Felsőbb Matematika Informatikusoknak – Sztochasztika
vázlatos jegyzet-vázlat a 2. részhez

Kivonat

Ez egy nagyon rövid vázlat sok hibával, következetlenséggel és elírással. Egyelőre az első három témakör anyaga van benne. Nem követi pontosan az előadáson elhangzottakat, főleg nem azok sorrendjét. Cserébe remélhetőleg tartalmazza a legfontosabb tudnivalókat. Nem törekszem az ismereteket tökéletesen szétválasztani definíciókra, tételekre és bizonyításokra. Sőt, bizonyításokat egyáltalán nem írok le. Szintén nem törekszem a felépítés logikai sorrendjének pontos betartására, így előfordul, hogy egy fogalmat hamarabb használok, mint ahogy definiálom. Ezért ez az iromány igazán matematikus lelkületű olvasóknak nem ajánlott.

Tartalomjegyzék

1. Ismétlés

1.1. Nevezetes valószínűségi változók

A valószínűségi változó egy „véletlen szám”, ami egy kísérlet kimenetelétől függ (de általában nem ad teljes információt a kísérlet kimeneteléről). Matematikailag egy

$$X : \Omega \to \mathbb{R}$$

függvény, ahol Ω a valószínűségi mező, vagyis egy kísérlet összes lehetséges kimeneteleinek halmaza. Így ha $B \subset \mathbb{R}$, akkor

$$\{X \in B\} := \{\omega \in \Omega \,:\, X(\omega(\in B\} \subset \Omega$$

egy **esemény**, vagyis a kísérlet azon lehetséges kimeneteleinek halmaza, amihez az X függvény B-beli értéket rendel. Így értelemes a $\mathbb{P}(X \in B)$ valószínűségről beszélni.

Legfontosabb jellemzője az „eloszlása”, ami azt a „tudást” jelenti, hogy milyen értékeket vehet fel, és ezeket milyen valószínűséggel veszi fel.

Ez az eloszlás általában jellemezhető az F *eloszlásfüggvénnyel*:

1. Definíció. *Az X valószínűség változó eloszlásfüggvénye az $F : \mathbb{R} \to [0,1]$, amit úgy definiálunk, hogy minden $x \in \mathbb{R}$-re*

$$F(x) := \mathbb{P}(X \le x) \quad \textit{(az angol nyelvű irodalomban), vagy}$$

$$F(x) := \mathbb{P}(X < x) \quad \textit{(a magyar nyelvű tankönyvekben.}$$

Tulajdonságai:

- F monoton növekvő
- $\lim\limits_{x \to -\infty} F(x) = 0$
- $\lim\limits_{x \to \infty} F(x) = 1$
- F folytonos, kivéve az olyan $x \in \mathbb{R}$-ekben, amikre $\mathbb{P}(X = x) > 0$.

A két definíció közötti árnyalatnyi különbségnek számunkra nincs jelentősége, mert csak annyi múlik rajta, hogy ha egy $x \in \mathbb{R}$ értéket az X val.változó pozitív valószínűséggel vesz fel (vagyis $\mathbb{P}(X = x) > 0$), és emiatt x-ben az F-nek ugrása van, akkor ezen az x helyen az F balról folytonos lesz vagy jobbról folytonos. Mi viszont az eloszlásfüggvényt

- vagy csak folytonos eloszlású val.változókra használjuk, ahol szakadási pontok nincsenek, vagyis F úgyis (mindkét irányból) folytonos, a két definíció megegyezik,
- vagy csak olyan pontokban nézzük, ahol folytonos (lásd gyenge konvergencia definíciója), és ott mindegy.

Alkalmazása: $a, b \in \mathbb{R}$, $a < b$-re

$$\mathbb{P}(X \in (a,b]) = F(b) - F(a) \quad \text{(az angol nyelvű irodalomban szokásos definícióval)}$$

$$\mathbb{P}(X \in [a,b)) = F(b) - F(a) \quad \text{(a magyar nyelvű tankönyvekben szokásos definícióval)}$$

Ha X aboszolut folytonos, akkor tökmindegy:

$$\mathbb{P}(X \in [a,b)) = \mathbb{P}(X \in [a,b]) = \mathbb{P}(X \in (a,b)) = \mathbb{P}(X \in (a,b]) = F(b) - F(a)$$

Az eloszlásfüggvény szép és jó, mert minden val.változó eloszlása jellemezhető vele, de a vele való számolás nehézkes. Ezért csak két speciális esettel foglalkozunk.

1.1.1. Nevezetes diszkrét valószínűségi változók

Ha az X val.változó csak véges, vagy megszámlálhatóan végtelen sok értéket vehet fel, akkor az eloszlás megadása annyit jelent, hogy felsoroljuk a lehetséges értékeket:$\{x_1, x_2, x_3, \dots\}$, illetve ezek valószínűségeit: $\{p_1, p_2, p_3, \dots\}$. Vagyis $p_i = \mathbb{P}(X = x_i)$.

Tulajdonságai: $p_i \geq 0$ és $\sum_i p_i = 1$.

2. Definíció. *A fenti X val.változó várható értéke*

$$\mathbf{E}X = \sum_i p_i x_i$$

(ha létezik, vagyis ha ez a sor abszolut konvergens, esetleg $\pm\infty$).

1. Tétel. *Ha X a fenti diszkrét val.változó, és $g : \mathbb{R} \to \mathbb{R}$ függvény, akkor $g(X)$ is diszkrét val.váltzó, és várható értéke*

$$\mathbf{E}(g(X)) = \sum_i p_i g(x_i)$$

(ha létezik).

Legfontosabb példa:

$$\mathbf{E}(X^2) = \sum_i p_i x_i^2.$$

3. Definíció. *Egy X val.változó varianciája, avagy szórásnégyzete*

$$\mathbf{Var}X = \mathbf{D}^2X := \mathbf{E}((X - \mathbf{E}X)^2),$$

ha ez létezik.

