

Valószínűségszámítás

Szűcs Gábor

Szegedi Tudományegyetem, Bolyai Intézet

2024/25 őszi félév

Tudnivalók a kurzusról

- Az előadás és a gyakorlat külön kreditelt: gyakorlati jegy és kollokvium.
- A gyakorlatokon a leadott anyag és a számonkérés egységes:
 - 6 db kisdolgozat, a legjobb 5 eredmény számít: $5 \cdot 10 = 50$ pont
 - 1 db nagydolgozat a félév végén: 50 pont
 - összesen: 100 pont
- Kollokviumjegy: a vizsgaidőszakban tartott írásbeli vizsgákon.
- A jó (4) és jeles (5) gyakorlati jegyeket megajánlom kollokviumjegynek.
- Egyéni tanulmányi rendes hallgatók: egyéni számonkérés.
- További információ, tematika, előadásanyag, órai és gyakorló feladatok:
<http://www.math.u-szeged.hu/~szucsg/>

Mi is az a valószínűségszámítás?

A valószínűségszámítás (másnéven sztochasztika) a matematikának a véletlen jelenségekkel foglalkozó területe.

Néhány alapvető fogalom

- **Véletlen kísérlet, valószínűségi kísérlet:** egy véletlen jelenség előidézése vagy megfigyelése.
- **Kimenetek:** a kísérlet lehetséges eredményei.
- **Eseménytér:** a lehetséges kimenetek halmaza. Jele: Ω .
- **Események:** a kísérlet eredményére vonatkozó állítások.
- Akkor mondjuk, hogy egy esemény **bekövetkezik**, ha az esemény, mint állítás, igaz.
- **Eseményalgebra:** a véletlen kísérlethez kapcsolódó események halmaza. Jele: $\mathcal{A}$.
- **Esemény valószínűsége:** az esemény bekövetkezésének az esélye.

Feladat. Feldobunk egy szabályos dobókockát, ez egy véletlen kísérlet.
 Eseménytér: $\Omega = \{1, 2, 3, 4, 5, 6\}$. Néhány esemény:

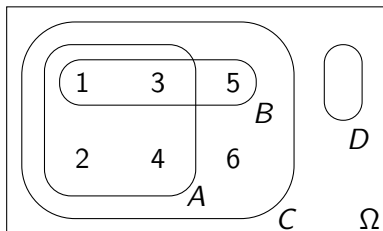
A = 5-nél kisebbet dobunk

C = egész számot dobunk

B = páratlan számot dobunk

D = negatív számot dobunk

Az eseményeket az Ω részhalmazzaival reprezentáljuk: összegyűjtjük azokat a kimeneteket, melyekre az egyes események bekövetkeznek.



$$A = \{1, 2, 3, 4\}$$

$$B = \{1, 3, 5\}$$

$$C = \{1, 2, 3, 4, 5, 6\} = \Omega$$

(biztos esemény)

$$D = \emptyset \quad (\text{lehetetlen esemény})$$

Két nevezetes esemény

Biztos esemény: minden lehetséges kimenetel esetén bekövetkezik.

Lehetetlen esemény: nincs olyan kimenetel, melyre bekövetkezne.

Mennyi ezeknek az eseményeknek a valószínűsége? Szabályos kockával dobunk, ezért minden kimenetelnek $1/6$ az esélye.

$$P(A) = P(5\text{-nél kisebbet dobunk}) = 4 \cdot 1/6 = 4/6 = 2/3$$

$$P(B) = P(\text{páratlan számot dobunk}) = 3 \cdot 1/6 = 3/6 = 1/2$$

$$P(C) = P(\text{egész számot dobunk}) = 6 \cdot 1/6 = 6/6 = 1 = 100\%$$

$$P(D) = P(\text{negatív számot dobunk}) = 0 \cdot 1/6 = 0/6 = 0$$

Az események valószínűsége most csak attól függ, hogy hány kimenetel esetén következnek be, attól már nem, hogy melyek ezek a kimenetek. Formálisan a következő szabályt alkalmaztuk:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{kedvező kimenetek száma}}{\text{összes kimenetel száma}}$$

Ez a formula a **valószínűség kombinatorikus kiszámítási módja**, és csak akkor lehet alkalmazni, ha teljesül a következő két feltétel:

- A kísérletnek csak véges sok kimenetele van.
- Minden kimenetelnek azonos az esélye.

Feladat. Feldobunk két szabályos dobókockát, egy pirosat és egy kéket. Adjuk meg az alábbi események valószínűségét:

A = a dobott számok összege 2, B = a dobott számok összege 7.

1. Megközelítés. A feladat a dobott értékek összegével foglalkozik. Legyenek a kimenetek a lehetséges összegek:

2	4	6	8	10	12
A	3	5	7	9	11
			B		Ω

$$\Omega = \{2, 3, \dots, 12\},$$

$$A = \{2\}, \quad B = \{7\}.$$

A kedvező/összes formula alkalmazásával: $P(A) = 1/11$, $P(B) = 1/11$.

Ez a megoldás nem lehet jó: a tapasztalatok szerint két kockával dobva a 7 sokkal gyakrabban jön ki, mint a 2. Tehát a két értéknek nem lehet azonos a valószínűsége. De mit rontottunk el?

A hiba oka az, hogy az egyes kimenetelnek nem azonos a valószínűsége, és emiatt nem használhatjuk a kedvező/összes formulát.

2. Megközelítés. Ne a dobott számok összegét nézzük, hanem azt, hogy mit dobunk a két kockával! Ebben a megközelítésben összesen $6 \cdot 6 = 36$ kimenetel van. A kedvező/összes formula alkalmazásával:

$$P(\text{a számok összege } 2) = \frac{1}{36},$$

$$P(\text{a számok összege } 7) = \frac{6}{36} = \frac{1}{6}.$$

	1	2	3	4	5	6
1	2					7
2					7	
3				7		
4			7			
5		7				
6	7					

Ezek már a helyes eredmények? Elfogadható, hogy mind a 36 kimenetelnek azonos a valószínűsége? (Igen, de erre majd még visszatérünk.)

A kérdésnek ezen megközelítése Galilei nevéhez fűződik. 1589-ben Toscana nagyhercege neki tette fel a következő kérdést: Miért van az, hogy három szabályos dobókockával dobva a 9 gyakrabban jön ki, mint a 10?

Feladat. Miben változik a megoldás, ha nem egy piros és egy kék, hanem két ugyanolyan színű kockát dobunk fel egyszerre?

1. Megközelítés. Összesen csak 21 kimenetel van, ugyanis a két dobókocka között nem tudunk különbséget tenni.

$$P(\text{a számok összege } 2) = \frac{1}{21},$$

$$P(\text{a számok összege } 7) = \frac{3}{21} = \frac{1}{7}.$$

Probléma: ez az eredmény nem lehet helyes, ugyanis ellentmond a tapasztalatnak.

	1	2	3	4	5	6
1	2					7
2					7	
3				7		
4			7			
5		7				
6	7					

2. Megközelítés. Anyagi tulajdonságaiban mindig van különbség a két kocka között, a kettő sosem megkülönböztethetetlen. Emiatt úgy kell eljárunk, mintha két különböző színű kockával végeznénk el a kísérletet. Ezt a gyakorlati tapasztalat is alátámasztja, ez a helyes megoldás.

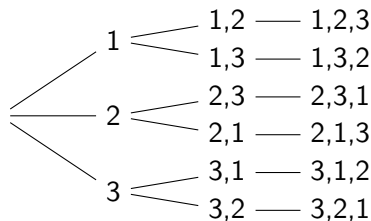
Összeszámlálási feladatok

Ahhoz, hogy a valószínűséget a kombinatorikus módszerrel ki tudjuk számolni, meg kell határoznunk a kedvező illetve az összes kimenetek számát. Az ilyen típusú feladatokat **összeszámlálási problémáknak** nevezzük, mi három fontosabb összeszámlálási problémával foglalkozunk. Mindegyik esetben egy véges (mondjuk n elemű) halmazt vizsgálunk.

- **Sorbaállítási feladatok:** Hányféleképpen állíthatjuk sorba a halmaz elemeit? A lehetséges sorbaállításokat **permutációknak** nevezzük, és a kérdés úgy is feltehető, hogy hány permutációja van a halmaznak.
- **Kiválasztási feladatok:** Hányféleképpen választhatunk ki k darabot az elemek közül? A kiválasztás szerint két esetet különböztetünk meg:
Kombinációk: A kiválasztást visszatevés nélkül végezzük, és a kiválasztás sorrendje nem számít.
Variációk: A kiválasztás visszatevéssel történik, és a sorrend számít.

Permutációk. Adott egy halmaz n különböző elemmel. Hányféleképpen állíthatjuk sorba a halmaz elemeit?

Például $n = 3$ esetén a lehetséges sorbaállítások száma $3 \cdot 2 \cdot 1 = 6$.



Az általános esetben a lehetséges sorbaállítások száma **n -faktoriális**:

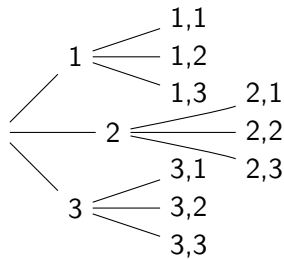
$$n! = n \cdot (n - 1) \cdot \dots \cdot 1$$

Konvenció: $0! = 1$

Pozíciók:				...		
Lehetőségek száma:	n	$n - 1$	$n - 2$	...	2	1

Variációk. Adott egy halmaz n különböző elemmel. Hányféleképpen választhatunk ki k elemet visszatevéssel, ha a kiválasztás sorrendjét is figyelembe vesszük?

Például $n = 3$ és $k = 2$ esetén $3^2 = 9$ lehetőségünk van.



Az általános esetben a lehetséges kiválasztások száma n^k .

Pozíciók:				...		
Lehetőségek száma:	n	n	n	...	n	n

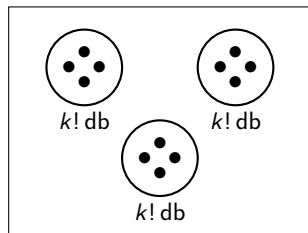
Kombinációk. Adott egy halmaz n különböző elemmel. Hányféleképpen választhatunk ki k elemet visszatevés nélkül, ha a sorrend nem számít?

Ha figyelembe vennénk a sorrendet is, akkor a lehetőségek száma:

$$n(n-1) \dots (n-k+2)(n-k+1) = n!/(n-k)!$$

Pozíciók:	$\square$	$\square$	$\square$	$\dots$	$\square$	$\square$
Lehetőségek száma:	n	$n-1$	$n-2$	$\dots$	$n-k+2$	$n-k+1$

Most felejtsük el, hogy milyen sorrendben húztuk ki az elemeket! Ha két kiválasztás csak a sorrendben különbözik, akkor a sorrend elfelejtése után ezek már azonosak lesznek. A k elemet $k!$ féleképpen lehet sorba állítani, tehát a sorrend nélküli kiválasztások száma:



$$\frac{\text{sorrendes minták száma}}{\text{hányféleképpen állíthatunk sorba } k \text{ elemet}} = \frac{n!/(n-k)!}{k!} = \frac{n!}{k! \cdot (n-k)!}.$$

Binomiális együttható

Binomiális együtthatónak nevezzük a következő szimbólumot:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n - k)!}.$$

A binomiális együttható megmutatja, hogy hányféleképpen választhatunk ki visszatevés nélkül egy n elemű halmazból k elemet, ha a kiválasztás sorrendjét nem vesszük figyelembe.

Feladat. Mik az ötöslottó sorsolás lehetséges kimenetelei? Hány lehetséges kimenetele van a kísérletnek?

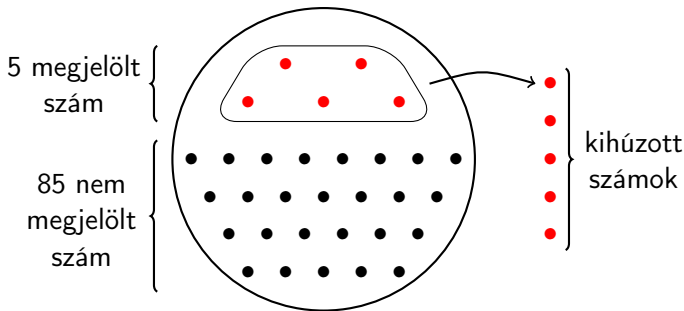
Az ötöslottó sorsolás lehetséges kimenetelei a kihúzható számötösök. A sorsolás egy visszatevés nélküli mintavételezés $n = 90$ és $k = 5$ paraméterekkel. A lehetséges kimenetek száma:

$$\binom{90}{5} = \frac{90!}{5! \cdot (90 - 5)!} = \frac{90!}{5! \cdot 85!} = \frac{90 \cdot 89 \cdot 88 \cdot 87 \cdot 86}{5!} = 43.949.268$$

Feladat. Egy szelvénnel játszunk az ötösloton. Mennyi az esélye annak, hogy telitalálatot érünk el?

Telitalálatot akkor érünk el, ha pontosan az általunk megjelölt öt számot húzzák ki. A telitalálat valószínűsége:

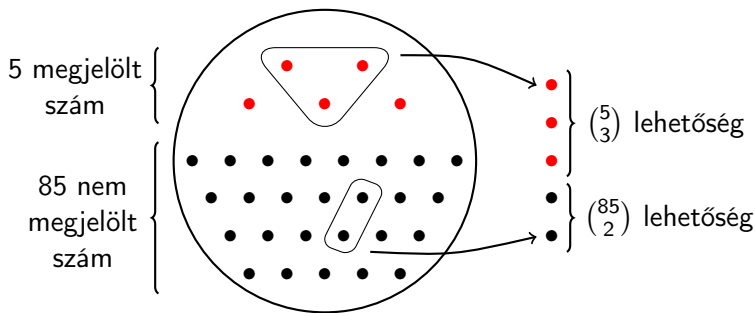
$$P(\text{telitalálat}) = \frac{\text{kedvező}}{\text{összes}} = \frac{1}{\binom{90}{5}} = \frac{1}{43.949.268}$$



Feladat. Mennyi annak az esélye, hogy pontosan 3 találatot érünk el?

Pontosan 3 találatot akkor érünk el, ha a mi számaink közül 3-at, a többi szám közül pedig 2-őt húznak ki. A kérdéses valószínűség:

$$P(\text{pontosan 3 találat}) = \frac{\text{kedvező}}{\text{összes}} = \frac{\binom{5}{3} \binom{85}{2}}{\binom{90}{5}} \approx \frac{1}{1231}.$$



Feladat. Mennyi annak az esélye, hogy pontosan k találatot érünk el?
Mennyi az esélye, hogy nyerünk valamennyi pénzt?

Az előző oldal gondolatmenetével:

$$P(\text{pontosan } k \text{ találat}) = \frac{\text{kedvező}}{\text{összes}} = \frac{\binom{5}{k} \binom{85}{5-k}}{\binom{90}{5}}, \quad k = 0, 1, 2, 3, 4, 5.$$

Találat	Valószínűség
5	$\approx 1 : 44.000.000$
4	$\approx 1 : 100.000$
3	$\approx 1 : 1000$
2	$\approx 2,25\%$
1	$\approx 23\%$
0	$\approx 74,6\%$

Pénzt akkor nyerünk, ha elérünk legalább 2 találatot:

$$P(\text{nyerünk pénzt}) = \frac{\binom{5}{2} \binom{85}{3} + \binom{5}{3} \binom{85}{2} + \binom{5}{4} \binom{85}{1} + 1}{\binom{90}{5}} \approx 2,33\%.$$

Műveletek eseményekkel

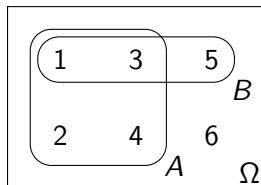
Az előző részben történt:

- Véletlen kísérlet: egy véletlen jelenség előidézése vagy megfigyelése.
- Kimenetelek: a véletlen kísérlet lehetséges eredményei.
- Eseménytér (Ω): az összes lehetséges kimenetel halmaza.
- Esemény: a kísérlet kimenetelével kapcsolatot állítás.
- A szövegesen megfogalmazott eseményeket halmazokkal reprezentáltuk: összegyűjtöttük azon kimeneteleket, melyre egy adott esemény bekövetkezik.

Példa. Dobjunk fel egy dobókockát, és tekintsük a következő eseményeket:

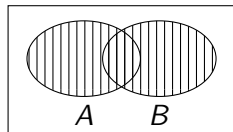
$A = 5\text{-nél kisebbet dobunk} = \{1, 2, 3, 4\},$

$B = \text{páratlan számot dobunk} = \{1, 3, 5\}.$

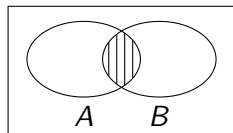


Két halmazon az alábbi **halmazelméleti műveleteket** hajthatjuk végre:

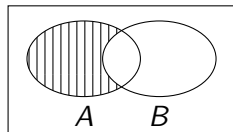
- A két halmaz **egyesítése** vagy **uniója** azon elemek halmaza, melyek a kettő közül valamelyik halmazban benne vannak. Jele: $A \cup B$.



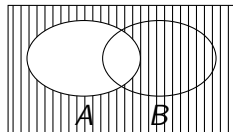
- A két halmaz **metSZete** azon elemek halmaza, melyek mindkét halmazban benne vannak. Jele: $A \cap B$.



- Az A és a B halmaz **különbsége** azon elemek halmaza, melyek benne vannak A -ban, de nincsenek benne B -ben. Jele: $A \setminus B$.



- Az A halmaz **komplementere** azon elemek halmaza, melyek nincsenek benne A -ban. Jele: $\overline{A}$.



Műveletek eseményekkel

- Események **összege**: az események „vagy” logikai művelettel való összekapcsolása. Ez azonos az események halmazelméleti uniójával:

$$A + B = A \text{ vagy } B = A \cup B$$

- Események **szorzata**: az események „és” logikai művelettel való összekapcsolása. Ez azonos az események halmazelméleti metszetével:

$$A \cdot B = A \text{ és } B = A \cap B$$

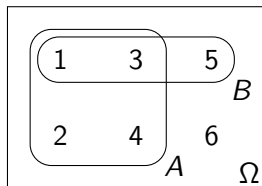
Példa. Dobókockával dobunk.

