

Diszkrét valószínűségi változók

Tartalom

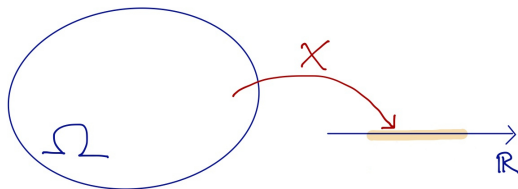
- ▶ Valószínűségi változók
- ▶ Diszkrét valószínűségi változók és súlyfüggvényeik
- ▶ Nevezetes diszkrét eloszlások és előfordulásuk:
indikátor/Bernoulli, binomiális, geometriai, Poisson, diszkrét egyenletes stb.
- ▶ Várható érték diszkrét esetben, a várható érték tulajdonságai, nevezetes diszkrét eloszlások várható értéke
- ▶ A nevezetes eloszlások tulajdonságai: binomiális eloszlás Poisson-approximációja, a geometriai eloszlás örökifjúsága
- ▶ Diszkrét valószínűségi változók transzformáltja, transzformált várható értéke
- ▶ Szórás(négyzet) és tulajdonságai, nevezetes diszkrét eloszlások szórása
- ▶ Valószínűségi változók függetlensége, diszkrét együttes eloszlások

Valószínűségi változók

Eddig: Véletlen kísérletek $\rightarrow$ eredmények/kimenetek, események, $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mező.

Az absztrakt valószínűségi mező helyett sokszor egyszerűbb **valós értékű függvényekkel** dolgozni.

Egy (egydimenziós) **valószínűségi változó** egy $X: \Omega \rightarrow \mathbb{R}$ függvény, amely az $\omega \in \Omega$ kimenetelhez az $X(\omega) \in \mathbb{R}$ valós számot rendeli hozzá. X -nek teljesítenie kell bizonyos feltételeket $\rightarrow$ pontos definíció: hamarosan.



Pl.: érmédobás modellezése, $\Omega = \{F, I\}$, $\mathcal{F} = \mathcal{P}(\Omega)$. $X(F) := 1$, $X(I) := 0$. Lesz: ekkor X egy valószínűségi változó, a dobott fejek száma a kísérletben.

Valószínűségi változók

Amikor csak véges sok lehetséges kimenetel van, sokszor gond nélkül tekinthetnénk akár minden $X: \Omega \rightarrow \mathbb{R}$ függvényt valószínűségi változónak.

Mégsem ezt tesszük, mert összetettebb Ω esetén fontos, hogy egyszerű $\mathbb{R}$ -beli halmazok (pl. intervallumok) Ω -beli *ősképei* **események** (vagyis az $\mathcal{F}$ σ -algebra elemei) legyenek.

Erre szolgálnak a következő definíciók.

(Def. 3.1.1. eleje.) Legyen adott egy $X: \Omega \rightarrow \mathbb{R}$ függvény. Adott $x \in \mathbb{R}$ esetén legyen

$$\{X < x\} \stackrel{\text{def.}}{=} \{\omega \in \Omega \mid X(\omega) < x\},$$

vagyis azon $\omega \in \Omega$ kimenetek halmaza, amelyre $X(\omega) < x$. Ezeket a halmazokat az X **nívóhalmazainak** hívjuk.

Vegyük észre: $\{X < x\}$ az $\mathbb{R}$ -beli $(-\infty, x)$ intervallum X általi ősképe (vagyis azon $\omega \in \Omega$ pontok halmaza, amelyek képe $(-\infty, x)$ -beli):

$$\{X < x\} = \{\omega \in \Omega \mid X(\omega) \in (-\infty, x)\} = X^{-1}((-\infty, x)) \subseteq \Omega.$$

Valószínűségi változók

Amikor csak véges sok lehetséges kimenetel van, sokszor gond nélkül tekinthetnénk akár minden $X: \Omega \rightarrow \mathbb{R}$ függvényt valószínűségi változónak.

Mégsem ezt tesszük, mert összetettebb Ω esetén fontos, hogy egyszerű $\mathbb{R}$ -beli halmazok (pl. intervallumok) Ω -beli *ősképei* **események** (vagyis az $\mathcal{F}$ σ -algebra elemei) legyenek.

Erre szolgálnak a következő definíciók.

(Def. 3.1.1. eleje.) Legyen adott egy $X: \Omega \rightarrow \mathbb{R}$ függvény. Adott $x \in \mathbb{R}$ esetén legyen

$$\{X < x\} \stackrel{\text{def.}}{=} \{\omega \in \Omega \mid X(\omega) < x\},$$

vagyis azon $\omega \in \Omega$ kimenetek halmaza, amelyre $X(\omega) < x$. Ezeket a halmazokat az X **nívóhalmazainak** hívjuk.

Vegyük észre: $\{X < x\}$ az $\mathbb{R}$ -beli $(-\infty, x)$ intervallum X általi ősképe (vagyis azon $\omega \in \Omega$ pontok halmaza, amelyek képe $(-\infty, x)$ -beli):

$$\{X < x\} = \{\omega \in \Omega \mid X(\omega) \in (-\infty, x)\} = X^{-1}((-\infty, x)) \subseteq \Omega.$$

Valószínűségi változók

$$\{X < x\} = X^{-1}((-\infty, x)) \subseteq \Omega.$$

Hasonlóképpen bevezethetők például az alábbi jelölések (ahol $x < y$ valós számok):

$$\{X \leq x\} = \{\omega \in \Omega \mid X(\omega) \leq x\} = X^{-1}((-\infty, x]),$$

$$\{X > x\} = \{\omega \in \Omega \mid X(\omega) > x\} = X^{-1}((x, \infty)),$$

$$\{X = x\} = \{\omega \in \Omega \mid X(\omega) = x\} = X^{-1}(\{x\}),$$

$$\{x < X \leq y\} = \{\omega \in \Omega \mid x < X(\omega) \leq y\} = X^{-1}((x, y]) \text{ stb.}$$

Sőt általában minden $A \subseteq \mathbb{R}$ halmazra

$$\{X \in A\} = \{\omega \in \Omega \mid X(\omega) \in A\} = X^{-1}(A).$$

Az $\{X < x\}$ jelölést felhasználva következzenek végre a valószínűségi változó definíciója.

Valószínűségi változók

(Def. 3.1.1. vége) Az $X: \Omega \rightarrow \mathbb{R}$ függvényt **valószínűségi változónak** nevezzük, ha $\forall x \in \mathbb{R}$ esetén $\{X < x\} \in \mathcal{F}$.

Vagyis egy $X: \Omega \rightarrow \mathbb{R}$ függvény akkor valószínűségi változó, ha bármely nívóhalmaza esemény.

- ▶ (3.1.2. Példa: példák valószínűségi változókra)
- ▶ Ha X val. változó, akkor bármely $x < y$ -ra $\{X \leq x\}, \{X \geq x\}, \{X > x\}, \{X = x\}, \{x \leq X \leq y\}, \{x < X < y\}, \{x \leq X < y\}, \{x < X \leq y\}$ szintén események. Ilyen alakú halmazok megszámlálható uniói szintén. (Ez a σ -algebra tulajdonságaiból adódik, de nem bizonyítjuk. Van közülük, ami könnyű, pl. $\{X \geq x\} = \overline{\{X < x\}}$.)

Valószínűségi változók

- ▶ Példa olyan $\Omega \rightarrow \mathbb{R}$ függvényre, ami egy adott eseménytérén/valószínűségi mezőn val. változó, és egy olyanra is, ami ott nem val. változó: össznépi gyak., 3.6. feladat.

Dobjunk fel egy szabályos érmét háromszor. Legyen az Ω eseménytér a 3 hosszú fej-írás sorozatok halmaza, és jelöljük az elemeit értelemszerűen: $FFF, FIF, \dots$ jelsorozatokkal. Definiáljuk az $X : \Omega \rightarrow \mathbb{R}$ függvényt az FFF kimenetelen 0-nak, és minden más kimenetel esetén az első „írás” jel sorszámának (pl. $X(FIF) = 2$).

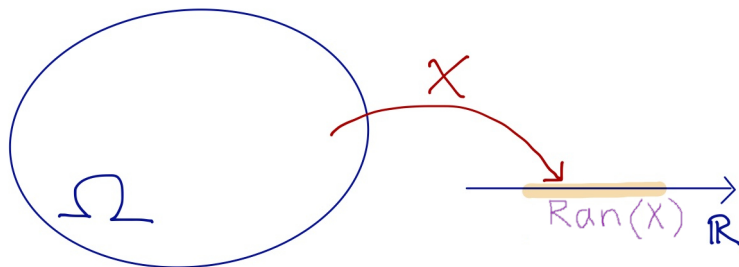
- (a) Mekkora az esélye, hogy X páratlan?
- (b) Definiáljuk Y -t ugyanúgy, mint X -et, azzal az eltéréssel, hogy $Y(FFF)$ véletlenszerűen vagy 0 vagy 1 értéket vesz fel.
Valószínűségi változó-e Y az Ω eseménytérén?

Valószínűségi változók

Egy X valószínűségi változó **értékkészlete** a lehetséges értékeinek halmaza:

$$\text{Ran}(X) = \{x \in \mathbb{R} \mid \exists \omega \in \Omega: X(\omega) = x\} \subseteq \mathbb{R}$$

(hasonlóan más függvényekéhez).



Diszkrét valószínűségi változók

(3.4.3. Definíció.) Egy X valószínűségi változó **diszkrét**, ha értékkészlete megszámlálható (azaz véges vagy megszámlálhatóan végtelen).

- ▶ Végtelen halmazok számossága $\rightarrow$ lásd BSz1 jegyzet!
Megszámlálhatóan végtelenek például az $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ halmazok és ezek végtelen részhalmazai, viszont az $\mathbb{R}, \mathbb{C}, \mathbb{R}^n, \mathcal{P}(\mathbb{N})$ halmazok nem.
- ▶ Nem diszkrét valószínűségi változó pl. az az X , amit „a $[0, 1]$ intervallumon találomra” választunk ($\rightarrow$ geometriai valószínűségi mező). Erről még lesz szó.
- ▶ X értékkészlete megszámlálhatóan végtelen $\Leftrightarrow$ létezik olyan páronként különböző számokból álló $(k_1, k_2, \dots)$ végtelen sorozat $\mathbb{R}$ -ben, hogy:
 - ▶ $\mathbb{P}(X = k_i) > 0$ minden i -re, és
 - ▶ $\sum_{i=1}^{\infty} \mathbb{P}(X = k_i) = 1$.
- ▶ Vegyük észre: az utóbbi esetben az $\{X = k_i\}$ események **teljes eseményrendszert alkotnak**, azzal az egy eltéréssel az eddigiektől, hogy nem véges sokan vannak.
A TVT és a Bayes-tétel igaz **megszámlálhatóan** végtelen sok eseményre is.

Diszkrét valószínűségi változók: súlyfüggvény

Egy diszkrét valószínűségi változót a **súlyfüggvény**ével a legegyszerűbb jellemezni, ami leírja, hogy a változó melyik lehetséges értékét mekkora valószínűséggel veszi fel.

(Definíció.) Legyen X diszkrét valószínűségi változó, amelynek értékkészletét jelölje $\text{Ran}(X)$. Ekkor X **súlyfüggvényének** nevezzük az alábbi $p_X : \text{Ran}(X) \rightarrow [0, 1]$ függvényt:

$$x \mapsto p_X(x) \stackrel{\text{def}}{=} \mathbb{P}(X = x)$$

- ▶ p_X kiterjeszthető $\mathbb{R} \rightarrow [0, 1]$ függvénnyé, ha $y \notin \text{Ran}(X)$ -re $p_X(y) = \mathbb{P}(X = y) = 0$ -ként definiáljuk.
- ▶ Fontos tulajdonság: $p_X(x) > 0$ minden $x \in \text{Ran}(X)$ -re,
 $\sum_{x \in \text{Ran}(X)} p_X(x) = \sum_{x \in \text{Ran}(X)} \mathbb{P}(X = x) = 1$.
- ▶ A súlyfüggvény csak diszkrét val. változó esetén hasznos! (Lásd később.)
- ▶ **Figyelem:** valószínűségi változónak nincs valószínűsége! Eloszlása (és eloszlásfüggvénye – lásd később) és diszkrét esetben súlyfüggvénye van.