Kiszámolása:

$$Var(X) = \mathbf{E}(X^2) - (\mathbf{E}X)^2.$$

A legfontosabb példák mindegyikében $p \in (0, 1)$ paraméter (esetleg lehet $p = 0$ vagy $p = 1$), és $q := 1 - p$ csak egy jelölés.

Bernoulli eloszlás:

X Bernoulli eloszlású p paraméterrel, vagyis $X \sim B(p)$, ha lehetséges értékei 0 és 1, valamint $\mathbb{P}(X = 0) = q$ és $\mathbb{P}(X = 1) = p$. Jelentése: $X = 1$, ha egy kísérlet sikeres, és $X = 0$, ha nem. Erre érdemes úgy tekinteni, hogy X a sikeres kísérlet„ek" száma egyetlen próbálkozásból.

Tulajdonságai:

$$\mathbf{E}(B(p)) = p \quad , \quad \mathbf{Var}(B(p)) = pq$$

Binomiális eloszlás:

Legyen $n \in \mathbb{N}$. Azt mondjuk, hogy X Bernoulli eloszlású (n, p) paraméterrel, vagyis $X \sim Bin(n, p)$, ha lehetséges értékei $\{0, 1, 2, \dots, n\}$ és ezekre

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

Jelentése: X a sikeres kísérletek száma, ha n *teljesen független* kísérletet végzünk, és a siker val.sége mindegyikre p. Tulajdonságai:

$$\mathbf{E}(Bin(n, p)) = np \quad , \quad \mathbf{Var}(Bin(n, p)) = npq$$

Geometriai eloszlás: X (optimista) geometriai eloszlású p paraméterrel, vagyis $X \sim Geom(p)$, ha lehetséges értékei $\{1, 2, 3, \dots\}$ és ezekre

$$\mathbb{P}(X = k) = pq^{k-1}.$$

Jelentése: X azt jelöli, hogy hanyadik lesz az *első sikeres* próbálkozás, ha *teljesen független* kísérletek sorozatát végezzük, és a siker val.sége mindegyikre p. Tulajdonságai:

$$\mathbf{E}(Geom(p)) = \frac{1}{p} \quad , \quad \mathbf{Var}(Geom(p)) \text{ pedig nekünk nem fontos.}$$

Példa: Ha X azt jelöli, hogy hanyadik kísérletre sikerül egy szabályos dobókockával 6-ost dobni, akkor $X \sim Geom(\frac{1}{6})$, és nem meglepő módon $\mathbf{E}X = 6$.

Pesszimista geometriai eloszlás: X pesszimista geometriai eloszlású p paraméterrel, vagyis $X \sim PessGeom(p)$, ha lehetséges értékei $\{0, 1, 2, 3, \dots\}$ és ezekre

$$\mathbb{P}(X = k) = pq^{k}.$$

Jelentése: X azt jelöli, hogy hány kudarc előzi meg az *első sikeres* próbálkozást, ha *teljesen független* kísérletek sorozatát végezzük, és a siker val.sége mindegyikre p. Tulajdonságai:

$$\mathbf{E}(PessGeom(p)) = \frac{1}{p} - 1 \quad , \quad \mathbf{Var}(PessGeom(p)) \text{ pedig nekünk nem fontos.}$$

Példa: Ha X azt jelöli, hogy egy szabályos dobókockával gurigázva az első 6-os előtt hányszor szomorkodunk, mert nem 6-os jött ki, akkor $X \sim PessGeom(\frac{1}{6})$, és nem meglepő módon $\mathbf{E}X = 5$.

Poisson eloszlás: Legyen $\lambda > 0$ valós szám. Azt mondjuk, hogy X Poisson eloszlású λ paraméterrel, vagyis $X \sim Poi(\lambda)$, ha lehetséges értékei $\{0, 1, 2, \dots, n\}$ és ezekre

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

Jelentése: X a sikeres kísérletek száma, ha *nagyon sok* teljesen független kísérletet végzünk, és a siker val.sége mindegyikre nagyon kicsi, de sem a kísérletek n számát, sem a siker p val.ségét nem ismerjük (vagy nem akarjuk közvetlenül használni), csak a kettő szorzatát (vagyis a várható értéket): $\lambda := np$. Tulajdonságai:

$$\mathbf{E}(Poi(\lambda)) = \lambda \quad , \quad \mathbf{Var}(Poi(\lambda)) = \lambda$$

A Poisson eloszláshoz tartozó meséből látszik, hogy a Poisson eloszlás valamilyen értelemben határesete a binomiálisnak, amint $n \to \infty$, $p \to 0$ és $np \to \lambda$. Lásd később, a „gyenge konvergencia” fejezetben.

1.1.2. Nevezetes abszolut folytonos valószínűségi változók

Egy X val.változót akkor nevezünk abszolut folytonosnak, ha van neki egy $f = \mathbb{R} \to \mathbb{R}^+$ sűrűségfüggvénye, ami azt tudja, hogy minden $[a, b] \subset \mathbb{R}$ intervallumra

$$\mathbb{P}(X \in [a, b]) = \int_a^b f(x)\,\mathrm{d}x.$$

Ebben az esetben minden konkrét $x \in \mathbb{R}$-re külön-külön $\mathbb{P}(X = x) = 0$, így mindegy, hogy nyílt vagy zárt intervallumok valószínűségeit nézzük.

Tulajdonságai: $f(x) \geq 0$ és $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$.

4. Definíció. *A fenti X val.változó várható értéke*

$$\mathbf{E}X = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$$

(ha létezik, vagyis ha ez az improprius integrál abszolut konvergens, esetleg $\pm\infty$*).*

2. Tétel. *Ha* X *a abszolut folytonos val.változó, és* $g : \mathbb{R} \to \mathbb{R}$ *függvény, akkor a* $g(X)$ *val.változó várható értéke*

$$\mathbf{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)\,\mathrm{d}x$$

(ha létezik).