$A = 5\text{-nél kisebbet dobunk} = \{1, 2, 3, 4\}$

$B = \text{páratlan számot dobunk} = \{1, 3, 5\}$

$A \text{ vagy } B = \{1, 2, 3, 4, 5\} = A \cup B$

$A \text{ és } B = \{1, 3\} = A \cap B$



Műveletek eseményekkel (folytatás)

- Két esemény **különbsége**: az „ A bekövetkezik, de B nem” logikai kifejezés. Ez megegyezik a két esemény halmazelméleti különbségével:

$$A - B = A \text{ igen, de } B \text{ nem} = A \setminus B$$

- Egy esemény **tagadása** megegyezik az esemény halmazelméleti komplementerével:

$$\neg A = \text{nem } A = \bar{A}$$

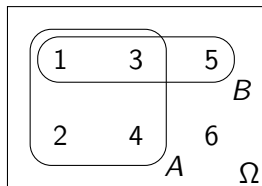
Példa. Dobókockával dobunk.

$A = 5\text{-nél kisebbet dobunk} = \{1, 2, 3, 4\}$

$B = \text{páratlan számot dobunk} = \{1, 3, 5\}$

$A \text{ igen, de } B \text{ nem} = \{2, 4\} = A \setminus B$

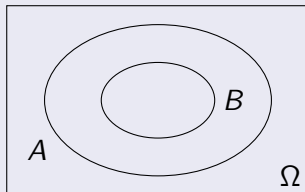
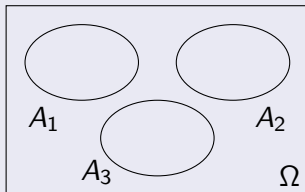
$\text{nem } A = \{5, 6\} = \bar{A}$



Kapcsolat az események között

Kizáró események: Legyen $A_1, A_2, \dots$ eseményeknek egy véges vagy végtelen sorozata. Azt mondjuk, hogy ezek **páronként kizáróak**, ha a halmazok **diszjunktak**, tehát bármely kettőt kiválasztva azoknak üres a metszete. Kizáró események közül legfeljebb egy következhet be egyszerre, hiszen nincs olyan kimenetel, melyet két vagy több esemény is tartalmazna.

Egymást tartalmazó események: Azt mondjuk, hogy a B esemény **maga után vonja** az A eseményt, ha $B \subset A$, tehát B minden eleme az A halmaznak is eleme. Ez azt jelenti, hogy ha a B esemény bekövetkezik, akkor az A esemény is feltétlenül bekövetkezik.



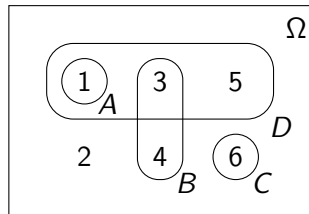
Feladat. Feldobunk egy dobókockát. Milyen kapcsolat áll fenn az alábbi események között?

$A = 1$ -est dobunk

$B = 3$ -ast vagy 4 -est dobunk

$C = 6$ -ost dobunk

$D =$ páratlan számot dobunk



- Az A, B, C események páronként kizáróak, semelyik kettő nem következhet be egyszerre.
- A C és a D esemény szintén kizárja egymást.
- Az A esemény maga után vonja D -t: ha A bekövetkezik, akkor D is bekövetkezik.
- A B és a D esemény között egyik tanult kapcsolat sem áll fenn: nem zárják ki egymást, és egyik sem vonja maga után a másikat.

A valószínűség fogalma és tulajdonságai

Mit jelent az, hogy egy esemény „valószínűsége”?

- valószínűség = esély: ez nem jó definíció, ezek szinonimák.
- valószínűség = a kísérletet sokszor elvégezve az esemény az ismétlések mekkora hányadában következik be.

Ez sem jó definíció: egy szabályos pénzérmét 99-szer feldobva nem kaphatunk pontosan az esetek felében fejet.

Bekövetkezési gyakoriság és relatív gyakoriság

Tekintsünk egy kísérletet, és egy ehhez kapcsolódó A eseményt. Hajtsuk végre a kísérletet egymás után n alkalommal!

- **Gyakoriság, bekövetkezési gyakoriság:** hányszor következett be az A esemény az n ismétlés során. Jele: $k_n(A)$.
- **Relatív gyakoriság:** az ismétlések mekkora hányadában következett be az A esemény. Jele: $r_n(A) = k_n(A)/n$.

A relatív gyakoriság nem egyenlő a valószínűséggel: $r_n(A) \neq P(A)$.

Tapasztalati tény, hogy a esemény relatív gyakoriság nagy n esetén egy meghatározott érték körül ingadozik. A köznyelv ezt az értéket nevezi az esemény valószínűségének:

$$r_n(A) = \frac{k_n(A)}{n} \approx P(A).$$

Ennél több is igaz, a relatív gyakoriság konvergál ehhez az értékhez:

$$r_n(A) \rightarrow P(A), \quad n \rightarrow \infty.$$

Miért nem ezzel a konvergenciával definiáljuk a valószínűséget? Mert ezzel a definícióval nagyon nehéz lenne dolgozni.

A valószínűségszámítás modern felépítése számolási szabályok segítségével definiálja a valószínűséget. Kérdés: milyen tulajdonságokat várunk el az események valószínűségétől?

- Egy A esemény relatív gyakorisága mindig 0 és 1 közé esik.
Elvárás: Legyen $0 \leq P(A) \leq 1$ minden A eseményre.
- A biztos esemény relatív gyakorisága:

$$r_n(\Omega) = k_n(\Omega)/n = n/n = 1$$

Elvárás: A biztos esemény valószínűsége legyen $P(\Omega) = 1$.

- Ha A és B kizáró események, akkor

$$r_n(A \cup B) = \frac{k_n(A \cup B)}{n} = \frac{k_n(A) + k_n(B)}{n} = r_n(A) + r_n(B)$$

Hasonló módon mutatható meg, hogy ha $A_1, A_2, \dots$ páronként kizáró eseményeknek véges vagy végtelen sorozata, akkor

$$r_n(A_1 \cup A_2 \cup \dots) = r_n(A_1) + r_n(A_2) + \dots$$

Elvárás: Páronként kizáró eseményekre legyen

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

Valószínűség, valószínűségi mező (Kolmogorov axiómák, 1933)

Tekintsük egy véletlen kísérletet! Ehhez tartozik egy Ω eseménytér és egy $\mathcal{A}$ eseményalgebra.

Azt mondjuk, hogy egy $P : \mathcal{A} \rightarrow [0, 1]$ függvény **valószínűség** vagy **valószínűségi mérték**, ha teljesíti az alábbi két tulajdonságot:

- A biztos esemény valószínűsége $P(\Omega) = 1$.
- **Additivitás:** Ha $A_1, A_2, \dots$ páronként kizáró eseményeknek egy véges vagy végtelen sorozata, akkor az egyesítésük valószínűsége

$$P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$$

Az A eseményhez rendelt $P(A)$ értéket úgy nevezzük, hogy az A **valószínűsége**. Az $(\Omega, \mathcal{A}, P)$ hármast **valószínűségi mezőnek** hívjuk.

A véletlen kísérleteket mindig egy megfelelő valószínűségi mezővel írjuk le:

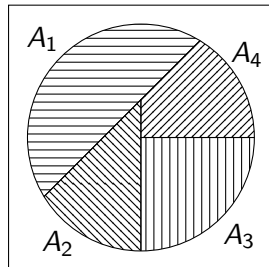
- eseménytér (Ω): a lehetséges kimeneteleinek a halmaza,
- eseményalgebra ($\mathcal{A}$): a kísérlethez kapcsolódó események,
- valószínűség (P).

Az additivitás azt fejezi ki, hogy ha egy eseményt felbontunk több kizáró eseményre, akkor

egész valószínűsége = részek valószínűségeinek az összege

A matematikában az additivitási szabály nem csak a valószínűségre teljesül, hanem más fogalmakra is. Például: tömeg, terület, térfogat. Ezeket a fogalmakat összefoglaló néven **mértékeknek** nevezzük.

$$\mu(A_1 \cup A_2 \cup \dots) = \mu(A_1) + \mu(A_2) + \dots$$



A nagy számok Borel-féle törvénye

Tekintsünk egy véletlen kísérletet, és legyen A egy tetszőleges esemény! Ekkor a relatív gyakoriság konvergál az esemény valószínűségéhez:

$$r_n(A) \rightarrow P(A), \quad n \rightarrow \infty.$$

A valószínűség tulajdonságai

- Komplementer esemény valószínűsége:

$$P(\bar{A}) = 1 - P(A).$$

- Lehetetlen esemény valószínűsége:

$$P(\emptyset) = P(\bar{\Omega}) = 1 - P(\Omega) = 0.$$

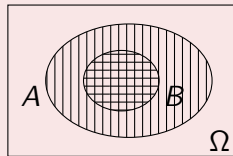
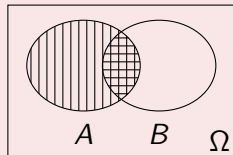
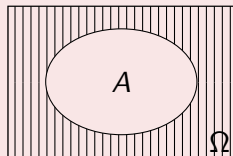
- Kivonási szabály:** tetszőleges A és B események mellett

$$P(A \setminus B) = P(A) - P(A \cap B).$$

Speciálisan ha B maga után vonja az A eseményt, akkor

$$P(A \setminus B) = P(A) - P(B).$$

- Monotonitás:** ha B maga után vonja az A eseményt, akkor $P(B) \leq P(A)$.



A valószínűség további tulajdonságai

- **Szubadditivitás:** Ha $A_1, A_2, \dots$ tetszőleges eseményeknek egy véges vagy végtelen sorozata, akkor

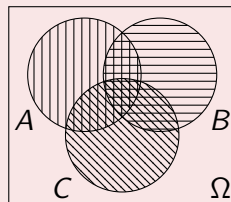
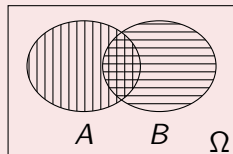
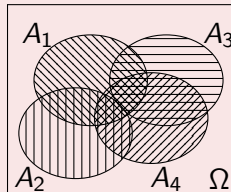
$$P(A_1 \cup A_2 \cup \dots) \leq P(A_1) + P(A_2) + \dots$$

- Két esemény összegének a valószínűsége: tetszőleges A és B esetén

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Három esemény összegének a valószínűsége: tetszőleges A , B és C esemény mellett

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\ &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &\quad + P(A \cap B \cap C) \end{aligned}$$



A valószínűség további tulajdonságai

- **Poincaré-formula** avagy **szitaformula**: tetszőleges $A_1, \dots, A_n$ események mellett

$$P(A_1 \cup \dots \cup A_n) = \sum_{k=1}^n (-1)^{k+1} \sum_{\substack{1 \leq i_1 < \dots < i_k \leq n \\ \text{különböző egészek}}} P(A_{i_1} \cap \dots \cap A_{i_k})$$

Ugyanez részletesebben:

$$P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n)$$

– ketteses metszetek valószínűsége

+ hármas metszetek valószínűsége

– négyes metszetek valószínűsége

⋮

$\pm P(A_1 \cap \dots \cap A_n)$

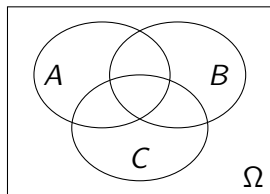
Feladat. Egy társaságban az emberek 40 százaléka beszél angolul, 35 százaléka németül, 30 százaléka franciául, 15 százaléka angolul és németül, 20 százaléka angolul és franciául, 20 százaléka németül és franciául, végül az emberek 10 százaléka beszél mindhárom nyelven. Tekintsük azt a kísérletet, hogy véletlenszerűen kiválasztunk egy embert a társaságból.

A kísérlet lehetséges kimenetelei az emberek a társaságban, tehát az eseménytér maga a társaság. Tekintsük a következő eseményeket:

A = a kiválasztott ember beszél angolul

B = a kiválasztott ember beszél németül

C = a kiválasztott ember beszél franciául



„Véletlenszerűen” választunk ki egy embert, tehát mindenkit azonos eséllyel választunk ki. Emiatt: valószínűség = arány a csoporton belül.

$$\begin{aligned} P(A) &= 0,4, & P(B) &= 0,35, & P(C) &= 0,3, & P(A \cap B \cap C) &= 0,1, \\ P(A \cap B) &= 0,15, & P(A \cap C) &= 0,2, & P(B \cap C) &= 0,2. \end{aligned}$$

Határozzuk meg annak a valószínűségét, hogy a kiválasztott ember...

- nem beszél angolul = nem $A = \bar{A}$

$$P(\bar{A}) = 1 - P(A) = 1 - 0,4 = 0,6$$

- beszél angolul, de nem beszél franciául = A teljesül, de C nem = $A \setminus C$

Most $C \not\subset A$, ezért $P(A \setminus C) \neq P(A) - P(C)$

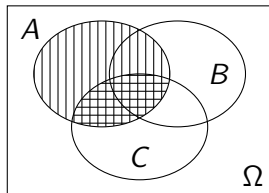
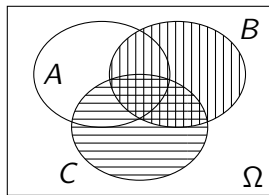
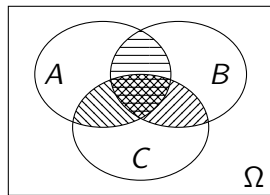
Helyes megoldás: $P(A \setminus C) = P(A) - P(A \cap C) = 0,4 - 0,2 = 0,2$

- beszél németül vagy franciául = $B \cup C$

$$P(B \cup C) = P(B) + P(C) - P(B \cap C) = 0,35 + 0,3 - 0,2 = 0,45$$

- D = legalább két nyelven beszél = $(A \cap B) \cup (A \cap C) \cup (B \cap C)$

$$P(D) = P(A \cap B) + P(A \cap C) + P(B \cap C) - 2P(A \cap B \cap C) = 0,35$$


 $A \setminus C$

 $B \cup C$

 D

Határozzuk meg annak a valószínűségét, hogy a kiválasztott ember...

- egyik nyelvet sem beszél = nem A és nem B és nem $C = \bar{A} \cap \bar{B} \cap \bar{C}$

Nincsen formulánk a metszetre. Térjünk át unióra!

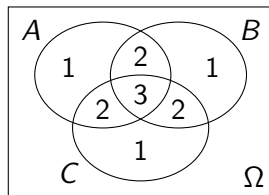
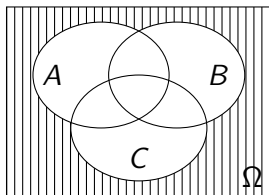
$$P(\bar{A} \cap \bar{B} \cap \bar{C}) = P(\overline{A \cup B \cup C}) = 1 - P(A \cup B \cup C)$$

Ez nem meglepő, hiszen:

$$P(\text{egyik nyelvet sem beszél}) = 1 - P(\text{beszéli valamelyik nyelvet})$$

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) \\ &\quad - P(B \cap C) + P(A \cap B \cap C) = 0,6 \end{aligned}$$

$$\text{Tehát: } P(\bar{A} \cap \bar{B} \cap \bar{C}) = 1 - P(A \cup B \cup C) = 1 - 0,6 = 0,4$$



Diszkrét valószínűségi mezők

Diszkrét valószínűségi mezők

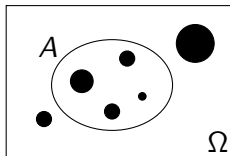
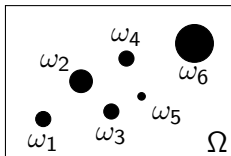
Azt mondjuk, hogy egy valószínűségi mező **diszkrét valószínűségi mező**, ha a kísérlet lehetséges kimenetelei egy véges vagy végtelen sorozatot alkotnak. Ekkor az eseménytér: $\Omega = \{\omega_1, \omega_2, \dots\}$.

Szemléletesen: szétosztunk 1 nagyságú súlyt a kimenetek között, és egy esemény valószínűsége azonos a benne található súlyok összegével.

Diszkrét valószínűségi mezőn a kimenetek valószínűségeinek az összege:

$$P(\omega_1) + P(\omega_2) + \dots = P(\{\omega_1, \omega_2, \dots\}) = P(\Omega) = 1$$

Egy tetszőleges A esemény valószínűsége: $P(A) = \sum_{\omega \in A} P(\omega)$.



Klasszikus valószínűségi mezők

Azt mondjuk, hogy egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező **klasszikus valószínűségi mező**, ha rendelkezik az alábbi tulajdonságokkal:

- a kísérletnek véges sok lehetséges kimenetele van, tehát $|\Omega| < \infty$;
- minden kimenetelnek azonos a valószínűsége.

Megjegyzések:

- A klasszikus valószínűségi mező egyben diszkrét valószínűségi mező is.
- Most minden ω kimenetelre $P(\omega) = 1/|\Omega|$.
- Bármely A eseményre $P(A) = \sum_{\omega \in A} P(\omega) = |A| \cdot 1/|\Omega|$.

Események valószínűsége klasszikus valószínűségi mezőn

Klasszikus valószínűségi mezőn egy tetszőleges A esemény valószínűsége:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{kedvező kimenetek száma}}{\text{összes kimenetek száma}}$$

Feladat. Egy szabályos pénzérmét feldobunk n alkalommal, ahol n egy rögzített pozitív egész szám. Milyen módon írható le a kísérlet?

A kísérlet lehetséges kimenetelei az n -hosszú fej-írás dobássorozatok:

$$\Omega = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in \{\text{fej}, \text{írás}\}\}, \quad |\Omega| = 2^n$$

A pénzérme szabályos, ezért minden kimenetelnek azonos a valószínűsége. Tehát ez egy klasszikus valószínűségi mező.

Feladat. Adjuk meg az alábbi események valószínűségét:

$$P(\text{sorban } n-1 \text{ fejet és végül } 1 \text{ írást kapunk}) = 1/2^n.$$

$$P(\text{dobunk legalább egy fejet}) = 1 - P(\text{csak írást dobunk}) = 1 - 1/2^n.$$

$$P(\text{pontosan } k \text{ darab írást kapunk}) = \binom{n}{k}/2^n, \quad k = 0, 1, \dots, n.$$

Feladat. Adjuk meg az $1 + x + x^2 + \dots$ összeg értékét, ahol $|x| < 1$.

