Valószínűségszámítás és statisztika
Valószínűségszámítás témakör jegyzet

Készült Zempléni András előadásai
és Kovács Ágnes gyakorlatai alapján

Sárközi Gergő, 2021-22-2. félév

Nincsen lektorálva!

Tartalomjegyzék

1. Előadás 1: Bevezetés

1.1. Alapfogalmak

- Eseménytér: Ω, összes lehetséges eredmény (elemi események összessége)
  - Elemi esemény: ω, egy lehetséges kimenet
  - Ω részhalmazai: események
  - Esemény akkor következik be, ha az egyik elemi eseménye bekötvetkezik
- Esemény: Ω részhalmaza
  - Speciális események: Ω (biztos), $\emptyset$ (lehetetlen)
  - Események összessége: $\mathcal{A}$ (halmazrendszer Ω részhalmazaiból)
- Műveletek eseményekkel = logikai műveletek = halmazműveletek
  - $A \cup B$: vagy (egyik vagy másik vagy mindkettő)
  - $A \cap B$: és (mindkettő)
  - $\overline{A}$, A^C: ellentett, komplementer
- Tulajdonságok:
  - $A \setminus B = A \cap \overline{B}$
  - De Morgan: $\overline{A \cup B} = \overline{A} \cap \overline{B}$
  - $\overline{\overline{A}} = A$, $\overline{\Omega} = \emptyset$

1.2. Valószínűség

- Valószínűség: végtelen próba esetén a relatív gyakoriság
- A valószínűségének jele: $P(A) \quad (P(A) \geq 0,\ P(\Omega) = 1,\ P(\emptyset) = 0)$
- Tulajdonságok:
  - Egymást kizáró események esetén additív:
    $A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B)$
  - $P(A \setminus B) = P(A) - P(A \cap B)$
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- A és B független, ha $P(A \cap B) = P(A) * P(B)$

1.2.1. Valószínűségi mező

- Véges valószínűségi mező:
  - $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$
  - $p_i = P(\omega_i)$
  - $P(A) = P(\bigcup_{\omega_i \in A} \omega_i) = \sum_{\omega_i \in A} p_i$
- Klasszikus valószínűségi mező:
  - $p_i = \frac{1}{n}$ (azonos valószínűségű elemi események)
  - $P(A) = \frac{k}{n}$ ahol k az A eseményszáma
  - Azaz $P(A) = \frac{\text{Kedvező esetek száma}}{\text{Összes eset száma}}$

1.2.2. Mintavétel

- N termékből M selejtes, n elemű a minta
- A: pontosan k selejtes van a mintában
- Visszatevéses mintavétel:
  - $P(A) = \binom{n}{k} (\frac{M}{N})^k (1 - \frac{M}{N})^{n-k}$
  - Ha $p = \frac{M}{N}$: $P(A) = \binom{n}{k} p^k (1-p)^{n-k}$
- Visszatevés nélküli mintavétel:
  - $P(A) = \binom{M}{k} \binom{N-M}{n-k} \ / \ \binom{N}{n}$

1.2.3. Feltételes valószínűség

- A esemény valószínűségét keressük; tudjuk, hogy B bekövetkezett
- Relatív gyakorisággal: $r_{A \cap B}$ / r_B
- $P(A|B) = \frac{P(A \cap B)}{P(B)}$ ahol $P(B) > 0$
- $P(A|B)P(B) = P(A,B) \quad (A \text{ és } B)$

1.2.4. Teljes eseményrendszer; tételek

- Teljes eseményrendszer: események halmaza, melyek egymást páronként kizárják és egyesítésük Ω.
  - $P(A_1) + P(A_2) + ... = 1$
- Teljes valószínűség tétele:
  - Legyen pozitív valószínűségű $B_1, B_2, ...$ teljes eseményrendszer
  - Legyen A tetszőleges esemény
  - Ekkor $P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + ...$
- Bayes tétel:
  - Legyen pozitív valószínűségű $B_1, B_2, ...$ teljes eseményrendszer
  - Legyen A tetszőleges pozitív valószínűségű esemény
  - $P(B_k|A) = \frac{P(A|B_k)P(B_k)}{\sum P(A|B_i)P(B_i)} = (\text{magyarázat}) = \frac{P(A \cap B_k)}{P(A)}$
  - Komplementerrel: $P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\overline{B})P(\overline{B})}$

