4. Folytonos valószínűségi változók

Eddig olyan valószínűségi változókról beszéltünk, amik értékei egy véges, vagy megszámlálhatóan végtelen halmazt alkotnak, tipikusan az egész számok valamely részhalmazát. Vannak viszont olyan véletlen mennyiségek, amiket célszerű úgy modellezni, hogy bármilyen valós értéket felvehessenek. Ilyen sok fizikai mennyiség, vagy valamely történés bekövetkezéséig eltelt idő. Most ezekről lesz szó.

4.1. Eloszlásfüggvény

A feladatokban már előfordult, hogy választottunk egy számot „egyenletesen véletlenszerűen” egy intervallumból, esetleg egy pontot egy kétdimenziós alakzatból. Ezek éppen olyan véletlen mennyiségek – a múlt heti szóhasználattal valószínűségi változók –, amik nem diszkrétek, azaz nem csak egy megszámlálhatóan végtelen halmazból vesznek fel értékeket.

Ez miért probléma? Amikor X egy diszkrét valószínűségi változó, akkor őt le tudjuk írni az **eloszlás**a segítségével, vagyis a

$$\mathbb{P}(X = k_1),\ \mathbb{P}(X = k_2),\ \dots$$

0 és 1 közti számokból álló sorozattal, ahol $k_1, k_2, \dots$ az X lehetséges értékei. Legyen most X egy egyenletesen véletlen szám a $[0, 1]$ intervallumból. Ez alatt azt értjük, hogy X olyan valószínűségi változó, amire $\mathbb{P}(X \leq t) = t$ ha $t \in [0, 1]$, például $\mathbb{P}(X \leq \frac{1}{2}) = \frac{1}{2}$. Ekkor $\mathbb{P}(X = k) = 0$ bármilyen k esetén, vagyis a fenti (diszkrét értelemben vett) eloszlás lényegében semmit nem mond a valószínűségi változóról.

Erre mondhatnánk, hogy „Oké, de a változót akkor is leírja, hogy a $[0, 1]$-ből veszi fel az értékeit, így elég, ha ezzel a tulajdonsággal hivatkozunk rá.” A következő példa mutatja, hogy miért nem.

4.1.1. Példa. Legyen X egy egyenletesen véletlen szám a $[0, 1]$ intervallumból, és nézzük az $Y = X^2$ valószínűségi változót. Ekkor Y értékei szintén a $[0, 1]$ intervallumból valók, de Y mégsem ugyanúgy működik, mint X. Hiszen egyrészt $\mathbb{P}(X \leq \frac{1}{2}) = \frac{1}{2}$, másrészt

$$\mathbb{P}\Big(Y \leq \frac{1}{2}\Big) = \mathbb{P}\Big(X^2 \leq \frac{1}{2}\Big) = \mathbb{P}\Big(X \leq \frac{1}{\sqrt{2}}\Big) = \frac{1}{\sqrt{2}}.$$

Vagyis Y nagyobb eséllyel vesz fel kisebb értékeket, az „eloszlása jobban koncentrálódik a 0 körül, mint X eloszlása”. Mindjárt tisztázzuk, hogy ez mit is jelent.

Általánosan, az X eloszlását az ún. eloszlásfüggvénye segítségével írhatjuk le.

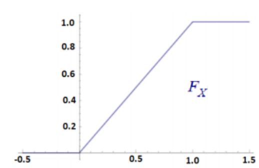
4.1.2. Definíció. Legyen X valószínűségi változó. Ekkor az

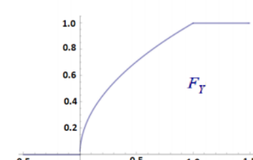
$$F_X : \mathbb{R} \to [0, 1] \qquad F_X(x) = \mathbb{P}(X < x)$$

függvényt az X **eloszlásfüggvény**ének hívjuk.