Diszkrét valószínűségi változók: súlyfüggvény

Egy diszkrét valószínűségi változót a **súlyfüggvény**ével a legegyszerűbb jellemezni, ami leírja, hogy a változó melyik lehetséges értékét mekkora valószínűséggel veszi fel.

(Definíció.) Legyen X diszkrét valószínűségi változó, amelynek értékkészletét jelölje $\text{Ran}(X)$. Ekkor X **súlyfüggvényének** nevezzük az alábbi $p_X: \text{Ran}(X) \rightarrow [0, 1]$ függvényt:

$$x \mapsto p_X(x) \stackrel{\text{def}}{=} \mathbb{P}(X = x)$$

- ▶ p_X kiterjeszthető $\mathbb{R} \rightarrow [0, 1]$ függvénnyé, ha $y \notin \text{Ran}(X)$ -re $p_X(y) = \mathbb{P}(X = y) = 0$ -ként definiáljuk.
- ▶ Fontos tulajdonság: $p_X(x) > 0$ minden $x \in \text{Ran}(X)$ -re,
 $\sum_{x \in \text{Ran}(X)} p_X(x) = \sum_{x \in \text{Ran}(X)} \mathbb{P}(X = x) = 1$.
- ▶ A súlyfüggvény csak diszkrét val. változó esetén hasznos! (Lásd később.)
- ▶ **Figyelem:** valószínűségi változónak nincs valószínűsége! Eloszlása (és eloszlásfüggvénye – lásd később) és diszkrét esetben súlyfüggvénye van.

A legegyszerűbb valószínűségi változó, amelynek értéke nem is véletlen

A legegyszerűbb X diszkrét valószínűségi változó az, amelyik 1 valószínűséggel konstans, vagyis valamilyen $x \in \mathbb{R}$ -re

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

X értéke tehát nem is véletlen, hanem 1 valószínűséggel mindig az egyetlen x értéket veszi fel.

Meglehetősen unalmas valószínűségi változó, de sokszor fog szerepelni speciális/elfajuló esetként.

Indikátorváltozó/Bernoulli-eloszlás

A 3.1.2. példában láthattuk: ha $A \in \mathcal{F}$ esemény, akkor az A eseményhez tartozó $\mathbb{1}_A$ indikátorváltozó

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \text{ha } \omega \in A, \\ 0, & \text{ha } \omega \notin A. \end{cases}$$

Mivel

$$\mathbb{P}(\mathbb{1}_A = 1) = \mathbb{P}(\{\omega \in \Omega \mid \mathbb{1}_A(\omega) = 1\}) = \mathbb{P}(\{\omega \mid \omega \in A\}) = \mathbb{P}(A)$$

és hasonlóan

$$\mathbb{P}(\mathbb{1}_A = 0) = \mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A),$$

ezért a $\mathbb{1}_A$ val. változó súlyfüggvényének értékei:

$$p_{\mathbb{1}_A}(1) = \mathbb{P}(A), \quad p_{\mathbb{1}_A}(0) = 1 - \mathbb{P}(A).$$

Ha $\mathbb{P}(A) = p \in [0, 1]$, akkor tehát $p_{\mathbb{1}_A}(1) = p$, $p_{\mathbb{1}_A}(0) = 1 - p$.

Ha egy X val. változó súlyfüggvénye $p_X(1) = p$ és $p_X(0) = 1 - p$, akkor azt mondjuk, hogy X **Bernoulli-eloszlású** p paraméterrel.

Könnyű belátni: minden ilyen X egy indikátorváltozó, mégpedig az $\{X = 1\} \in \mathcal{F}$ eseményé (amelynek valószínűsége p).

Binomális eloszlás

A Bernoulli-eloszlás tehát egy olyan val. változó eloszlása, amely egy $p \in [0, 1]$ valószínűséggel sikeres kísérlet eredményét adja meg. ($p = 1/2$ esetén a siker lehet pl. az, ha egy szabályos érmével fejet dobunk, $p = 1/6$ esetén pedig az, ha egy szabályos kockával 6-ost.)

Kérdés: ha ugyanazt a kísérletet n -szer (egymástól függetlenül) elvégezzük, milyen valószínűséggel tapasztalunk $0, 1, \dots, n$ sikert?

Jelölje X a sikerek számát, ekkor X egy val. változó

$\text{Ran}(X) = \{0, 1, \dots, n\}$ értékkészlettel.

Ahhoz, hogy pontosan $k \in \text{Ran}(X)$ sikert kapjunk, az n kísérletből ki kell választanunk k -t, amelyik sikeres $\rightarrow \binom{n}{k}$ lehetőség.

Annak a valószínűsége, hogy ez a k konkrét kísérlet sikeres, a maradék $n - k$ pedig nem, $p^k(1 - p)^{n-k}$, ezért

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \forall k \in \{0, 1, \dots, n\}. \quad (\star)$$

(3.2.5. Definíció) Egy X val. változó **binomiális eloszlású** $n \in \mathbb{N}$ és $p \in [0, 1]$ paraméterekkel, ha $(\star)$ teljesül rá.

Binomális eloszlás

A Bernoulli-eloszlás tehát egy olyan val. változó eloszlása, amely egy $p \in [0, 1]$ valószínűséggel sikeres kísérlet eredményét adja meg. ($p = 1/2$ esetén a siker lehet pl. az, ha egy szabályos érmével fejet dobunk, $p = 1/6$ esetén pedig az, ha egy szabályos kockával 6-ost.)

Kérdés: ha ugyanazt a kísérletet n -szer (egymástól függetlenül) elvégezzük, milyen valószínűséggel tapasztalunk $0, 1, \dots, n$ sikert?

Jelölje X a sikerek számát, ekkor X egy val. változó

$\text{Ran}(X) = \{0, 1, \dots, n\}$ értékkel.

Ahhoz, hogy pontosan $k \in \text{Ran}(X)$ sikert kapjunk, az n kísérletből ki kell választanunk k -t, amelyik sikeres $\rightarrow \binom{n}{k}$ lehetőség.

Annak a valószínűsége, hogy ez a k konkrét kísérlet sikeres, a maradék $n - k$ pedig nem, $p^k(1 - p)^{n-k}$, ezért

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \forall k \in \{0, 1, \dots, n\}. \quad (\star)$$

(3.2.5. Definíció) Egy X val. változó **binomiális eloszlású** $n \in \mathbb{N}$ és $p \in [0, 1]$ paraméterekkel, ha $(\star)$ teljesül rá.

Binomiális eloszlás

(3.2.5. Definíció) Egy X val. változó **binomiális eloszlású** $n \in \mathbb{N}$ és $p \in [0, 1]$ paraméterekkel, ha

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad \forall k \in \{0, 1, \dots, n\}.$$

Jelölés: $X \sim B(n; p)$.

Honnan látjuk, hogy ez a p_X valóban súlyfüggvény?

A nemnegativitás világos. Ezenkívül teljesül, hogy

$$\sum_{k=0}^n p_X(k) = \sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1,$$

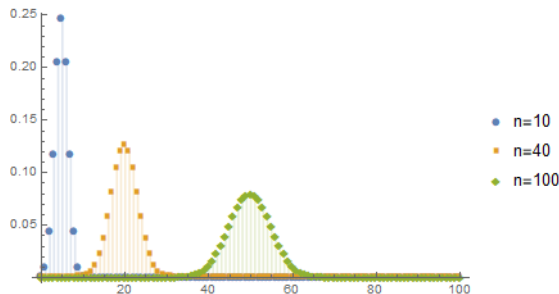
ahol az utolsó előtti egyenlőség a binomiális tétel miatt igaz.

Binomiális eloszlás

(3.2.5. Definíció) Egy X val. változó **binomiális eloszlású** $n \in \mathbb{N}$ és $p \in [0, 1]$ paraméterekkel, ha

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \forall k \in \{0, 1, \dots, n\}.$$

Jelölés: $X \sim B(n; p)$.



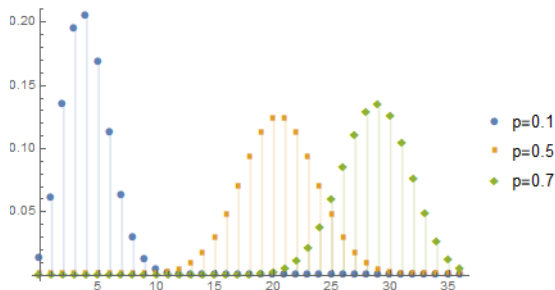
Ábra: n változtatása rögzített p mellett, $B(n; 1/2)$ eloszlás súlyfüggvénye $n = 10, 40, 100$ esetén. A súlyfüggvény np körül maximális.

Binomiális eloszlás

(3.2.5. Definíció) Egy X val. változó **binomiális eloszlású** $n \in \mathbb{N}$ és $p \in [0, 1]$ paraméterekkel, ha

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \forall k \in \{0, 1, \dots, n\}.$$

Jelölés: $X \sim B(n; p)$.



Ábra: p változtatása rögzített n mellett, $B(41; p)$ eloszlás súlyfüggvénye $p = 0,1, 0,5, 0,7$ esetén. A súlyfüggvény np körül maximális. $p = 0,5$ -re a súlyfüggvény szimmetrikus:

$$\mathbb{P}(k \text{ siker}) = \mathbb{P}(k \text{ kudarc}).$$

Binomiális eloszlás

(3.2.5. Definíció) Egy X val. változó **binomiális eloszlású** $n \in \mathbb{N}$ és $p \in [0, 1]$ paraméterekkel, ha

$$p_X(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad \forall k \in \{0, 1, \dots, n\}.$$

Jelölés: $X \sim B(n; p)$.

X tehát a sikerek száma n db egymástól független, p valószínűséggel sikeres kísérlet esetén.

Jelölje $k \in \{0, 1, \dots, n\}$ esetén A_k azt az eseményt, hogy a k -adik kísérlet sikeres. Ekkor $X = \mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n}$.

- ▶ Ahogy a függvények összeadása is pontonként történik, úgy a valószínűségi változók összeadását (összeszorozását, skalárral való szorzását stb.) is “ ω -nként” kell értelmezni, tehát:

$$X(\omega) = \mathbb{1}_{A_1}(\omega) + \dots + \mathbb{1}_{A_n}(\omega), \quad \omega \in \Omega.$$

- ▶ Így tehát $n = 1$ esetén a $B(1; p)$ eloszlás a p paraméterű Bernoulli-eloszlással egyezik meg.

Geometriai eloszlás

Ismételgessünk egy $p \in (0, 1)$ valószínűséggel sikeres kísérletet (pl. dobáljunk egy p valószínűséggel fejre mutató cinkelt pénzérmét, ahol a fej jelenti számunkra a sikert).

Láttuk: n kísérlet esetén a sikerek száma $\sim B(n; p)$.

Jelölje most X az **első sikeres kísérlet sorszámát** (ha a kísérlet ismétléseinek számát nem korlátozzuk). Mi X súlyfüggvénye?

Először is: $\text{Ran}(X) = \{1, 2, \dots\}$, a pozitív egész számok halmaza. X tehát megszámlálhatóan végtelen értékészletű val. változó.

Ahhoz, hogy pontosan a k -adik kísérlet legyen az 1. sikeres, ahol $k \geq 1$, az első $k - 1$ kísérletnek sikertelennek, a k -adiknak pedig sikeresnek kell lennie. A függetlenség miatt

$$p_X(k) = \mathbb{P}(X = k) = (1 - p)^{k-1} p, \quad \forall k \in \{1, 2, \dots\}. \quad (1)$$

(5.2.1. Definíció.) Az X valószínűségi változó **geometriai eloszlású** $p \in (0, 1)$ paraméterrel, ha (1) teljesül rá. **Jelölés:** $X \sim \text{Geo}(p)$.

Geometriai eloszlás

Ismételgessünk egy $p \in (0, 1)$ valószínűséggel sikeres kísérletet (pl. dobáljunk egy p valószínűséggel fejre mutató cinkelt pénzérmét, ahol a fej jelenti számunkra a sikert).

Láttuk: n kísérlet esetén a sikerek száma $\sim B(n; p)$.

Jelölje most X az **első sikeres kísérlet sorszámát** (ha a kísérlet ismétléseinek számát nem korlátozzuk). Mi X súlyfüggvénye?

Először is: $\text{Ran}(X) = \{1, 2, \dots\}$, a pozitív egész számok halmaza. X tehát megszámlálhatóan végtelen értékészletű val. változó.