1.3. Permutáció, variáció, kombináció

- Permutáció: n elem lehetséges sorrendje
- Kombináció (sorrend nem számít) és variáció (sorrend számít): n elemből k kiválasztása

-	Permutáció	Kombináció	Variáció
Ismétlés nélkül	$n!$	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$	$\frac{n!}{(n-k)!}$
Ismétléssel	k_i db egyezik: $\frac{n!}{k_1! * ... * k_r!}$	$\binom{n+k-1}{k}$	n^k

2. Előadás 2, 3: valószínűségi változók, eloszlások

2.1. Események függetlensége

- Nem összekeverni az egymást kizáró eseményekkel!
- A és B események függetlenek, ha $P(A \cap B) = P(A) * P(B)$
- Ha A és B függetlenek, akkor komplementereik is függetlenek
- Ha A és B diszjunktak, akkor csak triviális esetben függetlenek (P=0)
- Önmaguktól csak triviális események függetlenek
- Több páronkénti függetlenségből nem következik összevont függetlenség
- $A \subset B \implies$ csak akkor függetlenek, ha legalább az egyik triviális
- Példa függetlenségre: A az első, B a második kísérlet eredménye

2.2. Valószínűségi változók

- Számot rendel egy kísérlet eredményéhez (elemi eseményhez)
- $X : \Omega \to \mathbb{R}$ függvény, hogy $\forall a : \{\omega | X(\omega) < a\} \in \mathcal{A}$
- Eloszlásfüggvény: $F_X(x) = P(X < x)$
  - monoton növekvő, balról folytonos
  - $\lim\limits_{x \to -\infty} F_X(x) = 0$ és $\lim\limits_{x \to +\infty} F_X(x) = 1$ (tehát $0 \leq F_X(x) \leq 1$)
- $P(a \leq X < b) = F_X(b) - F_X(a)$
  $P(a < X \leq b) = F_X(b + 0) - F_X(a + 0)$
- Indikátorváltozó eloszlásfüggvénye: $F_X(x) = \begin{cases} 0, & \text{ha } x \leq 0 \\ 1 - p, & \text{ha } 0 < x \leq 1 \\ 1, & \text{ha } 1 < x \end{cases}$
- X folytonos eloszlású, ha eloszlásfüggvénye folytonos, pl:
  Egyenletes eloszlás $[a, b]$ intervallumon: $F_X(x) = \begin{cases} 0, & \text{ha } x \leq a \\ \frac{x-a}{b-a}, & \text{ha } a < x \leq b \\ 1, & \text{ha } b < x \end{cases}$
- Exponenciális eloszlás (λ paraméter): $F_X(x) = \begin{cases} 0, & \text{ha } x \leq 0 \\ 1 - e^{-\lambda x}, & \text{ha } 0 < x \end{cases}$

