Vizsgasegédlet

leíró és matematikai statisztika vizsgára
2016/2017 tavaszi félév

1. Nevezetes eloszlások néhány jellemzője

Jelölése	Eloszlása	EX	D^2X
Ind(p)	$P(X=1)=p$ $P(X=0)=1-p$	p	$p(1-p)$
Hipgeo(N,M,n)	$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$ $k=0,1,\dots,\min(n,M)$	$n\frac{M}{N}$	$n\frac{M}{N}\left(1-\frac{M}{N}\right)\left(1-\frac{n-1}{N-1}\right)$
Bin(n,p)	$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ $k=0,1,\dots,n$	np	$np(1-p)$
Geo(p)	$P(X=k)=p(1-p)^{k-1}$ $k=1,2,\dots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
NegBin(n,p)	$P(X=k)=\binom{k-1}{n-1}p^n(1-p)^{k-n}$ $k=n,n+1,\dots$	$\frac{n}{p}$	$\frac{n(1-p)}{p^2}$
Poi(λ)	$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}\quad k=0,1,\dots$	λ	λ

Jelölése	Eloszlásfüggvény	Sűrűségfüggvény	EX	D^2X
$E(a,b)$	$\begin{cases}0 & \text{ha } x\le a\\ \frac{x-a}{b-a} & \text{ha } a<x\le b\\ 1 & \text{ha } b<x\end{cases}$	$\begin{cases}\frac{1}{b-a} & \text{ha } a<x\le b\\ 0 & \text{különben}\end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$N(0,1^2)$	$\Phi(x)=\dots$	$\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\quad x\in\mathbb{R}$	0	1
$N(m,\sigma^2)$	$\dots$	$\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}}\quad x\in\mathbb{R}$	m	σ^2
Exp(λ)	$\begin{cases}1-e^{-\lambda x} & \text{ha } x\ge 0\\ 0 & \text{különben}\end{cases}$	$\begin{cases}\lambda e^{-\lambda x} & \text{ha } x\ge 0\\ 0 & \text{különben}\end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\Gamma(\alpha,\lambda)$	$\dots$	$\begin{cases}\frac{1}{\Gamma(\alpha)}\lambda^\alpha x^{\alpha-1}e^{-\lambda x} & \text{ha } x\ge 0\\ 0 & \text{különben}\end{cases}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$

Eloszlás neve	Jelölése	Eloszlásfüggvény	Sűrűségfüggvény	EX	D^{2}X
Cauchy	$Cauchy(a,b)$ $a\in\mathbb{R}, b>0$	$\frac{1}{\pi}\text{arctg}\left(\frac{x-a}{b}\right)+\frac{1}{2}$	$\frac{1}{\pi b\left[1+\left(\frac{x-a}{b}\right)^2\right]}\quad x\in\mathbb{R}$	$\nexists$	$\nexists$
Pareto*	$Pareto(\alpha,\beta)$ $\alpha,\beta>0$	$\begin{cases}1-\left(\frac{\beta}{x}\right)^\alpha & \text{ha } x\ge\beta\\ 0 & \text{ha } x<\beta\end{cases}$	$\begin{cases}\frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1} & \text{ha } x\ge\beta\\ 0 & \text{ha } x<\beta\end{cases}$	$\frac{\alpha\beta}{\alpha-1}$	$\frac{\beta^2\alpha}{(\alpha-1)^2(\alpha-2)}$

* A Pareto-eloszlásnak akkor van véges várható értéke a képletnek megfelelően, ha $\alpha>1$, szórásnégyzete pedig akkor, ha $\alpha>2$.