Legfontosabb példa:

$$\mathbf{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x.$$

Egyenletes eloszlás: Legyen a és b valós szám, $a < b$. X egyenletes elszlású az $[a, b]$ intervallumon, vagyis $X \sim Uni(a, b)$ vagy $X \sim E(a, b)$, ha súrúségfüggvénye

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

A sűrűségfüggvénynek az intervallum végpontjaiban felvett értékének nincs jelentősége.

Tulajdonságai:

$$\mathbf{E}(Uni(a, b)) = \frac{a+b}{2} \quad , \quad \mathbf{Var}(Uni(a, b)) = \frac{(b-a)^2}{12}$$

Exponenciális eloszlás: Legyen $\lambda > 0$ valós paraméter. X exponenciális eloszlású λ paraméterrel, vagyis $X \sim Exp(\lambda)$, ha súrúségfüggvénye

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

Momentumai:

$$\mathbf{E}(Exp(\lambda)) = \frac{1}{\lambda} \quad , \quad \mathbf{Var}(Exp(\lambda)) = \frac{1}{\lambda^2}$$

Eloszlásfüggvénye:

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

Sokkal fontosabb tulajdonság: örökifjúság. Ha X exponenciális eloszlású és $s, t > 0$ valós számok, akkor

$$\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t).$$

Ez azt jelenti, hogy ha egy kütyü élettartama exponenciális elsozlású val.változó, és a kütyü s ideig *még nem romlott el*, akkor az élettartamának hátralévő része pont olyan eloszlású, mint eredetileg volt. Vagyis a kütyü nem öregszik.

Normális eloszlás:

5. Definíció. *Az X val.változó* standard normális eloszlású, *vagyis* $X \sim \mathcal{N}(0,1)$, *ha sűrűségfüggvénye*

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

Tulajdonságai: $\mathbf{E}\mathcal{N}(0,1) = 0$, $\mathbf{Var}\mathcal{N}(0,1) = 1$.
Ennek eloszlásfüggvénye a *standard normális eloszlásfüggvény*, vagyis

$$\Phi(x) := \mathbb{P}(X \le x) = \int_{-\infty}^{x} \varphi(x)\,\mathrm{d}x$$

Ez sajnos nem írható fel elemi függvényekkel, ezért numerikusan kell számolni, vagy táblázatból kell kikeresni. A táblázatokban többnyire csak $x \ge 0$-ra van benne, mivel

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

6. Definíció. *Legyen m ás σ valós paraméter,* $\sigma > 0$. *Az X val.változó* normális eloszlású m és σ^2 paraméterekkel, *vagyis* $X \sim \mathcal{N}(m,\sigma^2)$, *ha sűrűségfüggvénye*

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

Tulajdonságai:

- $\mathbf{E}\mathcal{N}(m,\sigma^2) = m$
- $\mathbf{Var}\mathcal{N}(m,\sigma^2) = \sigma^2$

vagyis a paramétereknek természetes jelentése van: m a várható értéket, σ^2 a varianciát vagyis szórásnégyzetet jelöli, így σ a szórás. Magyar nyelvű tankönyvekben néha előfordul, hogy a normális eloszlás második paramétereként nem a szórásnégyzetet, hanem a szórást használják, de **nálunk a második paraméter mindig a szórás*négyzet***.

3. Tétel. *A fenti* $\mathcal{N}(m,\sigma^2)$ *eloszlás eloszlásfüggvénye*

$$F_{m,\sigma^2}(x) = \Phi\left(\frac{x-m}{\sigma}\right),$$

ahol Φ *a standard normális eloszlásfüggvény.*

Következmény: Ha $X \sim \mathcal{N}(m,\sigma^2)$, $a \le b \in \mathbb{R}$, akkor

$$\mathbb{P}(X \in [a,b]) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right).$$

4. Tétel. *Ha* $X \sim \mathcal{N}(0,1)$, $a,b \in \mathbb{R}$, $a > 0$ *és* $Y = aX + b$, *akkor* $X \sim \mathcal{N}(b,a^2)$.

Megjegyzés: ha $Y = aX$ ahol $a \in \mathbb{R}$, akkor $\mathbf{Var}Y = a^2\mathbf{Var}X$, vagyis $DY = |a|DX$ akkor is, ha $a < 0$.
Az előző tétel más szóval:

5. Tétel. *Ha* $Y \sim \mathcal{N}(m,\sigma^2)$, *akkor* $\frac{Y-m}{\sigma} \sim \mathcal{N}(0,1)$.

1.2. Függetlenség, feltételes valószínűség

1.2.1. Feltételes valószínűség, teljes valószínűség tétel

7. Definíció. *Legyenek A és B események (ugyanazon valószínűségi mezőn), és legyen $\mathbb{P}(A) > 0$. Ekkor a B esemény* feltételes valószínűsége *az A felétel mellett*

$$\mathbb{P}(B|A) := \frac{\mathbb{P}(A \text{ és } B)}{\mathbb{P}(A)}.$$

8. Definíció. *Legyen Ω valószínűségi mező (vagyis egy kísérlet összes lehetséges kimeneteleinek halmaza). Események egy $A_1, A_2, \dots \subset \Omega$ véges, vagy megszámlálhatóan végtelen rendszerét* teljes eseményrendszernek *nevezzük, ha*

- $\bigcup_k A_k = \Omega$, *vagyis az A_k-k közül valamelyik biztosan bekövetkezik, és*
- $\forall j \neq k \; A_k \cap A_j = \emptyset$, *vagyis az A_k-k közül semelyik kettő sem következik be egyszerre.*

Más szóval, az $A_1, A_2, \dots$ akkor és csak akkor teljes eseményrendszer, ha ezen események közül mindig pontosan egy következik be.

6. Tétel (Teljes valószínűség tétel). *Ha $A_1, A_2, \dots$ teljes eseményrendszer és B tetszőleges esemény, akkor*

$$\mathbb{P}(B) = \sum_k \mathbb{P}(A_k)\mathbb{P}(B|A_k).$$

A jobboldali összeget úgy kell érteni, hogy ha valamelyik k-ra $\mathbb{P}(A_k) = 0$, akkor a $\mathbb{P}(A_k)\mathbb{P}(B|A_k)$ szorzat is nulla (bár a második tényezőben a feltételes valószínűség nincs definiálva).