Ahhoz, hogy pontosan a k -adik kísérlet legyen az 1. sikeres, ahol $k \geq 1$, az első $k - 1$ kísérletnek sikertelennek, a k -adiknak pedig sikeresnek kell lennie. A függetlenség miatt

$$p_X(k) = \mathbb{P}(X = k) = (1 - p)^{k-1} p, \quad \forall k \in \{1, 2, \dots\}. \quad (1)$$

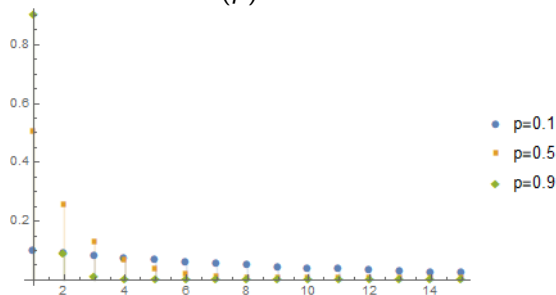
(5.2.1. Definíció.) Az X valószínűségi változó **geometriai eloszlású** $p \in (0, 1)$ paraméterrel, ha (1) teljesül rá. **Jelölés:** $X \sim \text{Geo}(p)$.

Geometriai eloszlás

(5.2.1. Definíció.) Az X valószínűségi változó **geometriai eloszlású** $p \in (0, 1)$ paraméterrel, ha

$$p_X(k) = \mathbb{P}(X = k) = (1 - p)^{k-1}p, \quad \forall k \in \{1, 2, \dots\}. \quad (2)$$

Jelölés: $X \sim \text{Geo}(p)$.



Ábra: a geometriai eloszlás sűrűfüggvénye $p = 0,1, 0,5, 0,9$ esetén. A sűrűfüggvény monoton csökkenő és $k = 1$ -ben pont p az értéke.

Geometriai eloszlás

(5.2.1. Definíció.) Az X valószínűségi változó **geometriai eloszlású** $p \in (0, 1)$ paraméterrel, ha

$$p_X(k) = \mathbb{P}(X = k) = (1 - p)^{k-1}p, \quad \forall k \in \{1, 2, \dots\}. \quad (2)$$

Jelölés: $X \sim \text{Geo}(p)$.

Megjegyzések:

- ▶ p_X tényleg súlyfüggvény, ez a végtelen mértani sor összegképletéből adódik (ellenőrzése házi feladat!).
- ▶ Az eloszlás a mértani (geometriai) sorozatról kapta a nevét. Nem összekeverendő a geometriai valószínűségi mezővel.
- ▶ Egyes források szokták X fenti eloszlását “optimista geometriai eloszlásnak” nevezni (kísérletek száma az 1. sikerig), $X - 1$ -ét pedig “pesszimista geometriai eloszlásnak” (kudarok száma az 1. siker előtt). Ebből a tárgyból geometriai eloszlás alatt mindig az “optimistát” értjük.

Geometriai eloszlás: örökifjúság

Egy lépkedős társasjátékban már nagyon régóta várunk arra, hogy először 6-ost dobjunk. Ez alapján azt reméljük, hogy a következő dobásunk most már tényleg biztos 6-os lesz. Megalapozott-e a reményünk? Kinek mi a tapasztalata ezzel kapcsolatban?

Jelölje X az első hatosdobás sorszámát, ekkor $X \sim \text{Geo}(p)$, ahol $p = 1/6$. Legyen $t, s \in \{1, 2, \dots\}$. Annak a valószínűsége, hogy t -edik kísérletig soha nem dobunk 6-ost:

$$\mathbb{P}(X > t) = \mathbb{P}(\text{első } t \text{ kísérlet sikertelen}) = (1 - p)^t.$$

És $s \geq 1$ -re annak a valószínűsége, hogy az $s + t$ -edik kísérletig soha nem dobunk 6-ost, feltéve, hogy az s -edikig nem dobtunk:

$$\mathbb{P}(X > s + t | X > s) = \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} = \frac{(1 - p)^{s+t}}{(1 - p)^s} = (1 - p)^t.$$

Vagyis a két valószínűség megegyezik:

$$\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t), \quad \forall s, t \in \{1, 2, \dots\}.$$

Vagyis: ha az első s kísérletből egyszer sem dobtunk 6-ost, akkor olyan, mintha most kezdődne előről a dobássorozatunk.

Geometriai eloszlás: örökifjúság

Egy lépkedős társasjátékban már nagyon régóta várunk arra, hogy először 6-ost dobjunk. Ez alapján azt reméljük, hogy a következő dobásunk most már tényleg biztos 6-os lesz. Megalapozott-e a reményünk? Kinek mi a tapasztalata ezzel kapcsolatban?

Jelölje X az első hatosdobás sorszámát, ekkor $X \sim \text{Geo}(p)$, ahol $p = 1/6$. Legyen $t, s \in \{1, 2, \dots\}$. Annak a valószínűsége, hogy t -edik kísérletig soha nem dobunk 6-ost:

$$\mathbb{P}(X > t) = \mathbb{P}(\text{első } t \text{ kísérlet sikertelen}) = (1 - p)^t.$$

És $s \geq 1$ -re annak a valószínűsége, hogy az $s + t$ -edik kísérletig soha nem dobunk 6-ost, feltéve, hogy az s -edikig nem dobtunk:

$$\mathbb{P}(X > s + t | X > s) = \frac{\mathbb{P}(X > s + t)}{\mathbb{P}(X > s)} = \frac{(1 - p)^{s+t}}{(1 - p)^s} = (1 - p)^t.$$

Vagyis a két valószínűség megegyezik:

$$\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t), \quad \forall s, t \in \{1, 2, \dots\}.$$

Vagyis: ha az első s kísérletből egyszer sem dobtunk 6-ost, akkor olyan, mintha most kezdődne előről a dobássorozatunk.

Geometriai eloszlás: örökifjúság

A geometriai eloszlás

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s), \quad \forall s, t \in \{1, 2, \dots\} \quad (3)$$

tulajdonságát **örökifjúságnak** (angolul: *memoryless property*) nevezzük (lásd később: 5.2.4. Definíció).

Annak a bizonyítását most láttuk, hogy a geometriai eloszlás örökifjú. Ennek egyfajta megfordítása is igaz:

(5.2.5. (1) Állítás.) Ha egy X nemkonstans valószínűségi változóra (3) teljesül és $\text{Ran}(X) = \{1, 2, \dots\}$, akkor $X \sim \text{Geo}(p)$, ahol $p = \mathbb{P}(X = 1)$.

(Bizonyítás)

A diszkrét eloszlások világában a geometriai az egyetlen örökifjú család. A folytonos eloszlásoknál fogunk találkozni az **exponenciális eloszlással**, amely szintén örökifjú és szorosan kapcsolódik a geometriai eloszláshoz is.

Diszkrét egyenletes eloszlás

(3.2.7. Definíció.) Egy X valószínűségi változó (diszkrét) egyenletes eloszlású egy n elemű $S \subset \mathbb{R}$ halmazon, ha

$$p_X(k) = \mathbb{P}(X = k) = \frac{1}{n}$$

minden $k \in S$ esetén.

Példák:

- ▶ $n = 6$, $S = \{1, 2, \dots, 6\}$: X egy szabályos kockadobás eredménye.
- ▶ $n = 2$, $S = \{0, 1\}$: X egy szabályos érmedobás eredménye fej=1, írás=0 elkódolással, azaz $X = \mathbb{1}_{\{\text{fej}\}}$.
(Általában a Bernoulli-eloszlás nem diszkrét egyenletes, csak ha $p = 1/2$!)
- ▶ $n = 1$, $S = \{x\}$: X egy 1 valószínűséggel konstans x értékű val. változó.

Poisson-eloszlás

A binomiális eloszlás pontosan modellezi azt a szituációt, ahol tudjuk, hogy hány kísérletet végzünk és azok milyen valószínűséggel sikeresek.

Vannak olyan szituációk, ahol egy bizonyos X véletlen mennyiség szintén tekinthető egy véletlen, sokszor egymástól **függetlenül** megismételt kísérletnél a sikerek számának, de:

- ▶ a kísérletek **száma nagy**,
- ▶ egy-egy kísérletnél **a siker valószínűsége kicsi**,
- ▶ és a kísérletek száma és a siker valószínűsége sem feltétlenül ismert pontosan.

Például:

- ▶ 100 évnél idősebb emberek száma Budapesten,
- ▶ telefonhívások száma Magyarországon ma délután 4 és 5 között,
- ▶ egy augusztusi estén egy óra alatt észlelt hullócsillagok száma,
- ▶ további (többé-kevésbé jó) példák → gyakorlaton.

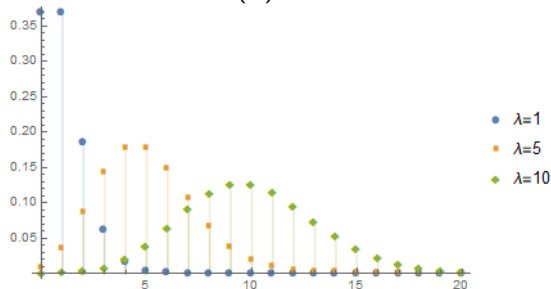
Ilyen esetekben X közelítőleg **Poisson-eloszlású**.

Poisson-eloszlás

(5.3.1. Definíció.) Az X valószínűségi változó Poisson-eloszlású $\lambda > 0$ paraméterrel, ha

$$p_X(k) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \forall k \in \{0, 1, 2, \dots\}$$

Jelölése: $X \sim \text{Pois}(\lambda)$.



Ábra: Poisson-eloszlás sűrűfüggvénye $\lambda = 1, 5, 10$ esetén. A sűrűfüggvény maximuma λ környékén van.

Poisson-eloszlás

(5.3.1. Definíció.) Az X valószínűségi változó Poisson-eloszlású $\lambda > 0$ paraméterrel, ha

$$p_X(k) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \forall k \in \{0, 1, 2, \dots\}$$

Jelölése: $X \sim \text{Pois}(\lambda)$.

- ▶ Ez tényleg egy súlyfüggvény. (Biz.: lásd előadáson)
- ▶ (5.3.2. Példa: kaszkádőr sérülései – a várható értéket még nem tanultuk, ezért kicsit módosítjuk a példát)

Intuíció (lásd előző dia): a Poisson-eloszlás olyan, mint egy binomiális eloszlás “nagy n -re és kis p -re”.

Ez nem csak heurisztika, hanem matematikailag is precízzé tehető (ideértve azt is, hogy n függvényében p -t hogyan kell megválasztani), amihez viszont segít, ha már ismerjük a **várható érték** fogalmát. Ezért erre a témára hamarosan visszatérünk.

Poisson-eloszlás

(5.3.1. Definíció.) Az X valószínűségi változó Poisson-eloszlású $\lambda > 0$ paraméterrel, ha

$$p_X(k) = \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad \forall k \in \{0, 1, 2, \dots\}$$

Jelölése: $X \sim \text{Pois}(\lambda)$.

- ▶ Ez tényleg egy súlyfüggvény. (Biz.: lásd előadáson)
- ▶ (5.3.2. Példa: kaszkádőr sérülései – a várható értéket még nem tanultuk, ezért kicsit módosítjuk a példát)

Intuíció (lásd előző dia): a Poisson-eloszlás olyan, mint egy binomiális eloszlás “nagy n -re és kis p -re”.

Ez nem csak heurisztika, hanem matematikailag is precízzé tehető (ideértve azt is, hogy n függvényében p -t hogyan kell megválasztani), amihez viszont segít, ha már ismerjük a **várható érték** fogalmát. Ezért erre a témára hamarosan visszatérünk.

