Valószínűségszámítás

Matematika M1 gépészmérnököknek

2017. április 3., 10., 24., május 8.

Alapok

- Tárgya: véletlen jelenségek matematikai modelljeinek vizsgálata
- Cél: olyan mérték bevezetése, mely a bizonytalanságot numerikusan méri
- Véletlen eseményekkel fogunk foglalkozni, pl. érmedobás, kockadobás, vízállás, stb...

Definíció 1.

Egy véletlen kísérlet kimeneteit **elemi események***nek* (ω)*, ezek összességét pedig* **eseménytér***nek* (Ω) *nevezzük.*

Definíció 2.

Legyen Ω *részhalmazainak egy* $\mathcal{F}$ *gyűjteménye olyan, melyre*

(i) $A, B \in \mathcal{F}$ *esetén* $A \cup B \in \mathcal{F}$ *(vagy)*

(ii) $A \in \mathcal{F}$ *esetén* $\overline{A} \in \mathcal{F}$ *(komplementer esemény)*

(iii) $\Omega \in \mathcal{F}$ *(biztos esemény)*

Ekkor $\mathcal{F}$ *elemeit* **események***nek,* $\mathcal{F}$*-et magát pedig* **eseményalgebrá***nak nevezzük.*

Világos, hogy (iii) és (iv) alapján $\emptyset \in \mathcal{F}$ (lehetetlen esemény).

Definíció 3.

Legyen hozzárendelve $\mathcal{F}$ minden A eleméhez egy olyan $P(A)$ szám, melyre

(i) $0 \leq P(A) \leq 1$

(ii) $P(\Omega) = 1$

(iii) $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$, *ha* $A_i \cap A_j = \emptyset$, $i \neq j$.

Ekkor P-t **valószínűségi mérték***nek nevezzük, $(\Omega, \mathcal{F}, P)$ pedig az ún.* **valószínűségi mező***.*

Példa:

- *Klasszikus valószínűségi mező:* Ω véges, és az elemi események egyenlő valószínűségűek. Ekkor $P(A) = |A|/|\Omega|$.
- *Geometriai valószínűségi mező:* $\Omega \subset \mathbb{R}^2$, melyen a valószínűség egyenletesen oszlik el. Ekkor $P(A) = t(A)/t(\Omega)$.

Tétel 1.

- $P(\emptyset) = 0$
- $P(\overline{A}) = 1 - P(A)$
- $A \subset B$ *esetén* $P(A) \leq P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Definíció 4.

Tegyük fel, hogy A *és* B *tetszőleges események, és* $P(B) \neq 0$. *Ekkor*

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

az ún. **feltételes valószínűség**.

Definíció 5.

A és B **független események**, *ha $P(A \cap B) = P(A) \cdot P(B)$.*

Tétel 2.

Ha $P(A) \neq 0 \neq P(B)$, akkor A és B pontosan akkor függetlenek, ha $P(A) = P(A|B)$ és $P(B) = P(B|A)$.

Definíció 6.

A és B **egymást kizáró események**, *ha $A \cap B = \emptyset$.*

Az $A_1, \ldots, A_n$ események **teljes eseményrendszert** *alkotnak, ha páronként egymást kizáró események, $P(A_i) \neq 0$, és $\bigcup_{i=1}^{n} A_i = \Omega$.*

Tétel 3 (Teljes valószínűség tétele).

Ha az $A_1, \dots, A_n$ *események teljes eseményrendszert alkotnak, akkor*

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i).$$

Tétel 4 (Bayes-tétel).

Ha az $A_1, \dots, A_n$ *események teljes eseményrendszert alkotnak, akkor*

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}.$$

VALÓSZÍNŰSÉGI VÁLTOZÓK

Definíció 7.

Az $\xi : \Omega \to \mathbb{R}$ *függvényeket* **valószínűségi változó***nak (v.v.) nevezzük.*

Definíció 8.