1.2.2. Páronkénti és teljes függetlenség

9. Definíció. *Legyenek A és B események (ugyanazon valószínűségi mezőn). A és B* függetlenek, *ha*

$$\mathbb{P}(A \text{ és } B) = \mathbb{P}(A)\mathbb{P}(B).$$

Megjegyzés: ha $\mathbb{P}(A) > 0$, akkor A és B függetlensége pontosan azt jelenti, hogy $\mathbb{P}(B|A) = \mathbb{P}(B)$, vagyis a B esemény bekövetkezési valószínűségére nézve az A bekövetkeztéből, mint feltételből semmi nem következik (A nem hordoz erre nézve információt). A függetlenség definíciója azonban ennél általánosabb, és $\mathbb{P}(A) = 0$ esetén is értelmes – igaz, nem túl izgalmas, hiszen ha $\mathbb{P}(A) = 0$, akkor A minden eseménytől független (beleértve önmagát is).

Legyen I tetszőleges inedxhalmaz (lehet véges, megszámlálhatóan végtelen, kontunuum számosságú, vagy akár annál is nagyobb), és legyen minden $i \in I$-re A_i egy esemény (ugyanazon valószínűségi mezőn).

10. Definíció. *Az $\{A_i\}_{i \in I}$ események* páronként függetlenek, *ha $\forall i, j \in I$, $i \neq j$-re A_i és A_j függetlenek, vagyis*

$$\forall i, j \in I,\ i \neq j\text{-re}\, \mathbb{P}(A_i A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j).$$

11. Definíció. *Az $\{A_i\}_{i \in I}$ események* teljesen függetlenek, *ha minden $K \subset I$* véges *részhalmazra*

$$\mathbb{P}\left(\bigcap_{k \in K} A_k\right) = \prod_{k \in K} \mathbb{P}(A_k).$$

Nyilvánvaló, hogy a teljes függetlenségből következik a páronkénti függetlenség. fordítva ez nem igaz. Tekintsük pl. azt a kísérletet, hogy kétszer feldobunk egy szabályos érmét, és a következő eseményeket:

- $A_1 := \{$az első dobás eredménye fej$\}$
- $A_2 := \{$a második dobás eredménye fej$\}$
- $A_3 := \{$a két dobás eredménye azonos$\}$

HF ellenőrizni, hogy $\{A_1, A_2, A_3\}$ páronként függetlenek , de nem teljesen függetlenek.

12. Definíció. *Legyenek X és Y valószínűségi változók (ugyanazon valószínűségi mezőn). X és Y* függetlenek, *ha tetszőleges $B_1, B_2 \subset R$ halmazokra az $\{X \in B_1\}$ és $\{Y \in B_2\}$ események függetlenek.*

Diszkrét val.változókra ez azt jelenti, hogy minden $x, y \in \mathbb{R} - re$ $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$. Folytoson val.változókra pedig azt jelenti, hogy az X és Y együttes sűrűségfüggvénye (amit most nem definiálok) $f(x, y) = f_X(x)f_Y(y)$ alakú.

13. Definíció. *Legyenek $X_1, X_2, \dots$ valószínűségi változók (ugyanazon valószínűségi mezőn). Ezek az X_k-k* teljesen függetlenek, *ha tetszőleges $B_1, B_2, \dots \subset \mathbb{R}$ halmazokra az $A_k := \{X_k \in B_k\}$ események teljesen függetlenek.*

Val.változók egy rendszerének páronkénti függetlensége hasonlóan definiálható, de nekünk a teljes függetlenség a fontos.

1.2.3. Valószínűségi változók összegének és szorzatának várható értéke, az összeg szórása

7. Tétel. *Legyen $X_1, X_2, \dots$ val.változók véges vagy megszámlálható sorozata. Ekkor*

$$\mathbf{E}(X_1 + X_2 + \dots) = \mathbf{E}X_1 + \mathbf{E}X_2 + \dots$$

(ha a jobb oldal létezik).

Hangsúlyozom, hogy a fenti tétel *tetszőleges* val.változókra igaz, és semmiféle függetlenséget nem kell feltenni.

8. Tétel. *Legyenek X és Y* független *valószínűségi változók. Ekkor $\mathbf{E}(X \cdot Y) = \mathbf{E}X \cdot \mathbf{E}Y$ (ha mindkét oldal létezik).*

Következmény:

9. Tétel. *Legyenek X és Y* független *valószínűségi változók. Ekkor*

$$\mathbf{Var}(X + Y) = \mathbf{Var}X + \mathbf{Var}Y,$$

vagyis

$$D(X + Y) = \sqrt{D^2X + D^2Y}.$$

Nagyon fontos, hogy független val.változóknak a szórás**négyzete** adódik össze, és nem pedig a szórása.

Ennek általánosítása a következő:

10. Tétel. *Legyen $X_1, X_2, \dots$ val.változók véges vagy megszámlálható sorozata. Ha az X_k-k* teljesen függetlenek, *akkor*

$$\mathbf{Var}\left(\sum_k X_k\right) = \sum_k \mathbf{Var}X_k,$$

vagyis

$$D\left(\sum_k X_k\right) = \sqrt{\sum_k D^2X_k}.$$

Mondanom se kell, hogy ebben a tételben a **teljes függetlenség**, mint feltétel, kulcsfontosságú.