2.2.1. Várható érték, szórás

- Legyen $X : \Omega \to \mathbb{R}$ valószínűségi változó
- Várható érték (mean): $E(X) = EX = \int_\Omega X dP$
- Szórásnégyzet (variance): $D^2(X) = E[(X - EX)]^2 = E(X^2) - E^2(X)$
  - Átrendezve: $E(X^2) = D^2(X) + E^2(X)$
- Szórás (standard deviation, σ): $D(X) = \sqrt{D^2(X)}$
- Várható érték lehet végtelen:
  $P(X = 2^k) = (1/2)^k \implies E(X) = 1 + 1 + 1 + ... = \infty$
- Legyen $E(X)$ és $E(Y)$ véges és $a, b \in \mathbb{R}$
  - Ekkor $E(aX + b) = aEX + b$
  - Ekkor $E(X + Y) = EX + EY$
- $X \geq 0$ és $E(X)$ véges $\implies E(X) \geq 0$
- $D^2(aX + b) = a^2 D^2(X) \quad$ (Tehát $D(aX) = |a|D(X)$)
- $D^2(X + Y) = D^2(X) + D^2(Y) + 2cov(X, Y)$
- Ha X nemnegatív egész értékű: $E(X) = P(X \geq 1) + P(X \geq 2) + ...$
  - pl. dobókocka: $E(X) = 3.5 = \frac{6}{6} + \frac{5}{6} + \frac{4}{6} + \frac{3}{6} + \frac{2}{6} + \frac{1}{6}$
- Attól még, hogy $E(X)$ véges, $D^2(X)$ nem feltétlenül véges:
  pl.: $P(X = k) = c/k^3$ ha eloszlás, attól még $E(X^2) = c(1/1 + 1/2 + ...)$
- X, Y függetlenek $\implies E(XY) = E(X)E(Y)$
- Z: p valószínűséggel X, $1 - p$ valószínűséggel Y:
  $E(Z) = p * E(X) + (1 - p) * E(Y)$

2.2.2. Valószínűségi változó: egyebek

- Vektor valószínűségi változó: ha $\underline{X}$ koordinátái függetlenek:
  $F_{\underline{X}}(\underline{z}) = P(X_1 < z_1, X_2 < z_2, ...) = F_1(z_1)F_2(z_2)...$
- *iid* (*i.i.d.*) jelentése: független és azonos eloszlású

2.3. Diszkrét valószínűségi változók

- Értékkészlet véges vagy megszámlálhatóan végtelen
  - Véges, vagy megszámlálhatóan végtelen valószínűségi mezőn minden valószínűségi változó diszkrét
- Legyen $X = (x_1, ..., x_n)$, ekkor $p_i = P(X = x_i)$ és $\sum_{i=1}^{n} P(X = x_i) = 1$
- Elfajult eloszlás: $\forall\omega : X(\omega) = c,\ \ P(X = c) = 1$
- Ha X diszkrét valószínűségi változó: $\{\omega | X(\omega) = x_i\}$ teljes eseményrendszer
- $X + Y$, mint diszkrét valószínűségi változó
  - $P(X + Y = z) = \sum_y P(X = z - y,\ Y = y)$
  - Ha X és Y függetlenek:
    $P(X + Y = z) = \sum_{x_i + y_j = z} P(X = x_i)P(Y = y_j)$

2.3.1. Valószínűségi változók függetlensége

- X és Y függetlenek, ha $\forall i, k$:
  $P(\{X = x_i\} \cap \{Y = Y_k\}) = P(X = x_i) * P(Y = y_k)$
  - (Azaz, ha X-hez és Y-hoz tartozó teljes eseményrendszerek függetlenek)
- Elfajult eloszlású valószínűségi változó miden val. változótól független
- Önmagától csak az elfajult eloszlású valószínűségi változó független

2.3.2. Várható érték, szórás

- $E(X) = \sum x_k P(X = x_k)$ ha a végtelen összeg abszolút konvergens
  - $E(X^2)$: az x_k értéket négyzetre kell emelni
  - Dobókocka: $E(X) = \frac{1}{6} * \sum_{i=1}^{6} i = 3.5$
    És a szórásnégyzet: $D^2(X) = 91/6 - 3.5^2 = 105/36$
- Elfajult eloszlás várható értéke: $E(X) = cP(X = c) = c$
- Egyenletes eloszlás várható értéke: számok számtani közepe
- Függvény várható értéke: $E(g(x)) = \sum P(X = x_i)g(x_i)$