Az **eloszlásfüggvény** *egy minden valószínűségi változóra értelmezhető, azt jól jellemző* $F_\xi : \mathbb{R} \to \mathbb{R}$ *függvény olyan, melyet az*

$$F_\xi(x) := P(\xi < x)$$

egyenlőség definiál, ahol ξ *v.v.,* x *pedig egy tetszőleges, rögzített, véletlentől nem függő érték.*

Tétel 5.

Minden ξ *v.v.* F_ξ *eloszlásfüggvényére*

(i) $0 \leq F_\xi \leq 1$ *bármely* $x \in \mathbb{R}$ *esetén*

(ii) F_ξ *monoton növő függvény*

(iii) $\lim\limits_{x\to-\infty} F_\xi(x) = 0$ *és* $\lim\limits_{x\to\infty} F_\xi(x) = 1$

(iv) F_ξ *balról folytonos függvény (azaz* $\lim\limits_{x\to a-0} F_\xi(x) = F_\xi(a)\ \forall a \in \mathbb{R}$*)*

A tétel megfordítása is igaz, azaz ha egy F függvény kielégíti a fenti 4 tulajdonság mindegyikét, akkor létezik olyan ξ v.v., melynek F az eloszlásfüggvénye.

Tehát ezek a tulajdonságok **karakterizálják** az eloszlásfüggvényeket.

Definíció 9.

Az olyan véletlen mennyiséget, mely csak véges vagy megszámlálhatóan végtelen sok különböző értéket vehet fel, **diszkrét v.v.***-nak nevezzük.*

Azaz egy diszkrét v.v. eloszlása a

$$p_k := P(\xi = k), \quad k = 1, 2, \ldots$$

értékekkel definiálható. Világos, hogy

- $0 \leq p_k \leq 1,\ k = 1, 2, \ldots$
- $\sum_{k=1}^{\infty} p_k = 1$
- maga az eloszlásfüggvény lépcsős függvény lesz.

Abszolút folytonos v.v.

Definíció 10.

Legyen ξ v.v. Ha létezik $f_\xi : \mathbb{R} \to \mathbb{R}_+$ függvény olyan, melyre bármely $I \subset \mathbb{R}$ esetén

$$P(\xi \in I) = \int_I f(x)\,\mathrm{d}x,$$

akkor ξ eloszlását **abszolút folytonos***nak nevezzük, f_ξ pedig ξ* **sűrűségfüggvény***e lesz.*

Tulajdonságok:

- $0 \le f_\xi$ bármely $x \in \mathbb{R}$ esetén
- $\displaystyle\int_{-\infty}^{\infty} f_\xi(x)\,\mathrm{d}x = 1$
- $\displaystyle F_\xi(x) = P(\xi \in (-\infty, x)) = \int_{-\infty}^{x} f_\xi(y)\,\mathrm{d}y$, azaz $F'_\xi(x) = f_\xi(x)$
- nem lesz feltétlenül folytonos függvény.

- A gyakorlatban általában nincs elég mérésünk ahhoz, hogy az eloszlás- és sűrűrégfüggvényeket pontosan meghatározzuk;
- helyettük elegendő lesz az átlagos érték és az e körüli ingadozás mértékének megadása.

Definíció 11.

Ha ξ diszkrét v.v. (x_i, p_i), $i \in \mathbb{N}$ eloszlással, akkor ξ **várható értéke**

$$\mathrm{E}\xi := \sum_{i=1}^{\infty} x_i p_i,$$

feltéve, hogy a fenti sor abszolút konvergens.
Folytonos eloszlású v.v. esetén

$$\mathrm{E}\xi = \int_{-\infty}^{\infty} x \cdot f_\xi(x)\,\mathrm{d}x,$$

feltéve, hogy $\int_{-\infty}^{\infty} |x| f_\xi(x)\,\mathrm{d}x < \infty$.