2. Kiegészítések az ismétléshez

2.1. Teljes várható érték tétel

Ha X valószínűségi változó, A pedig esemény és $\mathbb{P}(A) > 0$, akkor természetes módon definiálható X *feltételes eloszlása* az A feltétel mellett, úgy mint az az információ, hogy X az ő lehetséges értékei milyen *feltételes valószínűségekkel* veszi fel. Diszkrét X esetén ez a

$$p_{k|A} := \mathbb{P}(X = x_k|A)$$

valószínűségek ismeretét jelenti. Abszolut folytoson esetben pedig gondolhatunk egy $f(x|A)$ feltételes sűrűségfüggvényre, amit az definiál, hogy minden $a < b \in \mathbb{R}$-re

$$\mathbb{P}(X \in [a,b]|A) = \int_a^b f(x|A)\,\mathrm{d}x.$$

Ezek segítségével értelmes beszélni az X val.változó *feltételes várható értékéről* az A feltétel mellett:

- diszkrét esetben $\mathbf{E}(X_A) := \sum_k x_k \mathbb{P}(X = x_k|A)$
- absz.folytoson esetben $\mathbf{E}(X_A) := \int_{-\infty}^{\infty} x f(x|A)\,\mathrm{d}x$.

11. Tétel (Teljes várható érték tétel). *Legyen X val.változó és $A_1, A_2, \ldots$ teljes eseményrendszer. Ekkor*

$$\mathbf{E}X = \sum_k \mathbb{P}(A_k)\mathbf{E}(X|A_k)$$

(ha a jobb oldal létezik).

A jobboldali összeget úgy kell érteni, hogy ha valamelyik k-ra $\mathbb{P}(A_k) = 0$, akkor a $\mathbb{P}(A_k)\mathbf{E}(X|A_k)$ szorzat is nulla (bár a második tényezőben a feltételes várható érték nincs definiálva).

2.2. Poisson folyamat

Ez a fejezet a Poisson és az exponenciális eloszlás közötti meghitt jó viszonyról szól.

Móricka egy augusztusi éjszakán a hullócsillagokat nézi. Tízpercenként átlagosan egyet sikerül megcsodálnia. Minden $[t_1, t_2]$ időintevallumra jelölje $X_{[t_1,t_2]}$ az ezen időszakban látott hullócsillagok (véletlen) számát. Az időt mérjük percben, így $\lambda = \frac{1}{10}$ az időegységenként átlagosan látott hullócsillagok száma (mert „tízpercenként egy" az annyi mint „percenként $\frac{1}{10}$".

Rögzített t_1 és t_2 mellett $X_{[t_1,t_2]}$-t kézenfekvő Poisson eloszlással modellezni, hiszen nagyon sok meteor próbálkozik azzal, hogy pont ez idő alatt, és pont Móricka orra előtt égjen el a légkörben, de ez minden meteornak külön-külön csak nagyon kis val.séggel sikerül. $X_{[t_1,t_2]}$ pedig a sikeres próbálkozók száma. Ennek várható értéke $(t_2 - t_1) \cdot \lambda$, persze annál nagyobb, minél hosszabb intervallumról van szó. vagyis

$$X_{[t_1,t_2]} \sim Poi((t_2 - t_1)\lambda).$$

Ha nem csak egy rögzített $[t_1, t_2]$ időintervallumban nézzük a hullócsillagok számát, hanem *folyamatában* akarjuk látni az egymást követő megfigyelések időpontjait, akkor a **Poisson-folyamathoz** jutunk.

Jelölések:

- Legyenek $\sigma_1, \sigma_2, \sigma_3, \ldots$ az egymást követő hullócsillagok megfigyelési időpontjai, a $t = 0$ időponttól kezdve.

- Legyen $\tau_1 = \sigma_1$, $\tau_2 = \sigma_2 - \sigma_1$, $\tau_3 = \sigma_3 - \sigma_2$, $\dots$, $\tau_n = \sigma_n - \sigma_{n-1}$, vagyis τ_n azt jelöli, hogy mennyit kell még várni az n-edik hullócsillagra az $n-1$-edik után.
- $0 \le t \le s$-re legyen $X_{[t,s]} := \sharp\{k : \sigma_k \in [t,s]\}$, vagyis a $[t,s]$ időszakban megfigyelt hullócsillagok száma.

12. Tétel. *A következő két állítás ekvivalens:*

a.) $\tau_1, \tau_2, \tau_3, \dots$ *függetlenek és azonos exponenciális eloszlásúak* λ *paraméterrel:* $\tau_k \sim Exp(\lambda)$

b.) Minden $t \le s$*-re* $X_{[t,s]} \sim Poi((s-t)\lambda)$*, továbbá ha a* $[t_1, s_1]$, $[t_2, s_2]$, $\dots [t_n, s_n]$ *intervallumok diszjunktak, akkor az* $X_{[t_1,s_1]}$, $X_{[t_2,s_2]}$, $\dots$, $X_{[t_n,s_n]}$ *val.változók teljesen függetlenek.*

14. Definíció. *A fenti tulajdonságú rendszerben a* $t \mapsto X_{[0,t]}$ *véletlen függvényt* λ *intenzitású* Poisson folyamatnak *nevezzük. A* $\{\sigma_1, \sigma_2, \dots\}$ *véletlen ponthalmazt pedig* λ *intenzitású* Poisson pontfolyamatnak *nevezzük.*

2.3. Az exponenciális eloszlás rátájáról

Itt érdemes megemlékezni az exponenciális eloszlás paraméterének intuitív jelentéséről: Ha $\tau \sim Exp(\lambda)$, várakozási idő, akkor *kicsi* Δt-re a definíció miatt

$$\mathbb{P}(\tau \le \Delta t) \approx \lambda \Delta t,$$

és az örökifjúság miatt

$$\mathbb{P}(\tau \le t + \Delta t | \tau > t) \approx \lambda \Delta t.$$

Vagyis ha egy exponenciális eloszlású várakozási idő leteltére várunk, és az idő még nem jött el, akkor a következő rövid Δt idő alatt való eljövetel valószínűsége persze Δt-vel arányos, az arányossági tényező pedig éppen a λ paraméter.