2.3.3. Nevezetes diszkrét eloszlások

Név	Jelölés	Értékek	Valószínűségek	EX	D^2X
Indikátor vagy Karakterisztikus	$\text{Ind}(p) = \text{Bin}(1, p)$	$0, 1$	$P(X=1) = p$, $P(X=0) = 1-p$	p	$p(1-p)$
Binomiális	$\text{Bin}(n, p)$	$0, 1, ..., n$	$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Poisson	$\text{Poisson}(\lambda)$	$0, 1, ...$	$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$	λ	λ
Geometriai (Pascal)	$\text{Geo}(p) = \text{NegBin}(1, p)$	$1, 2, ...$	$P(X=k) = p(1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
Negatív binomiális	$\text{NegBin}(n, p)$	$n, n+1, ...$	$P(X=k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}$	$\frac{n}{p}$	$\frac{n(1-p)}{p^2}$
Hipergeometriai	$\text{HiperGeo}(N, M, n)$	$0, 1, ..., n$	$P(X=k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$	$n\frac{M}{N}$	$n\frac{M}{N}\left(1-\frac{M}{N}\right)\left(1-\frac{n-1}{N-1}\right)$

- Indikátor: p valószínűségű esemény bekövetkezik-e vagy sem
- Binomiális: visszatevéses mintavétel
- Poisson: időben lejátszódó folyamatnál adott $[a, b)$ intervallumba eső események száma ($X_{a,b}$) éppen $\lambda(b-a)$ paraméterű Poisson eloszlású
  - ha homogén az esemény: λ nem változik az idővel
  - ha utóhatás nélküli az esemény: a<b<c $\implies$ $X_{a,b}$ és $X_{b,c}$ függetlenek
  - nemelfajuló: $0 < P(X_{a,b} = 0) < 1$
- Geometriai (Pascal): hányadikra következik be először egy p valószínűségű
- Negatív binomiális: hányadikra következik be n. alkalommal egy p valószínűségű esemény
- Hipergeometriai: visszatevés nélküli mintavétel

2.4. Gyakorlat

- Diszkrét eloszlást a $P(X = ...) = ...$ értékekkel adhatjuk meg

3. Előadás 3: kovariancia, korreláció

3.1. Kovariancia

- $cov(X, Y) = E(\ (X - E(X)) * (Y - E(Y))\ ) =$
  $= E(\ XY - XE(Y) - YE(X) - E(X)E(Y)\ ) =$
  $= E(XY) - E(X)E(Y)$
- X, Y függetlenek $\implies$ $cov(X, Y) = 0$ (Visszafelé nem igaz)
- Szimmetrikus: $cov(X, Y) = cov(Y, X)$
- Kapcsolat szórásnégyzethez: $cov(X, X) = D^2(X)$
- Skálafüggő: $cov(a * X, b * Y) = ab * cov(X, Y)$

3.2. Korrelációs együttható

- Változók közötti lineáris kapcsolat erősségét méri:
  $R^2 \approx 1$ az erős kapcsolatot, $R^2 \approx 0$ a gyenge kapcsolatot jelenti
- $R(X, Y) = \frac{cov(X,Y)}{D(X)D(Y)}$
- X, Y függetlenek $\implies$ $R(X, Y) = 0$ (Visszafelé nem igaz)
- X vagy Y elfajult $\implies$ $R(X, Y) = 0$
- $R(X, aX + b) = 1$ ha $a > 0$

4. Előadás 4: abszolút folytonos eloszlások

- Sűrűségfüggvény eloszlásfüggvényből
  - Legyen F eloszlásfüggvény és létezzen f, hogy $F(z) = \int_{-\infty}^{z} f(t)dt$
  - Ekkor f sűrűségfüggvény és F abszolút folytonos
  - f nem egyértelmű $\implies$ legegyszerűbb szakaszonként folytonos változatát vesszük
- Eloszlásfüggvény sűrűségfüggvényből: (fentieken túl)
  - Ha $X \geq 0$, akkor $F(0) = 0$ (ebből integrálás konstans szerezhető)
- $P(a < X \leq b) = P(a \leq X < b) = F(b) - F(a)$
- Várható érték: $E(X) = \int_{-\infty}^{+\infty} xf(x)dx$ (ha az integrál létezik)
  - Függvény várható értéke: $E(g(x)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$