Tétel 6 (Tulajdonságok).

a.) $\mathrm{E}(\xi + \eta) = \mathrm{E}\xi + \mathrm{E}\eta$

b.) $\mathrm{E}(a \cdot \xi) = a \cdot \mathrm{E}\xi$, $a \in \mathbb{R}$

c.) $\mathrm{E}(b) = b$, *ha* $b \in \mathbb{R}$.

Definíció 12.

Ha $\mathrm{E}(\xi^2)$ *létezik, akkor a*

$$D^2(\xi) = \mathrm{E}[(\xi - \mathrm{E}\xi)^2]$$

mennyiséget ξ **szórásnégyzeté***nek nevezzük.*

Tétel 7 (Tulajdonságok).

a.) $D^2(\xi) = \mathrm{E}(\xi^2) - \mathrm{E}^2(\xi)$

b.) $D^2(a \cdot \xi + b) = a^2 \cdot D^2(\xi)$, $a, b \in \mathbb{R}$

VALÓSZÍNŰSÉGI VEKTORVÁLTOZÓK

Valószínűségi vektorváltozók

Mi most speciálisan csak a 2-dimenziós esettel foglalkozunk!

Definíció 13.

Legyen $\binom{\xi}{\eta} : \Omega \to \mathbb{R}^2$ *egy függvény. Ha* ξ *és* η *mindketten v.v.-k, akkor* $\binom{\xi}{\eta}$*-t* **2-dimenziós valószínűségi vektorváltozó***nak nevezzük.*

Ha őket egyszerre szeretnénk vizsgálni, akkor a közös eloszlásukra lesz szükségünk.

Definíció 14 (diszkrét eset).

Ha ξ *és* η *mindketten diszkrét eloszlásúak* $\{x_i\}$ *és* $\{y_j\}$ *értékekkel, akkor* **együttes eloszlásuk***on a*

$$p_{ij} = P(\xi = x_i \text{ és } \eta = y_j)$$

sorozatot értjük, $i, j \in \mathbb{N}$.

Definíció 15 (folytonos eset).

$\binom{\xi}{\eta}$ *abszolút folytonos eloszlású vektorváltozó, ha létezik olyan* $f : \mathbb{R}^2 \to \mathbb{R}_+$ *függvény, melyre az*

$$\int_I \int_J f_{(\xi,\eta)}(x, y)\,\mathrm{d}y\,\mathrm{d}x = P((\xi, \eta) \in I \times J)$$

integrál értelmes bármely $I \times J$ *téglalapra. Az* f *függvényt ekkor* ξ *és* η **együttes sűrűségfüggvényé***nek nevezzük.*

- $f \geq 0$
- $$\int_{\mathbb{R}} \int_{\mathbb{R}} f_{(\xi,\eta)}(x, y)\,\mathrm{d}y\,\mathrm{d}x = 1$$

Peremeloszlások

Kérdés: Visszanyerhető-e az együttes eloszlásból ξ s η egyenkénti eloszlása? A válasz igen, ezek lesznek az ún. **peremeloszlás**ok.

a.) *Diszkrét eset:*

$$p_{i\cdot} = P(\xi = x_i) = \sum_j P(\xi = x_i, \eta = y_j) = \sum_j p_{ij}$$

$$r_{\cdot j} = P(\eta = y_j) = \sum_i P(\xi = x_i, \eta = y_j) = \sum_i p_{ij}$$

b.) *Folytonos eset:*

$$f_\xi(x) = \int_{\mathbb{R}} f_{(\xi,\eta)}(x,y)\,\mathrm{d}y \quad \text{és} \quad f_\eta(y) = \int_{\mathbb{R}} f_{(\xi,\eta)}(x,y)\,\mathrm{d}x,$$

ugyanis

$$P(\xi \in I) = \int_I f_\xi(x)\,\mathrm{d}x = P(\xi \in I, \eta \in \mathbb{R}) = \int_I \underbrace{\int_{\mathbb{R}} f_{(\xi,\eta)}(x,y)\,\mathrm{d}y}_{f_\xi(x)}\,\mathrm{d}x,$$

Független v.v.-k

Definíció 16.