Ezért nevezzük az $Exp(\lambda)$ eloszlás λ paraméterét az exponenciális eloszlás *rátájának.*

3. Val.változók sorozatainak konvergenciája

15. Definíció (Erős konvergencia). *Legyenek* X *és* $X_1, X_2, \dots$ *val.változók ugyanazon az* Ω *valószínűségi mezőn. Azt mondjuk, hogy* X_n majdnem biztosan *konvergál* X*-hez, avagy* X_n 1 valószínűséggel *konvergál* X*-hez, avagy* erősen *konvergál* X*-hez, ha*

$$\mathbb{P}(\{X_n \to X\}) = 1.$$

Jelölése:

$$X_n \overset{m.b.}{\to} X$$

Ebben a definícióban $\{X_n \to X\}$ azt az eseményt jelenti, hogy az X_n véletlen számsorozat tart az X véletlen számhoz, vagyis

$$\{X_n \to X\} = \{\omega \in \Omega : X_n(\omega) \to X(\omega)\}.$$

Ebben rögzített $\omega \in \Omega$-ra $X_n(\omega)$ már egyszerű valós számsorozat, $X(\omega)$ pedig valós szám, így a konvergencia a régi analízis-beli számsorozat-konvergencia.

16. Definíció (Gyenge konvergencia 1). *Legyenek* X *és* $X_1, X_2, \dots$ *val.változók, nem feltétlen egyazon valószínűségi mezőn. Legyen* X *eloszlásfüggvénye* F*, illetve* X_n *eloszlásfüggvénye* F_n*. Azt modjuk, hogy* X_n gyengén konvergál X*-hez, avagy* eloszlásban konvergál X*-hez, ha*

$$F_n(x) \to F(x)$$

teljesül minden olyan $x \in \mathbb{R}$*-re, ahol* F *folytonos.*

Jelölése:

$$X_n \Rightarrow X$$

A „val.változók gyenge konvergenciája” elnevezés kicsit félrevezető, mert ez a konvergencia láthatóan csak az **eloszlások tulajdonsága**. Ahhoz, hogy értelmes legyen róla beszélni, nem hogy közös val-ségi mező nem szükséges, amin az X_n-ek és az X élnek, de még csak val.változókra sincs szükség: elég, ha megadjuk az eloszlásukat. Mondunk pl. olyat, hogy „val.változók egy sorozata gyengén konvergál egy eloszláshoz”, vagy hogy „eloszlások egy sorozata konvegál egy határeloszláshoz” Ezt mindenki értse jól - a pontos szóhasználatú definíciókat nem írom ide.

A gyenge konvergencia fenti definíciója ekvivalens az alábbival:

17. Definíció (Gyenge konvergencia 2). *Legyenek X és $X_1, X_2, \dots$ val.változók, nem feltétlen egyazon valószínűségi mezőn. Azt modjuk, hogy X_n* gyengén konvergál *X-hez, avagy* eloszlásban konvergál *X-hez, ha minden $g : \mathbb{R} \to \mathbb{R}$ korlátos és folytonos próbafüggvényre*

$$\mathbf{E}g(X-n) \to \mathbf{E}g(X).$$

Mivel g korlátos és folytonos, ezek a várható értékek mindig léteznek és végesek. Rögzített g mellett $\mathbf{E}g(X-n)$ valós számsorozat, $\mathbf{E}g(X)$ pedig valós szám.

13. Tétel. *Ha $X_n \overset{m.b.}{\to} X$, akkor $X_n \Rightarrow X$.*

Ennek megfordítása olyannyira nem igaz, hogy sokszor, amikor $X_n \Rightarrow X$, erős konvergenciáról nincs is értelme beszélni, mert nincs közös valószínűségi mező.

3.1. Nagy számok törvényei

Ebben a fejezetben legyen $X_1, X_2, \dots$ **teljesen független és azonos eloszlású** val.változók sorozata, és legyen

$$S_n = \frac{X_1 + \cdots + X_n}{n}$$

közülük az eslő n-nek az átlaga. A nagy számok törvényei azt mondják, hogy $\frac{S_n}{n}$ bizonyos feltételek mellett, bizonyos értelemben tart az X_k-k közös várható értékéhez.

3.1.1. Markov és Csebisev egyenlőtlenség

Az alábbi két tétel nagyon egyszerű, hasznos és univerzális eszköz val.változókkal kapcsolatos valószínűségek becslésére

14. Tétel (Markov egyenlőtlenség). *Legyen X* **nemnegatív** *val.változó, és $t \in \mathbb{R}^+$, Ekkor*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbf{E}X}{t}.$$

15. Tétel (Csebisev egyenlőtlenség). *Legyen X val.változó, aminek várható értéke létezik: $m = \mathbf{E}X \in \mathbb{R}$. Ekkor $t \in \mathbb{R}^+$-ra*

$$\mathbb{P}(|X-m| \geq t) \leq \frac{\mathbf{Var}X}{t^2}.$$

3.1.2. Nagy számok gyenge törvénye

A gyenge konvergencia definíciója átírható kicsit egyszerűbbre abban a speciális esetben, amikor a határérték val.változó konstans, vagyis egy determinisztikus szám:

1. Lemma. *Legyenek $Y_1, Y_2, \dots$ val.változók, és legyen $Y \equiv c$ determinisztikus. $Y_n \Rightarrow c$ akkor és csak akkor, ha minden $\varepsilon > 0$-ra*

$$\mathbb{P}(|Y_n - c| > \varepsilon) \to 0$$

amint $n \to \infty$.