Vegyük észre, hogy $\{X < x\}$ eleme $\mathcal{F}$-nek, azaz esemény (mivel X valószínűségi változó), ezért van értelme beszélni a $\mathbb{P}(X < x)$ valószínűségről.

A fenti példában szereplő X és Y eloszlásfüggvényei:





$$F_X(x) = \mathbb{P}(X < x) = \begin{cases} 0 & \text{ha } x \leq 0, \\ x & \text{ha } x \in (0, 1], \\ 1 & \text{ha } x > 1, \end{cases}$$

$$F_Y(x) = \mathbb{P}(Y < x) = \mathbb{P}(X^2 < x) = \mathbb{P}(X < \sqrt{x}) = \begin{cases} 0 & \text{ha } x \leq 0, \\ \sqrt{x} & \text{ha } x \in (0, 1], \\ 1 & \text{ha } x > 1. \end{cases}$$

Nézzük, mit tud általánosan egy eloszlásfüggvény. Egyrészt világos, hogy

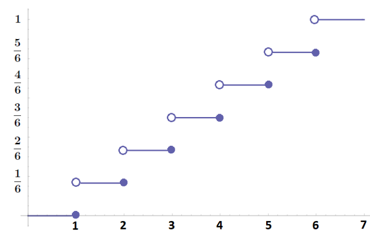
$$F_X(b) - F_X(a) = \mathbb{P}(X < b) - \mathbb{P}(X < a) = \mathbb{P}(a \leq X < b)$$

tetszőleges $a < b$ valós számokra, $\mathbb{P}$ additivitása miatt. Az eloszlásfüggvények karakterizálhatók is.

4.1.3. Állítás. *Egy $F : \mathbb{R} \to [0,1]$ függvény akkor és csak akkor eloszlásfüggvénye valamilyen valószínűségi változónak, ha*

(1) F (nem feltétlenül szigorúan) monoton növő,
(2) F balról folytonos, azaz minden x-re az F baloldali határértéke x-ben $F(x)$,
(3) $\lim_{x\mapsto-\infty} F(x) = 0$ és $\lim_{x\mapsto\infty} F(x) = 1$.

A balról folytonosság közel sem jelent folytonosságot. Például diszkrét valószínűségi változók eloszlásfüggvénye sosem folytonos. Igen, eloszlásfüggvény diszkrét esetben is mindig van, és ez is balról folytonos, de jobbról nem. Például egy kockadobás, mint valószínűségi változó, eloszlásfüggvényét lásd jobbra.



Bizonyítás. Legyen X valószínűségi változó, és $x < y$. Belátjuk, hogy F_X-re igaz a fenti három feltétel. Valóban, F_X monoton növő, hiszen $\{X < x\} \subseteq \{X < y\}$, ezért

$$F_X(x) = \mathbb{P}(X < x) \le \mathbb{P}(X < y) = F_X(y),$$

felhasználva a valószínűségi mezőről szóló 1.3 alfejezet következményét.

Nézzük a második tulajdonságot. Az F_X balról folytonossága ekvivalensen azt jelenti (átviteli elv), hogy bármely monoton növő $(x_n)_{n\in\mathbb{N}}$ sorozatra, amire $x_n \neq x$ és $x_n \to x$, teljesülnie kell, hogy $\lim_{n\to\infty} F_X(x_n) = F_X(x)$. Megmutatjuk, hogy ez valóban teljesül. Egyrészt

$$\lim_{n\to\infty} F_X(x_n) = \lim_{n\to\infty} \mathbb{P}(X < x_n) = \mathbb{P}(X < x_0) + \lim_{n\to\infty} \mathbb{P}\Big(\bigcup_{k=1}^{n} \{x_{k-1} \le X < x_k\}\Big),$$

felhasználva a valószínűség additivitását, illetve hogy $\{X < x_n\} = \{X < x_0\} \cup \{x_0 \le X < x_n\} = \{X < x_0\} \cup \bigcup_{k=1}^{n} \{x_{k-1} \le X < x_k\}$. A második tag a következőképp alakítható át:

$$\lim_{n\to\infty} \mathbb{P}\Big(\bigcup_{k=1}^{n} \{x_{k-1} \le X < x_k\}\Big) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(x_{k-1} \le X < x_k) = \mathbb{P}(x_0 \le X < x),$$

felhasználva a valószínűség σ-additivitását, illetve hogy $\cup_{k=1}^{\infty} \{x_{k-1} \le X < x_k\} = \{x_0 \le X < x\}$. Ezt visszahelyettesítve az előző egyenletbe azt kapjuk, hogy $\lim_{n\to\infty} F_X(x_n) = \mathbb{P}(X < x_0) + \mathbb{P}(x_0 \le X < x) = \mathbb{P}(X < x) = F_X(x)$.[16] Hasonló okból teljesül a harmadik feltétel is.[17]

Visszafelé, legyen adott F, ehhez keresünk egy megfelelő X valószínűségi változót, amire $F = F_X$. Legyen U egy egyenletesen véletlen szám a $[0,1]$ intervallumból.[18] Definiáljuk az X-et, mint

$$X = \inf\{y \in \mathbb{R} \mid U < F(y)\}.$$

Ekkor az X és az infimum definíciója miatt

$$\mathbb{P}(X < x) = \mathbb{P}\big(\inf\{y \in \mathbb{R} \mid U < F(y)\} < x\big) =$$
$$= \mathbb{P}\big(\text{van } y \in \mathbb{R},\ \text{amire } y < x \text{ és } U < F(y)\big).$$

Belátható, hogy ilyen y pontosan akkor létezik, ha $U < F(x)$. Valóban, ha van ilyen y, akkor $U < F(y) \le F(x)$, mivel F monoton növő. Megfordítva, ha $U < F(x)$, akkor F balról folytonossága miatt van x-hez elég közel egy y, amire $U < F(y)$ szintén teljesül. Következésképp:

$$\mathbb{P}(X < x) = \mathbb{P}\big(U < F(x)\big) = F(x),$$

hiszen $0 \le F(x) \le 1$, és a példában láttuk, hogy $\mathbb{P}(U < z) = z$ bármilyen $0 < z < 1$ szám esetén. Ezzel az állítást beláttuk. □

[16] Ezzel az érveléssel kaphatjuk, hogy az F_X jobboldali határértéke x-ben $\mathbb{P}(X \le x)$.

[17] A hasonló ok neve a valószínűség folytonossági tulajdonsága.

[18] Azt illenék bebizonyítani, hogy ilyen egyenletesen véletlen szám, mint valószínűségi változó, valóban létezik. Ennek a precíz leírása támaszkodik a Lebesgue-mérték fogalmára, így ettől most eltekintünk.

4.2. Sűrűségfüggvény

Most a valószínűségi változók másik fontos hozzárendelt függvényét vizsgáljuk: a sűrűségfüggvényét. Ennek többek közt az az oka, hogy az eloszlásfüggvény grafikonjáról nem feltétlenül könnyű leolvasni a valószínűségi változó tulajdonságait. Szó volt már például az X és X^2 változókról, ahol X egyenletesen véletlenszerű $[0,1]$-en. Meg tudjuk-e mondani az eloszlásfüggvény ábrája alapján, melyik szám 0,01 sugarú környezetében lesz a legnagyobb eséllyel X^2? Vagy hogy hányszor akkora eséllyel lesz az X^2 értéke az $\frac{1}{4}$ kis környezetében, mint $\frac{3}{4}$ kis környezetében?