ξ és η függetlenek, ha

a.) *diszkrét esetben $p_{ij} = p_{i\cdot} \cdot r_{\cdot j}$ bármely i, j párra*

b.) *folytonos esetben $f_{(\xi,\eta)}(x, y) = f_\xi(x) \cdot f_\eta(y)$ bármely $(x, y) \in \mathbb{R}^2$ esetén.*

Tétel 8.

Ha ξ és η függetlenek, akkor

$$\mathrm{E}(\xi\eta) = \mathrm{E}\xi \cdot \mathrm{E}\eta$$

és

$$D^2(\xi + \eta) = D^2\xi + D^2\eta,$$

feltéve, hogy a megfelelő várható értékek léteznek.

Kovariancia

A kölcsönös viszony mérésének eszközei

Definíció 17.

Ha $\mathrm{E}\xi^2$ *és* $\mathrm{E}\eta^2$ *léteznek, akkor a*

$$\mathrm{cov}(\xi, \eta) := \mathrm{E}[(\xi - \mathrm{E}\xi)(\eta - \mathrm{E}\eta)]$$

mennyiséget ξ *és* η ***kovarianciá****jának nevezzük.*

- Ha ξ és η olyanok, hogy $\mathrm{cov}(\xi, \eta) = 0$, akkor **korrelálatlanok**- nak nevezzük őket.
- Ha ξ és η függetlenek, akkor korrelálatlanok is, de a megfordítás már nem igaz!
- Bilineáris függvény, továbbá $\mathrm{cov}(\xi, \xi) = D^2\xi$
- $|\mathrm{cov}(\xi, \eta)| \leq D\xi \cdot D\eta$ bármely ξ, η v.v.-k esetén. (CSB $\leq$)

Probléma: nagysága függ a mértékegység megválasztásától!

Korreláció

Definíció 18.

Ha $\mathrm{E}\xi^2$ *és* $\mathrm{E}\eta^2$ *léteznek, akkor a*

$$\mathrm{corr}(\xi,\eta) := \frac{\mathrm{cov}(\xi,\eta)}{D\xi \cdot D\eta}$$

mennyiséget ξ *és* η **korreláció***jának nevezzük.*

- $|\mathrm{corr}(\xi,\eta)| \leq 1$
- Ha ξ és η függetlenek, akkor $\mathrm{corr}(\xi,\eta) = 0$, de ez sem fordítható meg.
- $|\mathrm{corr}(\xi,\eta)| = 1$ pontosan akkor, ha létezik $a \neq 0$ és $b \in \mathbb{R}$ olyanok, hogy
$$P(\eta = a\xi + b) = 1.$$

NEVEZETES ELOSZLÁSOK

Definíció 19.

ξ (p) *paraméterű* **geometriai eloszlás***ú v.v., ha lehetséges értékei a* $k = 0, 1, 2, \ldots$ *számok, eloszlása pedig*

$$p_k = P(\xi = k) = (1-p)^{k-1} \cdot p, \quad k = 0, 1, 2, \ldots$$

- ugyanazt a kísérletet egymástól függetlenül ismételjük az első sikerig, p egy kísérlet sikerének valószínűsége
- könnyen ellenőrizhető, hogy $\sum_k p_k = 1$, azaz tényleg eloszlást definiál
- $\mathrm{E}\xi = \frac{1}{p}$, $D\xi = \sqrt{\frac{1-p}{p^2}}$

Definíció 20.

