fejek száma 85% eséllyel a $[279, 321]$ intervallumba esik.



17. ábra. A 15. és 16. ábrák készítéséhez hasonlóan jártunk el, de itt gyakoriságok helyett relatív gyakoriságot ábrázoltunk (a szimuláció során feltételeztük, hogy az igazi arány 0.5), illetve a megbízhatósági intervallumok megbízhatósági szintjeit ebben az esetben 90%, 95%, 96%, 98% és 99% esetére számítottuk ki. Például, az ábra alapján azt mondhatjuk, hogy amennyiben az igazi arány 0.5, úgy a mérésünkkor 96% eséllyel ettől legfeljebb 1 százalékponttal fogunk eltérni.



18. ábra. A 15. és 16. ábrák készítéséhez hasonlóan jártunk el, de itt gyakoriságok helyett relatív gyakoriságot ábrázoltunk (a szimuláció során feltételeztük, hogy az igazi arány 0.2), illetve a megbízhatósági intervallumok megbízhatósági szintjeit ebben az esetben 90%, 95%, 96%, 98% és 99% esetére számítottuk ki. Például, az ábra alapján azt mondhatjuk, hogy amennyiben az igazi arány 0.2, úgy a mérésünkkor 90% eséllyel ettől legfeljebb 1.5 százalékponttal fogunk eltérni.