Zipf-eloszlás és skálafüggetlen hálózatok (kiegészítő anyag)

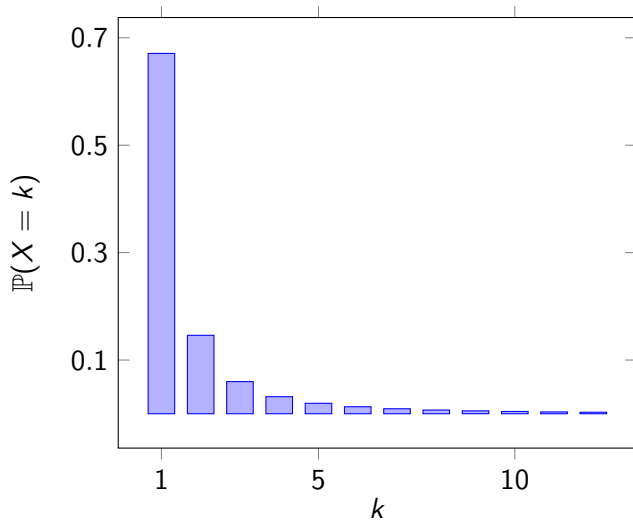
(Def.) Legyen $a > 1$. Egy X val. változó **Zipf-eloszlású** (avagy **Dzéta-eloszlású**) a paraméterrel, ha $\text{Ran}(X) = \{1, 2, \dots\}$ és

$$p_X(k) = \mathbb{P}(X = k) = \frac{k^{-a}}{Z(a)},$$

ahol $Z(a) := \sum_{k=1}^{\infty} k^{-a}$.

- ▶ p_X valóban egy valószínűségi eloszlás (mi $Z(a)$ szerepe?)
- ▶ az a paraméter nem valószínűség – más a szerepe, mint p -nek a geometriai eloszlás esetén
- ▶ (Mi a probléma az $a \leq 1$ esettel?)
- ▶ $k \mapsto p_X(k)$ „hatványtörvény szerint cseng le”
- ▶ (Összehasonlítás a geometriai eloszlással)

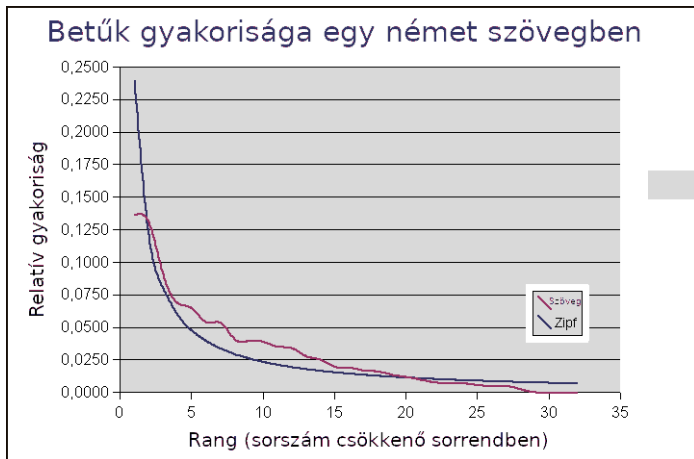
Zipf-eloszlás $a = 2,2$ esetén



A Zipf-eloszlás előfordulása

(Példa: $a = 1$)

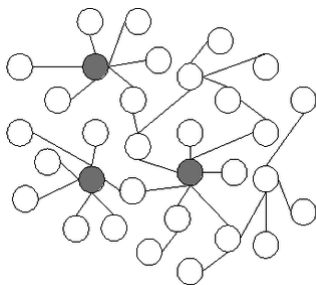
(Betűk gyakorisága egy német nyelvű szövegben)



Kép: „Zipf-Verteilung-Buchstaben” Antontól a német nyelvű Wikipediából. Licenz: CC BY-SA 3.0, Wikimedia

Commons. A szövegeket magyarra fordítottam.

Alkalmazás: skálafüggetlen hálózatok/*preferential attachment* modell



Kép: A „Scale-free network deutsch” feldolgozása Noemi Kurt által, CollectiveStupidity - az angol nyelvű eredeti feldolgozása. Licenz: CC BY-SA 3.0, Wikipedia

A véletlen irányítatlan gráf létrehozásának (egyszerűsített) szabálya: a kezdetben egyetlen csúcsból álló, él nélküli gráfhoz rendre egy-egy új csúcsot adunk hozzá, amelyet pontosan egy korábbi csúccsal kötünk össze. Ezt a csúcsot a meglévő csúcsok **fokszámával** arányos valószínűséggel választjuk.

Skálafüggetlen hálózatok és a Zipf-eloszlás

Legyen X egy véletlenszerűen választott csúcs fokszáma.

Ha a csúcsok száma **nagy**, akkor X közelítőleg Zipf-eloszlású. Az a paraméter értéke a hálózat pontos konstrukciójától függ. Az előző dián leírt esetben $a \approx 3$.

Szimulációs program: <http://ccl.northwestern.edu/netlogo/models/PreferentialAttachment>

Skálafüggetlen hálózatok

Skálafüggetlen (nem feltétlenül *preferential attachment*, de Zipfhez hasonló fokszámeloszlású) hálózatok (feltételezett) előfordulási helyei:

- ▶ Internet
- ▶ Szociális hálózatok
- ▶ Tudományos együttműködések (pl. társszerzők hálózatai)
- ▶ Elektromos hálózatok
- ▶ Fehérjemolekula-fajták kapcsolódási hálózatai
- ▶ stb.

Matematikailag rokon modellek

- ▶ Biológia: élőlényfajok kialakulása
- ▶ Városok méretének eloszlása
- ▶ Léteznek hasonló urnamodellek is

Összefoglalás: nevezetes diszkrét eloszlások

Név	Paraméter(ek)	$\text{Ran}(X)$	$\mathbb{P}(X = k)$
Diszkrét egyenletes	S	S	$1/ S $
Bernoulli/indikátor	$p \in [0, 1]$	$\{0, 1\}$	p , ha $k = 1$, $1 - p$, ha $k = 0$
Binomiális	$p \in [0, 1], n \in \mathbb{N}$	$\{0, \dots, n\}$	$\binom{n}{k} p^k (1 - p)^{n-k}$
Geometriai	$p \in [0, 1]$	$\{1, 2, \dots\}$	$(1 - p)^{k-1} p$
Poisson	$\lambda > 0$	$\{0, 1, \dots\}$	$\frac{e^{-\lambda} \lambda^k}{k!}$

Jelölés: $X = \mathbb{1}_A$, $X \sim B(n; p)$, $X \sim \text{Geo}(p)$, $X \sim \text{Pois}(\lambda)$...

Melyik mit modellez?

Diszkrét egyenletes	Pl. kockadobás bárhány oldalú kockával
Indikátor (Bernoulli)	Egy p valószínűségű esemény indikátora (1 kísérlet, siker $\rightarrow 1$, kudarc $\rightarrow 0$, siker valószínűsége p)
Binomiális	Sikerek száma egy p valószínűséggel sikeres kísérlet n -szeri független ismétlésekor
Geometriai	Próbálkozások száma az 1. sikerig (azt is beleértve) egy p valószínűséggel sikeres kísérlet független ismétléseinél
Poisson	Sikerek száma egy nagyon sokszor ismételt, nagyon kis valószínűséggel sikeres kísérletnél; ritka események száma egy adott időintervallumban

Összefoglalás: nevezetes diszkrét eloszlások

Név	Paraméter(ek)	$\text{Ran}(X)$	$\mathbb{P}(X = k)$
Diszkrét egyenletes	S	S	$1/ S $
Bernoulli/indikátor	$p \in [0, 1]$	$\{0, 1\}$	p , ha $k = 1$, $1 - p$, ha $k = 0$
Binomiális	$p \in [0, 1], n \in \mathbb{N}$	$\{0, \dots, n\}$	$\binom{n}{k} p^k (1 - p)^{n-k}$
Geometriai	$p \in [0, 1]$	$\{1, 2, \dots\}$	$(1 - p)^{k-1} p$
Poisson	$\lambda > 0$	$\{0, 1, \dots\}$	$\frac{e^{-\lambda} \lambda^k}{k!}$

Jelölés: $X = \mathbb{1}_A$, $X \sim B(n; p)$, $X \sim \text{Geo}(p)$, $X \sim \text{Pois}(\lambda)$...

Mi segíthet még felismerni őket (a nevezetes eloszlás táblázaton kívül)?

Diszkrét egyenletes	véges sok lehetséges érték, mind azonos valószínűséggel
Indikátor	2 lehetséges érték: 0 és 1 („siker” és „kudarc”)
Binomiális	véges sok lehetséges érték, nem feltétlenül mind azonos valószínűségű*
Geometriai	a val. változó értéke bármilyen pozitív egész szám lehet, de 0 nem, és intuitíve teljesül az örökifjúság
Poisson	a változó értéke bármely természetes szám lehet, 0 is*

*Ez nem csak a táblázatban szereplő eloszlásra igaz!

Várható érték – egyszerű valószínűségi változók

A várható érték azt adja meg, hogy “átlagosan” mekkora egy valószínűségi változó.

(3.2.1. Definíció.) Egy X valószínűségi változó egyszerű, ha csak véges sok értéket vehet fel. Egy egyszerű valószínűségi változó várható értéke:

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \sum_{k \in \text{Ran}(X)} k \mathbb{P}(X = k) = \sum_{k \in \text{Ran}(X)} k p_X(k).$$

- ▶ Intuíció: a várható érték a val. változó értékeinek a valószínűségeikkel súlyozott átlaga.
- ▶ (3.2.2. Példa (1)–(3))
- ▶ Össznépi gyakorlat: várható nyeremény a lottón, 4.2. feladat

Várható érték – egyszerű valószínűségi változók

A kockadobás speciális esete a diszkrét egyenletes eloszlásnak: ha X egyenletes eloszlású az n elemű $S = \{1, 2, \dots, n\}$ halmazon, akkor

$$\mathbb{E}(X) = \frac{1}{n}(1 + 2 + \dots + n) = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

(Lásd 3.2.7. Definíció)

Indikátorváltozó várható értéke: ha $A \in \mathcal{F}$ esemény, akkor

$$\mathbb{E}(1_A) = 1 \cdot \mathbb{P}(A) + 0\mathbb{P}(\overline{A}) = \mathbb{P}(A).$$

(Lásd 3.2.2. Példa (4))

A várható érték linearitása

(3.2.3. Állítás [A várható érték linearitása]) Legyenek X és Y egyszerű valószínűségi változók és $c \in \mathbb{R}$. Ekkor $\mathbb{E}(cX) = c\mathbb{E}(X)$ és $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

- ▶ Bár a bizonyítás (lásd jegyzet) már egyszerű val. változók esetén sem triviális, az állítás bármilyen X és Y (nem is feltétlenül diszkrét) val. változóra igaz, amelyeknek definiált a várható értéke.
- ▶ Ez alapján a várható érték úgy viselkedik, mint a vektorok mátrixszal való szorzása. Ahogyan használjuk a zárójelek nélküli $A\underline{x}$ jelölést egy A mátrixra és neki megfelelő méretű $\underline{x}$ vektorra, itt is szokás $\mathbb{E}(X)$ helyett $\mathbb{E}X$ -et vagy $\mathbb{E}[X]$ -et írni. A zh-n, vizsgán ezek mindegyikét elfogadjuk.
- ▶ Fontos azonban, hogy a jelölés értelmes legyen. Pl.: $\mathbb{E}(X^2) = \mathbb{E}X^2 = \mathbb{E}[X^2]$ az X^2 val. változó várható értéke, míg $\mathbb{E}(X)^2 = (\mathbb{E}X)^2 = \mathbb{E}[X]^2$ az X val. változó várható értékének négyzete – látni fogjuk, hogy a kettő általában nem ugyanaz.

A várható érték linearitása

(3.2.3. Állítás [A várható érték linearitása]) Legyenek X és Y egyszerű valószínűségi változók és $c \in \mathbb{R}$. Ekkor $\mathbb{E}(cX) = c\mathbb{E}(X)$ és $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

- ▶ Bár a bizonyítás (lásd jegyzet) már egyszerű val. változók esetén sem triviális, az állítás bármilyen X és Y (nem is feltétlenül diszkrét) val. változóra igaz, amelyeknek definiált a várható értéke.
- ▶ Ez alapján a várható érték úgy viselkedik, mint a vektorok mátrixszal való szorzása. Ahogyan használjuk a zárójelek nélküli $A\underline{x}$ jelölést egy A mátrixra és neki megfelelő méretű $\underline{x}$ vektorra, itt is szokás $\mathbb{E}(X)$ helyett $\mathbb{E}X$ -et vagy $\mathbb{E}[X]$ -et írni. A zh-n, vizsgán ezek mindegyikét elfogadjuk.
- ▶ Fontos azonban, hogy a jelölés értelmes legyen. Pl.:
 $\mathbb{E}(X^2) = \mathbb{E}X^2 = \mathbb{E}[X^2]$ az X^2 val. változó várható értéke, míg
 $\mathbb{E}(X)^2 = (\mathbb{E}X)^2 = \mathbb{E}[X]^2$ az X val. változó várható értékének négyzete – látni fogjuk, hogy a kettő általában nem ugyanaz.