Eloszlás neve	Jelölése	Sűrűségfüggvény	EX	D^{2}X
Lognormális	LN(m,σ^2) $m\in\mathbb{R}, \sigma>0$	$\begin{cases}\frac{1}{x\sqrt{2\pi}\sigma}e^{-\frac{(\log x-m)^2}{2\sigma^2}} & \text{ha } x\le 0\\ 0 & \text{ha} x<0\end{cases}$	$e^{m+\sigma^2/2}$	$(e^{\sigma^2}-1)e^{2m+\sigma^2}$
Gamma	$\Gamma(\alpha,\lambda)$ $\alpha,\lambda>0$	$\begin{cases}\frac{1}{\Gamma(\alpha)}\lambda^\alpha e^{-\lambda x}x^{\alpha-1} & \text{ha } x\ge 0\\ 0 & \text{ha } x<0\end{cases}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
Béta	$Beta(\alpha,\beta)$ $\alpha,\beta>0$	$\begin{cases}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1} & x\in[0,1]\\ 0 & \text{különben}\end{cases}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Khí-négyzet	$\chi^2_k\quad k\in\mathbb{N}$	$\frac{1}{2^{k/2}\Gamma(k/2)}x^{k/2-1}e^{-x/2}\quad x\in\mathbb{R}$	k	$2k$
Student	$t_\nu\quad \nu>0$	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$	0 (ha $\nu>1$)	$\frac{\nu}{\nu-2}$ (ha $\nu>2$)
Fisher	F_{d_1,d_2} $d_1,d_2>0$	$\frac{\Gamma\left(\frac{d_1+d_2}{2}\right)}{\Gamma\left(\frac{d_1}{2}\right)\Gamma\left(\frac{d_2}{2}\right)}\left(\frac{d_1}{d_2}\right)^{\frac{d_1}{2}}x^{\frac{d_1}{2}-1}\left(1+\frac{d_1}{d_2}x\right)^{-\frac{d_1+d_2}{2}}$	$\frac{d_2}{d_2-2}$ (ha $d_2>2$)	$\frac{2d_2^2(d_1+d_2-2)}{d_1(d_2-2)^2(d_2-4)}$ (ha $d_2>2$)

2. Koncentráció

$$G=\frac{1}{n(n-1)}\sum_{i=1}^k\sum_{j=1}^k f_if_j|x_i-x_j| \qquad L=\frac{G}{2\overline{x}} \qquad HI=\sum_{i=1}^k z_i^2$$

3. Alapstatisztikák számítása tapasztalati mintából

$$\text{Mo}=x_{mo,a}+\frac{d_a}{d_a+d_f}\cdot h_{mo}$$

$$\text{Me}=\begin{cases}x^*_{\frac{n+1}{2}}, & \text{ha } n \text{ páratlan}\\ \frac{x^*_{\frac{n}{2}}+x^*_{\frac{n}{2}+1}}{2}, & \text{ha } n \text{ páros}\end{cases} \qquad \text{Me}=x_{i,a}+\frac{\frac{n}{2}-f'_{i-1}}{f_i}\cdot h_i$$

$$q_y=x^*_e+t(x^*_{e+1}-x^*_e) \qquad q_y=x_{i,a}+\frac{s-f'_{i-1}}{f_i}h_i$$

$R=x^*_n-x^*_1$ (R=range)

$IQR=Q_3-Q_1$

$$s_n=\sqrt{\frac{\sum\limits_{i=1}^n(x_i-\overline{x})^2}{n}} \qquad s_n=\sqrt{\frac{\sum\limits_{i=1}^k f_i(x_i-\overline{x})^2}{n}}$$

$$s^*_n=\sqrt{\frac{\sum\limits_{i=1}^n(x_i-\overline{x})^2}{n-1}} \qquad s^*_n=\sqrt{\frac{\sum\limits_{i=1}^k f_i(x_i-\overline{x})^2}{n-1}}$$

$V=\frac{s^*_n}{\overline{x}}$ vagy $V=\frac{s_n}{\overline{x}}$

4. Becsléselmélet

$$f_n(x)=\frac{1}{n\cdot h_n}\sum_{i=1}^n K\left(\frac{x-X_i}{h_n}\right)$$

$$D^2_\vartheta(T(\mathbf{X}))\ge\frac{(g'(\vartheta))^2}{I_n(\vartheta)}$$

$$\overline{X}\pm u_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \qquad \overline{X}\pm t_{n-1,\frac{\alpha}{2}}\frac{S^*_n}{\sqrt{n}} \qquad \left[\frac{(n-1)\cdot(S^*_n)^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}}\ ,\ \frac{(n-1)\cdot(S^*_n)^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$$

$$\left[-\infty\ ,\ \overline{X}+u_\alpha\frac{\sigma}{\sqrt{n}}\right] \qquad \left[-\infty\ ,\ \overline{X}+t_{n-1,\alpha}\frac{S^*_n}{\sqrt{n}}\right] \qquad \left[-\infty\ ,\ \frac{(n-1)\cdot(S^*_n)^2}{\chi^2_{n-1,\alpha}}\right]$$