A fenti összeget **mértani sornak** nevezzük, és a következő módon határozható meg:

$$\begin{aligned} 1 + x + x^2 + \dots &= \lim_{n \rightarrow \infty} (1 + x + \dots + x^n) \\ &= \lim_{n \rightarrow \infty} (1 + x + \dots + x^n) \frac{1 - x}{1 - x} = \lim_{n \rightarrow \infty} \frac{1 - x^{n+1}}{1 - x} = \frac{1 - 0}{1 - x} \end{aligned}$$

Speciálisan $x = 1/2$ esetén:

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots = 1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \dots = \frac{1}{1 - 1/2} = 2.$$

$x = 1/4$ esetén:

$$1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \dots = 1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^3 + \dots = \frac{1}{1 - 1/4} = \frac{4}{3}.$$

Feladat. Addig dobunk egy szabályos pénzérmével, míg írást nem kapunk. Milyen valószínűségi mezővel írható le a kísérlet?

A kísérletnek két fajta kimenetele van:

- Véges sok dobás után kapunk egy írást, és utána nem dobunk többet:

$$\omega_n = (\underbrace{\text{fej}, \dots, \text{fej}}_{n-1 \text{ darab}}, \text{írás}), \quad n = 1, 2, \dots$$

- Végtelen sokat dobunk, és mindig fejet kapunk: $\omega_\infty = (\text{fej}, \text{fej}, \dots)$.

A kimenetek egy végtelen sorozatot alkotnak, az eseménytér:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_\infty\}.$$

Megjegyzések:

- A kísérletet leíró valószínűségi mező egy diszkrét valószínűségi mező.
- De ez NEM egy klasszikus valószínűségi mező, hiszen végtelen sok kimenetel van.
- A kimenetek valószínűsége: $P(\omega_n) = 1/2^n$, $P(\omega_\infty) = ?$

Feladat. Mennyi annak az esélye, hogy sosem kapunk írást?

$$\begin{aligned} P(\text{előbb-utóbb dobunk írást}) &= P(\omega_1) + P(\omega_2) + P(\omega_3) + \dots \\ &= \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \dots = \frac{1}{2} \left[1 + \frac{1}{2} + \frac{1}{4} + \dots \right] = \frac{1}{2} \cdot 2 = 1. \end{aligned}$$

Emiatt: $P(\omega_\infty) = P(\text{sosem dobunk írást}) = 1 - 1 = 0$.

Feladat. Melyiknek nagyobb az esélye? Annak, hogy az első írást páros sorszámú, vagy annak, hogy első írást páratlan sorszámú dobásra kapjuk?

A = az első írást páratlan sorszámú dobásra kapjuk = $\{\omega_1, \omega_3, \omega_5, \dots\}$,

B = az első írást páros sorszámú dobásra kapjuk = $\{\omega_2, \omega_4, \omega_6, \dots\}$.

$$\begin{aligned} P(A) &= P(\omega_1) + P(\omega_3) + P(\omega_5) + \dots = \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \dots \\ &= \frac{1}{2} \left[1 + \frac{1}{4} + \frac{1}{4^2} + \dots \right] = \frac{1}{2} \cdot \frac{4}{3} = \frac{2}{3}. \end{aligned}$$

Továbbá: $P(A) + P(B) = P(A \cup B) = P(\text{valamikor írást dobunk}) = 1$.

Emiatt: $P(B) = 1 - P(A) = 1/3$.

Feladat. Egy vállalat úgy próbálja népszerűsíteni az általa forgalmazott chipset, hogy a zacskókba a Micimackó című mese figuráit rejtje el: Micimackót, Malackát, Tigrist és Fülest. Minden zacskóban pontosan egy figura található, és minden figurának azonos a gyakorisága. Mi addig vásároljuk a terméket, míg meg nem kapjuk a teljes kollekciót, tehát míg végül mindegyik figurából lesz legalább egy darab. Mennyi annak az esélye, hogy ehhez elég 4 darab chipset megvenni? Mi annak a valószínűsége, hogy elég 6 zacskót megvásárolni?

A feladat egy általános témakör speciális esete. Adott egy kísérlet, melynek véges sok lehetséges kimenetele van. A kérdés az, hogy a kísérletet többször elvégezve mekkora valószínűséggel fog az összes lehetséges kimenetel megjelenni. Ezt a problémakört nevezzük a **kupongyűjtő problémájának**. Az eredeti feladat több különböző szöveggel is megfogalmazható. Például:

- Tízszer feldobok egy szabályos dobókockát. Mennyi az esélye, hogy a dobások között az 1, 2, 3, 4, 5, 6 értékek mindegyike megjelenik?
- Anna minden nap véletlenszerűen választ egyet a 3 pár cipője közül. Mennyi az esélye, hogy 5 nap alatt mindhárom párat hordani fogja?

Kérdés. Mennyi annak az esélye, hogy elég 4 darab chipset megvenni?

$$P(\text{a 4 zacskóban 4 különböző figura lesz}) = \frac{\text{kedvező}}{\text{összes}} = \frac{4!}{4^4} = 0,09.$$

Kérdés. Mi annak a valószínűsége, hogy elég 6 zacskót megvásárolni?

Vezessük be a következő eseményeket:

A_i = az i -dik figurát egyik zacskóban sem találjuk meg, $i = 1, 2, 3, 4$

Ekkor a kérdéses valószínűség:

$$\begin{aligned} P(\text{a 6 zacskóból összeáll a teljes kollekció}) &= P(\bar{A}_1 \cap \bar{A}_2 \cap \bar{A}_3 \cap \bar{A}_4) \\ &= 1 - P(\text{valamelyik figurát nem kapjuk meg}) = 1 - P(A_1 \cup A_2 \cup A_3 \cup A_4) \end{aligned}$$

A Poincaré-formulával az unió valószínűsége:

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3 \cup A_4) &= P(A_1) + P(A_2) + P(A_3) + P(A_4) \\ &\quad - \text{kettes metszetek} + \text{hármass metszetek} - P(A_1 \cap A_2 \cap A_3 \cap A_4) \end{aligned}$$

A felírt metszetek valószínűségei:

$$P(A_1) = P(6 \text{ zacskóból nem kapunk Micimackót}) = 3^6/4^6,$$

$$P(A_1 \cap A_2) = P(\text{nem kapunk Micimackót és Malackát}) = 2^6/4^6,$$

$$P(A_1 \cap A_2 \cap A_3) = P(\text{csak Fülest kapunk}) = 1/4^6,$$

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(\text{semmit sem kapunk}) = 0/4^6 = 0.$$

Most 4 darab eseményünk van. Ezekből $\binom{4}{2}$ darab kettes és $\binom{4}{3}$ darab hármas metszet képezhető. Emiatt

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3 \cup A_4) &= P(A_1) + P(A_2) + P(A_3) + P(A_4) \\ &\quad - \text{kettes metszetek} + \text{hármas metszetek} - P(A_1 \cap A_2 \cap A_3 \cap A_4) \\ &= 4 \frac{3^6}{4^6} - \binom{4}{2} \frac{2^6}{4^6} + \binom{4}{3} \frac{1}{4^6} - 0 \approx 0,62 \end{aligned}$$

Tehát a kérdéses valószínűség:

$$P(6 \text{ zacskóból teljes kollekció}) = 1 - P(A_1 \cup A_2 \cup A_3 \cup A_4) \approx 0,38$$

Geometriai valószínűségi mezők

Feladat. Adott egy 5×5 méteres négyzet alakú kert, melynek egyik sarkában egy harapós kutya van kikötve egy $r = 3$ méteres láncon. A kert mellett gyerekek játszanak egy gumilabdával, és véletlenszerűen beejtik a labdát a kertbe. Mennyi a valószínűsége annak, hogy a kutya eléri, és ezáltal kiharapja a labdát?

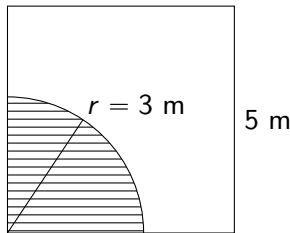
Észrevételek:

- A lehetséges kimenetek a kert (négyzet) pontjai.
- Végtelen sok kimenetel van, ez nem egy klasszikus valószínűségi mező.

Összes terület: $5 \cdot 5 = 25$

Kedvező terület: $3^2\pi/4 = 7,07$

$P(\text{a kutya eléri a labdát}) = 7,07/25 = 0,28$



A hosszúságot, a területet, és a térfogatot összefoglaló néven **mértéknek** nevezzük. Egy A geometriai alakzat mértékére (hosszúságára, területére vagy térfogatára) a $\mu(A)$ jelölést szokás alkalmazni.

Geometriai valószínűségi mezők

Azt mondjuk, hogy egy $(\Omega, \mathcal{A}, P)$ valószínűségi mező **geometriai valószínűségi mező**, ha rendelkezik az alábbi tulajdonságokkal:

- Az Ω eseménytér egy geometriai alakzat, és $0 < \mu(\Omega) < \infty$.
- **Egyenletességi hipotézis:** Az események valószínűsége egyenesen arányos az események mértékével. Tehát minden $A \subset \Omega$ eseményre

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{\text{kedvező hosszúság/terület/térfogat}}{\text{összes hosszúság/terület/térfogat}}$$

Az egyenletességi hipotézis azt fogalmazza meg, hogy egy A esemény valószínűsége csak attól függ, hogy mekkora az A halmaz nagysága, de attól már nem, hogy ez a halmaz hol helyezkedik el az eseménytéren belül.

Kérdés. Mennyi annak a valószínűsége, hogy a gyerekek pontosan az 5×5 méteres négyzet alakú kert geometriai középpontjára ejtik a labdát?

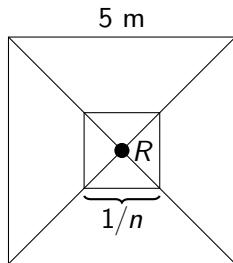
A geometriai középpont a két átló R metszéspontja. Tetszőleges n pozitív egész szám esetén a pont lefedhető egy $1/n$ oldalú kis négyzettel.

$$0 \leq \mu(R) \leq \text{a kis négyzet területe} = 1/n^2 \rightarrow 0, \quad n \rightarrow \infty.$$

Ekkor a rendőr-elv miatt: $\mu(R) = 0$.

$$P(\text{a labda az } R \text{ pontra esik}) = 0/25 = 0$$

Bekövetkezik ez az esemény?



Hasonló módon bizonyítható, hogy

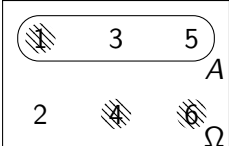
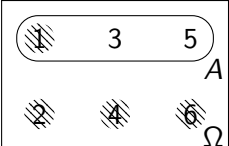
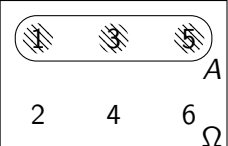
- minden pontnak 0 a hosszúsága, a területe és a térfogata;
- minden szakasznak illetve egyenesnek 0 a területe és a térfogata;
- minden síkidomnak 0 a térfogata. (Például: mértani testek lapjai.)

A feltételes valószínűség

Feladat. Feldobunk egy szabályos dobókockát, és legyen A az az esemény, hogy páratlan számot dobunk. Ekkor: $P(A) = 1/2$.

Adjuk meg A valószínűségét, ha rendelkezünk az alábbi információval!

- Ha tudjuk, hogy prímszámot dobtunk: $P(A) = 2/3$.
- Ha tudjuk, hogy 5-nél kisebbet dobtunk: $P(A) = 2/4 = 1/2$.
- Ha tudjuk, hogy 3-ast vagy 5-öst dobtunk: $P(A) = 2/2 = 1$.
- Ha tudjuk, hogy páros számot dobtunk: $P(A) = 0/3 = 0$.

			
--	---	--	---

A háttérinformáció nagy mértékben befolyásolja a valószínűséget. Hogyan számoljunk feltételes valószínűséget nem klasszikus valószínűségi mezőn?

Feltételes valószínűség

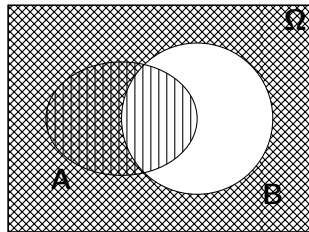
Tegyük fel, hogy $P(B) > 0$. Ekkor az A eseménynek a B eseményre vett **feltételes valószínűsége**

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

A feltételes valószínűség megmutatja, hogy mennyi az A esemény valószínűsége, ha tudjuk, hogy a B esemény bekövetkezik.

Formálisan a következő történik:

- kidobjuk a B -n kívüli kimeneteket, hiszen ezek most nem következhetnek be;
- a B halmaz egy „szűkített” eseménytér, ezen belül nézzük az A súlyarányát.



Feladat. Jelölje A azt az eseményt, hogy egy szabályos dobókockát feldobva páratlan számot dobunk. Adjuk meg A feltételes valószínűségét az alábbi B feltételekre nézve:

- $B =$ prímszámot dobunk

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\text{páratlan és prím})}{P(\text{prím})} = \frac{2/6}{1/2} = \frac{4}{6} = \frac{2}{3}$$

- $B =$ 5-nél kisebbet dobunk

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(\text{páratlan és 5-nél kisebb})}{P(\text{5-nél kisebb})} = \frac{2/6}{4/6} = \frac{2}{4} = \frac{1}{2}$$

Ezek az értékek azonosak a korábban kapott valószínűségekkel.

Hogyan számolhatunk feltételes valószínűséget, ha nem akarjuk a definíciót alkalmazni?

Klasszikus valószínűségi mező: dobjuk ki a B -n kívüli kimeneteket.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{|A \cap B|/|\Omega|}{|B|/|\Omega|} = \frac{|A \cap B|}{|B|} = \frac{\text{megmaradt kedvező}}{\text{megmaradt összes}}$$

Geometriai valószínűségi mező: dobjuk ki a B -n kívüli kimeneteket.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\mu(A \cap B)/\mu(\Omega)}{\mu(B)/\mu(\Omega)} = \frac{\mu(A \cap B)}{\mu(B)} = \frac{\text{megmaradt kedvező}}{\text{megmaradt összes}}$$

Két esemény függetlensége

Két esemény függetlensége

Legyenek A és B tetszőleges események! Azt mondjuk, hogy a két esemény **független** egymástól, ha $P(A \cap B) = P(A)P(B)$.

A függetlenséggel ekvivalens tulajdonságok

Ha A és B pozitív valószínűségű, akkor az alábbiak ekvivalensek:

- A és B függetlenek egymástól.
- $P(A|B) = P(A)$.
- $P(B|A) = P(B)$.

Bizonyítás. Ha A és B pozitív valószínűségű események, akkor

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

Feladat. Feldobunk egy szabályos dobókockát. Döntsük el, hogy az alábbi események függetlenek-e egymástól!

- A = páratlan számot dobunk, B = 5-nél kisebbet dobunk.

Most: $P(A) = 1/2$, $P(B) = 4/6$, $P(A \cap B) = 2/6$.

Tehát $P(A \cap B) = P(A)P(B)$, az események függetlenek.

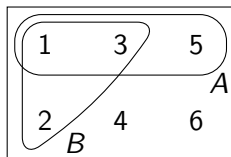
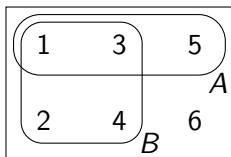
Alternatív megoldás: $P(A|B) = 2/4 = P(A)$, tehát függetlenek.

- A = páratlan számot dobunk, B = 4-nél kisebbet dobunk.

Most: $P(A) = 1/2$, $P(B) = 1/2$, $P(A \cap B) = 2/6$.

Tehát $P(A \cap B) \neq P(A)P(B)$, az események nem függetlenek.

Alternatív megoldás: $P(A|B) = 2/3 \neq P(A)$, nem függetlenek.



Feladat. Ha feldobunk két szabályos dobókockát, akkor feltehető, hogy a dobott értékek függetlenek egymástól. Mutassuk meg, hogy ezen feltevés mellett minden kimenetelnek azonos a valószínűsége!

Tetszőleges $i, j \in \{1, \dots, 6\}$ esetén legyen

A = az 1. kockával i -t dobunk, B = a 2. kockával j -t dobunk.

A feltevés szerint ezek független események, tehát

$$\begin{aligned} P(\text{az 1. kockával } i\text{-t, a 2. kockával } j\text{-t dobunk}) \\ = P(A \cap B) = P(A)P(B) = 1/6 \cdot 1/6 = 1/36. \end{aligned}$$

Feladat. Mutassunk példát olyan kísérletre, mikor a két kockadobás nem független! Ekkor is azonos a kimenetek valószínűsége?

A kísérlet: feldobok egy szabályos dobókockát, felveszem filmre, majd ezt lejátszom újra. Ekkor két szabályos kockadobást kapok, de ezek nem függetlenek. A kimenetek valószínűsége:

$$P(\text{az 1. kockával } i\text{-t, a 2. kockával } j\text{-t dobunk}) = \begin{cases} 1/6, & \text{ha } i = j, \\ 0, & \text{ha } i \neq j. \end{cases}$$

Feladat. Mi a kapcsolat a kizáróság (=diszjunktság) és a függetlenség között? Azonos a jelentésük, vagy az egyikből következik a másik?

Nem azonos a jelentésük, és általában egyik tulajdonságból sem következik a másik. Legyenek például A és B pozitív valószínűségű események.

- Ha a két esemény kizáró, akkor $P(A \cap B) = 0$. Ekkor az események nem lehetnek függetlenek:

$$0 = P(A \cap B) \neq P(A)P(B) > 0.$$

Ugyanez másfajta gondolatmenettel:

$$P(A|B) = P(A \cap B)/P(B) = 0/P(B) = 0 \neq P(A) > 0.$$

- Ha az események függetlenek, akkor nem lehetnek kizáróak, ugyanis

$$P(A \cap B) = P(A)P(B) > 0.$$

A fenti gondolatmenet csak pozitív valószínűségű eseményekre vonatkozott. Ha az egyik esemény valószínűsége 0, akkor lehetnek egyszerre kizáróak és függetlenek.