A várható érték linearitása

(3.2.3. Állítás [A várható érték linearitása]) Legyenek X és Y egyszerű valószínűségi változók és $c \in \mathbb{R}$. Ekkor $\mathbb{E}(cX) = c\mathbb{E}(X)$ és $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

(3.2.6. Példa alatt.) Ez alapján számoljuk ki egy $X \sim B(n; p)$ binomiális eloszlású val. változó várható értékét.

Volt: jelölje $k \in \{0, 1, \dots, n\}$ esetén A_k azt az eseményt, hogy a k . kísérlet sikeres. Ekkor $X = \mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n}$. Így $\mathbb{E}(X) = p + \dots + p = np$.

A binomiális eloszlású val. változó várható értékét definíció szerint is egyszerűen kiszámolhatjuk (lásd előadás).

(3.2.4. Példa: 3 halmazra vonatkozó Poincaré-formula bizonyítása)

(Össznépi gyakorlat: 4.8. feladat, algel style, vö. 1.13. feladat $\rightarrow$ következő dia)

4.8. feladat (megoldás: lásd össznépi gyakorlat)

Egy kezdetben üres, 1-től $2n$ -ig indexelt tömbbe (ahol $n \geq 1$) beszúrunk $2n$ elemet úgy, hogy bármely két elem különböző helyre kerüljön (vö. 1. feladatsor, 13. feladat). Tegyük fel most, hogy az elemek fele páros, a másik fele pedig páratlan szám. A $2n - 1$ szomszédos elempár közül várhatóan hány olyan lesz, amely

- a) két páros számból áll?
- b) egy páros és egy páratlan számból áll (tetszőleges sorrendben)?

Várható érték – diszkrét valószínűségi változók

Egy egyszerű valószínűségi változó várható értéke:

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \sum_{k \in \text{Ran}(X)} k \mathbb{P}(X = k) = \sum_{k \in \text{Ran}(X)} k p_X(k). \quad (4)$$

Kérdés: jó-e ugyanez a képlet bármely X diszkrét valószínűségi változóra? Válasz: igen, ha a várható érték létezik!

- ▶ Ha $\mathbb{P}(X \geq 0) = 1$: $\mathbb{E}(X) \in [0, \infty]$ jóldefiniált a fenti képlettel.
- ▶ Ha $\mathbb{P}(X \leq 0) = 1$: $\mathbb{E}(X) \in [-\infty, 0]$ jóldefiniált.
- ▶ (4.3.2. Definíció.) Legyen X *tetszőleges* valószínűségi változó. Jelölje $X^+ = \max(X, 0)$ az X **pozitív részét**, és $X^- = \max(-X, 0)$ az X **negatív részét**. Ekkor X^+ és X^- nemnegatív valószínűségi változók, továbbá

$$X = X^+ - X^- \text{ és } |X| = X^+ + X^-$$

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.

Várható érték – diszkrét valószínűségi változók

Egy egyszerű valószínűségi változó várható értéke:

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \sum_{k \in \text{Ran}(X)} k \mathbb{P}(X = k) = \sum_{k \in \text{Ran}(X)} k p_X(k). \quad (4)$$

Kérdés: jó-e ugyanez a képlet bármely X diszkrét valószínűségi változóra? Válasz: igen, ha a várható érték létezik!

- ▶ Ha $\mathbb{P}(X \geq 0) = 1$: $\mathbb{E}(X) \in [0, \infty]$ jóldefiniált a fenti képlettel.
- ▶ Ha $\mathbb{P}(X \leq 0) = 1$: $\mathbb{E}(X) \in [-\infty, 0]$ jóldefiniált.
- ▶ (4.3.2. Definíció.) Legyen X *tetszőleges* valószínűségi változó. Jelölje $X^+ = \max(X, 0)$ az X **pozitív részét**, és $X^- = \max(-X, 0)$ az X **negatív részét**. Ekkor X^+ és X^- nemnegatív valószínűségi változók, továbbá

$$X = X^+ - X^- \text{ és } |X| = X^+ + X^-$$

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.

Várható érték – diszkrét valószínűségi változók

A következő definíció az előző speciális esete, arra az esetre vonatkozóan, amikor $\mathbb{E}(|X|)$ véges, azaz amikor $\mathbb{E}(X^+)$ és $\mathbb{E}(X^-)$ is véges (a nemnegatív val. változók várható értékének előbbi képlete szerint).

(Definíció.) Legyen X diszkrét valószínűségi változó, amelyre

$$\sum_{k \in \text{Ran}(X)} |k| p_X(k) < \infty.$$

Ekkor

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \sum_{k \in \text{Ran}(X)} k \mathbb{P}(X = k) = \sum_{k \in \text{Ran}(X)} k p_X(k) \in \mathbb{R}$$

- ▶ A fenti definíciónál a Mészáros Szabolcs-féle jegyzetet kissé egyszerűsítve a 3.4.1., 3.4.3. Definíciót, valamint a 3.4.2. Állítást egy definícióvá vontam össze, és használtam a 4.3.2. Definíciót is.

Viszont fontos, hogy lássuk, hogy egy végtelen értékű készletű val. változó várható értéke nem mindig jóldefiniált.

Várható érték – geometriai és Poisson-eloszlás

(Állítás.) Legyen $X \sim \text{Geo}(p)$, $p \in (0, 1)$. Ekkor $\mathbb{E}(X) = \frac{1}{p}$.

Intuíció: a kockadobás átlagosan 6. kísérletre lesz hatos.

(Bizonyítás – lásd 5.2.2. Példa alatt)

(Állítás.) Legyen $X \sim \text{Pois}(\lambda)$, $\lambda > 0$. Ekkor $\mathbb{E}(X) = \lambda$.

(Bizonyítás – lásd 5.3.1. Definíció alatt)

(Kérdés, előadáson nem részletezzük.) Legyen X Zipf-eloszlású $a > 1$ paraméterrel. Milyen a esetén véges, illetve végtelen X várható értéke?

Binomiális eloszlás közelítése Poisson-eloszlással

Volt: sok azonos sikervalószínűségű, független kísérlet esetén, ha n kísérlet van és mind p valószínűséggel sikerül, akkor a sikerek száma $\sim B(n; p)$.

Ha n nagy, p kicsi és ismeretlenek, akkor a sikerek száma jó közelítéssel $\text{Pois}(\lambda)$ alkalmas $\lambda > 0$ -ra.

Kérdés:

- ▶ Nagy n és alkalmas kicsi, n -től függő $p = p_n$ esetén tényleg közelíthető a $B(n; p_n)$ eloszlás egy $\text{Pois}(\lambda)$ eloszlással?
Ez nagyban segítené a számolásokat, mert nem kellene óriási binomiális együtthatókat meghatározni (még akkor sem, ha n -t pontosan ismerjük).
- ▶ Ha igen: hogyan kell ehhez p_n értékét megválasztani, hogyan tartson p_n nullához, amint $n \rightarrow \infty$?
- ▶ Mi lesz (a p_n -ektől függően) λ értéke?

Láttuk: $X_n \sim B(n; p_n) \Rightarrow \mathbb{E}(X) = np_n$.

$X \sim \text{Pois}(\lambda) \Rightarrow \mathbb{E}(X) = \lambda$.

Intuíció: a $p_n = \lambda/n$ választás jó lesz (a várható érték ekkor változatlan marad).

Binomiális eloszlás közelítése Poisson-eloszlással

(5.3.3. Állítás.) Legyen n pozitív egész, $\lambda \in (0, \infty)$, és jelölje $p_n = \lambda/n$ -et. Ekkor

$$\lim_{n \rightarrow \infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k \in \{0, 1, 2, \dots\}.$$

(Bizonyítás)

A bizonyítás hasonlóan működik akkor is, ha p_n nem pontosan λ/n , de $np_n \rightarrow \lambda$. Pl. $p_n = \frac{\lambda}{n-42}$ vagy $p_n = \frac{\lambda}{n+\ln n}$ esetén is igaz az állítás.

A nevezőben lévő n -hez bármit hozzáadhatunk vagy bármit kivonhatunk belőle, ami n -nel osztva nullához tart, amint $n \rightarrow \infty$.

(5.3.4. Példa: magyarérettségi)

A várható érték egyéb tulajdonságai

A várható érték **linearitása** ($\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$, $\mathbb{E}(cX) = c\mathbb{E}(X)$, vö. 3.2.3 Állítás) tetszőleges (nem feltétlenül diszkrét) X, Y valószínűségi változókra is igaz, amelyeknek létezik a várható értéke.

Folytonos valószínűségi változók $\rightarrow$ a várható érték kiszámításának módja más lesz, mint a diszkrét esetben (lásd később), de a linearitás érvényes marad.

Egyéb könnyű tulajdonságok: egy (nem feltétlenül diszkrét) X val. változó és $-\infty < a < b < \infty$ számok esetén

- ▶ ha $\mathbb{P}(a \leq X \leq b) = 1$, akkor $a \leq \mathbb{E}(X) \leq b$,
- ▶ ha $\mathbb{P}(X \geq a) = 1$, akkor $\mathbb{E}(X)$ definiálható (legrosszabb esetben $+\infty$ értékkel) és $\mathbb{E}(X) \geq a$.

Speciális eset: ha X nemnegatív, akkor $\mathbb{E}(X) \in [0, \infty]$,

- ▶ ha $\mathbb{P}(X \leq b) = 1$, akkor $\mathbb{E}(X)$ definiálható (legrosszabb esetben $-\infty$ értékkel) és $\mathbb{E}(X) \leq b$.

Zh-n, vizsgán várható értékkel kapcsolatos feladatoknál érdemes pl.

ellenőrizni, hogy egy $[0, 1]$ -értékű val. változó várható értékére 0 és 1 közötti számot kaptunk-e. Sajnos tavaly elég sok ellennpéldát láttunk... :(

A várható érték egyéb tulajdonságai

A várható érték **linearitása** ($\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$, $\mathbb{E}(cX) = c\mathbb{E}(X)$, vö. 3.2.3 Állítás) tetszőleges (nem feltétlenül diszkrét) X, Y valószínűségi változókra is igaz, amelyeknek létezik a várható értéke.

Folytonos valószínűségi változók $\rightarrow$ a várható érték kiszámításának módja más lesz, mint a diszkrét esetben (lásd később), de a linearitás érvényes marad.

Egyéb könnyű tulajdonságok: egy (nem feltétlenül diszkrét) X val. változó és $-\infty < a < b < \infty$ számok esetén

- ▶ ha $\mathbb{P}(a \leq X \leq b) = 1$, akkor $a \leq \mathbb{E}(X) \leq b$,
- ▶ ha $\mathbb{P}(X \geq a) = 1$, akkor $\mathbb{E}(X)$ definiálható (legrosszabb esetben $+\infty$ értékkel) és $\mathbb{E}(X) \geq a$.

Speciális eset: ha X nemnegatív, akkor $\mathbb{E}(X) \in [0, \infty]$,

- ▶ ha $\mathbb{P}(X \leq b) = 1$, akkor $\mathbb{E}(X)$ definiálható (legrosszabb esetben $-\infty$ értékkel) és $\mathbb{E}(X) \leq b$.

Zh-n, vizsgán várható értékkel kapcsolatos feladatoknál érdemes pl.

ellenőrizni, hogy egy $[0, 1]$ -értékű val. változó várható értékére 0 és 1 közötti számot kaptunk-e. Sajnos tavaly elég sok ellenpéldát láttunk... :(

Transzformált várható értéke diszkrét esetben

Egy X valószínűségi változó egy *transzformáltján* valamely $g: \text{Ran}(X) \subseteq \mathbb{R} \rightarrow \mathbb{R}$ függvény esetén a $g(X): \Omega \rightarrow \mathbb{R}$ függvényt értjük.

A transzformáltat is “ ω -nként” definiáljuk:

$$(g(X))(\omega) = g(X(\omega)).$$

Például ha X egy kockadobás eredménye és tudjuk, hogy $X = 5$, akkor $X^2 = 25$.