4.1. Sűrűségfüggvény

- Legyen X abszolút folytonos eloszlású
- $f(x) = F'(x)$ (esetleg néhány pont kivételével)
- $f(x) \geq 0$ és $\int_{-\infty}^{+\infty} f(x)dx = 1$
  - Megengedett: $\exists x : f(x) > 1$
- F folytonos, tehát $\forall x : P(X = x) = 0$

4.2. Standard normális eloszlás

- Sűrűségfüggvény: $f(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$
  - Haranggörbe, csúcs $x = 0$-ban
- Normális eloszlás standardizálása:
  - Legyen $X \sim N(m, \sigma^2)$
  - Ekkor $\frac{X-m}{\sigma} \sim N(0, 1)$

4.3. Nevezetes abszolút folytonos eloszlások

Név (paraméterek)	Értékek	Eloszlásfüggvény (F)	Sűrűségfüggvény (f)	EX	D^2X
Standard normális $N(0,1)$	$(-\infty,\infty)$	$\Phi(x)=$ táblázatban	$\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \quad x\in\mathbb{R}$	0	1
Normális N(m,σ^2)	$(-\infty,\infty)$	visszavezethető $\Phi(x)$-re	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}} \quad x\in\mathbb{R}$	m	σ^2
Egyenletes E$[a,b]$	$[a,b]$	$\begin{cases}0 & \text{ha } x\le a\\ \frac{x-a}{b-a} & \text{ha } a<x\le b\\ 1 & \text{ha } b<x\end{cases}$	$\begin{cases}\frac{1}{b-a} & \text{ha } a<x\le b\\ 0 & \text{különben}\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponenciális Exp(λ)	$(0,\infty)$	$\begin{cases}1-e^{-\lambda x} & \text{ha } x\ge 0\\ 0 & \text{különben}\end{cases}$	$\begin{cases}\lambda e^{-\lambda x} & \text{ha } x\ge 0\\ 0 & \text{különben}\end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma $\Gamma(\alpha,\lambda)$	$(0,\infty)$	nincs zárt elemi képlet	$\begin{cases}\frac{1}{\Gamma(\alpha)}\lambda^\alpha x^{\alpha-1}e^{-\lambda x} & \text{ha } x\ge 0\\ 0 & \text{különben}\end{cases}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$

- A "geometriai modell" az egyenletes eloszlást jelent
- Chi-négyzet eloszlás (Ki^2): $X \sim \chi^2_n$ ha $X = \sum Z_i$ ahol $Z_i \sim N(0,1)$
- t eloszlás: $X \sim t_n$ ha $X = \frac{Z}{\sqrt{Y_n/n}}$ ahol $Z \sim N(0,1)$ és $Y_n \sim \chi^2_n$
- F eloszlás: $X \sim F_{n_1,n_2}$ ha $X = \frac{U_1/n_1}{U_2/n_2}$ ahol $U_i \sim \chi^2_{n_i}$