A feltételes valószínűség tulajdonságai

A feltételes valószínűség valóban valószínűség, azaz valószínűségi mérték?
A valószínűség definíciója alapján az alábbi tulajdonságokat kell ellenőrizni:

- A valószínűség egy olyan függvény, mely az eseményekhez 0 és 1 közötti értékeket rendel hozzá: $P(A|B) \in [0, 1]$.
- A biztos esemény valószínűsége: $P(\Omega|B) = 1$.
- Teljesül az additivitás: $A_1, A_2, \dots$ kizáró események mellett

$$P(A_1 \cup A_2 \cup \dots | B) = P(A_1|B) + P(A_2|B) + \dots$$

Könnyen ellenőrizhető, hogy ezek a tulajdonságok teljesülnek a feltételes valószínűségre, így igaz az alábbi tétel.

A feltételes valószínűség valószínűség

Legyen B pozitív valószínűségű esemény. Ekkor a B eseményre vett feltételes valószínűség valószínűségi mérték.

A tétel értelmében teljesülnek a következő azonosságok is:

- A lehetetlen esemény valószínűsége: $P(\emptyset|B) = 0$.
- Monotonitás: Ha $A_1 \subset A_2$, akkor $P(A_1|B) \leq P(A_2|B)$.
- A komplementer esemény valószínűsége: $P(\bar{A}|B) = 1 - P(A|B)$.
- Két esemény uniójának a valószínűsége:

$$P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B) - P(A_1 \cap A_2 | B).$$

- Poincaré-formula a feltételes valószínűsége...

Fontos: a fenti azonosságokban a B esemény rögzített! A feltételes valószínűségben a feltételt nem lehet ilyen könnyen manipulálni.

Példa. Feldobunk egy szabályos dobókockát.

Legyen $A = 6$ -ost dobunk, $B =$ páros számot dobunk.

Ekkor $P(A|B) = 1/3$, $P(A|\bar{B}) = 0$.

Tehát $P(A|\bar{B}) \neq 1 - P(A|B)$.

Láncszabály

Legyenek $A_1, \dots, A_n$ olyan események, melyekre $P(A_1 \cap \dots \cap A_n) > 0$.
Ekkor

$$P(A_1 \cap \dots \cap A_n) = P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2) \\ \dots P(A_n | A_1 \cap \dots \cap A_{n-1}).$$

Bizonyítás. A feltételes valószínűség definícióját többször alkalmazva:

$$\begin{aligned} & P(A_1)P(A_2 | A_1)P(A_3 | A_1 \cap A_2) \dots P(A_n | A_1 \cap \dots \cap A_{n-1}) \\ &= P(A_1) \frac{P(A_1 \cap A_2)}{P(A_1)} \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \dots \frac{P(A_1 \cap \dots \cap A_{n-1} \cap A_n)}{P(A_1 \cap \dots \cap A_{n-1})} \\ &= P(A_1 \cap \dots \cap A_{n-1} \cap A_n). \end{aligned}$$

Feladat. Adott n ember, és ki akarnak sorsolni maguk közül egy embert. Fognak n darab egyforma gyufaszálat, egynek letörlik a fejét, és egymás után húznak. Igazságos-e ez a sorsolás, tehát függ-e annak az esélye, hogy valaki a rövidet húzza attól, hogy hanyadikként húz?

Tekintsük az alábbi eseményeket:

$$A_i = \text{az } i. \text{ ember húzza a rövid szálat.}, \quad i = 1, \dots, n.$$

A sorsolás akkor igazságos, ha $P(A_1) = \dots = P(A_n) = 1/n$.

1. Megközelítés. Az első húzó esetében:

$$P(A_1) = \text{kedvező/összes} = 1/n.$$

A második húzó esetében:

$$P(A_2) = \text{kedvező/összes} = 1/(n-1).$$

Ezek szerint a sorsolás nem igazságos? Itt valami nem stimmel.

Probléma: mi van akkor, ha a második ember nem is húz?

2. Megközelítés. Az első húzó esetében: $P(A_1) = 1/n$.

A második húzónál: $A_2 \subset \bar{A}_1$, ezért a láncszabály szerint:

$$P(A_2) = P(\bar{A}_1 \cap A_2) = P(\bar{A}_1) \cdot P(A_2 | \bar{A}_1) = \frac{n-1}{n} \cdot \frac{1}{n-1} = \frac{1}{n}.$$

Az n . ember esetében: $A_n \subset \bar{A}_1 \cap \dots \cap \bar{A}_{n-1}$, ezért

$$\begin{aligned} P(A_n) &= P(\bar{A}_1 \cap \dots \cap \bar{A}_{n-1} \cap A_n) \\ &= P(\bar{A}_1) P(\bar{A}_2 | \bar{A}_1) P(\bar{A}_3 | \bar{A}_1 \cap \bar{A}_2) \\ &\quad \dots P(\bar{A}_{n-1} | \bar{A}_1 \cap \dots \cap \bar{A}_{n-2}) P(A_n | \bar{A}_1 \cap \dots \cap \bar{A}_{n-1}) \\ &= \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot \frac{n-3}{n-2} \cdot \dots \cdot \frac{1}{2} \cdot 1 = \frac{1}{n}. \end{aligned}$$

Házi feladat: $P(A_i) = 1/n$ minden i esetén.

Tehát a sorsolás szabályos.

Bayes-formula

Ha A és B pozitív valószínűségi események, akkor

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

Bizonyítás. A láncszabály alkalmazásával:

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

Feladat. Egy gyárban anyaghibás és mérethibás gyártmányok is készülnek, és ezek a teljes mennyiség 25 illetve 40 százalékát teszik ki. A mérethibás darabok ötöde anyaghibás is egyben. Az anyaghibás gyártmányos mekkora hányada mérethibás?

$$\begin{aligned} P(\text{mérethibás} \mid \text{anyaghibás}) &= \frac{P(\text{anyaghibás} \mid \text{mérethibás}) \cdot P(\text{mérethibás})}{P(\text{anyaghibás})} \\ &= \frac{0,2 \cdot 0,4}{0,25} = 0,32 = 32\% \end{aligned}$$

A teljes valószínűség tétele

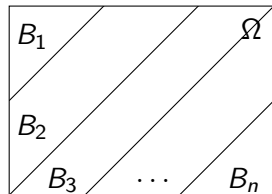
Teljes eseményrendszer

Azt mondjuk, hogy a $B_1, \dots, B_n$ események **teljes eseményrendszert** (t.e.r.) alkotnak, ha teljesítik az alábbi tulajdonságokat:

- páronként diszjunktak, tehát tetszőleges $i \neq j$ esetén $B_i \cap B_j = \emptyset$;
- együttesen lefedik az eseményteret: $B_1 \cup \dots \cup B_n = \Omega$;
- mindegyik eseménynek pozitív a valószínűsége.

- A teljes eseményrendszer az eseménytér egy partíciója, felbontása egymást kizáró pozitív valószínűségű eseményekre.
- A kísérletet végrehajtva az események közül pontosan egy következik be.
- A valószínűségek összege:

$$P(B_1) + \dots + P(B_n) = P(B_1 \cup \dots \cup B_n) = P(\Omega) = 1.$$



A teljes valószínűség tétele és a Bayes-tétel

Legyen $B_1, \dots, B_n$ teljes eseményrendszer, és tekintsünk egy tetszőleges A eseményt. Ekkor teljesülnek az alábbi összefüggések:

- **Teljes valószínűség tétele:**

$$P(A) = \sum_{i=1}^n P(A|B_i)P(B_i) = P(A|B_1)P(B_1) + \dots + P(A|B_n)P(B_n)$$

- **Bayes-tétel:** Ha $P(A) > 0$, akkor

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^n P(A|B_i)P(B_i)}$$

Bizonyítás. Az additivitás és a láncszabály alkalmazásával:

$$\begin{aligned} P(A) &= P(A \cap B_1) + \dots + P(A \cap B_n) \\ &= P(A|B_1)P(B_1) + \dots + P(A|B_n)P(B_n). \end{aligned}$$

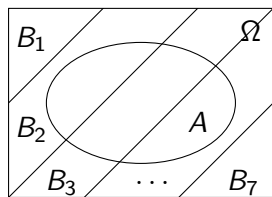
A Bayes-formulával és a teljes valószínűség tételével:

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^n P(A|B_i)P(B_i)}.$$

Feladat. Szeretnénk meghatározni az országos munkanélküliségi rátát, de sajnos a foglalkoztatási adatok csak regionális szinten állnak rendelkezésre. Hogyan kapható meg ezekből az országos ráta? Arra is kíváncsiak vagyunk, hogy a munkanélküliek mekkora hányada él az egyes régiókban.

Kísérlet: véletlenszerűen kiválasztunk egy munkaképes korú állampolgárt.

A = a kiválasztott ember munkanélküli,
 B_i = a kiválasztott ember az i . régióban él,
 $B_1, \dots, B_7$ teljes eseményrendszer.



A valószínűségek jelentése:

- $P(A)$ = országos munkanélküliségi ráta,
- $P(B_i)$ = az állampolgárok ekkora hányada él az i . régióban, (ismert)
- $P(A|B_i)$ = munkanélküliségi ráta az i . régióban, (ismert)
- $P(B_i|A)$ = a munkanélküliek ekkora hányada él az i . régióban.

Feladat. Tekintsük az alábbi fiktív adatokat:

i	1	2	3	4	5	6	7
$P(A B_i)$	5%	8%	10%	5%	12%	8%	5%
$P(B_i)$	40%	10%	10%	10%	10%	10%	10%

- Hogyan aránylik a munkanélküliek száma az egyes régiókban a teljes munkaképes lakossághoz?

Az i . régióban élő munkanélküliek aránya a teljes lakosságon belül:

$$P(\text{munkanélküli és az } i\text{-edik régióban él}) = P(A \cap B_i) = P(A|B_i)P(B_i).$$

i	1	2	3	4	5	6	7
$P(A B_i)$	5%	8%	10%	5%	12%	8%	5%
$P(B_i)$	40%	10%	10%	10%	10%	10%	10%
$P(A \cap B_i)$	2%	0,8%	1%	0,5%	1,2%	0,8%	0,5%

- Adjuk meg az országos munkanélküli rátát!

A teljes valószínűség tételével:

$$P(A) = \sum_{i=1}^7 P(A|B_i)P(B_i) = \sum_{i=1}^7 P(A \cap B_i) = 6,8\%$$

- A munkanélküliek mekkora hányada él az egyes régiókban?

$$P(\text{az } i\text{-edik régióban él} \mid \text{munkanélküli}) = P(B_i|A) = \frac{P(A \cap B_i)}{P(A)}$$

i	1	2	3	4	5	6	7
$P(A B_i)$	5%	8%	10%	5%	12%	8%	5%
$P(B_i)$	40%	10%	10%	10%	10%	10%	10%
$P(A \cap B_i)$	2%	0,8%	1%	0,5%	1,2%	0,8%	0,5%
$P(B_i A)$	29,4%	11,7%	14,7%	7,4%	17,6%	11,7%	7,4%

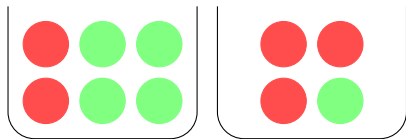
Feladat. Adott két urna. Az egyikben 2 piros és 4 zöld, a másikban 3 piros és 1 zöld golyót találhatók. Véletlenszerűen kiválasztunk egy urnát, és kiveszünk belőle egy golyót. Mennyi annak az esélye, hogy pirosat húzunk?

Legyen:

A = pirosat húzunk,

B_1 = az 1. urnát választjuk,

B_2 = az 2. urnát választjuk.



Ekkor: $P(B_1) = P(B_2) = 1/2$, $P(A|B_1) = 2/6$, $P(A|B_2) = 3/4$.

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) = \frac{2}{6} \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{1}{2} = \frac{13}{24}.$$

Feladat. Tegyük fel, hogy a húzás során zöld golyót kaptunk! Mennyi az esélye, hogy az 1. urnából húztunk?

A Bayes-formula alkalmazásával

$$P(B_1|\bar{A}) = \frac{P(\bar{A}|B_1)P(B_1)}{P(\bar{A})} = \frac{4/6 \cdot 1/2}{11/24} = \frac{6}{11} > \frac{1}{2} = P(B_1).$$

Feladat. Összeöntjük a golyókat egy urnába, majd abból húzunk egyet. Mennyi az esélye, hogy pirosat kapunk?

Az összeöntés után 5 piros és 5 zöld golyó van egy urnában, így

$$P(A) = 5/10 = 1/2.$$

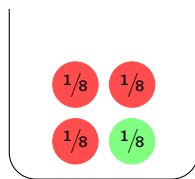
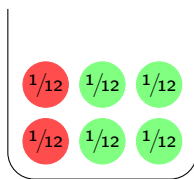
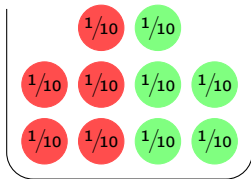
Miért változott meg a valószínűség az előző elrendezéshez viszonyítva?

- Az egyurnás elrendezésnél minden golyónak $1/10$ az esélye, ezért

$$P(A) = 5 \cdot 1/10 = 5/10.$$

- A kéturnás elrendezésnél az első urnában minden golyónak $1/12$, a másodikban $1/8$ az esélye, ezért

$$P(A) = 2 \cdot 1/12 + 3 \cdot 1/8 = 13/24.$$



Feladat. Egy urnában 5 piros és 5 zöld golyó található. Visszatevés nélkül kihúzunk két golyót. Mennyi a valószínűsége, hogy a második golyó piros?

Vezessük be a következő eseményeket:

A_1 = az első golyó piros, A_2 = a második golyó piros.

Ekkor: $P(A_1) = 1/2$, $P(A_2|A_1) = 4/9$, $P(A_2|\bar{A}_1) = 5/9$,

$$P(A_2) = P(A_2|A_1)P(A_1) + P(A_2|\bar{A}_1)P(\bar{A}_1) = \frac{4}{9} \cdot \frac{1}{2} + \frac{5}{9} \cdot \frac{1}{2} = \frac{1}{2}.$$

Feladat. Mennyi az esélye, hogy az első golyó piros volt, ha tudjuk, hogy a második piros?

A Bayes-formula alkalmazásával

$$P(A_1|A_2) = \frac{P(A_2|A_1)P(A_1)}{P(A_2)} = \frac{4/9 \cdot 1/2}{1/2} = \frac{4}{9} < \frac{1}{2} = P(A_1).$$

Tehát az a tény, hogy a második piros, csökkenti annak az esélyét, hogy az első piros volt. Ez hogyan lehetséges? Időutazás?

Több esemény függetlensége

Több esemény függetlensége

Legyen $A_1, A_2, \dots$ eseményeknek egy véges vagy végtelen sorozata.

- Azt mondjuk, hogy az események **páronként függetlenek**, ha bármely kettő független egymástól, tehát tetszőleges A_i és A_j különböző eseményeket kiválasztva

$$P(A_i \cap A_j) = P(A_i)P(A_j).$$

- Azt mondjuk, hogy az események **(teljesen) függetlenek**, ha közülük tetszőleges $A_{i_1}, \dots, A_{i_n}$ különböző eseményeket kiválasztva

$$P(A_{i_1} \cap \dots \cap A_{i_n}) = P(A_{i_1}) \dots P(A_{i_n}).$$

A feladatokban, ha csak az ellenkezőjét külön nem hangsúlyozzuk, akkor függetlenség alatt teljes függetlenséget értünk.

Mi a kétféle függetlenség szemléletes jelentése?

- Páronkénti függetlenség: tetszőleges A_i és A_j különböző eseményekre

$$P(A_j|A_i) = P(A_j).$$

Ez azt fejezi ki, hogy az $A_1, A_2, \dots$ sorozatban egyik esemény sem tartalmaz információt semelyik másik eseményre sem.

- Teljes függetlenség: tetszőleges $A_{i_1}, \dots, A_{i_n}$ különböző eseményeket kiválasztva

$$\begin{aligned} P(A_j | A_{i_1} \cap \dots \cap A_{i_n}) &= \frac{P(A_j \cap A_{i_1} \cap \dots \cap A_{i_n})}{P(A_{i_1} \cap \dots \cap A_{i_n})} \\ &= \frac{P(A_j)P(A_{i_1}) \dots P(A_{i_n})}{P(A_{i_1}) \dots P(A_{i_n})} = P(A_j) \end{aligned}$$

Ez több, mint a páronkénti függetlenség.

Kapcsolat a kétféle függetlenség között

Ha az $A_1, A_2, \dots$ események teljesen függetlenek, akkor páronként is függetlenek. Az állítás megfordítása nem igaz, tehát a páronkénti függetlenségből nem következik a teljes függetlenség.

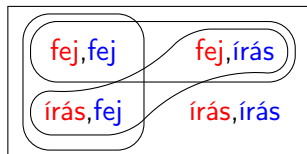
Bizonyítás. Ha az események teljesen függetlenek, akkor $n = 2$ mellett azonnal következik a páronkénti függetlenség. A következő példa pedig illusztrálja, hogy az állítás nem megfordítható.

Feldobunk két szabályos pénzérmét, egy pirosat és egy kékét. Legyen

A = a piros érmével fejet dobunk,

B = a kék érmével fejet dobunk,

C = összesen 1 fejet dobunk.



Ekkor megmutatható, hogy a három esemény páronként független. Ezzel szemben a három esemény nem teljesen független, ugyanis

$$P(A \cap B \cap C) = 0 \neq P(A)P(B)P(C) > 0.$$

Feladat. Tegyük fel, hogy a családokban egymástól független a gyerekek neme, továbbá az egyes gyerekek $1/2$ – $1/2$ eséllyel születnek fiúnak illetve lánynak! Véletlenszerűen kiválasztunk egy háromgyerekes családot. Mutassuk meg, hogy ekkor minden családtípusnak azonos az esélye!

A kísérletnek összesen $2^3 = 8$ lehetséges kimenetele van:

$$\Omega = \{FFF, FFL, FLF, LFF, FLL, LFL, LLF, LLL\}$$

Legyen: $A_i =$ az i . gyerek fiú

Ekkor

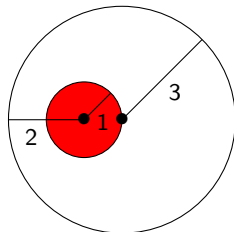
- $P(A_1) = P(A_2) = P(A_3) = 1/2$,
- az A_1, A_2, A_3 események teljesen függetlenek.