Ha X **diszkrét** val. változó, akkor bármilyen transzformáltja szintén val. változó (azaz $\{g(X) < x\} \in \mathcal{F}$ teljesül minden $x \in \mathbb{R}$ -re) és szintén diszkrét. Általános val. változóknál ennél bonyolultabb a helyzet.

Kérdés: hogyan számoljuk ki a diszkrét val. változó transzformáltjának várható értékét (ha az létezik)?

Transzformált várható értéke diszkrét esetben

Kockadobásos példa (3.2.2. Példa (2)) $\rightarrow$ kiszámoltuk X^2 súlyfüggvényét és ez alapján $\mathbb{E}(X^2)$ -et:

$$\mathbb{E}(X^2) = \sum_{k \in \{1,4,9,16,25,36\} = \text{Ran}(X^2)} k \mathbb{P}(X^2 = k).$$

Kérdés: tényleg muszáj ehhez X^2 súlyfüggvényét kiszámolni?

Alternatíva: tekintsük X súlyfüggvényét és súlyozzunk k^2 -tel k helyett:

$$\mathbb{E}(X^2) = \sum_{k \in \{1,2,\dots,6\} = \text{Ran}(X)} k^2 \mathbb{P}(X = k).$$

Általában igaz (lásd 4.3.5. Állítás, nem bizonyítjuk):

Ha X diszkrét valószínűségi változó, akkor bármely $g(X)$ transzformáltjára, amelynek a várható értéke létezik:

$$\mathbb{E}(g(X)) = \sum_{k \in \text{Ran}(X)} g(k) \mathbb{P}(X = k).$$

Valószínűségi változók momentumai

Ha X diszkrét valószínűségi változó, akkor bármely $g(X)$ transzformáltjára, amelynek a várható értéke létezik:

$$\mathbb{E}(g(X)) = \sum_{k \in \text{Ran}(X)} g(k) \mathbb{P}(X = k).$$

A korábbi definíció szerint $\mathbb{E}(g(X))$ akkor létezik, ha

$$\infty > \sum_{k \in \text{Ran}(g(X))} |k| \mathbb{P}(g(X) = k) = \sum_{l \in \text{Ran}(X)} |g(l)| \mathbb{P}(X = l),$$

a jobb oldal pedig a transzformált várható értékére vonatkozó képlet szerint éppen $\mathbb{E}(|g(X)|)$. (Példa: össznépi 4.13. feladat)

Speciális eset: $g(x) = x^n$, $n = 1, 2, \dots$

(Definíció.) $\mathbb{E}[|X|^n]$ -t az X val. változó n . **abszolút momentumának** nevezzük. (Ez mindig jóldefiniált, ha megengedjük a $+\infty$ értéket.)

Ha X -nek véges az n . abszolút momentuma, akkor $\mathbb{E}(X^n)$ -t az X val. változó n . **momentumának** nevezzük.

Az 1. momentum tehát a várható érték (ha létezik).

(Példa: indikátorváltozó momentumai)

Valószínűségi változók momentumai

Ha X diszkrét valószínűségi változó, akkor bármely $g(X)$ transzformáltjára, amelynek a várható értéke létezik:

$$\mathbb{E}(g(X)) = \sum_{k \in \text{Ran}(X)} g(k) \mathbb{P}(X = k).$$

A korábbi definíció szerint $\mathbb{E}(g(X))$ akkor létezik, ha

$$\infty > \sum_{k \in \text{Ran}(g(X))} |k| \mathbb{P}(g(X) = k) = \sum_{l \in \text{Ran}(X)} |g(l)| \mathbb{P}(X = l),$$

a jobb oldal pedig a transzformált várható értékére vonatkozó képlet szerint éppen $\mathbb{E}(|g(X)|)$. (Példa: össznépi 4.13. feladat)

Speciális eset: $g(x) = x^n$, $n = 1, 2, \dots$

(Definíció.) $\mathbb{E}[|X|^n]$ -t az X val. változó n . **abszolút momentumának** nevezzük. (Ez mindig jóldefiniált, ha megengedjük a $+\infty$ értéket.)

Ha X -nek véges az n . abszolút momentuma, akkor $\mathbb{E}(X^n)$ -t az X val. változó n . **momentumának** nevezzük.

Az 1. momentum tehát a várható érték (ha létezik).

(Példa: indikátorváltozó momentumai)

Monte Carlo- és Las Vegas-algoritmus

Emlékeztető: a véletlent használó algoritmusok két nagy csoportja:

- ▶ **Monte Carlo-algoritmus:** A lépésszám rögzített, nem függ a véletlentől, de az algoritmus bizonyos pozitív valószínűséggel hibás eredményt ad.
Példa: Fermat-prímteszt, RSA és variánsaik, Karger-algoritmus.
- ▶ **Las Vegas-algoritmus:** Az algoritmus mindig helyes eredményt ad, de a lépésszám véletlen.
Példa: Randomized Quicksort.

A várható érték alkalmazásai diszkrét esetben

(3.3. fejezet, Randomized Quicksort algoritmus)

Emlékeztető algeből: minden összehasonlítás alapú rendezés használ legalább $\Omega(n \log n)$ összehasonlítást (legrosszabb esetben, ahol n a tömb rendezni kívánt elemeinek száma).

Példák rendező algoritmusokra, amelyek $O(n \log n)$ lépésben rendezik a tömböt (lásd algel):

- ▶ összefésüléssel rendezés (mergesort),
- ▶ kupacos rendezés (heapsort).

Randomized Quicksort: egy olyan, véletlent használó algoritmus, amelynek lépésszáma legrosszabb esetben $\Theta(n^2)$, de **várható értékben** $O(n \log n)$, és az előbbi kettőnél könnyebben implementálható.

Randomized Quicksort

Legyenek $x_1, x_2, \dots, x_n$ a rendezni kívánt tömb elemei (x_i áll az i indexű helyen), ahol feltesszük, hogy ezek páronként különbözőek. Jelöljük $y_1, \dots, y_n$ -nel a rendezés végeredményét, azaz ez az x_i -k egy olyan felsorolása, ahol $y_1 < \dots < y_n$.

Algoritmus: Ha a tömb egy elemű, visszatérési érték a tömb.

Egyébként választunk egy p elemet (neve: pivot elem/particionáló elem), a többieket pedig szétválogatjuk két tömbre: p -nél kisebbek és p -nél nagyobbak (ez $n - 1$ összehasonlítást jelent).

Ennek konkrét implementációját két mozgó pointerrel idén tavasszal láttuk algebrából, ezzel itt nem foglalkozunk.

Alkalmazzuk rekurzívan a quicksort algoritmust a kapott két tömbre, majd adjuk vissza az ebből összekonkatenált eredményt: p -nél kisebbek rendezve, aztán p , végül p -nél nagyobbak rendezve.

- ▶ Las Vegas-algoritmus (biztosan jó eredmény, véletlen lépésszám)
- ▶ legrosszabb esetben mindenkit mindenkivel össze kell hasonlítanunk $\rightarrow \binom{n}{2} = \Theta(n^2)$ összehasonlítás.

Például ha már előre sorba van rendezve a tömb, és mindig a

Randomized Quicksort

Legyenek $x_1, x_2, \dots, x_n$ a rendezni kívánt tömb elemei (x_i áll az i indexű helyen), ahol feltesszük, hogy ezek páronként különbözőek. Jelöljük $y_1, \dots, y_n$ -nel a rendezés végeredményét, azaz ez az x_i -k egy olyan felsorolása, ahol $y_1 < \dots < y_n$.

Algoritmus: Ha a tömb egy elemű, visszatérési érték a tömb.

Egyébként választunk egy p elemet (neve: pivot elem/particionáló elem), a többieket pedig szétválogatjuk két tömbre: p -nél kisebbek és p -nél nagyobbak (ez $n - 1$ összehasonlítást jelent).

Alkalmazzuk rekurzívan a quicksort algoritmust a kapott két tömbre, majd adjuk vissza az ebből összekonkatenált eredményt: p -nél kisebbek rendezve, aztán p , végül p -nél nagyobbak rendezve.

- ▶ Las Vegas-algoritmus (biztosan jó eredmény, véletlen lépésszám)
- ▶ legrosszabb esetben mindenkit mindenkivel össze kell hasonlítanunk $\rightarrow \binom{n}{2} = \Theta(n^2)$ összehasonlítás.
Például ha már eleve sorba van rendezve a tömb, és mindig a legelső elemet választjuk pivot elemnek.

Randomized Quicksort

Kérdés: **átlagosan** mennyi idő alatt fut le az algoritmus, ha a pivot elemeket véletlenszerűen és egymástól függetlenül választjuk?

(3.3.1. Állítás.) Jelölje X_n a quicksort algoritmusban elvégzett összehasonlítások (véletlen) számát, ha a bemenet hossza n . Ekkor $\mathbb{E}(X_n) \leq 2n \ln n + 2n = O(n \log n)$.

(Bizonyítás: lásd előadás. Ehhez használjuk a $\sum_{k=1}^n \frac{1}{k} \leq \ln n + 1$ állítást, amit nem bizonyítunk.)

(Mészáros Szabolcs jegyzetében $\mathbb{E}(X_n) \leq 2n \ln n$ van, itt csak $\mathbb{E}(X_n) \leq 2n \ln n + 2n$ -t bizonyítok be, hogy a bizonyítás egy kicsit kellemesebb legyen. Mivel a $2n$ lassabban nő, mint bármilyen konstansszor $n \ln n$, ez nem nagy veszteség.)

A gyorsrendezés előnye az egyszerű implementálhatóságán kívül, hogy „helyben rendez”, nincs további memóriaigénye.

Megjegyzés: rendezés C++-ban

A C++ szabvány szerint a rendezésnek $O(n \log n)$ lépésbe bele kell férnie legrosszabb esetben is.

Algoritmus: valamilyen konkrét $c > 0$ -ra $c \log n$ lépésig megpróbáljuk lefuttatni a randomized quicksort algoritmust. Mivel ez átlagosan $O(n \log n)$ ideig tart, jó eséllyel $c \log n$ lépés elég lesz, ha c elég nagy.

De a véletlen miatt időről időre előfordulhat, hogy nem készül el ennyi idő alatt, hiszen a lépésszám legrosszabb esetben $\Theta(n^2)$.

Ekkor $c \log n$ lépés után a gyorsrendezést abbahagyjuk és kupacos rendezéssel $O(n \log n)$ lépésben rendezzük a kapott tömböt.

A kapott algoritmus még mindig egy Las Vegas-algoritmus, de determinisztikusan (a legrosszabb esetre vonatkozóan) is $O(n \log n)$ lépésszámú.

Ha c -t megfelelően választjuk, akkor ez várható értékben még mindig jobb lépésszámú, mint a sima kupacos rendezés (ami a leggyorsabb ismert determinisztikus rendező algoritmus, lásd algel).

Legrosszabb esetben persze rosszabb nála, de még mindig $O(n \log n)$ lépésszámú.

Szórás, szórásnégyzet

Azt szokták mondani, hogy a várható érték „a legtöbb információ, amit egy eloszlásról/valószínűségi változóról egyetlen számmal ki lehet fejezni”.

Ez az információ nyilvánvalóan nem lehet teljes. (Pl.: „a dobókocka átlagosan 3,5-et mutat” → ez igaz, de ebből semmi nem derül ki pl. a dobókocka értékkészletére vonatkozóan.)

Példa: egy nyereményjátékban való részvételhez 1 eurót kell befizetnünk. A játékvezető feldob egy pénzérmét, és ha az eredmény fej, akkor átad nekünk 2 eurót, különben buktuk az 1 eurónkat is.

Ekkor a nyereményünk várható értéke 0 euró.

Most emeljük a befizetendő összeget 1000 euróra, a lehetséges nyereményt pedig 2000-re. A várható érték nem változik, de érezzük a különbséget.

Általános kérdés: Hogyan tudjuk a valószínűségi változó és a várható értéke távolságát mérni?

Szórás, szórásnégyzet

Azt szokták mondani, hogy a várható érték „a legtöbb információ, amit egy eloszlásról/valószínűségi változóról egyetlen számmal ki lehet fejezni”.

Ez az információ nyilvánvalóan nem lehet teljes. (Pl.: „a dobókocka átlagosan 3,5-et mutat” → ez igaz, de ebből semmi nem derül ki pl. a dobókocka értékkészletére vonatkozóan.)

