

Sztochasztikus konvergencia, majdnem biztos konvergencia

Az X_n véletlen változó sorozat sztochasztikusan (valószínűségben) konvergál az X véletlen változóhoz ($X_n \xrightarrow{sz} X$), ha bármely ε -hoz van olyan N , hogy minden $n > N$ -re

$$P(|X_n - X| > \varepsilon) < \varepsilon.$$

Az X_n véletlen változó sorozat majdnem biztosan (1 valószínűséggel) konvergál az X véletlen változóhoz ($X_n \xrightarrow{mb} X$), ha

$$P\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1.$$

Diszkrét eloszlások

- Bernoulli-eloszlás (egyszerű alternatíva):

$$P(X = 1) = p, \quad P(X = 0) = q, \quad p + q = 1.$$

$$E(X) = p, \quad D^2(X) = p \cdot q, \quad G_X(s) = q + p \cdot s.$$

- Binomiális eloszlás (n független Bernoulli összege):

$$P(X = k) = \binom{n}{k} p^k q^{n-k}, \quad p + q = 1, \quad k = 0, 1, \dots, n.$$

$$E(X) = n \cdot p, \quad D^2(X) = n \cdot p \cdot q, \quad G_X(s) = (q + p \cdot s)^n.$$

- Poisson-eloszlás (binomiális eloszlás limesze, ha $n \rightarrow \infty$ és $p \cdot n = \lambda$):

$$P(X = k) = \frac{1}{k!} \lambda^k \cdot e^{-\lambda}, \quad \lambda > 0, \quad k = 0, 1, \dots$$

$$E(X) = \lambda, \quad D^2(X) = \lambda, \quad G_X(s) = e^{\lambda \cdot (s-1)}.$$

- Geometriai eloszlás (az egyszerű alternatíva független ismétléseinek száma az első 1-es megjelenéséig):

$$P(X = k) = p \cdot q^{k-1}, \quad p + q = 1, \quad k = 1, 2, \dots$$

$$E(X) = \frac{1}{p}, \quad D^2(X) = \frac{q}{p^2}, \quad G_X(s) = \frac{p \cdot s}{1 - q \cdot s}.$$

- Negatív binomiális eloszlás (r darab geometriai összege):

$$P(X = r + k) = \binom{k+r-1}{r-1} p^r q^k, \quad p + q = 1, \quad k = 0, 1, \dots$$

$$E(X) = \frac{r}{p}, \quad D^2(X) = \frac{r \cdot q}{p^2}, \quad G_X(s) = \left(\frac{p \cdot s}{1 - q \cdot s}\right)^r.$$

- Hipergeometrikus eloszlás (visszatevés nélküli mintavétel):

$$P(X = k) = \frac{\binom{M}{k} \cdot \binom{N-M}{n-k}}{\binom{N}{n}} M < N, \quad n \leq N, \quad k = 0, 1, \dots, n.$$

$$E(X) = n \cdot \frac{M}{N}, \quad D^2(X) = n \cdot \frac{M}{N} \cdot \left(1 - \frac{M}{N}\right) \cdot \left(1 - \frac{n-1}{N-1}\right).$$

Abszolút folytonos eloszlások:

- Normális (Gauss-) eloszlás:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad 0 < \sigma.$$

$$E(X) = m, \quad D^2(X) = \sigma^2 \quad \psi_X(t) = e^{i \cdot m \cdot t - \frac{\sigma^2}{2} t^2}.$$

- Lognormális eloszlás (e^X eloszlása, ahol X Gauss):

$$f_X(x) = \frac{1}{x \cdot \sqrt{2\pi}\sigma} e^{-\frac{(\ln x - m)^2}{2\sigma^2}} \quad 0 < \sigma.$$

$$E(X) = e^{m+\sigma^2/2}, \quad D^2(X) = e^{2m+\sigma^2} \cdot (e^{\sigma^2} - 1).$$

- Exponenciális eloszlás:

$$f_X(x) = \lambda \cdot e^{-\lambda \cdot x}, \quad 0 < x, \quad 0 < \lambda.$$

$$E(X) = \frac{1}{\lambda}, \quad D^2(X) = \frac{1}{\lambda^2} \quad \psi_X(t) = \frac{1}{1 - \frac{it}{\lambda}}.$$

- Általánosított exponenciális eloszlás:

$$f_X(x) = \lambda \cdot e^{-\lambda(x-a)}, \quad a < x, \quad 0 < \lambda.$$

$$E(X) = \frac{1}{\lambda} + a, \quad D^2(X) = \frac{1}{\lambda^2}.$$

- Gamma-eloszlás:

$$f_X(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x \geq 0 \quad (\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx)$$

$$E(X) = \frac{\alpha}{\lambda} \quad D^2(X) = \frac{\alpha}{\lambda^2} \quad \psi_X(t) = (1 - \frac{it}{\lambda})^{-\alpha}.$$

- χ^2 eloszlás n szabadságfokkal:

$$f_X(x) = \frac{x^{n/2-1} e^{-x/2}}{2^{n/2} \Gamma(n/2)} \quad x \geq 0$$

$$E(X) = n \quad D^2(X) = 2n \quad \psi_X(t) = (1 - \frac{it}{2})^{-n/2}.$$

- Béta eloszlás a, b paraméterrel:

$$f_X(x) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} \quad 0 < x < 1, \quad B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$E(X) = \frac{a}{a+b}, \quad D^2(X) = \frac{ab}{(a+b)^2(a+b+1)}$$

- Egyenletes eloszlás (az (a, b) intervallumon):

$$f_X(x) = \frac{1}{b-a}, \quad \text{ha } a < x < b, \quad 0 \text{ különben.}$$

$$E(X) = \frac{a+b}{2}, \quad D^2(X) = \frac{1}{12}(b-a)^2 \quad \text{ha } a = -b: \quad \psi_X(t) = \frac{\sin bt}{b \cdot t}.$$