A függetlenség alkalmazásával:

$$P(FLF) = P(A_1 \cap \overline{A_2} \cap A_3) = P(A_1)P(\overline{A_2})P(A_3) = (1/2)^3 = 1/8.$$

Feladat. Egy kör alakú és 3 méter sugarú kerti tó felszínén, a tó szélétől 2 méterre egy molnárka áll lesben. A víz felszínére eső apró rovarok közül azokat veszi észre, amik tőle legfeljebb 1 méterre vannak. Ha egy rovar véletlenszerűen esik a vízbe, akkor mennyi az esélye, hogy ezt a molnárka észreveszi?

$$P(\text{a molnárka észreveszi a rovar}) = \frac{\text{kedvező terület}}{\text{összes terület}} = \frac{1^2\pi}{3^2\pi} = \frac{1}{9} \approx 11\%$$



Feladat. Ha 3 rovar esik egymástól függetlenül a vízre, akkor mennyi az esélye, hogy a molnárka mindhármat észreveszi?

Legyen: $A_i =$ a molnárka észreveszi az i . rovar, $i = 1, 2, 3$.

$P(A_1) = P(A_2) = P(A_3) = 1/9$, és A_1, A_2, A_3 független események.

$P(\text{mindhármat észreveszi}) = P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3) = (1/9)^3$

Feladat. Ha 3 rovar esik egymástól függetlenül a vízre, akkor mennyi a valószínűsége az alábbi eseményeknek?

- A molnárka észreveszi az elsőt, de a harmadikat nem.

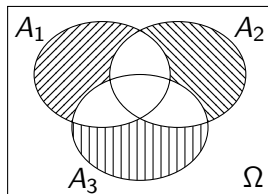
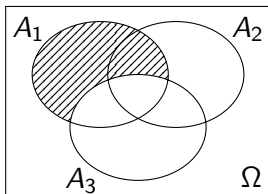
$$P(A_1 \setminus A_3) = P(A_1) - P(A_1 \cap A_3) = P(A_1) - P(A_1)P(A_3) = 1/9 - (1/9)^2$$

Alternatív megoldás:

$$P(A_1 \setminus A_3) = P(A_1 \cap \bar{A}_3) = P(A_1)P(\bar{A}_3) = 1/9 \cdot 8/9 \approx 10\%$$

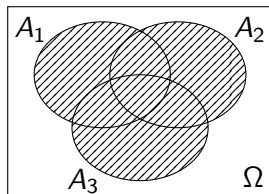
- A molnárka a három közül pontosan egy rovart vesz észre.

$$\begin{aligned} &P(A_1 \cap \bar{A}_2 \cap \bar{A}_3) + P(\bar{A}_1 \cap A_2 \cap \bar{A}_3) + P(\bar{A}_1 \cap \bar{A}_2 \cap A_3) \\ &= P(A_1)P(\bar{A}_2)P(\bar{A}_3) + P(\bar{A}_1)P(A_2)P(\bar{A}_3) + P(\bar{A}_1)P(\bar{A}_2)P(A_3) \\ &= 1/9 \cdot 8/9 \cdot 8/9 + 8/9 \cdot 1/9 \cdot 8/9 + 8/9 \cdot 8/9 \cdot 1/9 \approx 26\% \end{aligned}$$



- A molnárka észrevesz legalább egy rovar a három közül.

$$\begin{aligned}
 P(A_1 \cup A_2 \cup A_3) &= 1 - P(\overline{A_1 \cup A_2 \cup A_3}) \\
 &= 1 - P(\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}) = 1 - P(\overline{A_1})P(\overline{A_2})P(\overline{A_3}) \\
 &= 1 - (8/9)^3 \approx 30\%
 \end{aligned}$$



Feladat. Hány rovarnak kell egymástól függetlenül a vízre esnie ahhoz, hogy a molnárka 90% eséllyel észrevegyen közülük legalább egyet?

Ha n rovar esik a víz felszínére, akkor

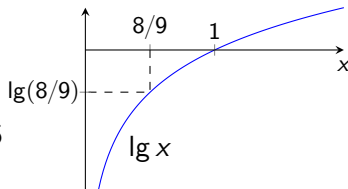
$$\begin{aligned}
 P(\text{a molnárka észrevesz legalább egyet}) &= 1 - P(\text{egyet sem vesz észre}) \\
 &= 1 - P(\overline{A_1} \cap \dots \cap \overline{A_n}) = 1 - P(\overline{A_1}) \dots P(\overline{A_n}) = 1 - (8/9)^n \geq 0,9
 \end{aligned}$$

$$(8/9)^n \leq 0,1$$

$$n \lg(8/9) = \lg(8/9)^n \leq \lg 0,1$$

$$n \geq \lg 0,1 / \lg(8/9) = 19,5$$

$$n \geq 20$$



Valószínűségi változók

Számos esetben előfordul, hogy nem a véletlen kísérlet kimeneteleire, hanem egy véletlen nagyságú mennyiségre vagyunk kíváncsiak. Ezek a véletlen számok technikailag $\xi : \Omega \rightarrow \mathbb{R}$ függvények. Például:

- Feldobunk három pénzérmét, jelölje ξ a kapott fejek számát!

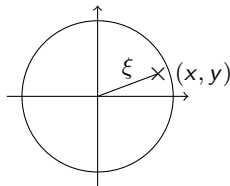
$$\Omega = \{(x_1, x_2, x_3) : x_1, x_2, x_3 \in \{\text{fej}, \text{írás}\}\}$$

$\xi : (x_1, x_2, x_3) \mapsto$ a fejek száma az x_1, x_2, x_3 eredmények között

- Véletlenszerűen választunk egy pontot egy origó középpontú körlapon, jelölje ξ a középponttól való távolságot!

$\Omega =$ a körlap (x, y) pontjai

$$\xi : (x, y) \mapsto \sqrt{x^2 + y^2}.$$



Valószínűségi változók

Legyen $(\Omega, \mathcal{A}, P)$ egy tetszőleges véletlen kísérletet leíró valószínűségi mező! A $\xi : \Omega \rightarrow \mathbb{R}$ függvényeket **valószínűségi változóknak** vagy **véletlen változóknak** nevezzük. Az **értékkészlet** a valószínűségi változó lehetséges értékeinek a halmaza. Jele: R_ξ .

Megjegyzések:

- A valószínűségi változó egy olyan mennyiség, melynek értéke csak a kísérlet kimenetelétől függ, más tényezőktől nem.
- A valószínűségi változó nem azért véletlen, mert a hozzárendelési szabály véletlen lenne, hanem azért, mert a kimenetel véletlen.
- A hozzárendelési szabály nagyon bonyolult is lehet, nem mindig tudjuk matematikai formulákkal felírni.
- A fenti definíció nem teljesen pontos: a valószínűségi változóknak teljesíteniük kell a **mérhetőségi feltételt** is. A mérhetőségi feltétel egy technikai követelmény, ami minden épeszű alkalmazásban teljesül.
- Jelölés a valószínűségi változókra: ξ (kszi), η (éta)

Diszkrét valószínűségi változók

Diszkrét valószínűségi változók

A ξ valószínűségi változó **diszkrét**, ha értékkészlete egy megszámlálható halmaz, tehát egy véges vagy végtelen sorozat: $R_\xi = \{x_1, x_2, \dots\}$.

Egy diszkrét valószínűségi változó **valószínűségeloszlása** a lehetséges értékek valószínűségei:

$$p_{x_k} = P(\xi = x_k) \quad x_k \in R_\xi.$$

Egy diszkrét valószínűségi változó **móduszai** azok az értékek, melyeket a legnagyobb valószínűséggel vesz fel.

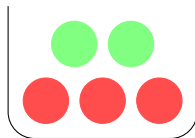
Ha az alábbi feltételek közül valamelyik teljesül, akkor a valószínűségi változó diszkrét:

- a változónak csak véges sok lehetséges értéke van;
- a lehetséges értékek mind egész számok.

Feladat. Szükségünk van egy működőképes biztosítékra. Egy fiókban van 5 biztosítékunk, de ezek közül 3 kiégett, tehát csak 2 használható. Egymás után megvizsgáljuk a biztosítékokat egészen addig, míg nem találunk egy használhatót. Jelölje a ξ valószínűségi változó azt, hogy hanyadik húzásra kapjuk meg az első működőképes biztosítékot. Adjuk meg ξ értékkészletét, valószínűségeloszlását és móduszait!

A változó lehetséges értékei: $R_\xi = \{1, 2, 3, 4\}$.

Mivel a változónak csak véges sok lehetséges értéke van, a ξ diszkrét változó.



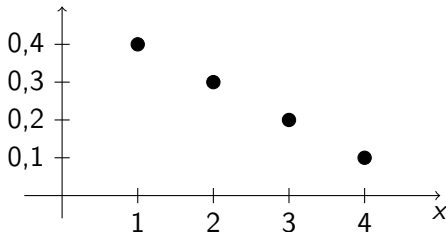
A valószínűségeloszlás a láncszabállyal írható fel. Például:

$$\begin{aligned} P(\xi = 3) &= P(\text{két rossz és utána egy jó biztosíték}) \\ &= P(\text{az első rossz}) \\ &\quad \cdot P(\text{a második rossz} \mid \text{az első rossz}) \\ &\quad \cdot P(\text{a harmadik jó} \mid \text{az első és a második rossz}) \\ &= \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = 0,2 \end{aligned}$$

$$p_1 = P(\xi = 1) = \frac{2}{5} = 0,4, \quad p_2 = P(\xi = 2) = \frac{3}{5} \cdot \frac{2}{4} = 0,3,$$

$$p_3 = P(\xi = 3) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = 0,2, \quad p_4 = P(\xi = 4) = \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot \frac{2}{2} = 0,1.$$

x	1	2	3	4
p_x	0,4	0,3	0,2	0,1



A valószínűségi változónak egy módusza van, az 1 érték.

Feladat. Mennyi annak a valószínűsége, hogy a 2-nél több biztosítékot kell megvizsgálnunk? Mennyi az esélye, hogy a megvizsgált biztosítékok száma legalább 2, de kisebb, mint 4?

$$P(\xi > 2) = P(\xi = 3) + P(\xi = 4) = 0,2 + 0,1 = 0,3,$$

$$P(2 \leq \xi < 4) = P(\xi = 2) + P(\xi = 3) = 0,3 + 0,2 = 0,5.$$

A valószínűségeloszlások karakterizációja

Egy $p_1, p_2, \dots$ véges vagy végtelen elemszámú sorozat pontosan akkor egy diszkrét valószínűségi változó eloszlása, ha teljesül az alábbi két feltétel:

- a sorozat elemei nemnegatívok: $p_1, p_2, \dots \geq 0$;
- a sorozat elemeinek az összege 1: $p_1 + p_2 + \dots = 1$.

Feladat. Mennyi legyen az a paraméter, hogy az alábbi értékek valószínűségeloszlást alkossanak?

$$p_1 = 0,15a, \quad p_2 = 0,5, \quad p_3 = a^2, \quad p_4 = 0,25a, \quad p_5 = 0,18.$$

Az összeget ellenőrizve az alábbi egyenletet kapjuk:

$$1 = p_1 + p_2 + p_3 + p_4 + p_5 = a^2 + 0,4a + 0,68,$$

Az egyenlet két gyöke:

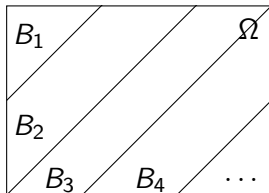
- $a = -0,8$: ez nem megoldás, ugyanis $p_1 < 0$.
- $a = 0,4$: ez már megoldás, ugyanis

$$p_1 = 0,06, \quad p_2 = 0,5, \quad p_3 = 0,16, \quad p_4 = 0,1, \quad p_5 = 0,18.$$

Bizonyítás. Először tegyük fel, hogy létezik olyan ξ nemnegatív egész értékű diszkrét változó, melyre $p_k = P(\xi = x_k)$ minden k esetén. Ekkor:

- $p_k = P(\xi = x_k) \geq 0$ minden k esetén,
- $p_1 + p_2 + \dots = P(\xi = x_1) + P(\xi = x_2) + \dots = 1$.

Fordított irány: tekintsünk egy $p_1, p_2, \dots$ számsorozatot, melyre teljesülnek a fenti tulajdonságok! Bontsuk fel az Ω halmazt $B_1, B_2, \dots$ kizáró eseményekre úgy, hogy $P(B_k) = p_k$ teljesüljön minden k -ra.



Definiáljuk a ξ nemnegatív egész értékű változót a következő módon:

$$\xi(\omega) = k \text{ pontosan akkor, ha } \omega \in B_k.$$

Ekkor $P(\xi = k) = P(B_k) = p_k$.

Adott egy ξ diszkrét valószínűségi változó, $R_\xi = \{x_1, x_2, \dots\}$.

Kérdés: mennyi a ξ változó átlagos értéke?

Megismételjük a kísérletet n alkalommal, a megfigyelt értékek: $\xi_1, \dots, \xi_n$.
Ekkor a megfigyelt értékek átlaga:

$$\begin{aligned}\frac{\xi_1 + \dots + \xi_n}{n} &= \frac{x_1 k_n(\xi = x_1) + x_2 k_n(\xi = x_2) + \dots}{n} \\ &= x_1 r_n(\xi = x_1) + x_2 r_n(\xi = x_2) + \dots \\ &\approx x_1 P(\xi = x_1) + x_2 P(\xi = x_2) + \dots\end{aligned}$$

Diszkrét valószínűségi változók várható értéke

Legyen ξ diszkrét valószínűségi változó, és legyen $R_\xi = \{x_1, x_2, \dots\}$.
Ekkor a ξ változó **várható értéke**:

$$E(\xi) = \sum_{x \in R_\xi} x P(\xi = x) = x_1 P(\xi = x_1) + x_2 P(\xi = x_2) + \dots$$

Megjegyzések:

- Szokásos jelölések a várható értékre: E (expected value); M (mean).
- A várható érték jelentése: a ξ változó hosszútávú átlagos értéke.
- Jellemzően: várható érték $\neq$ módusz.
- A várható érték nem mindig eleme az értékkészletnek.
- A várható érték lehet végtelen is!

A nagy számok Kolmogorov-féle törvénye (1933)

Legyen ξ tetszőleges valószínűségi változó, és legyenek $\xi_1, \xi_2, \dots$ a ξ egymástól független megfigyelései! Ekkor

$$\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow E(\xi), \quad n \rightarrow \infty.$$

Feladat. Adjuk meg a várható értéket a biztosítékos feladatban!

$$\begin{aligned} E(\xi) &= 1 \cdot P(\xi = 1) + 2 \cdot P(\xi = 2) + 3 \cdot P(\xi = 3) + 4 \cdot P(\xi = 4) \\ &= 1 \cdot 0,4 + 2 \cdot 0,3 + 3 \cdot 0,2 + 4 \cdot 0,1 = 2 \end{aligned}$$

Szóródási mutatószámok

Legyen ξ olyan valószínűségi változó, melynek véges a várható értéke!

- **Várható értéktől vett átlagos eltérés:** $E(|\xi - E(\xi)|)$
- **Variancia** avagy **szórásnégyzet:**

$$\text{Var}(\xi) = D^2(\xi) = E\left([\xi - E(\xi)]^2\right)$$

- **Szórás:** $D(\xi) = \sqrt{\text{Var}(\xi)}$

Megjegyzések:

- A várható értéktől vett átlagos eltérés egy egyszerű és szemléletes mutatószám, de nincsenek szép matematikai tulajdonságai.
- A szórás definíciójában

$$D(\xi) = \sqrt{E\left([\xi - E(\xi)]^2\right)} \approx E\left(\sqrt{[\xi - E(\xi)]^2}\right) = E(|\xi - E(\xi)|)$$

Tehát: szórás $\approx$ várható értéktől vett átlagos eltérés.

Feladat. Határozzuk meg az előző oldalon bevezetett mutatószámokat a biztosítékos feladatban!

ξ	1	2	3	4
$ \xi - E(\xi) $	1	0	1	2
$[\xi - E(\xi)]^2$	1	0	1	4
p_x	0,4	0,3	0,2	0,1

Várható értéktől való átlagos eltérés:

$$\begin{aligned}
 E(|\xi - E(\xi)|) &= |1 - 2| \cdot p_1 + |2 - 2| \cdot p_2 + |3 - 2| \cdot p_3 + |4 - 2| \cdot p_4 \\
 &= 1 \cdot 0,4 + 0 \cdot 0,3 + 1 \cdot 0,2 + 2 \cdot 0,1 = 0,8
 \end{aligned}$$

Variancia:

$$\begin{aligned}
 \text{Var}(\xi) &= [1 - 2]^2 \cdot p_1 + [2 - 2]^2 \cdot p_2 + [3 - 2]^2 \cdot p_3 + [4 - 2]^2 \cdot p_4 \\
 &= 1 \cdot 0,4 + 0 \cdot 0,3 + 1 \cdot 0,2 + 4 \cdot 0,1 = 1
 \end{aligned}$$

Szórás: $D(\xi) = \sqrt{\text{Var}(\xi)} = 1.$

Diszkrét valószínűségi változók transzformáltjai

Tekintsünk egy ξ diszkrét valószínűségi változót és egy $h : \mathbb{R} \rightarrow \mathbb{R}$ függvényt! Ekkor a $h(\xi)$ transzformált változó várható értéke:

$$E(h(\xi)) = \sum_{x \in R_\xi} h(x)P(\xi = x) = h(x_1)P(\xi = x_1) + h(x_2)P(\xi = x_2) + \dots$$

Néhány speciális eset:

- Várható értéktől való átlagos eltérés: $h(x) = |x - E(\xi)|$,

$$E(|\xi - E(\xi)|) = E(h(\xi)) = \sum_{x \in R_\xi} |x - E(\xi)|P(\xi = x)$$

- Variancia: $h(x) = [x - E(\xi)]^2$,

$$\text{Var}(\xi) = E([\xi - E(\xi)]^2) = E(h(\xi)) = \sum_{x \in R_\xi} [x - E(\xi)]^2 P(\xi = x)$$

- Második momentum: $h(x) = x^2$,

$$E(\xi^2) = E(h(\xi)) = \sum_{x \in R_\xi} x^2 P(\xi = x).$$

A variancia meghatározása

Ha ξ olyan valószínűségi változó, melynek véges a várható értéke, akkor

$$\text{Var}(\xi) = E(\xi^2) - (E(\xi))^2.$$

Bizonyítás. Később még visszatérünk rá.