ξ (n,p) paraméterű **binomiális eloszlás***ú v.v., ha lehetséges értékei a $k=0,1,2,\dots,n$ számok, eloszlása pedig*

$$p_k = P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0,1,2,\dots,n.$$

- ugyanazt a kísérletet n-szer egymástól függetlenül elvégezve a sikerek (ξ) számának eloszlását vizsgálja, ahol p a siker egyszeri bekövetkezésének valószínűsége (pl. selejtek vagy meghibásodások száma)
- könnyen ellenőrizhető, hogy $\sum_k p_k = 1$, azaz tényleg eloszlást definiál
- $\mathrm{E}\xi = np$, $D\xi = \sqrt{np(1-p)}$

Definíció 21.

Legyen $\lambda > 0$ *rögzített. Az* η *v.v.* **λ-paraméterű Poisson eloszlású** *v.v., ha értéke tetszőleges* $k \in \mathbb{N}$ *lehet, eloszlása pedig*

$$p_k = P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \ldots .$$

- pl. sajtóhibák, beérkező telefonhívások, balesetek száma
- sok, egymástól független, egyenként kis valószínűségű esemény bekövetkezésének eloszlása
- könnyen ellenőrizhető, hogy $\sum_k p_k = 1$, azaz tényleg eloszlást definiál
- $\mathrm{E}\eta = \lambda$, $D\eta = \sqrt{\lambda}$

Tétel 9.

Ha $np = \lambda > 0$ *konstans és* $0 \leq k \leq n$ *rögzített, akkor*

$$\binom{n}{k} p^k (1-p)^{n-k} \xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!} e^{-\lambda},$$

azaz ha n *elég nagy, akkor a binomiális eloszlás közelíthető a Poisson eloszlással.*

Definíció 22.

Legyenek $a < b$ valós paraméterek. A ξ v.v. az $[a, b]$ intervallumon **egyenletes eloszlású**, *ha sűrűségfüggvénye*

$$f(x) = \begin{cases} \frac{1}{b-a} & \textit{ha } a \leq x \leq b \\ 0 & \textit{egyébként.} \end{cases}$$

Eloszlásfüggvénye:

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \begin{cases} 0 & \text{ha } x < a \\ \frac{x-a}{b-a} & \text{ha } a \leq x \leq b \\ 1 & \text{ha } x > b. \end{cases}$$

Továbbá

$$\mathrm{E}\xi = \frac{a+b}{2} \quad \text{és} \quad D\xi = \frac{b-a}{2\sqrt{3}}$$

Definíció 23.

Legyen $\lambda > 0$ rögzített. Az η v.v. **λ-paraméterű exponenciális eloszlású** *v.v., ha sűrűségfüggvénye*

$$f(x) = \begin{cases} 0 & \textit{ha } x \le 0 \\ \lambda e^{-\lambda x} & \textit{ha } x > 0. \end{cases}$$

Eloszlásfüggvénye:

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \begin{cases} 0 & \text{ha } x \le 0 \\ 1 - e^{-\lambda x} & \text{ha } x > 0 \end{cases}$$

Továbbá

$$\mathrm{E}\xi = \frac{1}{\lambda} \quad \text{és} \quad D\xi = \frac{1}{\lambda}$$

- pl. élettartamok, várakozási idők esetén használjuk
- igaz rá az ún. örökifjú tulajdonság, azaz az eloszlásnak nincs memóriája, a további várakozási igényeket nem befolyásolja az, hogy addig mennyit vártunk:

$$P(\xi < x + y | \xi \geq y) = \frac{P(y \leq \xi < x + y)}{P(y \leq \xi)} =$$

$$= \frac{\overbrace{1 - e^{-\lambda(x+y)}}^{F(x+y)} - (\overbrace{1 - e^{-\lambda y}}^{F(y)})}{\underbrace{e^{-\lambda y}}_{1-F(y)}} = \frac{e^{-\lambda y}(1 - e^{-\lambda x})}{e^{-\lambda y}} =$$

$$= 1 - e^{-\lambda x} = P(\xi < x).$$

Definíció 24.