Az első kérdésre hamar rávághatjuk, hogy 0-nál (egész pontosan 0,01-nél), hiszen ott nő a legmeredekebben az eloszlásfüggvény, más szavakkal az ottani x-ek járulnak hozzá leginkább az $F_{X^2}(x) = \mathbb{P}(X^2 < x)$ valószínűség növekedéséhez. A második kérdés valamivel trükkösebb, a válasz $\sqrt{3}$, amihez az eloszlásfüggvény érintőinek meredekségeit kell összehasonlítsuk, lásd lejjebb. Vegyük észre, hogy mindkét esetben a meredekségeket kell vizsgáljuk.

Ha valakinek rossz előérzete lenne az „érintő meredeksége” szókapcsolat hallatán, megnyugtatnám, deriválni fogunk. Hiszen az első példa is éppen azt mutatja, hogy az eloszlásfüggvény meredekségeinek függvénye, azaz deriváltfüggvénye lenne hasznunkra. Már amennyiben létezne, csakhogy 0-ban és 1-ben F_{X^2} nem deriválható. Ezt a problémát megkerülendő, a deriválás egyik általánosítását fogjuk használni, ami a „jó-lesz-az-úgy” filozófiát követve egyszerűen nem foglalkozik azzal, hogy a függvény egy-egy pontban nem deriválható (ha a függvény legalább folytonos). Ez nem fogja elrontani a számításainkat.

⚠ **4.2.1. Definíció.** Egy X valószínűségi változót **folytonos**nak nevezünk, ha létezik olyan nemnegatív, valós $f_X : \mathbb{R} \to \mathbb{R}$ függvény, amire az $\int_{-\infty}^{\infty} f_X(z)\mathrm{d}z$ improprius Riemann-integrál véges[19], és minden $x \in \mathbb{R}$ esetén

$$\int_{-\infty}^{x} f_X(z)\mathrm{d}z = F_X(x),$$

ahol az integrál improprius Riemann-integrál. Az f_X függvényt az X **sűrűségfüggvény**ének hívjuk.

A definíció nem túl konstruktív: a feltételként adott integrálokból nehéz kiszámolni f_X-et. Sőt, valójában azt sem biztosítja, hogy a sűrűségfüggvény egyértelmű legyen, hiszen ha f sűrűségfüggvénye X-nek, akkor az is sűrűségfüggvény, ha f-et megváltoztatjuk egy ponton (hiszen az integrálok nem változnak). Hogyan lehet akkor kiszámolni valamit, ami nem is egyértelmű?

4.2.2. Állítás. *Ha F_X folytonos és végessok pont kivételével mindenhol deriválható, akkor X folytonos valószínűségi változó, és az*

$$f(x) = \begin{cases} F'_X(x) & \text{ha } F_X \text{ deriválható } x\text{-ben,} \\ 0 & \text{egyébként} \end{cases} \qquad (x \in \mathbb{R})$$

függvény sűrűségfüggvénye X-nek.

4.2.3. Példa. Az állítás szerint a fenti $[0,1]$-en egyenletes eloszlású X esetében

$$f_X(x) = \begin{cases} 1 & \text{ha } 0 < x < 1, \\ 0 & \text{egyébként,} \end{cases} \qquad \text{és} \qquad f_{X^2}(x) = \begin{cases} \frac{1}{2\sqrt{x}} & \text{ha } 0 < x < 1, \\ 0 & \text{egyébként} \end{cases} \qquad (x \in \mathbb{R})$$

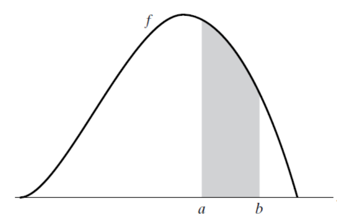
függvények sűrűségfüggvényei X-nek és X^2-nek. Szemléletesen a sűrűségfüggvény értékei annak az esélyét jelölik, hogy az X változó a x érték kis környezetébe esik[20]. Így az alfejezet elején feltett második kérdés válasza $f_{X^2}\left(\frac{1}{4}\right)/f_{X^2}\left(\frac{3}{4}\right) = \frac{1}{1}/\frac{1}{\sqrt{3}} = \sqrt{3}$.