Feladat. A biztosítékos feladatban határozzuk meg a varianciát a második momentum segítségével!

A második momentum:

$$E(\xi^2) = \sum_{x \in R_\xi} x^2 P(\xi = x) = 1^2 \cdot 0,4 + 2^2 \cdot 0,3 + 3^2 \cdot 0,2 + 4^2 \cdot 0,1 = 5.$$

A variancia: $\text{Var}(\xi) = E(\xi^2) - (E(\xi))^2 = 5 - 2^2 = 1.$

Ez azonos a korábban megkapott értékkel.

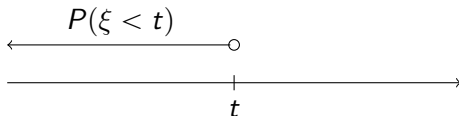
Eloszlásfüggvény

Egy tetszőleges ξ valószínűségi változó eloszlásfüggvénye:

$$F_{\xi} : \mathbb{R} \rightarrow [0, 1], \quad F_{\xi}(t) = P(\xi < t), \quad t \in \mathbb{R}.$$

Megjegyzések:

- Az eloszlásfüggvény jelentése: a ξ változó mekkora valószínűséggel esik a számegyenesen a t értéktől balra.



- Egyes könyvekben a definíció: $F_{\xi}(t) = P(\xi \leq t)$.

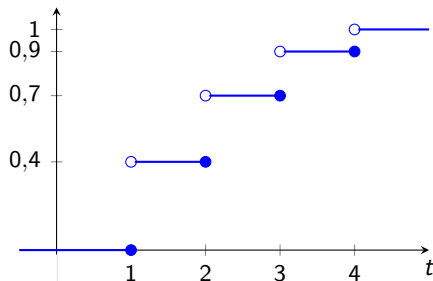
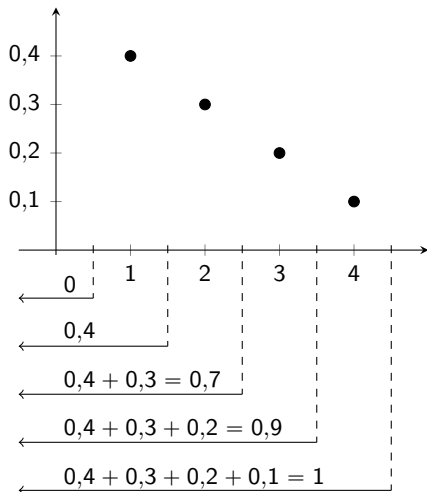
Diszkrét valószínűségi változók eloszlásfüggvénye

Egy diszkrét valószínűségi változó eloszlásfüggvénye egy olyan lépcsős függvény, mely pontosan a változó lehetséges értékeiben ugrik, és egy $t \in R_{\xi}$ helyen az ugrás nagysága $P(\xi = t)$.

Feladat. Korábbi diszkrét valószínűségi változó: $R_\xi = \{1, 2, 3, 4\}$,

$$P(\xi = 1) = 0,4, \quad P(\xi = 2) = 0,3, \quad P(\xi = 3) = 0,2, \quad P(\xi = 4) = 0,1.$$

Határozzuk meg a változó eloszlásfüggvényét!



$$F_\xi(t) = \begin{cases} 0, & t \leq 1, \\ 0,4, & 1 < t \leq 2, \\ 0,7, & 2 < t \leq 3, \\ 0,9, & 3 < t \leq 4, \\ 1, & 4 < t. \end{cases}$$

Folytonos valószínűségi változók

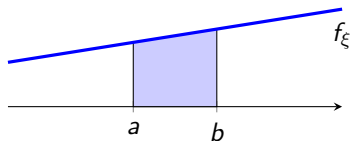
Folytonos valószínűségi változók, sűrűségfüggvény

Egy ξ valószínűségi változó **folytonos**, ha létezik olyan $f_\xi : \mathbb{R} \rightarrow \mathbb{R}$ függvény, hogy tetszőleges $a \leq b$ valós számok esetén

$$P(a \leq \xi \leq b) = \int_a^b f_\xi(x) dx.$$

Ekkor az f_ξ függvényt a ξ változó **sűrűségfüggvényének** nevezzük.

A definíció értelmezése: valószínűség = függvénygörbe alatti terület.



Megmutatható, hogy az egyenlőség $a = -\infty$ és $b = +\infty$ esetén is igaz.

Néhány fontosabb függvény deriválási és integrálási szabálya, nekünk a félév folyamán csak ezekre lesz szükségünk (a rögzített valós szám):

Hatvány- és konstansfüggvények deriváltja:

$$(x^a)' = ax^{a-1}, \quad (a)' = 0.$$

Hatványfüggvények primitív függvénye:

$$\int x^a dx = \frac{x^{a+1}}{a+1}, \quad \text{ha } a \neq -1, \quad \int x^{-1} dx = \ln |x|.$$

Exponenciális függvények deriváltja és primitív függvénye:

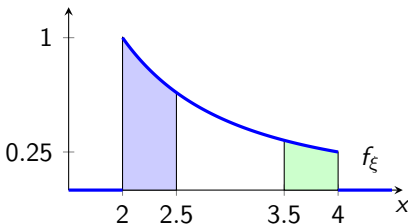
$$(a^x)' = a^x \ln a, \quad \int a^x dx = \frac{a^x}{\ln a}, \quad a > 0.$$

Speciálisan:

$$(e^x)' = e^x, \quad \int e^x dx = e^x.$$

Feladat: Egy ξ folytonos valószínűségi változó sűrűségfüggvénye az f_ξ függvény. Határozzuk meg az alábbi valószínűségeket!

$$f_\xi(x) = \begin{cases} 4/x^2, & \text{ha } 2 \leq x \leq 4, \\ 0, & \text{különben.} \end{cases}$$



$$\begin{aligned} P(\xi \leq 2,5) &= P(-\infty \leq \xi \leq 2,5) = \int_{-\infty}^{2,5} f_\xi(x) dx = \int_{-\infty}^2 0 dx + 4 \int_2^{2,5} x^{-2} dx \\ &= 0 + 4 \left[\frac{x^{-1}}{-1} \right]_2^{2,5} = 4 \left[-\frac{1}{x} \right]_2^{2,5} = 4 \left[\left(-\frac{1}{2,5} \right) - \left(-\frac{1}{2} \right) \right] = 4 \cdot 0,1 = 0,4 \\ P(\xi \geq 3,5) &= P(3,5 \leq \xi \leq +\infty) = \int_{3,5}^{+\infty} f_\xi(x) dx = 4 \int_{3,5}^4 x^{-2} dx + \int_4^{+\infty} 0 dx \\ &= 4 \left[\frac{x^{-1}}{-1} \right]_{3,5}^4 + 0 = 4 \left[-\frac{1}{x} \right]_{3,5}^4 = 4 \left[\left(-\frac{1}{4} \right) - \left(-\frac{1}{3,5} \right) \right] \approx 0,14. \end{aligned}$$

Folytonos valószínűségi változók értékkészlete

Egy ξ folytonos valószínűségi változó értékkészlete azon x valós számoknak a halmaza, ahol az f_ξ sűrűségfüggvény értéke pozitív:

$$R_\xi = \{x \in \mathbb{R} : f_\xi(x) > 0\}$$

Bizonyítás. Ha $f_\xi(x) = 0$ minden $x \in [a, b]$ esetén, akkor

$$P(a \leq \xi \leq b) = \int_a^b f_\xi(x) dx = \int_a^b 0 dx = 0.$$

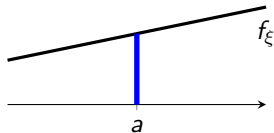
Ha $f_\xi(x) > 0$ az $[a, b]$ intervallumon, akkor

$$P(a \leq \xi \leq b) = \int_a^b f_\xi(x) dx > 0.$$

Kérdés: Mit mutat meg a sűrűségfüggvény egy adott pontban? Teljesül a következő egyenlőség: $f(a) = P(\xi = a)$?

Ez az egyenlőség nem mindig teljesül! A definíció szerint tetszőleges a valós szám esetén:

$$P(\xi = a) = P(a \leq \xi \leq a) = \int_a^a f_{\xi}(x) dx = 0$$



Valószínűségek folytonos változók esetén

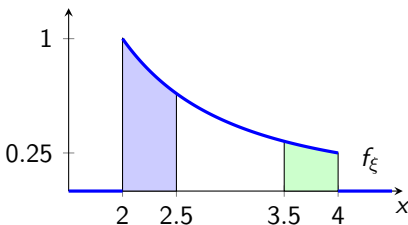
Legyen ξ folytonos valószínűségi változó f_{ξ} a sűrűségfüggvénnyel. Ekkor minden a valós szám esetén $P(\xi = a) = 0$. Ebből következik, hogy tetszőleges $a \leq b$ valós számok mellett

$$P(a < \xi < b) = P(a < \xi \leq b) = P(a \leq \xi < b) = \int_a^b f_{\xi}(x) dx.$$

Fontos: ez az állítás diszkrét változók esetén nem igaz!

Feladat. Egy ξ folytonos valószínűségi változó sűrűségfüggvénye az alábbi f_ξ függvény:

$$f_\xi(x) = \begin{cases} 4/x^2, & \text{ha } 2 \leq x \leq 4, \\ 0, & \text{különben.} \end{cases}$$



Adjuk meg a ξ valószínűségi változó értékkészletét!

Megoldás: $R_\xi = [2, 4]$.

Határozzuk meg az alábbi valószínűségeket!

- $P(\xi = 2,5) = 0 = P(\xi = 3,5)$
- $P(\xi < 2,5) = P(\xi \leq 2,5) - P(\xi = 2,5) = P(\xi \leq 2,5) - 0 = 0,4$
- $P(\xi > 3,5) = P(\xi \geq 3,5) - P(\xi = 3,5) = P(\xi \geq 3,5) - 0 \approx 0,14$

A sűrűségfüggvények karakterizációja

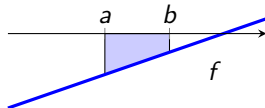
Egy $f : \mathbb{R} \rightarrow \mathbb{R}$ függvény akkor és csak akkor egy folytonos valószínűségi változó sűrűségfüggvénye, ha teljesíti az alábbi két feltételt:

- $f(x) \geq 0$ minden x valós szám esetén;
- $\int_{-\infty}^{+\infty} f(x) dx = 1$.

Bizonyítás. Tegyük fel, hogy az f egy ξ folytonos valószínűségi változó sűrűségfüggvénye!

- Ha valamilyen $[a, b]$ intervallumon $f(x) < 0$ teljesülne, akkor

$$P(a \leq \xi \leq b) = \int_a^b f(x) dx < 0.$$



Ez viszont nem lehetséges.

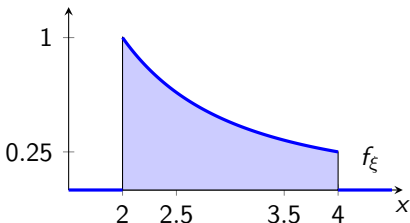
- A sűrűségfüggvény definícióját alkalmazva

$$\int_{-\infty}^{+\infty} f(x) dx = P(-\infty \leq \xi \leq +\infty) = 1.$$

A bizonyítás másik iránya sokkal nehezebb...

Feladat. Ellenőrizzük le, hogy az alábbi függvény valóban sűrűségfüggvény!

$$f_{\xi}(x) = \begin{cases} 4/x^2, & \text{ha } 2 \leq x \leq 4, \\ 0, & \text{különben.} \end{cases}$$



- $f_{\xi}(x) \geq 0$ teljesül minden x valós számra.
- Teljes görbe alatti terület:

$$\begin{aligned} \int_{-\infty}^{+\infty} f_{\xi}(x) dx &= \int_{-\infty}^2 0 dx + 4 \int_2^4 x^{-2} dx + \int_4^{+\infty} 0 dx \\ &= 0 + 4 \left[\frac{x^{-1}}{-1} \right]_2^4 + 0 = 4 \left[-\frac{1}{x} \right]_2^4 = 4 \left[\left(-\frac{1}{4} \right) - \left(-\frac{1}{2} \right) \right] = 4 \cdot \frac{1}{4} = 1 \end{aligned}$$

Tehát f_{ξ} valóban sűrűségfüggvény.

Hogyan definiáljuk egy ξ folytonos valószínűségi változó várható értékét?
 Ötlet: közelítsünk egy diszkrét valószínűségi változóval!

- Osszuk be a számegyenest $\delta \approx 0$ hosszúságú intervallumokra:

$$\dots \quad [-2\delta, -\delta), \quad [-\delta, 0), \quad [0, \delta), \quad [\delta, 2\delta), \quad [2\delta, 3\delta), \quad \dots$$

- Ha $\xi \in [k\delta, (k+1)\delta)$, akkor legyen $\eta = k\delta$.
- Ez egy pontos közelítés, hiszen $|\xi - \eta| \leq \delta$.
- Ekkor η diszkrét valószínűségi változó, és a valószínűségeloszlása:

$$P(\eta = k\delta) = P(k\delta \leq \xi < (k+1)\delta) = \int_{k\delta}^{(k+1)\delta} f_{\xi}(x) dx.$$

- Ebből következik, hogy

$$\begin{aligned} E(\eta) &= \sum_{k \in \mathbb{Z}} (k\delta) P(\eta = k\delta) = \sum_{k \in \mathbb{Z}} \int_{k\delta}^{(k+1)\delta} (k\delta) f_{\xi}(x) dx \\ &\approx \sum_{k \in \mathbb{Z}} \int_{k\delta}^{(k+1)\delta} x f_{\xi}(x) dx = \int_{-\infty}^{+\infty} x f_{\xi}(x) dx. \end{aligned}$$

Folytonos valószínűségi változók várható értéke és szórása

Legyen ξ folytonos valószínűségi változó f_ξ sűrűségfüggvénnyel!

- **Várható érték:**

$$E(\xi) = \int_{-\infty}^{\infty} x f_\xi(x) dx$$

Jelentése: átlagos érték.

- **Varianciája és szórás:** ha a várható érték véges, akkor

$$\text{Var}(\xi) = E\left([\xi - E(\xi)]^2\right), \quad D(\xi) = \sqrt{\text{Var}(\xi)}.$$

A szórás jelentése: várható értéktől való átlagos eltérés.

A nagy számok Kolmogorov-féle törvénye (1933)

Legyen $\xi_1, \xi_2, \dots$ a ξ valószínűségi változó egymástól független megfigyelései! Ekkor

$$\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow E(\xi), \quad n \rightarrow \infty.$$

Folytonos valószínűségi változók transzformáltjai

Legyen ξ folytonos valószínűségi változó f_ξ sűrűségfüggvénnyel, és tekintsünk egy $h: \mathbb{R} \rightarrow \mathbb{R}$ függvényt! Ekkor a $h(\xi)$ transzformált változó várható értéke:

$$E(h(\xi)) = \int_{-\infty}^{+\infty} h(x) f_\xi(x) dx.$$

Speciálisan a változó **második momentuma**:

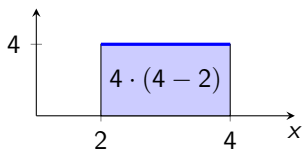
$$E(\xi^2) = \int_{-\infty}^{+\infty} x^2 f_\xi(x) dx.$$

Ennek segítségével a variancia az alábbi formulával is meghatározható:

$$\text{Var}(\xi) = E(\xi^2) - (E(\xi))^2.$$

Feladat. Mennyi a ξ változó várható értéke és szórása a jelen feladatban?

$$f_{\xi}(x) = \begin{cases} 4/x^2, & \text{ha } 2 \leq x \leq 4, \\ 0, & \text{különben.} \end{cases}$$



- Várható érték:

$$\begin{aligned} E(\xi) &= \int_{-\infty}^{\infty} x f_{\xi}(x) dx = \int_{-\infty}^2 x \cdot 0 dx + \int_2^4 x \cdot \frac{4}{x^2} dx + \int_4^{\infty} x \cdot 0 dx \\ &= 0 + 4 \int_2^4 x^{-1} dx + 0 = 4 \left[\ln |x| \right]_2^4 = 4 [\ln 4 - \ln 2] \approx 2,77, \end{aligned}$$

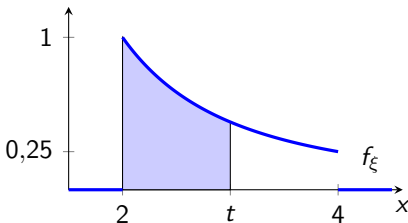
- Második momentum:

$$E(\xi^2) = \int_{-\infty}^{\infty} x^2 f_{\xi}(x) dx = \int_2^4 x^2 \frac{4}{x^2} dx = \int_2^4 4 dx = 4 \cdot (4 - 2) = 8,$$

- Variancia: $\text{Var}(\xi) = E(\xi^2) - (E(\xi))^2 = 8 - (2,77)^2 \approx 0,33,$
- Szórás: $D(\xi) = \sqrt{\text{Var}(\xi)} = \sqrt{0,33} = 0,57.$

Feladat: Határozzuk meg a ξ változó eloszlásfüggvényét!

$$f_{\xi}(x) = \begin{cases} 4/x^2, & \text{ha } 2 \leq x \leq 4, \\ 0, & \text{különben.} \end{cases}$$



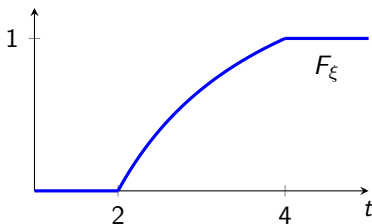
Eloszlásfüggvény definíciója: $F_{\xi}(t) = P(\xi < t)$, $t \in \mathbb{R}$.