5. Hipotézisvizsgálat

$$u=\sqrt{n}\frac{\overline{X}-m_0}{\sigma} \overset{H_0 \text{ esetén}}{\sim} N(0,1) \qquad t=\sqrt{n}\frac{\overline{X}-m_0}{S_n^*} \overset{H_0 \text{ esetén}}{\sim} t_{n-1}$$

$$u=\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}} \overset{H_0 \text{ esetén}}{\sim} N(0,1) \qquad t=\sqrt{\frac{nm}{n+m}}\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{(n-1)(S_1^*)^2+(m-1)(S_2^*)^2}{n+m-2}}} \overset{H_0 \text{ esetén}}{\sim} \mathrm{t}_{n+m-2}$$

$$h=\frac{(n-1)(S_n^*)^2}{\sigma_0^2} \overset{H_0 \text{ esetén}}{\sim} \chi^2_{n-1} \qquad F=\frac{(S_1^*)^2}{(S_2^*)^2} \overset{H_0 \text{ esetén}}{\sim} F_{n-1,m-1}$$

$$T_n=\sum_{i=1}^{r}\frac{(N_i-np_i)^2}{np_i} \xrightarrow[n\to\infty]{H_0 \text{ esetén}} \chi^2_{r-1-s} \text{ eloszlásban}$$

$$T_{n,m}=nm\sum_{i=1}^{r}\frac{\left(\frac{N_i}{n}-\frac{M_i}{m}\right)^2}{N_i+M_i} \xrightarrow[n\to\infty]{H_0 \text{ esetén}} \chi^2_{r-1} \text{ eloszlásban}$$

$$T_n=n\left(\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{N_{i,j}^2}{N_{i\bullet}N_{\bullet j}}-1\right) \xrightarrow[n\to\infty]{H_0 \text{ esetén}} \chi^2_{(r-1)(s-1)} \text{ eloszlásban}$$

$$T_n=n\cdot\frac{(N_{11}N_{22}-N_{12}N_{21})^2}{N_{1\bullet}N_{2\bullet}N_{\bullet 1}N_{\bullet 2}} \xrightarrow[n\to\infty]{H_0 \text{ esetén}} \chi^2_1 \text{ eloszlásban}$$

6. Lineáris modell

$$\widehat{\mathbf{b}}=(X^TX)^{-1}X^T\mathbf{y} \qquad \widehat{\mathbf{y}}=X\widehat{\mathbf{b}} \qquad \widehat{\varepsilon}=\mathbf{y}-\widehat{\mathbf{y}}$$

$$\text{RNÖ}=\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2 \qquad \text{NÖ}=\sum_{i=1}^{n}(y_i-\overline{y})^2$$

$$R^2=1-\frac{\text{RNÖ}}{\text{NÖ}}$$

$$R^2_{\text{adj}}=1-\frac{n-1}{n-r-1}\frac{SSR}{SS} \qquad AIC=2(p+1)-2\log\widehat{L}$$

$$F=\frac{\frac{S_B-\text{RNÖ}}{q}}{\frac{\text{RNÖ}}{n-p}} \qquad S_B=\min_{\mathbf{b}:\ B\mathbf{b}=\mathbf{0}}\|\mathbf{y}-X\mathbf{b}\|^2$$

7. Asszociáció

$$C=\sqrt{\frac{\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(f_{ij}-f^*_{ij})^2}{f^*_{ij}}}{n\cdot(\min(r,s)-1)}}, \quad \text{ahol } f^*_{ij}=\frac{f_{i\bullet}\cdot f_{\bullet j}}{n} \qquad Y=\frac{f_{11}\cdot f_{22}-f_{12}\cdot f_{21}}{f_{11}\cdot f_{22}+f_{12}\cdot f_{21}}$$

8. ANOVA

$$SST=\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_{\bullet\bullet})^2 \qquad SSK=\sum_{i=1}^{p}n_i(\overline{y}_{i\bullet}-\overline{y}_{\bullet\bullet})^2 \qquad SSB=\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_{i\bullet})^2$$

$$SSI=m\sum_{i=1}^{p}\sum_{j=1}^{r}(\overline{y_{ij\bullet}}-\overline{y_{i\bullet\bullet}}-\overline{y_{\bullet j\bullet}}+\overline{y_{\bullet\bullet\bullet}})^2 \qquad H^2=\frac{SSK}{SST}$$

$$MSK=\frac{SSK}{p-1} \qquad MSI=\frac{SSI}{(p-1)(r-1)} \qquad MSB=\frac{SSB}{n-p} \qquad F=\frac{\frac{SSK}{p-1}}{\frac{SSB}{n-p}}$$

$$\overline{y}_{i\bullet}\pm t_{n-p;\alpha/2}\sqrt{\frac{MSB}{n_i}} \qquad \overline{y}_{i\bullet}-\overline{y}_{j\bullet}\pm t_{n-p;\alpha/2}\sqrt{MSB}\sqrt{\frac{n_i+n_j}{n_in_j}}$$