A val.változók összegének várható értékéről és szórásnégyzatéről szóló tételek azonnali következménye, hogy

2. Lemma. *Ha $X_1, X_2, \ldots, X_n$ teljesen függetlenek és azonos eloszlásúak, $m = \mathbf{E}X_k \in \mathbb{R}$ a közös várható értékük és $\sigma^2 = \mathbf{Var}X_k$ a közös szórásnégyzetük és $S_n = X_1 + X_2 + \cdots + X_n$, akkor*

$$\mathbf{E}\frac{S_n}{n} = m$$

és

$$\mathbf{Var}\frac{S_n}{n} = \frac{\sigma^2}{n}.$$

A fenti két lemmából és a Csebisev egyenlőtlenségből kijön a nagy számok gyenge törvényének könnyebbik változata:

16. Tétel (Nagy számok gyenge törvénye, véges szórású eset). *Legyen $X_1, X_2, \ldots$* ***teljesen független és azonos eloszlású*** *val.változók sorozata, $m = \mathbf{E}X_k \in \mathbb{R}$ a közös várható értékük és $\sigma^2 = \mathbf{Var}X_k$ a közös szórásnégyzetük. Legyen $S_n = \frac{X_1+\cdots+X_n}{n}$.* ***Tegyük fel továbbá,*** *hogy $\sigma^2 < \infty$. Ekkor*

$$\frac{S_n}{n} \Rightarrow m.$$

Később, sokkal erősebb eszközökkel (lásd karakterisztikus függvény módszer), be fogjuk látni az általános esetet:

17. Tétel (Nagy számok gyenge törvénye, általános eset). *Legyen $X_1, X_2, \ldots$* ***teljesen független és azonos eloszlású*** *val.változók sorozata, $m = \mathbf{E}X_k \in \mathbb{R}$ a közös várható értékük. Legyen $S_n = \frac{X_1+\cdots+X_n}{n}$. Ekkor*

$$\frac{S_n}{n} \Rightarrow m.$$

3.1.3. Borel-Cantelli lemmák

Legyen $A_1, A_2, \ldots$ események végtelen sorozata egyazon Ω valószínűségi mezőn. Vezessük be a következő eseményt:

$$\limsup_n A_n := \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} A_k.$$

Ez pont azt az eseményt jelölni, hogy az A_k-k közül végtelen sok bekövetkezik, hiszen $\bigcup_{k \geq n} A_k$ azt jelöli, hogy az n-ediktől kezdve is bekövetkezik még legalább egy. Vagyis

$$\limsup_n A_n = \{\omega \in \Omega \,:\, \exists \text{ végtelen sok } k, \text{ hogy } \omega \in A_k\}.$$

Megjegyzés: Ez az esemény az elnevezését onnan kapta, hogy ha χ_A-val jelöljük egy $A \subset \Omega$ esemény indikátorfüggvényét, vagyis $\chi_A : \Omega \to \mathbb{R}$ olyan, hogy $\omega \in \Omega$-ra

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

akkor minden $\omega \in \Omega$-ra

$$\chi_{\limsup_n A_n}(\omega) = \limsup_n \chi_{A_n}(\omega).$$

Ha valaki ezt nem érti, nem baj. Hogy mit jelent a $\limsup$ fogalma számsorozatok esetén, lásd az elemi analízis könyveket vagy egy lexikont.

A Borel-Cantelli lemmák arról szólnak, hogy az A_n-ek közül végtelen soknak bekövetkezése vagy be nem következése azon múlik, hogy a $\mathbb{P}(A_n)$ valószínűségek elég gyorsan tartanak-e nullához. Ha nem „elég gyorsan”, akkor majdnem biztosan csak véges sok A_n következhet be, ellenkező esetben pedig végtelen sok is hajlamos bekövetkezni.

18. Tétel (Első Borel-Cantelli lemma). *Ha* $\sum_{n=1}^{\infty}\mathbb{P}(A_n)<\infty$*, akkor* $\mathbb{P}(\limsup_n A_n)=0$*.*

Ezen tétel „ellenkezője" csak azon igen erős plussz-feltevéssel igaz, hogy az A_n-ek függetlenek:

19. Tétel (Második Borel-Cantelli lemma). *Ha az* A_n *események teljesen függetlenek és* $\sum_{n=1}^{\infty}\mathbb{P}(A_n)=\infty$*, akkor* $\mathbb{P}(\limsup_n A_n)=1$*.*

3.1.4. Nagy számok erős törvénye

Az első Borel-Cantelli lemmára és aMarkov egyenlőtlenségre építve, némi számolással, de különösebb trökk nélkül bebizonyítható a következő tétel:

20. Tétel (Nagy számok erős törvénye, véges negyedik momentum esete). *Legyen* $X_1, X_2, \ldots$ ***teljesen független és azonos eloszlású*** *val.változók sorozata,* $m=\mathbf{E}X_k\in\mathbb{R}$ *a közös várható értékük. Legyen* $S_n=\frac{X_1+\cdots+X_n}{n}$*.* ***Tegyük fel továbbá****, hogy* $\mathbf{E}(X_k^4)<\infty$*. Ekkor*

$$\frac{S_n}{n}\overset{m.b.}{\to} m.$$

Lényegesen nehezebben belátható az általános eset, amit csak kinyilatkoztatok:

21. Tétel (Nagy számok erős törvénye, általános eset). *Legyen* $X_1, X_2, \ldots$ ***teljesen független és azonos eloszlású*** *val.változók sorozata,* $m=\mathbf{E}X_k\in\mathbb{R}$ *a közös várható értékük. Legyen* $S_n=\frac{X_1+\cdots+X_n}{n}$*. Ekkor*

$$\frac{S_n}{n}\overset{m.b.}{\to} m.$$

3.2. Karakterisztikus függvény módszer

3.2.1. Definíció, tulajdonságok

18. Definíció. *Az* X *val.változó karakterisztikus függvénye a* $\Psi_{\mathbb{R}}\to\mathbb{C}$ *függvény, amit úgy definiálunk, hogy*

$$\Psi_X(t):=\mathbf{E}\left(e^{itX}\right).$$

Itt $i=\sqrt{-1}$ *a komplex egységgyök.*

A karakterisztikus függvény mindig létezik minden $t\in\mathbb{R}$-re

Elemi tuladonságok:

1.) A karakterisztikus függvény igazából az *eloszlás* tulajdonsága: a val.változótól csak annak eloszlásán keresztül függ.

2.) A karakterisztikus függvény tényleg karakterizálja az eloszlást, vagyis ha az X és Y val.változók karakterisztikus függvénye azonos, akkor ők azonos eloszlásúak.

3.) $\psi_X(0)=1$ és ψ_x folytonos $\mathbb{R}$-en.

4.) ha $c\in\mathbb{R}$, akkor $\Psi_{X+c}(t)=e^{itc}\Psi_X(t)$.