4.4. Függvény eloszlása

- X egy valószínűségi változó $\implies$ $g(X)$ is az
- X folytonosságából NEM következik $g(X)$ folytonossága
- Legyen X abszolút folytonos
- Legyen $g:\mathbb{R}\to\mathbb{R}$, folytonosan differenciálható, $g'\neq 0$
- $Y = g(x)$ eloszlásfüggvénye:
  $$F_Y(y) = F_{g(x)}(y) = \begin{cases} F_X(g^{-1}(y)) & \text{ha } g \text{ szig. mon. nő} \\ 1 - F_X(g^{-1}(y)) & \text{ha } g \text{ szig. mon. csökken} \end{cases}$$
  - Példa: $F_{aX+b} = F_X((z-b)/a)$ ha $a>0$
  - Példa: $F_{aX+b} = 1 - F_X((z-b)/a)$ ha $a<0$
- $Y = g(X)$ sűrűségfüggvénye:
  $f_Y(y) = f_{g(x)}(y) = F_X(g^{-1}(y)) * |(g^{-1}(y))'| = f_X(g^{-1}(y)) \;/\; |g'(g^{-1}(y))|$
  Lényeg: $f_{g(x)}(y) = f_X(g^{-1}(y)) \;/\; |g'(g^{-1}(y))|$

4.5. Együttes eloszlásfüggvények elnevezései

- Együttes eloszlásfüggvény: $F_{X,Y}(x,y) = P(X < x, Y < y)$
- Peremeloszlásfüggvények: $F_X(x) = P(X < x)$, $F_Y(y) = P(Y < y)$
- Együttes sűrűségfüggvény: $f_{X,Y}(x,y)$
- Peremsűrűségfüggvények: $f_X(x)$, $f_Y(y)$

4.6. Együttes eloszlásfüggvények kapcsolata

- $F_X(x) = \lim\limits_{y\to\infty} F_{X,Y}(x,y)$
- $F_{X,Y}(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u,v)dudv$
- $f_{X,Y}(x,y) = \partial_y \partial_x F_{X,Y}(x,y) = \partial_x \partial_y F_{X,Y}(x,y)$
- $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dxdy = 1$
- $F_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy$
- X, Y függetlenek $\Leftrightarrow F_{X,Y}(x,y) = F_X(x) * F_Y(y)$
- X, Y függetlenek $\Leftrightarrow f_{X,Y}(x,y) = f_X(x) * f_Y(y)$
- Diszkrét esetben:
  X, Y függetlenek $\Leftrightarrow P(X = x, Y = y) = P(X = x) * P(Y = y)$
- X, Y függetlenek $\Leftrightarrow E(XY) = E(X) * E(Y)$

5. Előadás 5: máshová írtam be az anyagot

6. Előadás 6: egyenlőtlenségek, aszimptotikus tulajdonságok

6.1. Markov-típusú egyenlőtlenségek

- Legyen $X \geq 0$ val. változó, legyen $g : \mathbb{R}_+ \to \mathbb{R}_+$ monoton növő
- Ekkor $\epsilon > 0$ esetén $P(X \geq \epsilon) \leq \frac{E(g(X))}{g(\epsilon)}$
- Bizonyítás: $E(g(X)) \geq g(\epsilon)P(X \geq \epsilon)$ mert $X \geq \epsilon \implies g(X) \geq g(\epsilon)$

6.1.1. Alkalmazásai

- Gyakorlatban nem kellően pontos becslést adnak, konkrét eloszlás ismeretével pontosabb becslés szerezhető
- Markov egyenlőtlenség:
  - $g(x) = x$ és $X \geq 0$ val. változó esete, ekkor $P(X \geq \epsilon) \leq \frac{E(X)}{\epsilon}$
  - Máshogy: $P(X \geq k * E(X)) \leq \frac{E(X)}{k*E(X)} = \frac{1}{k}$
- Csebisev egyenlőtlenség:
  - $g(x) = x$ és X helyett $(X - EX)^2$ esete
  - $P((X-EX)^2 \geq \epsilon^2) \leq \frac{E(X-EX)^2}{\epsilon^2} \implies P(|X-EX| \geq \epsilon) \leq \frac{D^2(X)}{\epsilon^2}$

6.2. Konvergencia típusai

- Legyen X_i val. változók sorozata, ekkor $X_n \to (n \to +\infty) \to X$...
- ... 1 valószínűséggel: $P(\{\omega : X_n(\omega) \to (n \to +\infty) \to X(\omega)\}) = 1$
- ... gyengén: ha eloszlásfüggvényeikre $F_n(x) \to (n \to +\infty) F(x)$
  - F minden folytonossági pontjában