9. Néhány kvantilis

Jelölje $u_\alpha=\Phi^{-1}(1-\alpha)$

$u_{0,01}=2,326$	$u_{0,02}=2,054$	$u_{0,03}=1,881$	$u_{0,04}=1,751$	$u_{0,05}=1,645$

Jelölje $t_{n;\alpha}=F^{-1}_{t_n}(1-\alpha)$

$t_{4;0,01}=3,747$	$t_{4;0,02}=2,999$	$t_{4;0,03}=2,601$	$t_{4;0,04}=2,333$	$t_{4;0,05}=2,132$
$t_{5;0,01}=3,365$	$t_{5;0,02}=2,757$	$t_{5;0,03}=2,422$	$t_{5;0,04}=2,191$	$t_{5;0,05}=2,015$
$t_{6;0,01}=3,143$	$t_{6;0,02}=2,612$	$t_{6;0,03}=2,313$	$t_{6;0,04}=2,104$	$t_{6;0,05}=1,943$
$t_{7;0,01}=2,998$	$t_{7;0,02}=2,517$	$t_{7;0,03}=2,241$	$t_{7;0,04}=2,046$	$t_{7;0,05}=1,895$
$t_{8;0,01}=2,896$	$t_{8;0,02}=2,449$	$t_{8;0,03}=2,189$	$t_{8;0,04}=2,004$	$t_{8;0,05}=1,860$
$t_{9;0,01}=2,821$	$t_{9;0,02}=2,398$	$t_{9;0,03}=2,150$	$t_{9;0,04}=1,973$	$t_{9;0,05}=1,833$
$t_{10;0,01}=2,764$	$t_{10;0,02}=2,359$	$t_{10;0,03}=2,12$	$t_{10;0,04}=1,948$	$t_{10;0,05}=1,812$
$t_{11;0,01}=2,718$	$t_{11;0,02}=2,328$	$t_{11;0,03}=2,096$	$t_{11;0,04}=1,928$	$t_{11;0,05}=1,796$
$t_{12;0,01}=2,681$	$t_{12;0,02}=2,303$	$t_{12;0,03}=2,076$	$t_{12;0,04}=1,912$	$t_{12;0,05}=1,782$
$t_{13;0,01}=2,65$	$t_{13;0,02}=2,282$	$t_{13;0,03}=2,06$	$t_{13;0,04}=1,899$	$t_{13;0,05}=1,771$

Jelölje $\chi^2_{n;\alpha}=F^{-1}_{\chi^2_n}(\alpha)$

$\chi^2_{1;0,99}=6,635$	$\chi^2_{1;0,98}=5,412$	$\chi^2_{1;0,97}=4,709$	$\chi^2_{1;0,96}=4,218$	$\chi^2_{1;0,95}=3,841$
$\chi^2_{2;0,99}=9,210$	$\chi^2_{2;0,98}=7,824$	$\chi^2_{2;0,97}=7,013$	$\chi^2_{2;0,96}=6,438$	$\chi^2_{2;0,95}=5,991$
$\chi^2_{3;0,99}=11,345$	$\chi^2_{3;0,98}=9,837$	$\chi^2_{3;0,97}=8,947$	$\chi^2_{3;0,96}=8,311$	$\chi^2_{3;0,95}=7,815$
$\chi^2_{4;0,99}=13,277$	$\chi^2_{4;0,98}=11,668$	$\chi^2_{4;0,97}=10,712$	$\chi^2_{4;0,96}=10,026$	$\chi^2_{4;0,95}=9,488$

Jelölje $F_{n;m;\alpha}=F^{-1}_{F_{n;m}}(\alpha)$. Ekkor $F_{n;m;\alpha}=\frac{1}{F_{m;n;1-\alpha}}$