- Ha $t < 2$: $F_{\xi}(t) = P(\xi < t) = 0$.
- Ha $t > 4$: $F_{\xi}(t) = P(\xi < t) = 1$.
- Ha $2 \leq t \leq 4$:

$$\begin{aligned} F_{\xi}(t) &= P(\xi < t) = P(-\infty \leq \xi \leq t) = \int_{-\infty}^t f_{\xi}(x) dx = 4 \int_2^t x^{-2} dx \\ &= 4 \left[\frac{x^{-1}}{-1} \right]_2^t = 4 \left[-\frac{1}{x} \right]_2^t = 4 \left[\left(-\frac{1}{t} \right) - \left(-\frac{1}{2} \right) \right] = 2 - \frac{4}{t} \end{aligned}$$

Tehát az eloszlásfüggvény:

$$F_{\xi}(t) = \begin{cases} 0, & t < 2, \\ 2 - 4/t, & 2 \leq t \leq 4, \\ 1, & t > 4. \end{cases}$$



Folytonos valószínűségi változók eloszlásfüggvénye

Egy folytonos valószínűségi változó eloszlásfüggvénye mindehol folytonos. Az eloszlásfüggvény és a sűrűségfüggvény kapcsolata:

$$F_{\xi}(t) = \int_{-\infty}^t f_{\xi}(x) dx \quad \text{és} \quad f_{\xi}(t) = F'_{\xi}(t).$$

A feladatban $2 \leq t \leq 4$ esetén:

$$F'_{\xi}(t) = (2 - 4 \cdot t^{-1})' = (2)' - 4 \cdot (t^{-1})' = 0 - 4 \cdot (-1)t^{-2} = 4/t^2.$$

Az eloszlásfüggvények további alkalmazásai

Valószínűségek számolása az eloszlásfüggvény segítségével

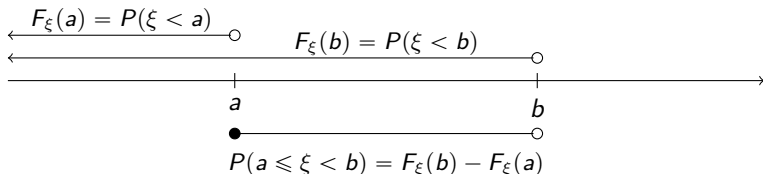
Legyen ξ valószínűségi változó, és legyen F_ξ az eloszlásfüggvénye!
Ekkor tetszőleges $a \leq b$ valós számok esetén

$$P(a \leq \xi < b) = F_\xi(b) - F_\xi(a).$$

Speciálisan ha a ξ változó folytonos eloszlású, akkor

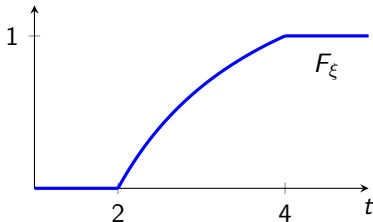
$$P(a \leq \xi \leq b) = P(a < \xi < b) = P(a < \xi \leq b) = F_\xi(b) - F_\xi(a).$$

Bizonyítás.



Feladat. Egy ξ folytonos valószínűségi változó eloszlásfüggvénye:

$$F_{\xi}(t) = \begin{cases} 0, & t < 2, \\ 2 - 4/t, & 2 \leq t \leq 4, \\ 1, & t > 4. \end{cases}$$



Határozzuk meg a következő valószínűségeket:

- $P(\xi < 2,5) = F_{\xi}(2,5) = 2 - 4/2,5 = 0,4$
- $P(\xi \geq 3,5) = 1 - P(\xi < 3,5) = 1 - F_{\xi}(3,5) = 1 - (2 - 4/3,5) \approx 0,14$
- $P(1 \leq \xi < 3) = F_{\xi}(3) - F_{\xi}(1) = (2 - 4/3) - 0 \approx 0,67$
- $P(1 < \xi \leq 3) = P(1 \leq \xi < 3) \approx 0,67$

Kvantilisek

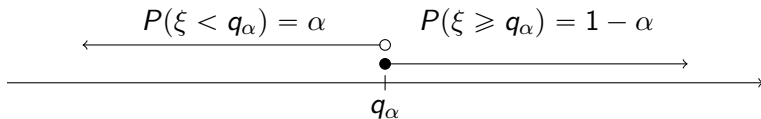
Legyen ξ tetszőleges valószínűségi változó, és legyen $\alpha \in (0, 1)$. A q_α valós számot a ξ változó **α -kvantilisének** nevezzük, ha $P(\xi < q_\alpha) = \alpha$.

Nevezetes kvantilisek:

- **Medián:** $q_{50\%}$
- **Alsó és felső kvartilis:** $q_{25\%}$ és $q_{75\%}$
- **Decilisek:** $q_{10\%}, q_{20\%}, \dots, q_{90\%}$

Megjegyzések:

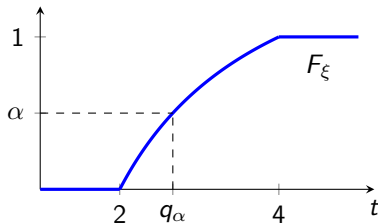
- Az α -kvantilis jelentése:



- Az α -kvantilis nem mindig létezik, és ha létezik, akkor nem feltétlenül egyértelmű.
- Az α -kvantilis meghatározása: $F_\xi(q_\alpha) = \alpha$.

Feladat: Adjuk meg a mediánt valamint az alsó és a felső kvartilist a jelen feladatban!

$$F_{\xi}(t) = \begin{cases} 0, & t < 2, \\ 2 - 4/t, & 2 \leq t \leq 4, \\ 1, & t > 4. \end{cases}$$

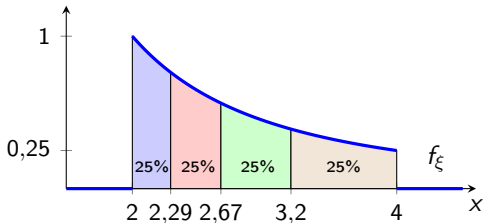


Most tetszőleges $\alpha \in (0, 1)$ esetén az α -kvantilis létezik és egyértelmű:

$$\alpha = P(\xi < q_{\alpha}) = F_{\xi}(q_{\alpha}) = 2 - 4/q_{\alpha}.$$

Tehát: $q_{\alpha} = 4/(2 - \alpha)$.

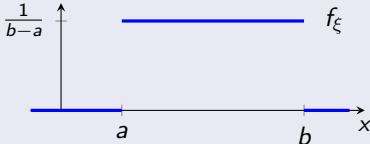
α	25%	50%	75%
q_{α}	2,29	2,67	3,2



Az egyenletes eloszlás és az exponenciális eloszlás

Az egyenletes eloszlás

Azt mondjuk, hogy egy ξ folytonos valószínűségi változó **egyenletes eloszlást** követ az $[a, b]$ intervallumon, ha a sűrűségfüggvénye:

$$f_{\xi}(x) = \begin{cases} 1/(b-a), & a \leq x \leq b, \\ 0, & \text{különben.} \end{cases}$$


Az egyenletes eloszlás várható értéke illetve szórása:

$$E(\xi) = \frac{a+b}{2}, \quad D(\xi) = \frac{b-a}{\sqrt{12}}.$$

Megjegyzések:

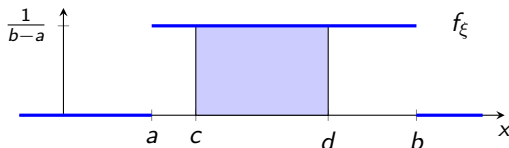
- Az f_{ξ} függvény valóban sűrűségfüggvény.
- Az egyenletes eloszlás értékkeszlete: $R_{\xi} = [a, b]$.

Az egyenletes eloszlás előállítása

Ha egy ξ változó egyenletes eloszlású az $[a, b]$ intervallumon, akkor teljesíti az egyenletességi hipotézist.

Bizonyítás. Legyenek $c, d \in [a, b]$, $c \leq d$, tetszőleges valós számok!

$$P(c \leq \xi \leq d) = \int_c^d \frac{1}{b-a} dx = (d-c) \frac{1}{b-a} = \frac{\text{kedvező hosszúság}}{\text{összes hosszúság}}$$



Megjegyzés: A nagy számok Kolmogorov-féle törvénye szerint a ξ egyenletes eloszlású változó értékét többször megfigyelve:

$$\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow E(\xi) = \frac{a+b}{2}, \quad n \rightarrow \infty.$$

Feladat. Adott egy befektetés, melyből a jövőben egy véletlen nagyságú jövedelem fog majd származni. Ez a jövedelem egyenletes eloszlású 0 és 1000 euro között. Adjuk meg a jövedelem várható értékét és szórását!

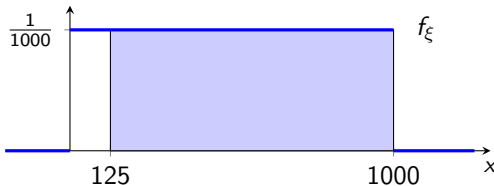
ξ = a jövedelem nagysága, egyenletes eloszlású, $a = 0$, $b = 1000$

Várható érték és szórás:

$$E(\xi) = \frac{0 + 1000}{2} = 500, \quad D(\xi) = \frac{1000 - 0}{\sqrt{12}} = 289.$$

Feladat. Mennyi annak az esélye, hogy a befektetésből legalább 125 euro jövedelem származik?

$$P(\xi \geq 125) = P(125 \leq \xi \leq 1000) = \frac{1000 - 125}{1000 - 0} = 0,875 = 87,5\%$$

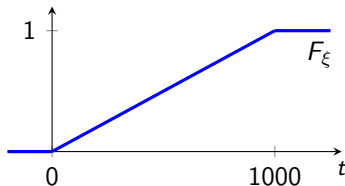
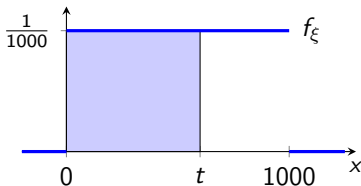


Feladat. A ξ valószínűségi változó egyenletes eloszlású a $[0, 1000]$ intervallumon. Írjuk fel a változó eloszlásfüggvényét!

Eloszlásfüggvény: $F_{\xi}(t) = P(\xi < t)$, $t \in \mathbb{R}$.

- Ha $t < 0$: $F_{\xi}(t) = 0$.
- Ha $t > 1000$: $F_{\xi}(t) = 1$.
- Ha $0 \leq t \leq 1000$:

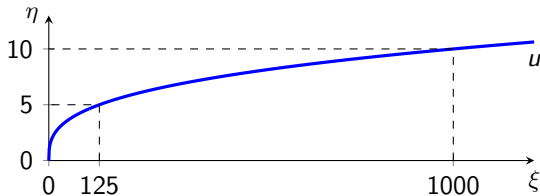
$$F_{\xi}(t) = P(\xi < t) = P(0 \leq \xi \leq t) = \frac{t - 0}{1000 - 0} = \frac{t}{1000}$$



Feladat. Tegyük fel, hogy a jövedelmünket az $u(x) = \sqrt[3]{x}$ hasznossági függvénnyel értékeljük. Mennyi annak az esélye, hogy az elért hasznosság legalább 5? Az elért hasznosság egyenletes eloszlású?

Legyen: η = elért hasznosság $= u(\xi) = \sqrt[3]{\xi}$, $R_\eta = [0, 10]$,

$$P(\eta \geq 5) = P(\sqrt[3]{\xi} \geq 5) = P(\xi \geq 5^3) = P(\xi \geq 125) = 87,5\%$$



Ha az elért hasznosság egyenletes eloszlású lenne a $[0, 10]$ intervallumon, akkor teljesülne a következő:

$$P(\eta \geq 5) = P(5 \leq \eta \leq 10) = \frac{\text{kedvező}}{\text{összes}} = \frac{10 - 5}{10 - 0} = 50\%.$$

Ez most nem teljesül, tehát η nem egyenletes eloszlású.

Feladat. Adjuk meg az elért hasznosság várható értékét és szórását!

ξ = elért jövedelem, egyenletes eloszlású, $a = 0$, $b = 1000$

η = elért hasznosság = $u(\xi) = \sqrt[3]{\xi}$

$$\begin{aligned} E(\eta) &= E(u(\xi)) = \int_{-\infty}^{+\infty} u(x) f_{\xi}(x) dx = \int_0^{1000} \sqrt[3]{x} \frac{1}{1000} dx \\ &= \frac{1}{1000} \int_0^{1000} x^{1/3} dx = \frac{1}{1000} \left[\frac{x^{4/3}}{4/3} \right]_0^{1000} = \frac{1}{1000} \left[\frac{1000^{4/3}}{4/3} - \frac{0^{4/3}}{4/3} \right] = 7,5, \end{aligned}$$

$$\begin{aligned} E(\eta^2) &= E(u^2(\xi)) = \int_{-\infty}^{+\infty} u^2(x) f_{\xi}(x) dx = \int_0^{1000} (\sqrt[3]{x})^2 \frac{1}{1000} dx \\ &= \frac{1}{1000} \int_0^{1000} x^{2/3} dx = \frac{1}{1000} \left[\frac{x^{5/3}}{5/3} \right]_0^{1000} = \frac{1}{1000} \left[\frac{1000^{5/3}}{5/3} - \frac{0^{5/3}}{5/3} \right] = 60, \end{aligned}$$

$$\text{Var}(\eta) = E(\eta^2) - (E(\eta))^2 = 60 - 7,5^2 = 3,75,$$

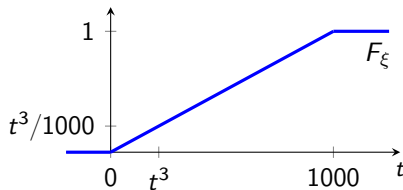
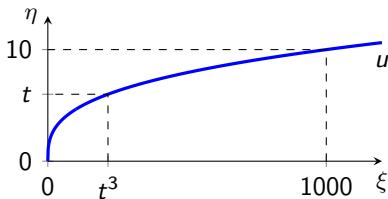
$$D(\eta) = \sqrt{\text{Var}(\eta)} = 1,94.$$

Feladat. Írjuk fel és ábrázoljuk az elért hasznosság eloszlásfüggvényét!

Eloszlásfüggvény: $F_{\eta}(t) = P(\eta < t), \quad t \in \mathbb{R}.$

- Ha $t < 0$: $F_{\eta}(t) = 0.$
- Ha $t > 10$: $F_{\eta}(t) = 1.$
- Ha $0 \leq t \leq 10$:

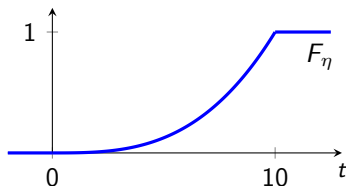
$$F_{\eta}(t) = P(\eta < t) = P(\sqrt[3]{\xi} < t) = P(\xi < t^3) = F_{\xi}(t^3) = t^3/1000$$



Feladat. Határozzuk meg az elért hasznosság sűrűségfüggvényét!

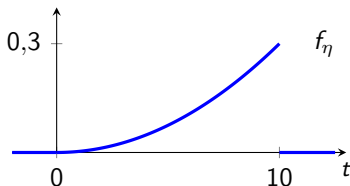
Eloszlásfüggvény: $F_{\eta}(t) = P(\eta < t)$.

$$F_{\eta}(t) = \begin{cases} 0, & t < 0, \\ t^3/1000, & 0 \leq t \leq 10, \\ 1, & t > 10. \end{cases}$$



Sűrűségfüggvény: $f_{\eta}(t) = F'_{\eta}(t)$.

$$f_{\eta}(t) = \begin{cases} 0, & t < 0, \\ 3t^2/1000, & 0 \leq t \leq 10, \\ 0, & t > 10. \end{cases}$$

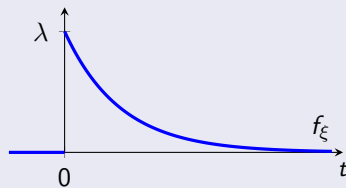


Házi feladat. Számoljuk ki az η valószínűségi változó várható értékét és szórását a sűrűségfüggvény alkalmazásával!

Az exponenciális eloszlás

Egy ξ folytonos valószínűségi változó **exponenciális eloszlást** követ $\lambda > 0$ paraméterrel, ha a sűrűségfüggvénye

$$f_{\xi}(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$

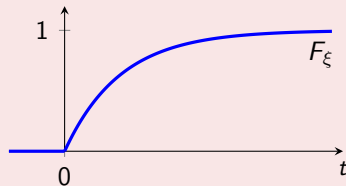


Az exponenciális eloszlás tulajdonságai

Várható érték és szórás: $E(\xi) = D(\xi) = 1/\lambda$.

Eloszlásfüggvény:

$$F_{\xi}(t) = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0, \\ 0, & t < 0. \end{cases}$$



Feladat. Egy adott típusú izzó élettartama exponenciális eloszlást követ ismeretlen λ paraméterrel. A tapasztalatok szerint az izzók negyede megy tönkre ezer óra használat során. Mekkora az izzók átlagos élettartama?

Legyen ξ egy izzó ezer órában számolt élettartama! Ekkor

$$0,25 = P(\xi < 1) = F_{\xi}(1) = 1 - e^{-\lambda \cdot 1}.$$

Átrendezés után:

$$e^{-\lambda \cdot 1} = 0,75, \quad \text{vagyis} \quad \lambda = \frac{\ln 0,75}{-1} = 0,29.$$

Átlagos élettartam: $E(\xi) = 1/\lambda = 3,48$ ezer óra.