[19] Általánosabban Lebesgue-integrálható sűrűségfüggvényről is beszélhetnénk. Most nem fogunk.

[20] Feltéve, hogy a sűrűségfüggvény épp folytonos. Ha egy-egy pontban megszüntethető szakadása van f_{X^2}-nek, akkor a pontbeli értékének nincs jelentéstartalma.

Megjegyzés. Mi most nem fogunk ilyen esetekkel foglalkozni, de egy valószínűségi változó lehet olyan, ami nem diszkrét, de nem is folytonos (például egy indikátor változó és egy folytonos szorzata). Sőt az sem igaz, hogy ha F_X folytonos, abból következne, hogy X is folytonos (bár a megfordítottja teljesül: egy folytonos változó eloszlásfüggvénye folytonos). Az ilyen kényelmetlen eseteket most tegyük félre.

A sűrűségfüggvény praktikus haszna jóval nagyobb annál, mint hogy a fenti kis környezetekről információval szolgál. Nézzük tehát a tulajdonságait.



4.2.4. Állítás. *Legyen X folytonos valószínűségi változó. Ekkor minden $a < b$ valós szám esetén*

$$\mathbb{P}(a < X < b) = \int_a^b f_X(x)\mathrm{d}x.$$

Bizonyítás. A valószínűség additivitása miatt

$$\mathbb{P}(a < X < b) = \mathbb{P}(X < b) - \mathbb{P}(X < a) - \mathbb{P}(X = a) =$$

$$= \int_{-\infty}^b f_X(x)\mathrm{d}x - \int_{-\infty}^a f_X(x)\mathrm{d}x - 0 = \int_a^b f_X(x)\mathrm{d}x,$$

ahol felhasználtuk, hogy az integrálás additív az integrálási intervallumban. □

Az eloszlásfüggvényhez hasonlóan a sűrűségfüggvények is karakterizálhatók.

4.2.5. Állítás. *Egy nemnegatív $f : \mathbb{R} \to \mathbb{R}$ függvényhez akkor és csak akkor létezik X folytonos valószínűségi változó, aminek az f sűrűségfüggvénye, ha f Riemann-integrálható és*

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1.$$

Az világos, hogy ha X folytonos valószínűségi változó, akkor f_X-re teljesül az egyenlet. A visszafelé irányt nem bizonyítjuk.

Feladat. Jelölje egy alkatrész élettartamát Z (órákban számolva). Ha Z eloszlásfüggvénye

$$F_Z(x) = \begin{cases} 0 & \text{ha } x \le 100, \\ 1 - \frac{100}{x} & \text{ha } x > 100, \end{cases}$$

akkor mi a valószínűsége, hogy az alkatrész nem romlik el az első 150 órában?

4.3. Várható érték, folytonos eset

Előző előadáson definiáltuk nemnegatív valószínűségi változók várható értékét:

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \sup_{\substack{Z \text{ egyszerű}, \\ Z \le X}} \mathbb{E}(Z), \qquad \text{ahol} \qquad \mathbb{E}(Z) = \sum_{k \in \mathrm{Ran}(X)} k \cdot \mathbb{P}(Z = k).$$

Vegyük észre, hogy az első definíció nem csak diszkrét esetre értelmes, folytonos valószínűségi változókra is ad valamit, csak nehezen látszik, hogy mit. Zavaró viszont benne a nemnegativitási feltétel. Ahhoz, hogy megszabaduljunk ettől a feltételtől, felhasználjuk, hogy a várható érték a fenti definícióval is additív, ahogy egyszerű valószínűségi változókra ezt már belláttuk.

4.3.1. Állítás. *Legyenek X és Y nemnegatív valószínűségi változók. Ekkor $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.*

Ezen azonosság segítségével definiálható nem feltétlenül nemnegatív (és nem is feltétlenül egyszerű) valószínűségi változókra is a várható érték.