5.) ha $c\in\mathbb{R}$, akkor $\Psi_{cX}(t)=\Psi_X(ct)$.

Kiszámolása:

- Ha X diszkrét és $p_k=\mathbb{P}(X=x_k)$ az elsozlása, akkor $\Psi_X(t)=\sum_k p_k e^{itx_k}$
- Ha X abszolut folytonos és sűrűségfüggvénye f, akkor $\Psi_X(t)=\int_{-\infty}\infty f(x)e^{itx}\,\mathrm{d}x$

Az alábbi tételek mind egy-egy rettentően fontos tulajdonságról szólnak:

22. Tétel (Karakterisztikus függvény és várható érték). *Ha* $\mathbf{E}X$ *létezik és véges, akkor* Ψ_X *folytonosan differenciálható és* $\Psi'_X(0) = i\mathbf{E}X$.

23. Tétel (Karakterisztikus függvény és magasabb momentumok). *Legyen* $k \in \mathbb{N}$. *Ha* $\mathbf{E}(X^k)$ *létezik és véges, akkor* Ψ_X k*-szor folytonosan differenciálható és a* k*-adik deriválta nulla helyen* $\Psi_X^{(k)}(0) = i^k \mathbf{E}X^k$.

24. Tétel (Karakterisztikus függvény és konvolúció). *Ha* X *és* Y *függetlenek, akkor*

$$\Psi_{X+Y}(t) = \Psi_X(t) \cdot \Psi_Y(t).$$

Következmény: ha $X_1, X_2, \ldots, X_n$ *függetlenek és azonos eloszlásúak, továbbá* $S_n = X_1 + \cdots + X_n$, *akkor*

$$\Psi_{S_n}(t) = (\Psi_X(t))^n$$

A legfontosabb tulajdonság viszont az, hogy a karakterisztikus függvények konvergenciája ekvivalens a gyenge konvergenciával:

25. Tétel (Folytonossági tétel). *Legyek* X *valamint* $X_1, X_2, \ldots$ *val.változók, melyek karakterisztikus függvényei* Ψ *illetve* $\Psi_1, \Psi_2, \ldots$. *Ezekkel a jelölésekkel* $X_n \Rightarrow X$ *akkor és csak akkor, ha* $\Psi_n \to \Psi$ *pontonként, vagyis minden* $t \in \mathbb{R}$*-re* $\Psi_n(t) \to \Psi(t)$.

3.2.2. Nevezetes eloszlások karakterisztikus függvénye

- Ha $X \sim B(p)$, akkor $\Psi_X(t) = q + pe^{it}$
- Ha $X \sim Bin(n,p)$, akkor $\Psi_X(t) = (q + pe^{it})^n$
- Ha $X \sim Geom(p)$, akkor $\Psi_X(t) = \frac{pe^{it}}{1-qe^{it}}$
- Ha $X \sim PesszGeom(p)$, akkor $\Psi_X(t) = \frac{p}{1-qe^{it}}$
- Ha $X \sim Poi(\lambda)$, akkor $\Psi_X(t) = e^{\lambda(e^{it}-1)}$
- Ha $X \sim Uni(a,b)$, akkor $\Psi_X(t) = \frac{e^{itb}-e^{ita}}{it(b-a)}$
- Ha $X \sim Exp(\lambda)$, akkor $\Psi_X(t) = \frac{\lambda}{\lambda - it}$
- Ha $X \sim \mathcal{N}(0,1)$, akkor $\Psi_X(t) = e^{-\frac{t^2}{2}}$
- Ha $X \sim \mathcal{N}(m, \sigma^2$, akkor $\Psi_X(t) = e^{itm}e^{-\frac{\sigma^2 t^2}{2}}$

3.2.3. Nagy számok gyenge törvénye, újra

A folytonossági tételből – felhasználva a karakterisztikus függvény kapcsolatát a konvolúcióval és a várható értékkel, igyakorlás képpen kijön a nagy számok gyenge törvényének általános esetének bizonyítása (amit most nem írok le):

26. Tétel (Nagy számok gyenge törvénye, általános eset). *Legyen* $X_1, X_2, \ldots$ ***teljesen független és azonos eloszlású*** *val.változók sorozata,* $m = \mathbf{E}X_k \in \mathbb{R}$ *a közös várható értékük. Legyen* $S_n = \frac{X_1 + \cdots + X_n}{n}$. *Ekkor*

$$\frac{S_n}{n} \Rightarrow m.$$

3.2.4. Centrális határeloszlás tétel

A fenti gyakorlás után, lényegében ugyanazt a számolást megismételve kijön a fejezet fénypontja:

27. Tétel (Centrális határeloszlás tétel). *Legyen* $X_1, X_2, \dots$ ***teljesen független és azonos eloszlású*** *val.változók sorozata,* $m = \mathbf{E}X_k \in \mathbb{R}$ *a közös várható értékük és* $\sigma^2 = \mathbf{Var}X_k$ *a közös szórásnégyzetük. Tegyük fel, hogy* $\sigma^2 < \infty$*. Legyen* $S_n = \frac{X_1+\dots+X_n}{n}$*. Ekkor*

$$\frac{S_n - nm}{\sqrt{n}} \Rightarrow \mathcal{N}(0,1).$$

Ebbe a tételbe beírva a gyenge konvergencia eloszlásfüggvényes definícióját, azt kapjuk, hogy a tétel feltételei mellett, minden rögzített $x \in \mathbb{R} - re$

$$\mathbb{P}\left(\frac{S_n - nm}{\sqrt{n}} \leq x\right) \to \Phi(x)$$

amint $n \to \infty$, ahol Ψ a standard normális eloszlásfüggvény.

Alkalmazása: ha $X_1, X_2, \dots, X_n$ független azonos eloszlásúak m várható értékkel és σ szórással és S_n az összegük, akkor

$$\mathbb{P}\left(S_n \in [a,b]\right) \approx \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right).$$

Hogy a közelítés milyen $[a,b]$-re és milyen n-re mennyire jó, az más kérdés.