$F_{4;4;0,99}=15,98$	$F_{4;4;0,98}=10,9$	$F_{4;4;0,97}=8,65$	$F_{4;4;0,96}=7,31$	$F_{4;4;0,95}=6,39$
$F_{4;5;0,99}=11,39$	$F_{4;5;0,98}=8,23$	$F_{4;5;0,97}=6,75$	$F_{4;5;0,96}=5,83$	$F_{4;5;0,95}=5,19$
$F_{4;6;0,99}=9,15$	$F_{4;6;0,98}=6,86$	$F_{4;6;0,97}=5,74$	$F_{4;6;0,96}=5,04$	$F_{4;6;0,95}=4,53$
$F_{5;4;0,99}=15,52$	$F_{5;4;0,98}=10,62$	$F_{5;4;0,97}=8,44$	$F_{5;4;0,96}=7,14$	$F_{5;4;0,95}=6,26$
$F_{5;5;0,99}=10,97$	$F_{5;5;0,98}=7,95$	$F_{5;5;0,97}=6,54$	$F_{5;5;0,96}=5,66$	$F_{5;5;0,95}=5,05$
$F_{5;6;0,99}=8,75$	$F_{5;6;0,98}=6,58$	$F_{5;6;0,97}=5,53$	$F_{5;6;0,96}=4,86$	$F_{5;6;0,95}=4,39$
$F_{5;7;0,99}=7,46$	$F_{5;7;0,98}=5,76$	$F_{5;7;0,97}=4,91$	$F_{5;7;0,96}=4,37$	$F_{5;7;0,95}=3,97$
$F_{5;5;0,99}=10,97$	$F_{5;5;0,98}=7,95$	$F_{5;5;0,97}=6,54$	$F_{5;5;0,96}=5,66$	$F_{5;5;0,95}=5,05$
$F_{6;4;0,99}=15,21$	$F_{6;4;0,98}=10,42$	$F_{6;4;0,97}=8,3$	$F_{6;4;0,96}=7,03$	$F_{6;4;0,95}=6,16$
$F_{6;5;0,99}=10,67$	$F_{6;5;0,98}=7,76$	$F_{6;5;0,97}=6,39$	$F_{6;5;0,96}=5,54$	$F_{6;5;0,95}=4,95$
$F_{6;6;0,99}=8,47$	$F_{6;6;0,98}=6,39$	$F_{6;6;0,97}=5,38$	$F_{6;6;0,96}=4,74$	$F_{6;6;0,95}=4,28$
$F_{6;7;0,99}=7,19$	$F_{6;7;0,98}=5,58$	$F_{6;7;0,97}=4,77$	$F_{6;7;0,96}=4,24$	$F_{6;7;0,95}=3,87$
$F_{6;8;0,99}=6,37$	$F_{6;8;0,98}=5,04$	$F_{6;8;0,97}=4,35$	$F_{6;8;0,96}=3,91$	$F_{6;8;0,95}=3,58$
$F_{7;5;0,99}=10,46$	$F_{7;5;0,98}=7,61$	$F_{7;5;0,97}=6,28$	$F_{7;5;0,96}=5,45$	$F_{7;5;0,95}=4,88$
$F_{7;6;0,99}=8,26$	$F_{7;6;0,98}=6,25$	$F_{7;6;0,97}=5,27$	$F_{7;6;0,96}=4,65$	$F_{7;6;0,95}=4,21$
$F_{7;7;0,99}=6,99$	$F_{7;7;0,98}=5,44$	$F_{7;7;0,97}=4,65$	$F_{7;7;0,96}=4,15$	$F_{7;7;0,95}=3,79$
$F_{7;8;0,99}=6,18$	$F_{7;8;0,98}=4,9$	$F_{7;8;0,97}=4,24$	$F_{7;8;0,96}=3,81$	$F_{7;8;0,95}=3,5$
$F_{8;6;0,99}=8,1$	$F_{8;6;0,98}=6,14$	$F_{8;6;0,97}=5,19$	$F_{8;6;0,96}=4,58$	$F_{8;6;0,95}=4,15$
$F_{8;7;0,99}=6,84$	$F_{8;7;0,98}=5,33$	$F_{8;7;0,97}=4,57$	$F_{8;7;0,96}=4,08$	$F_{8;7;0,95}=3,73$
$F_{8;8;0,99}=6,03$	$F_{8;8;0,98}=4,79$	$F_{8;8;0,97}=4,16$	$F_{8;8;0,96}=3,74$	$F_{8;8;0,95}=3,44$