- Pareto eloszlás:

$$f_X(x) = p \frac{a^p}{x^{p+1}}, \quad \text{ha } x \geq a \quad 0 \text{ különben}$$

$$E(X) = \frac{pa}{p-1} \quad \text{ha } p > 1 \text{ és } \infty \text{ különben}$$

$$D^2(X) = \frac{a^2 p}{(p-1)(p-1)} \quad \text{ha } p > 2 \text{ és } \infty, \text{ ha } 1 < p \leq 2.$$

- d -dimenziós normális eloszlás:

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-d/2} |\mathbf{C}|^{-1/2} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})^\top \mathbf{C}^{-1}(\mathbf{x}-\mathbf{m})}.$$

$$E(\mathbf{X}) = \mathbf{m}, \quad D^2(\mathbf{X}) = \mathbf{C}$$

Nevezetes próbák statisztikái, H_0 melletti eloszlásai

- Egymintás t -próba: $t = \frac{\bar{X} - \mu_0}{S_n^*/\sqrt{n}} \sim t(n-1)$
- Kétmintás t -próba: $t = \frac{\bar{X} - \bar{Y}}{\sqrt{(n_1-1)S_X^{*2} + (n_2-1)S_Y^{*2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}} \sim t(n_1 + n_2 - 2)$
- Welch próba: $t = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^{*2}}{n_1} + \frac{S_Y^{*2}}{n_2}}} \sim t(df)$, $df \approx \left(\frac{c^2}{n_1-1} + \frac{(1-c)^2}{n_2-1} \right)^{-1}$ és $c = \frac{S_X^{*2}/n_1}{S_X^{*2}/n_1 + S_Y^{*2}/n_2}$
- χ^2 próba (szórásra): $\chi^2 = \frac{nS^{*2}}{\sigma_0^2} \sim \chi^2(n-1)$
- F -próba: $F = \frac{S_X^{*2}}{S_Y^{*2}} \sim \mathcal{F}(n_1-1, n_2-1)$
- χ^2 -próba: $\chi^2 = \sum_{i=1}^r \frac{(\varphi_i - np_i)^2}{np_i} \rightarrow \chi^2(r-1)$, $n \rightarrow \infty$
- Homogenitásvizsgálat χ^2 -próbával:

$$\chi^2 = n_1 n_2 \sum_{i=1}^r \frac{1}{\mu_i + \nu_i} \left(\frac{\mu_i}{n_1} - \frac{\nu_i}{n_2} \right)^2 \rightarrow \chi^2(r-1), \min(n_1, n_2) \rightarrow \infty$$
- Függetlenségvizsgálat χ^2 -próbával:

$$\chi^2 = n \sum_{i=1}^r \sum_{j=1}^s \frac{(\nu_{ij} - \frac{\nu_{i.} \nu_{.j}}{n})^2}{\nu_{i.} \nu_{.j}} \rightarrow \chi^2((r-1)(s-1)), n \rightarrow \infty$$
- Kolmogorov–Szmirnov-próba: $\Delta_n = \sup_{x \in \mathbb{R}} |F_n^*(x) - F_0(x)|$, $\sqrt{n}\Delta_n$ aszimptotikusan Kolmogorov-eloszlású
- Kétmintás Kolmogorov–Szmirnov-próba: $\Delta_{n_1, n_2} = \sup_{x \in \mathbb{R}} |F_{n_1}^*(x) - G_{n_2}^*(x)|$, $\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \Delta_{n_1, n_2}$ aszimptotikusan Kolmogorov-eloszlású
- Wilcoxon-próba: $T^+ = \sum_{i: X_i > m_0} R_i$, ahol R_i a $|X_i - m_0|$ rangszáma,

$$\frac{T^+ - E(T^+)}{D(T^+)} = \frac{T^+ - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}} \rightarrow \mathcal{N}(0, 1), n \rightarrow \infty$$
- Kétmintás Wilcoxon-próba: $R_Y = \sum_{i=1}^{n_1} R_i$, ahol Y_i rangja R_i az egyesített mintában,

$$\frac{R_Y - E(R_Y)}{D(R_Y)} = \frac{R_Y - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}} \rightarrow \mathcal{N}(0, 1), \min(n_1, n_2) \rightarrow \infty$$
- Hotelling-féle T^2 -próba: $T^2 = n(\bar{\mathbf{X}} - \mathbf{m})^\top \hat{\mathbf{C}}^{*-1}(\bar{\mathbf{X}} - \mathbf{m})$; $\frac{n-d}{d} \cdot \frac{T^2}{n-1} \sim \mathcal{F}(d, n-d)$
- Kétmintás Hotelling teszt: $T^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{\mathbf{X}} - \bar{\mathbf{Y}})^\top \hat{\mathbf{C}}^{*-1}(\bar{\mathbf{X}} - \bar{\mathbf{Y}})$; $\frac{n_1 + n_2 - d - 1}{d} \cdot \frac{T^2}{n_1 + n_2 - 2} \sim \mathcal{F}(d, n_1 + n_2 - d - 1)$
- Egyszempontos varianciaanalízis: $F = \frac{Q_a}{Q_e} \cdot \frac{n-r}{r-1} \sim \mathcal{F}(r-1, n-r)$, ahol

$$Q_a = \sum_{i=1}^r n_i (\bar{X}_i - \bar{X})^2, \quad Q_e = \sum_{i=1}^r \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$
- Regresszióanalízis: $F = \frac{Q_r}{Q_e} (n-2) \sim \mathcal{F}(1, n-2)$, ahol

$$Q_r = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2, \quad Q_e = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$