A ξ *v.v.* **standard normális eloszlású**, *ha sűrűségfüggvénye*

$$f_\xi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$$

- A természetben és a műszaki életben leggyakrabban előforduló eloszlás.
- Eloszlásfüggvényére nincs explicit formula, értékeit táblázatos formában, numerikusan adják meg. A sűrűségfüggvény párossága miatt könnyne látható, hogy

$$\Phi(x) = 1 - \Phi(-x).$$

- $\mathrm{E}\xi = 0$ és $D\xi = 1$
- Jelölés: $\xi \sim N(0,1)$

Definíció 25.

Az η v.v. m **várható értékű és σ szórású normális eloszlás**t *követ, ha sűrűségfüggvénye*

$$f_\eta(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-m)^2}{2\sigma^2}}$$

- Jelölés: $\xi \sim N(m,\sigma)$. Világos, hogy $\mathrm{E}\xi = m$ és $D\xi = \sigma$.
- Ha $\xi \sim N(0,1)$, akkor

$$\sigma\xi + m = \eta \sim N(m,\sigma)$$

- és ha $\xi \sim N(m,\sigma)$, akkor

$$\frac{\xi - m}{\sigma} = \eta \sim N(0,1)$$

- Ha $\xi \sim N(m_1,\sigma_1)$ és $\eta \sim N(m_2,\sigma_2)$, akkor

$$\xi + \eta \sim N(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$$

Példa

A TRAGACS személyautó csomagtartója szabvány szerint 70 cm széles, a RÉMÁLOM gyerekágyakat pedig lapra szerelve, szabványosan 65 cm széles csomagban árulják. Valójában mindkét szélesség normális eloszlású valószínűségi változó, a csomagtartó esetében 3, a lapra szerelt csomag esetében 4 cm szórással (a két szélesség függetlennek tekinthető). Veszek egy RÉMÁLOM gyerekágyat, és szeretném a TRAGACSom-mal hazavinni. Mi a valószínűsége annak, hogy bele fog férni a csomagtartóba?

A NAGY SZÁMOK TÖRVÉNYE ÉS A CENTRÁLIS HATÁRELOSZLÁS-TÉTEL

A mérnöki gyakorlatban is fontos feladat olyan mennyiségek ingadozásának pontos mérése, melyek sok, egymásra rakódó és egymástól függetlennek tekinthető véletlen jelenség hatására alakulnak ki. Az ilyen ingadozásokat szeretnénk megérteni és becsülni.

Bevezetésül két olyan egyenlőtlenséget tekintünk, melyek v.v.-k ingadozásait általában jellemzik.

Tétel 10 (Markov-egyenlőtlenség).

Legyen ξ egy nemnegatív v.v. (azaz $P(\xi < 0) = 0$). Ekkor tetszőleges $a > 0$ számra

$$P(\xi \geq a) \leq \frac{\mathrm{E}\xi}{a}$$

Azaz kicsi annak a valószínűsége, hogy egy pozitív véletlen mennyiség a várható értékénél lényegesen nagyobb legyen.

Tétel 11 (Csebisev-egyenlőtlenség).

Legyen ξ *olyan v.v., melyre* $\mathrm{E}\xi = m$ *és* $D\xi = \sigma$ *véges. Ekkor tetszőleges* $\varepsilon > 0$ *esetén*

$$P(|\xi - m| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Azaz a véletlen mennyiségek a várható értékük körül ingadoznak, a szórásnál lényegesen nagyobb kilengések azonban csak kis valószínűséggel fordulnak elő.

Feladat: végezzük el ugyanazt a kísérletet egymástól függetlenül n-szer. Jelöljük X-szel az adott kísérlet szempontjából a "sikerek" számát, ezek bekövetkezési valószínűségét pedig jelölje p. (Pl. érmedobás, kockadobás, selejtek száma, sb...) A cél éppen p becslése a mintavétel segítségével.