4.3.2. Definíció. Legyen X tetszőleges valószínűségi változó. Jelölje $X^+ = \max(X, 0)$ az X **pozitív rész**ét, és $X^- = \max(-X, 0)$ az X **negatív rész**ét. Ekkor X^+ és X^- nemnegatív valószínűségi változók, továbbá $X = X^+ - X^-$. Ha $\mathbb{E}(X^+) < \infty$ vagy $\mathbb{E}(X^-) < \infty$, akkor legyen

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \mathbb{E}(X^+) - \mathbb{E}(X^-),$$

ami vagy egy valós szám, vagy $+\infty$, vagy $-\infty$. Ha $\mathbb{E}(X^+) = \mathbb{E}(X^-) = \infty$, akkor a várható értéket nem értelmezzük.

Bár ezek a definíciók értelmesek folytonos valószínűségi változókra is, nem konstruktívak, konkrét valószínűségi változó várható értékét nehéz így meghatározni. A következő állítás ezt hidalja át.

4.3.3. Állítás. *Legyen X folytonos valószínűségi változó, amire*

$$\int_{-\infty}^{\infty} |t| \cdot f_X(t) \mathrm{d}t < \infty.$$

Ekkor

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t \cdot f_X(t) \mathrm{d}t. \tag{2}$$

Az állítás feltételére azért van szükség, hogy kizárjuk azt az esetet, amikor $\mathbb{E}(X)$ nem definiált.

4.3.4. Definíció. Egy X valószínűségi változó **egyenletes eloszlás**ú az (a, b) intervallumon, ha sűrűségfüggvénye

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

Jelölése: $X \sim U(a; b)$.

Ez valóban sűrűségfüggvény, hiszen

$$\int_{-\infty}^{\infty} f(x) \mathrm{d}x = \int_a^b \frac{1}{b-a} \mathrm{d}x = \Big[\frac{x}{b-a}\Big]_a^b = \frac{b-a}{b-a} = 1.$$

Egy egyenletes eloszlású valószínűségi változó várható értéke

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) \mathrm{d}x = \int_a^b \frac{x}{b-a} \mathrm{d}x = \Big[\frac{x^2}{2(b-a)}\Big]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2},$$

ami intuitívan is világos: átlagosan az intervallum közepét kapjuk értékül.

Vegyük észre, mennyire hasonlítanak a diszkrét esetre vonatkozó (1) egyenlet és a folytonos esetre vonatkozó (2) egyenlet. Ez azért van, mert mindkettő a várható érték általános fogalmának a realizációja. Ez a lenti állításban is megnyilvánul, ahol párhuzamosan tárgyalhatjuk a két esetet.

4.3.5. Állítás (Transzformált várható értéke)**.** *Legyen X valószínűségi változó, és $g : \mathbb{R} \to \mathbb{R}$ függvény. Tegyük fel, hogy $\mathbb{E}(g(X))$ létezik. Ha X diszkrét, akkor*

$$\mathbb{E}(g(X)) = \sum_{j=1}^{\infty} g(k_j) \cdot \mathbb{P}(X = k_j),$$

ahol $\mathrm{Ran}(X) = \{k_1, k_2, \dots\}$. *Ha pedig X folytonos, akkor*

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) \mathrm{d}x.$$

4.3.6. Példa. Legyen X olyan valószínűségi változó, aminek $f_X : \mathbb{R} \to \mathbb{R}$ sűrűségfüggvényére teljesül, hogy $f(x) = 2e^{-2x}$ ha $x \in [0, \infty)$, és 0 egyébként. (Ellenőrizzük le, hogy ez valóban sűrűségfüggvény.) Ekkor

$$\mathbb{E}(e^X) = \int_{-\infty}^{\infty} e^x \cdot f_X(x) \mathrm{d}x = \int_0^{\infty} e^x 2e^{-2x} \mathrm{d}x = 2 \int_0^{\infty} e^{-x} \mathrm{d}x = 2.$$