Példa: egy nyereményjátékban való részvételhez 1 eurót kell befizetnünk. A játékvezető feldob egy pénzérmét, és ha az eredmény fej, akkor átad nekünk 2 eurót, különben buktuk az 1 eurónkat is.

Ekkor a nyereményünk várható értéke 0 euró.

Most emeljük a befizetendő összeget 1000 euróra, a lehetséges nyereményt pedig 2000-re. A várható érték nem változik, de érezzük a különbséget.

Általános kérdés: Hogyan tudjuk a valószínűségi változó és a várható értéke távolságát mérni?

Szórás, szórásnégyzet

Egy X valószínűségi változó várható értéktől való (abszolút) eltérése: $|X - \mathbb{E}(X)|$.

Vizsgálhatnánk ennek a várható értékét, de az abszolút értékkel sokszor kellemetlen dolgozni.

Helyette ezen távolság négyzetét, vagyis az

$$\mathbb{E}[(X - \mathbb{E}(X))^2]$$

mennyiséget fogjuk vizsgálni. Ennek a mennyiségnek számos érdekes és hasznos tulajdonsága van – amellet, hogy könnyebb vele számolni.

(6.4.1. Definíció második fele.) Legyen X valószínűségi változó, amelyre $\mathbb{E}(X^2) < \infty$. Ekkor a

$$\mathbb{D}^2(X) := \mathbb{E}[(X - \mathbb{E}(X))^2]$$

mennyiséget az X val. változó szórásnégyzetének nevezzük, a $\mathbb{D}(X) = \sqrt{\mathbb{D}^2(X)}$ mennyiséget pedig X szórásának.

Szórás, szórásnégyzet

$$\mathbb{D}^2(X) = \mathbb{E}[(X - \mathbb{E}(X))^2], \mathbb{D}(X) = \sqrt{\mathbb{D}^2(X)}.$$

Megjegyzések:

- ▶ A szórásnégyzet (mint egy nemnegatív mennyiség várható értéke) mindig nemnegatív, így lehet belőle négyzetgyököt vonni.
- ▶ Ha $\mathbb{D}^2(X) = 0$, akkor $(X - \mathbb{E}(X))^2$ egy nemnegatív val. változó, amelynek várható értéke 0, így $(X - \mathbb{E}(X))^2 = 0$ teljesül 1 valószínűséggel, vagyis $\mathbb{P}(X = \mathbb{E}(X)) = 1$.
Ez az állítás fordítva is igaz, tehát:
 X szórása/szórásnégyzete pontosan akkor 0, ha X egy valószínűséggel konstans.
- ▶ A szórásnégyzetet szokás **varianciának** is nevezni (később látni fogjuk, hogy miért). A szórás angol neve *standard deviation*, ami arra utal, hogy ez a várható értéktől való átlagos eltérés mérőszáma.

Szórás, szórásnégyzet

$$\mathbb{D}^2(X) = \mathbb{E}[(X - \mathbb{E}(X))^2], \quad \mathbb{D}(X) = \sqrt{\mathbb{D}^2(X)}.$$

A szórásnégyzet kiszámítása a gyakorlatban a következő formula alapján történik (bár néha egyszerűbb a definíció alapján számolni):

$$\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

(Bizonyítás:

$$\begin{aligned}\mathbb{D}^2(X) &= \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2X\mathbb{E}(X) + \mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2,\end{aligned}$$

a várható érték linearitása miatt.)

A szórásnégyzet előbbi tulajdonságainak következménye:
ha $\mathbb{E}(X^2) < \infty$, akkor $\mathbb{E}(X^2) \geq \mathbb{E}(X)^2$, és egyenlőség akkor és csak akkor áll fent, ha X egy valószínűséggel konstans.

A szórás(négyzet) elemi tulajdonságai

(Első példa a szórásra: dobókocka, 6.4.4. Példa (1))

(Példa szórás kiszámítására: össznépi 4.14. „szarvasos” feladat)

(Állítás.) Ha $\mathbb{E}(X^2) < \infty$ és $a, b \in \mathbb{R}$, akkor

$$\mathbb{D}^2(aX + b) = a^2 \mathbb{D}^2(X).$$

(Bizonyítás)

Speciális esetek: $c \in \mathbb{R}$ esetén

- ▶ $\mathbb{D}^2(cX) = c^2 \mathbb{D}^2(X)$
- ▶ $\mathbb{D}^2(X + c) = \mathbb{D}^2(X)$.

Következmény (6.4.3. Állítás): ha $\mathbb{E}(X^2) < \infty$ és $c \in \mathbb{R}$, akkor

- ▶ $\mathbb{D}(cX) = |c| \mathbb{D}(X)$ (a szórás abszolút homogén)
- ▶ $\mathbb{D}(X + c) = \mathbb{D}(X)$ (a szórás eltolás-invariáns).

A szórás(négyzet) elemi tulajdonságai

$$\mathbb{D}^2(aX + b) = a^2\mathbb{D}^2(X).$$

Figyelem: a várható értékkel ellentétben a szórásnégyzet általában nem additív!

Pl. legyen X nemkonstans val. változó, amelynek szórása létezik, és legyen $Y = -X$. Ekkor

$$\mathbb{D}^2(X + Y) = \mathbb{D}^2(X - X) = 0,$$

de

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

Lesz: Ha X és Y független val. változók, akkor

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

Ehhez még meg kell ismerkednünk a val. változók függetlenségének fogalmával – hamarosan következik.

A szórás(négyzet) elemi tulajdonságai

$$\mathbb{D}^2(aX + b) = a^2\mathbb{D}^2(X).$$

Figyelem: a várható értékkel ellentétben a szórásnégyzet általában nem additív!

Pl. legyen X nemkonstans val. változó, amelynek szórása létezik, és legyen $Y = -X$. Ekkor

$$\mathbb{D}^2(X + Y) = \mathbb{D}^2(X - X) = 0,$$

de

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

Lesz: Ha X és Y független val. változók, akkor

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

Ehhez még meg kell ismerkednünk a val. változók függetlenségének fogalmával – hamarosan következnek.

Nevezetes diszkrét eloszlások szórásnégyzete

(Lásd 6.4.4. Példa (2)–(5), a (3)-at még nem kell érteni)

Indikátor: Legyen $X = \mathbb{1}_A$ egy A esemény indikátora. Ekkor $\mathbb{D}^2(X) = \mathbb{P}(A)(1 - \mathbb{P}(A))$.

(Bizonyítás)

Más szavakkal: ha X Bernoulli-eloszlású $p \in [0, 1]$ paraméterrel, akkor $\mathbb{D}^2(X) = p(1 - p)$.

Binomiális: Legyen $X \sim B(n; p)$. Ekkor $\mathbb{D}^2(X) = np(1 - p)$.

Később fogjuk bizonyítani, a zh-n elég a képletet tudni vagy az eloszlástáblázatból kiolvasni.

Geometriai: Legyen $X \sim \text{Geo}(p)$, ekkor $\mathbb{D}^2(X) = (1 - p)/p^2$.

(Részleges bizonyítás – csak úgy, ahogy Mészáros Szabolcs jegyzetében van)

Poisson: Legyen $X \sim \text{Pois}(\lambda)$, ekkor $\mathbb{D}^2(X) = \lambda$.

(Bizonyítás: gyakorlaton! 4. gyakorlat, 11. feladat.)

Megjegyzés: Poisson-eloszlásnál $\mathbb{D}^2(X) = \mathbb{E}(X)$, ez más ismert eloszláscsaládra nem jellemző.

Diszkrét valószínűségi változók függetlensége

Eddig alapvetően 1 darab (többnyire diszkrét) valószínűségi változó eloszlásáról (súlyfüggvényéről, várható értékéről, szórásáról stb.) volt szó.

Most: Szeretnénk valamit mondani arról, hogy több (általában 2) diszkrét valószínűségi változó egymástól hogyan függ, illetve egymás viselkedését hogyan befolyásolja, és ez hogy írható le valamilyen egyszerű, kompakt formában.

Ezen belül szeretnénk a **függetlenség** fogalmát, amelyet eddig csak események esetén definiáltunk, val. változókra is kiterjeszteni.

Példa: Egy dobókockát kétszer feldobunk, jelölje X az első és Y a második dobás eredményét. Ekkor X értéke ismeretében semmivel nem tudunk többet Y értékéről, mint ezen információ nélkül, és viszont. Mindjárt látni fogjuk: ebben az esetben X és Y valóban független valószínűségi változók.

Valószínűségi változók függetlensége

Emlékeztető: az A, B események függetlenek, ha $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

(6.1.1. Definíció.) Legyenek $X, Y: \Omega \rightarrow \mathbb{R}$ valószínűségi változók az $(\Omega, \mathcal{F}, \mathbb{P})$ val. mezőn. Azt mondjuk, hogy X és Y függetlenek, ha minden $x, y \in \mathbb{R}$ -re az $\{X < x\}$ és az $\{Y < y\}$ események függetlenek, azaz $\mathbb{P}(X < x, Y < y) = \mathbb{P}(X < x)\mathbb{P}(Y < y)$.

Ez a definíció jó lesz akkor is, ha nem feltétlenül diszkrét val. változókkal foglalkozunk. Diszkrét esetben van ennél egyszerűbb/intuitívabb ekvivalens definíciója is a függetlenségnek.

(6.1.3. Állítás.) Két diszkrét valószínűségi változó, X és Y pontosan akkor független, ha minden $x, y \in \mathbb{R}$ -re az $\{X = x\}$ és $\{Y = y\}$ események függetlenek, azaz

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Megjegyzések:



Valószínűségi változók függetlensége

Emlékeztető: az A, B események függetlenek, ha $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

(6.1.1. Definíció.) Legyenek $X, Y: \Omega \rightarrow \mathbb{R}$ valószínűségi változók az $(\Omega, \mathcal{F}, \mathbb{P})$ val. mezőn. Azt mondjuk, hogy X és Y függetlenek, ha minden $x, y \in \mathbb{R}$ -re az $\{X < x\}$ és az $\{Y < y\}$ események függetlenek, azaz $\mathbb{P}(X < x, Y < y) = \mathbb{P}(X < x)\mathbb{P}(Y < y)$.

(6.1.3. Állítás.) Két diszkrét valószínűségi változó, X és Y pontosan akkor független, ha minden $x, y \in \mathbb{R}$ -re az $\{X = x\}$ és $\{Y = y\}$ események függetlenek, azaz

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Megjegyzések:

- ▶ $\{X = x, Y = y\} \stackrel{\text{def}}{=} \{X = x \text{ és } Y = y\} = \{X = x\} \cap \{Y = y\}$.
Hasonlóan definiáljuk pl. az $\{X < x, Y < y\}$ eseményt is.
- ▶ (A függetlenség „váratlan” esetekben is előfordulhat, lásd 6.1.2. Példa második fele)

Valószínűségi változók függetlensége

Emlékeztető: az A, B események függetlenek, ha $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

(6.1.1. Definíció.) Legyenek $X, Y: \Omega \rightarrow \mathbb{R}$ valószínűségi változók az $(\Omega, \mathcal{F}, \mathbb{P})$ val. mezőn. Azt mondjuk, hogy X és Y függetlenek, ha minden $x, y \in \mathbb{R}$ -re az $\{X < x\}$ és az $\{Y < y\}$ események függetlenek, azaz $\mathbb{P}(X < x, Y < y) = \mathbb{P}(X < x)\mathbb{P}(Y < y)$.

(6.1.3. Állítás.) Két diszkrét valószínűségi változó, X és Y pontosan akkor független, ha minden $x, y \in \mathbb{R}$ -re az $\{X = x\}$ és $\{Y = y\}$ események függetlenek, azaz

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Megjegyzések:

- ▶ Általában igaz: X és Y pontosan akkor függetlenek, ha bármilyen “csak X -től függő esemény” (pl. $\{X^2 > 5\}$) és bármilyen “csak Y -től függő esemény” (pl. $\{\sin(Y) = 1/2\}$) függetlenek. (Pontosabb, de kicsit ijesztőbb megfogalmazás: lásd a 6.1.3. Állítás alatt.)

Példa: diszkrét együttes eloszlás és függetlenség (6.2. fejezet)

Példa: Egy szabályos pénzérmét kétszer feldobunk. Legyen $X = \mathbb{1}_{\{\text{a 2. dobás fej}\}}$ és legyen Y a dobott fejek száma.