Persze tudjuk, hogy X binomiális eloszlású, így várható értéke és szórása számunkra ismert lehetne, ha tudnánk p értékét. Ennek becslésére tekintsük az X/n mutatót, azaz a sikerek arányát. Azt várjuk, hogy ha n elég nagy, akkor ez p értékét, az egyedi kísérletben a siker esélyét jól fogja közelíteni.

Alapvetően ezt a sejtést formalizálja a **Nagy Számok Törvénye**.

Tétel 12 (Nagy számok gyenge törvénye - Bernoulli-féle törvény).

Legyen $\delta > 0$ *tetszőleges rögzített szám. Ekkor*

$$P\left(\left|\frac{X}{n} - p\right| \geq \delta\right) \to 0, \quad \textit{ha} \quad n \to \infty.$$

Bizonyítás: Legyen $\varepsilon = n\delta$, és alkalmazzuk a Csebisev-egyenlőt- lenséget. Ekkor $\mathrm{E}X = np$ és $D^2X = np(1-p)$ felhasználásával

$$P\left(\left|\frac{X}{n} - p\right| \geq \delta\right) = P\left(|X - np| \geq n\delta\right) \leq \frac{D^2X}{n^2\delta^2} =$$

$$= \frac{np(1-p)}{n^2\delta^2} = \frac{p(1-p)}{n\delta^2} \to 0, \quad \text{ha} \quad n \to \infty.$$

Azaz a kísérletek számának növelésével a sikerek aránya egyre pontosabban, egyre nagyobb valószínűséggel közelíti meg p-t.

Egy fontos általánosítás

A fenti X v.v.-ra vonatkozó becsléseinket annak binomiális eloszlása alapján kaptuk. Ez a feltétel azonban gyengíthető.

Tétel 13.

Ha ξ_i, $i \in \mathbb{N}$ független, azonos eloszlású v.v-k olyanok, hogy $\sigma^2 = D^2\xi_i$ véges, akkor bármely $\delta \geq 0$ rögzített szám esetén

$$P\left(\left|\frac{\sum_{i=1}^{n}\xi_i}{n} - \mathrm{E}\xi_1\right| \geq \delta\right) \to 0, \quad \textit{ha} \quad n \to \infty.$$

Pl. kíváncsiak vagyunk valamilyen mennyiség (pl. szakítószilárdság) tényleges értékére, amelyre azonban számos, véletlennek tekinthető fluktuáció rakódik rá (pl. hőmérsékletből vagy páratartalomból adódóan). Ha a mennyiséget sokszor megmérjük, egymástól független mérésekkel, akkor a mért értékek átlaga egyre pontosabban fogja közelíteni a mennyiség tényleges értékét.

Centrális határeloszlás-tétel - CHT

A NSZT-nek fenti alakjai úgy is értelmezhetők, mint hogy az n független mennyiség összegeként kapott v.v.-ból levonva a várható értékét, majd ezt elosztva a szórásnál lényegesen nagyobb n-nel, a véletlen fluktuációk kiskálázódnak a véletlen mennyiségből. Mi történne akkor, ha n helyett a $\sigma\sqrt{n}$ mennyiséggel osztanánk?

Tétel 14 (Centrális határeloszlás tétel).

Ha ξ_i, $i \in \mathbb{N}$ *független, azonos eloszlású v.v-k olyanok, hogy* $\mathrm{E}\xi_i = m$ *és* $\sigma^2 = D^2\xi_i$ *véges, akkor bármely* $x \in \mathbb{R}$ *esetén*

$$P\left(\frac{\sum_{i=1}^{n} \xi_i - nm}{\sigma\sqrt{n}} < x\right) \to P(\eta < x), \quad \text{ha} \quad n \to \infty,$$

ahol $\eta \sim N(0,1)$. *Azaz* $\sum_{i=1}^{n} \xi_i \sim N(nm, \sigma\sqrt{n})$.

Lényegesen jobb becslést ad, mint a NSZT, kevesebb mintavétel is elegendő ugyanahhoz a pontossághoz. A matematikai statisztika legalapvetőbb tétele.