6.3. Nagy számok törvényei

- Legyen $X_1, X_2, ..., X_n$ független, azonos eloszlású valószínűségi változók
- Legyen a várható értékük (ami azonos, hiszen azonos az eloszlás) véges
- $\frac{X_1+...+X_n}{n} \to m = E(X_i) = E(X_1)$ ha $n \to +\infty$ (1 valószínűséggel)

6.4. Centrális határelosztás tétel

- Legyen $X_1, X_2, ..., X_n$ független, azonos eloszlású valószínűségi változók
- Legyen a szórásnégyzetük ($\sigma^2 = D^2(X_i)$) véges
  - Nulla sem lehet, nem? Hiszen osztunk vele
- Szórásuk, várható értékük ($m = E(X_i)$) egyezik, mert azonos az eloszlásuk
- $\frac{X_1+...+X_n-n*m}{\sqrt{n}*\sigma} \to N(0,1)$ ha $n \to +\infty$ (gyengén)
  - Máshogy: $P(\frac{X_1+...+X_n-n*m}{\sqrt{n}*\sigma} < x) \to \Phi(x)$
  - Máshogy: n kísérlet összegének szórása $\frac{1}{\sqrt{n}}$-nel arányos
- Legyen $\overline{X}$ átlag: $\overline{X} = \frac{X_1+...X_n}{n}$, ekkor $\overline{X} \sim N(\mu, (\frac{\sigma}{\sqrt{n}})^2)$ $(E(X_i) = \mu)$
- Felhasználás: 0.9 valószínűséggel legyen $X < 99.5$ ahol $X \sim N(100, 3^2)$
  - $\overline{X} \sim N(100, (\frac{3}{\sqrt{n}})^2)$ a tétel szerint
  - Legyen S std normál: $S \sim N(0,1)$
  - $0.9 = P(\overline{X} < 99.5) = P(S < \frac{99.5-100}{3/\sqrt{n}}) = \Phi(\frac{99.5-100}{3/\sqrt{n}})$
  - Megoldás: $0.9 = \Phi(\frac{99.5-100}{3/\sqrt{n}}) = \Phi(\frac{\sqrt{n}}{-6}) = \Phi(\frac{\sqrt{n}}{6})$
    $n = (qnorm(0.9) * 6)^2 \approx 59.12$

7. R jegyzet

- choose(x, y); `cat("X:", x, "\n", "valami"))`; round(x, 4)
- diszkrét
  - ppois, pbinom, phyper
  - sum(dbinom(0:4, 12, 0.2)) = pbinom(4, 12, 0.2)
  - ```
    barplot(dbinom(0:6, 6, 527/1000),
            col="red",
            xlab="k értéke",
            ylab="Valószínűség",
            main="Binomiális eloszlás",
            names.arg = 0:6)
    ```
- abszolút folytonos
  - pnorm
  - d{dist}(a, ...) = f(a)
  - p{dist}(a, ...) = F(X < a)
  - q{dist}(p, ...) = a $\implies$ F(X < a) = p
  - ```
    valx <- seq(90, 110, 0.01)
    valy <- pnorm(valx, 100, 2)
    plot(valx, valy, col = "red", type = "l", lwd = 2,
         xlab = "x", ylab = "Fi(x)",
         main = "Eloszlásfüggvény")
    abline(h = c(0, 1), lty = 2)
    ```
- plot-olás
  - lwd: line width (2-3 szokott lenni)
  - lty: line type, (1=folytonos, 2=dashed, 3=dotted, ...)
  - ```
    my_plot <- function(f, from_x, to_x) {
      x <- seq(from_x, to_x, 1 / 100)
      y <- sapply(x, f)
      plot(x, y, type = "l", col = "black", lwd = 3)
      abline(h = c(0, 1))
    }
    ```
    ```