1. Adjuk meg X és Y **együttes eloszlását**, azaz a $\mathbb{P}(X = k, Y = l)$ valószínűségeket minden olyan k, l értékre, amire ez nem 0.
2. Adjuk meg X és Y **peremeloszlásait (marginális eloszlásait)** is, azaz a p_X és p_Y súlyfüggvényt.
3. Független-e X és Y ?
4. Számoljuk ki $\mathbb{E}(XY)$ -t. (Emlékeztető: definíció szerint $XY(\omega) = X(\omega)Y(\omega)$, $\omega \in \Omega$.)

(Mielőtt elkezdenénk számolni, gondoljuk meg, hogy függetlenek-e, illetve mik lesznek X és Y perem-súlyfüggvényei.)

Példa: diszkrét együttes eloszlás és függetlenség

$X = \mathbb{1}_{\{\text{a 2. dobás fej}\}}$ $Y = \text{a dobott fejek száma}$.

Az együttes eloszlást táblázatban adjuk meg (számolás: előadás):

$Y \backslash X$	0	1
0	$\frac{1}{4}$	0
1	$\frac{1}{4}$	$\frac{1}{4}$
2	0	$\frac{1}{4}$

A táblázatban szereplő valószínűségek összege 1.

Hol látszanak a táblázatban a peremeloszlások?

Példa: diszkrét együttes eloszlás és függetlenség

$X = \mathbb{1}_{\{\text{a 2. dobás fej}\}}$, $Y = \text{a dobott fejek száma}$.

Az együttes eloszlást táblázatban adjuk meg (számolás: előadás):

$Y \backslash X$	X		p_Y
	0	1	
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
2	0	$\frac{1}{4}$	$\frac{1}{4}$
p_X	$\frac{1}{2}$	$\frac{1}{2}$	1

A táblázat fő részében lévő valószínűségek összege 1.

Hol látszanak a táblázatban a peremeloszlások?

- ▶ $p_X(k) = \mathbb{P}(X = k) = \sum_{l \in \text{Ran}(Y)} \mathbb{P}(X = k, Y = l)$, ezért a táblázat oszlopösszegei adják p_X megfelelő értékeit.
- ▶ $p_Y(l) = \mathbb{P}(Y = l) = \sum_{k \in \text{Ran}(X)} \mathbb{P}(X = k, Y = l)$, ezért a táblázat sorösszegei adják p_Y megfelelő értékeit.

$\mathbb{E}(XY) = \frac{3}{4}$ (lásd előadás). X és Y nem függetlenek, mert pl.

$\mathbb{P}(X = 0, Y = 2) = 0 \neq \mathbb{P}(X = 0)\mathbb{P}(Y = 2)$.

„Tippek, trükkök” együttes eloszlásokhoz

Y \ X	X		p_Y
	0	1	
0	$\frac{1}{4}$	0	$\frac{1}{4}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
2	0	$\frac{1}{4}$	$\frac{1}{4}$
p_X	$\frac{1}{2}$	$\frac{1}{2}$	1

Bármilyen X, Y diszkrét val. változók esetén:

- ▶ Ahhoz, hogy belássuk, hogy X és Y **nem függetlenek**, elég egy olyan (k, l) párt találni, amelyre $\mathbb{P}(X = k, Y = l) \neq \mathbb{P}(X = k)\mathbb{P}(Y = l)$.

A táblázatban könnyen látszó speciális eset: ha $0 = \mathbb{P}(X = k, Y = l)$, $\mathbb{P}(X = k) \neq 0, \mathbb{P}(Y = l) \neq 0$.

- ▶ Viszont ahhoz, hogy belássuk, hogy X és Y **függetlenek, minden** (k, l) párra végig kell nézni, hogy $\mathbb{P}(X = k, Y = l) = \mathbb{P}(X = k)\mathbb{P}(Y = l)$ teljesül-e.
- ▶ Fontos X és Y sorrendje. Ha fordított sorrendben írjuk fel X -et és Y -t, akkor p_X lesz a sorokban és p_Y az oszlopokban.

Független valószínűségi változók szorzatának várható értéke

Ha X és Y függetlenek, akkor a szorzatuk várható értéke a várható értékük szorzata:

(6.1.4. Állítás.) Ha X és Y független valószínűségi változók, és $\mathbb{E}(XY)$, $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ létezik, akkor

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y). \quad (5)$$

(Bizonyítás egyszerű val. változók esetén – az általános diszkrét eset bizonyítása hasonló, határátmenettel)

Független valószínűségi változók szorzatának várható értéke

(6.1.4. Állítás.) Ha X és Y független valószínűségi változók, és $\mathbb{E}(XY)$, $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ létezik, akkor

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y). \quad (5)$$

Az állítás megfordítása nem igaz! Lehetséges, hogy X és Y nem függetlenek, de (5) mégis teljesül. Példa: később.

(Megjegyzés.) Az X és Y val. változók pontosan akkor függetlenek, ha minden olyan $f, g: \mathbb{R} \rightarrow [0, \infty)$ nemnegatív valós függvényre, amelyekre $f(X)$, $g(Y)$ és $f(X)g(Y)$ is valószínűségi változók,

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y)) \in [0, \infty]. \quad (6)$$

Pl. $\mathbb{E}(X^2(Y^4 + 1)) = \mathbb{E}(X^2)(\mathbb{E}(Y^4) + 1) \in [0, \infty]$.

Jogos kérdés: a narancssárgával írt feltétel milyen f és g esetén teljesül? Hogyan ellenőrizzük, hogy (6) igaz-e minden ilyen f -re és g -re?

Ezekre a kérdésekre mértékelméleti alapozás nélkül nem tudunk teljes választ adni. Részleges válasz: ha X, Y diszkrét, akkor minden transzformáltjuk val. változó, és ezek szorzata is.

Független valószínűségi változók szorzatának várható értéke

(6.1.4. Állítás.) Ha X és Y független valószínűségi változók, és $\mathbb{E}(XY)$, $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ létezik, akkor

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y). \quad (5)$$

Következmény:

(6.4.2. Állítás.) Ha X és Y független valószínűségi változók és $\mathbb{E}(X^2), \mathbb{E}(Y^2) < \infty$, akkor

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

- ▶ (Bizonyítás – lásd előadás, nem egészen ugyanaz, mint a jegyzetben)
- ▶ (Emlékeztető: nem független X, Y esetén általában nem igaz!)

(Példa: Ha X, Y függetlenek, $\mathbb{D}^2(X) = 4$, $\mathbb{D}^2(Y) = 9$, mennyi $\mathbb{D}(2X + Y + 9)$?)

Több valószínűségi változó függetlensége

Emlékeztető: $n \geq 2$ darab esemény, $A_1, \dots, A_n$ függetlensége azt jelentette, hogy bármely $2 \leq k \leq n$ bármelyik k darab eseményt kiválasztva a k -as metszet valószínűsége a k esemény valószínűségének szorzata.

Formálisan: $A_1, \dots, A_n$ függetlenek $\Leftrightarrow \forall k \in \{2, \dots, n\}$,
 $\forall 1 \leq i_1 < i_2 < \dots < i_k \leq n$,

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \cdot \mathbb{P}(A_{i_2}) \cdot \dots \cdot \mathbb{P}(A_{i_k}).$$

Ehhez $n \geq 3$ esemény esetén nem volt elég pl. bármely két esemény függetlenségét vagy az n -es metszetre vonatkozó tulajdonságot megkövetelni.

Ugyanígy, ha $X_1, \dots, X_n$ ($n \geq 2$) valószínűségi változók az $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mezőn, ezek függetlensége definíció szerint azt jelenti, hogy bármely $k \in \{2, \dots, n\}$,

$1 \leq i_1 < i_2 < \dots < i_k \leq n$ és $x_{i_1}, \dots, x_{i_k} \in \mathbb{R}$ esetén

$$\mathbb{P}(X_{i_1} < x_{i_1}, \dots, X_{i_k} < x_{i_k}) = \mathbb{P}(X_{i_1} < x_{i_1}) \cdot \dots \cdot \mathbb{P}(X_{i_k} < x_{i_k}).$$

Diszkrét esetben ez megint ekvivalens azzal, ha a „<” jelek helyett mindenhol „ $\leq$ ”-t írunk (a fenti feltételekkel).

Több valószínűségi változó függetlensége

Emlékeztető: $n \geq 2$ darab esemény, $A_1, \dots, A_n$ függetlensége azt jelentette, hogy bármely $2 \leq k \leq n$ bármelyik k darab eseményt kiválasztva a k -as metszet valószínűsége a k esemény valószínűségének szorzata.

Formálisan: $A_1, \dots, A_n$ függetlenek $\Leftrightarrow \forall k \in \{2, \dots, n\}$,
 $\forall 1 \leq i_1 < i_2 < \dots < i_k \leq n$,

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \cdot \mathbb{P}(A_{i_2}) \cdot \dots \cdot \mathbb{P}(A_{i_k}).$$

Ugyanígy, ha $X_1, \dots, X_n$ ($n \geq 2$) valószínűségi változók az $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mezőn, ezek függetlensége definíció szerint azt jelenti, hogy bármely $k \in \{2, \dots, n\}$,
 $1 \leq i_1 < i_2 < \dots < i_k \leq n$ és $x_{i_1}, \dots, x_{i_k} \in \mathbb{R}$ esetén

$$\mathbb{P}(X_{i_1} < x_{i_1}, \dots, X_{i_k} < x_{i_k}) = \mathbb{P}(X_{i_1} < x_{i_1}) \cdot \dots \cdot \mathbb{P}(X_{i_k} < x_{i_k}).$$

Diszkrét esetben ez megint ekvivalens azzal, ha a “<” jelek helyett mindenhova “=”-ket írunk, azaz (a fenti feltételekkel)

$$\mathbb{P}(X_{i_1} = x_{i_1}, \dots, X_{i_k} = x_{i_k}) = \mathbb{P}(X_{i_1} = x_{i_1}) \cdot \dots \cdot \mathbb{P}(X_{i_k} = x_{i_k}).$$

A diszkrét együttes eloszlás mint kétdim. súlyfüggvény

Két val. változó, X és Y **diszkrét együttes eloszlása** nem más, mint egy *kétdimenziós súlyfüggvény*: a $(x, y) \mapsto p_{(X,Y)}(x, y)$ súlyfüggvény, ahol definíció szerint

$$p_{(X,Y)}(x, y) := \mathbb{P}(X = x, Y = y) = \mathbb{P}(\{X = x\} \cap \{Y = y\}).$$

Ezt X és Y **együttes súlyfüggvényének** is szokás nevezni.

Y \ X	X	
	0	1
0	$\frac{1}{4}$	0
1	$\frac{1}{4}$	$\frac{1}{4}$
2	0	$\frac{1}{4}$

A példában: $p_{(X,Y)}(0, 1) = \frac{1}{4}$.

(6.1.3. Állítás átfogalmazása.) Két diszkrét valószínűségi változó, X és Y pontosan akkor független, ha az együttes súlyfüggvényük a perem-súlyfüggvényeik szorzata, azaz minden $x, y \in \mathbb{R}$ esetén

$$p_{(X,Y)}(x, y) = \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y) = p_X(x)p_Y(y).$$

Konvolúció diszkrét esetben

(Kiegészítő anyag: Mészáros Szabolcs jegyzetében szerepel, de az új TAD-ban már nem)

Ha X, Y független valószínűségi változók, mi $X + Y$ eloszlása?

(7.3.3. Állítás.) Legyenek X és Y független, **diszkrét** valószínűségi változók, amik értékei nemnegatív egész számok. Ekkor

$$\mathbb{P}(X + Y = k) = \sum_{i=0}^k \mathbb{P}(X = i) \mathbb{P}(Y = k - i)$$

minden $k \in \{0, 1, 2, \dots\}$ esetén.

(Bizonyítás) (**Konvolúció**: független val. változók összegének eloszlása)
(Általános diszkrét X, Y esetén a képlet hasonló)

(7.3.4. Példa.) Legyenek X és Y független valószínűségi változók, amire $X \sim \text{Pois}(\lambda)$ és $Y \sim \text{Pois}(\mu)$. Ekkor $X + Y \sim \text{Pois}(\lambda + \mu)$, azaz bármely $k \in \{0, 1, 2, \dots\}$ -ra

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

(Bizonyítás: lásd előadáson)