Feladat. Az izzók mekkora hányada éri el a 10 ezer órás élettartamot?

$$\begin{aligned} P(\xi \geq 10) &= 1 - P(\xi < 10) = 1 - F_{\xi}(10) \\ &= 1 - [1 - e^{-\lambda \cdot 10}] = e^{-0,29 \cdot 10} = 0,055 = 5,5\% \end{aligned}$$

Örökifjú tulajdonság

Legyen ξ exponenciális eloszlású valószínűségi változó tetszőleges $\lambda > 0$ paraméterrel! Ekkor bármely $x, y > 0$ valós számok esetén

$$P(\xi \geq x + y \mid \xi \geq x) = P(\xi \geq y).$$

Bizonyítás.

$$\begin{aligned} P(\xi \geq x + y \mid \xi \geq x) &= \frac{P(\xi \geq x + y \text{ és } \xi \geq x)}{P(\xi \geq x)} = \frac{P(\xi \geq x + y)}{P(\xi \geq x)} \\ &= \frac{1 - F_{\xi}(x + y)}{1 - F_{\xi}(x)} = \frac{1 - [1 - e^{-\lambda(x+y)}]}{1 - [1 - e^{-\lambda x}]} = \frac{e^{-\lambda(x+y)}}{e^{-\lambda x}} = e^{-\lambda y} = P(\xi \geq y) \end{aligned}$$

Feladat. Tekintsük csak azon izzókat, melyek élettartama eléri a 10 ezer órát! Ezek mekkora hányada fog még legalább ezer órán át működni?

Az örökifjú tulajdonság alkalmazásával:

$$P(\xi \geq 10 + 1 \mid \xi \geq 10) = P(\xi \geq 1) = 1 - P(\xi < 1) = 75\%.$$

Az exponenciális eloszlást főleg élettartamok modellezésére szoktuk alkalmazni. Az örökifjú tulajdonság jelentése: az eddigi élettartam nincsen hatással a hátralevő élettartamra, nincs öregedés vagy fiatalodás.

Milyen élettartamok nem modellezhetők exponenciális eloszlással:

- élőlények élettartama;
- mechanikus berendezések élettartama.

Milyen élettartamok modellezhetők (aránylag pontosan) exponenciális eloszlással:

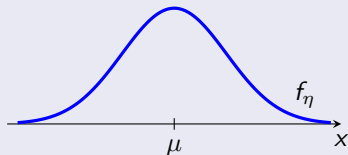
- egyszerűbb elektronikus berendezések élettartama;
- biztosítóknál az egymást követő káresemények közötti idő;
- rádióaktív atomok élettartama.

A normális eloszlás

Normális (normál, Gauss-) eloszlás

Az η valószínűségi változó **normális eloszlást** követ $\mu \in \mathbb{R}$ (mű) és $\sigma > 0$ (szigma) paraméterekkel, ha a sűrűségfüggvénye:

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



A sűrűségfüggvény neve: **Gauss-görbe, haranggörbe.**

A normális eloszlás néhány alkalmazása:

- Mérési hibák modellezése: mért érték = igazi érték + mérési hiba, ahol a mérési hiba normális eloszlást követ.
- Élettudományok: számos mennyiség (testmagasság, vérnyomás, IQ) normális vagy a normálisból származtatott eloszlást követ.
- Black–Scholes-modell: a részvényárak logaritmusa normális eloszlású.

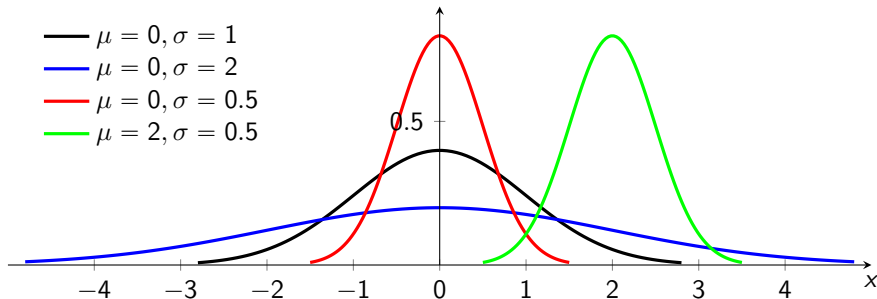
A normális eloszlás értékészlete, várható értéke és szórása

Legyen η normális eloszlású μ és σ paraméterekkel! Ekkor:

- $f_{\eta}(x) > 0$ minden x valós számra, ezért $R_{\eta} = \mathbb{R}$.
- $E(\eta) = \mu$ és $D(\eta) = \sigma$.

Hogyan hat a két paraméter a sűrűségfüggvényre:

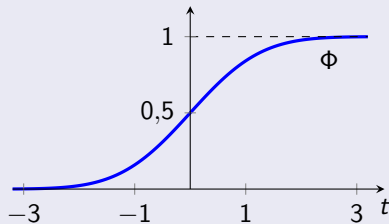
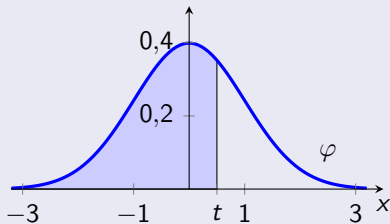
- σ : a sűrűségfüggvény alakját határozza meg,
- μ : eltolás, a sűrűségfüggvény szimmetriatengelye.



Standard normális eloszlás (sztenderd normál eloszlás)

A $\mu = 0$ és $\sigma = 1$ paraméteres normális eloszlást **standard normális eloszlásnak** nevezzük. Jelölésben: $\eta_{0,1}$. Sűrűség- és eloszlásfüggvénye:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad \Phi(t) = P(\eta_{0,1} < t) = \int_{-\infty}^t \varphi(x) dx.$$

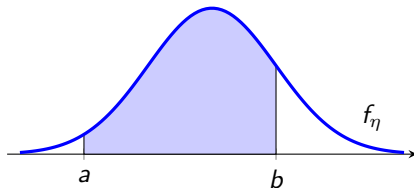


Standardizálás

Legyen η normális eloszlású μ várható értékkel és σ szórással! Ekkor az $(\eta - \mu)/\sigma$ valószínűségi változót η **standardizáltjának** nevezzük. Megmutatható, hogy ez az új változó standard normális eloszlás követ.

Ha η normális eloszlású, akkor tetszőleges $a \leq b$ valós számokra:

$$P(a \leq \eta \leq b) = \int_a^b f_{\eta}(x) dx = \int_a^b \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$



Probléma: ezt az integrált nem tudjuk papíron kiszámolni.

Megoldás: standardizálással áttérünk a standard normális eloszlásra. A Φ függvény értékeit táblázatból tudjuk kikeresni.

$$P(a \leq \eta \leq b) = P\left(\frac{a - \mu}{\sigma} \leq \underbrace{\frac{\eta - \mu}{\sigma}}_{\eta_{0,1}} \leq \frac{b - \mu}{\sigma}\right) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right).$$

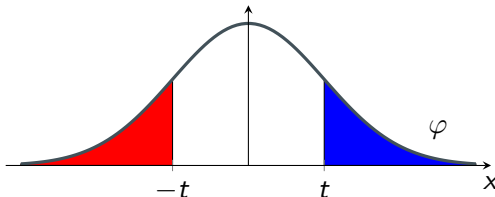
A standard normális eloszlás eloszlásfüggvénye

A Φ eloszlásfüggvény rendelkezik az alábbi tulajdonságokkal.

- 1 Szigorúan monoton növekvő és mindenhol folytonos.
- 2 $\Phi(0) = 0,5$.
- 3 Tetszőleges t valós szám esetén $\Phi(-t) = 1 - \Phi(t)$.

Bizonyítás. (1), (2): lásd a grafikont. (3):

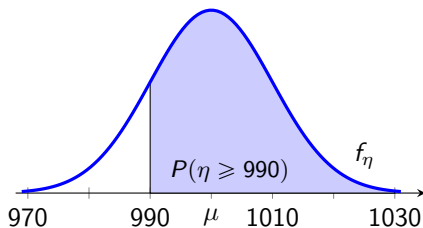
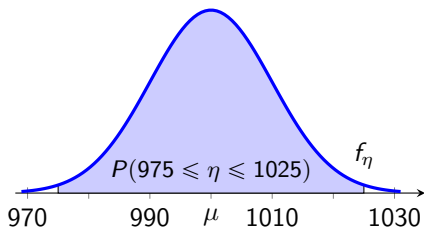
$$\Phi(-t) = \int_{-\infty}^{-t} \varphi(x) dx = \int_t^{+\infty} \varphi(x) dx = P(\eta_{0,1} \geq t) = 1 - \Phi(t).$$



Feladat. Egy tejgyárban az 1 literes dobozos tej csomagolását automata töltőberendezés végzi. A dobozokba töltött mennyiség egy normális eloszlású valószínűségi változó, melynek várható értéke a névleges tartalom és szórása $\sigma = 10$ ml. Véletlenszerűen kiválasztunk egy dobozt.

- Mennyi annak a valószínűsége, hogy a doboz legfeljebb 2,5%-kal tér el a névleges tartalomtól?
- Mennyi annak az esélye, hogy a doboz legalább 990 ml tejet tartalmaz?

Legyen η a kiválasztott dobozban található mennyiség! Az η változó normális eloszlású $\mu = 1000$ ml várható értékkel és $\sigma = 10$ ml szórással. A következő valószínűségekre (=területekre) vagyunk kíváncsiak:



Feladat. Mennyi annak a valószínűsége, hogy a doboz legfeljebb 2,5%-kal tér el a névleges tartalomtól?

$$\begin{aligned} P(975 \leq \eta \leq 1025) &= P\left(\frac{975 - 1000}{10} \leq \frac{\eta - \mu}{\sigma} \leq \frac{1025 - 1000}{10}\right) \\ &= P(-2,5 \leq \eta_{0,1} \leq 2,5) = \Phi(2,5) - \Phi(-2,5) = 0,9938 - 0,0062 = 0,9876 \end{aligned}$$

Tehát a tejesdobozok 98,76%-a tartalmaz 975 ml és 1025 ml közötti tejet. Itt felhasználtuk azt, hogy

$$\Phi(-2,5) = 1 - \Phi(2,5) = 1 - 0,9938 = 0,0062$$

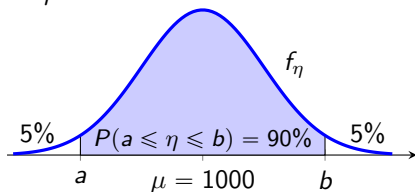
Feladat. Mennyi az esélye, hogy a doboz legalább 990 ml tejet tartalmaz?

$$\begin{aligned} P(\eta \geq 990) &= P\left(\frac{\eta - \mu}{\sigma} \geq \frac{990 - 1000}{10}\right) = P(\eta_{0,1} \geq -1) \\ &= 1 - P(\eta_{0,1} < -1) = 1 - \Phi(-1) = 1 - [1 - \Phi(1)] \\ &= 1 - [1 - 0,84] = 0,84 \end{aligned}$$

Feladat. Adjunk meg egy olyan $[a, b]$ intervallumot, amire teljesül, hogy a tejesdobozok 90%-a ebbe az intervallumba esik: $P(a \leq \eta \leq b) = 0,9$.

Legyenek $a = q_{5\%}$ és $b = q_{95\%}$ az η változó két kvantilise! Ekkor

$$\begin{aligned} P(a \leq \eta \leq b) &= P(q_{5\%} \leq \eta < q_{95\%}) \\ &= P(\eta < q_{95\%}) - P(\eta < q_{5\%}) \\ &= 0,95 - 0,05 = 0,9 \end{aligned}$$



Ezt a két kvantilist kell meghatároznunk. Ismét csak standardizálással:

$$0,95 = P(\eta < b) = P\left(\frac{\eta - \mu}{\sigma} < \frac{b - 1000}{10}\right) = \Phi\left(\frac{b - 1000}{10}\right)$$

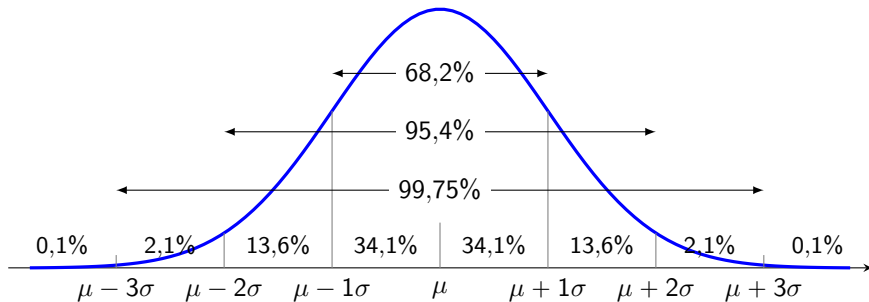
Mivel $\Phi(1,64) = 0,95$, ezért $(b - 1000)/10 = 1,64$, tehát $b = 1016,4$.

$$0,05 = P(\eta < a) = \Phi\left(\frac{a - 1000}{10}\right), \quad \text{tehát} \quad \Phi\left(-\frac{a - 1000}{10}\right) = 1 - 0,05 = 0,95$$

Ismét csak $\Phi(1,64) = 0,95$, ezért $a = 983,6$.

Tehát a kérdéses intervallum: $[983,6, 1016,4]$.

Az alábbi ábra azt mutatja meg, hogy egy η normális eloszlású változó mekkora eséllyel esik a várható érték két oldalára felmért intervallumokba:



Intervallumok a normális eloszlásra

Legyen η normális eloszlású változó μ és σ paraméterekkel! Ekkor:

- 1σ -szabály: $P(\mu - \sigma \leq \eta \leq \mu + \sigma) \approx 68\%$,
- 2σ -szabály: $P(\mu - 2\sigma \leq \eta \leq \mu + 2\sigma) \approx 95\%$,
- 3σ -szabály: $P(\mu - 3\sigma \leq \eta \leq \mu + 3\sigma) \approx 99,75\%$.

A binomiális eloszlás és a de Moivre–Laplace-tétel

Binomiális eloszlás

Azt mondjuk, hogy egy ξ valószínűségi változó **binomiális eloszlású** $n \in \{1, 2, \dots\}$ és $p \in (0, 1)$ paraméterekkel, ha

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

A binomiális eloszlás várható értéke és szórása:

$$E(\xi) = np, \quad D(\xi) = \sqrt{np(1 - p)}.$$

Bernoulli-kísérlet

Egy kísérletet **Bernoulli-kísérletnek** nevezünk, ha csak kétfajta kimenetele van, egy kedvező és egy kedvezőtlen.

A binomiális eloszlás előállítása

Legyen adva egy Bernoulli-kísérlet, ahol p a kedvező kimenetel esélye! Ha ezt a kísérletet egymástól függetlenül n alkalommal megismételjük, akkor az összesen kapott kedvező kimenetek száma binomiális eloszlást követ n és p paraméterrel.

Bizonyítás. Legyen ξ a kedvező kimenetek száma!

$R_\xi = \{0, 1, \dots, n\}$, tehát ξ diszkrét valószínűségi változó.

$$P(\text{az első } k \text{ kísérletre kedvező, a többire kedvezőtlen kimenetelt kapunk}) \\ = \underbrace{p \cdot \dots \cdot p}_{k \text{ darab}} \cdot \underbrace{(1-p) \cdot \dots \cdot (1-p)}_{(n-k) \text{ darab}} = p^k (1-p)^{n-k},$$

$$P(\xi = k) = P(\text{összesen } k \text{ kedvező kimenetelt kapunk}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Feladat. A foci világbajnokság döntőjében a hosszabbítások után még mindig döntetlen az eredmény, ezért büntetőrúgásokra kerül sor. Az egyik csapatnál a játékosok az 5 büntetőt egymástól függetlenül rendre 0,8 valószínűséggel értékesítik. Jelölje ξ a gólok számát ennél a csapatnál! Adjuk meg a ξ valószínűségi változó eloszlását, várható értékét és szórását! Mennyi az esélye, hogy pontosan 4 gól születik?

Értékkészlet: $R_\xi = \{0, 1, 2, 3, 4, 5\}$.

Bernoulli-kísérlet: egy büntetőrúgás, kedvező kimenetel: gól, $p = 0,8$.

$n = 5$ darab független ismételés, ξ a kedvező kimenetek száma.

Tehát ξ binomiális eloszlású $n = 5$ és $p = 0,8$ paraméterekkel.

$$P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{5}{k} 0,8^k 0,2^{5-k}, \quad k = 0, 1, \dots, 5.$$

Behelyettesítéssel: $P(\xi = 4) = \binom{5}{4} 0,8^4 0,2^{5-4} \approx 41\%$

Továbbá: $E(\xi) = np = 4$, $D(\xi) = \sqrt{np(1-p)} = \sqrt{0,8} \approx 0,89$.

Feladat. A statisztikai adatok szerint Magyarországon az emberek 30%-a dohányzik rendszeresen. Véletlenszerűen kiválasztunk 100 embert. Mennyi az esélye, hogy közöttük legalább 25, de legfeljebb 35 dohányos lesz?

Legyen ξ = a dohányosok száma a megkérdezettek között.

A ξ binomiális eloszlást követ, paraméterek: $n = 100$, $p = 0,3$.

Várható érték és szórás:

$$E(\xi) = np = 30, \quad D(\xi) = \sqrt{np(1-p)} = 4,6.$$

A kérdéses valószínűség:

$$P(25 \leq \xi \leq 35) = \sum_{k=25}^{35} P(\xi = k) = \sum_{k=25}^{35} \binom{100}{k} 0,3^k 0,7^{100-k}$$

A kapott formula nehezen számolható:

- egyrészt a formula sok tagból áll;
- másrészt a binomiális együtthatók meghatározása macerás.

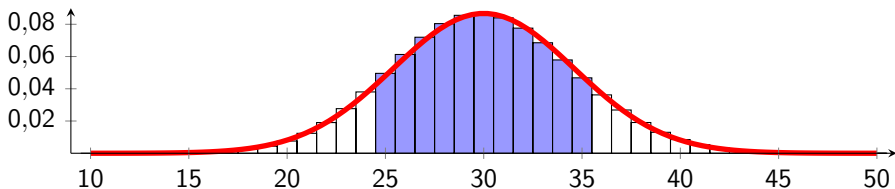
Az ábrán a ξ változó eloszlása látható oszlopdiagrammal ábrázolva.

- A k érték feletti oszlop magassága $P(\xi = k)$, szélessége 1. Ekkor a kék tartomány összterülete:

$$\sum_{k=25}^{35} 1 \cdot P(\xi = k) = \sum_{k=25}^{35} P(\xi = k) = P(25 \leq \xi \leq 35)$$

- Legyen η normális eloszlású valószínűségi változó $\mu = 30$ és $\sigma = 4,6$ paraméterrel. A piros színű görbe az η sűrűségfüggvénye.
- A 25 és a 35 értékek között a sűrűségfüggvény alatti terület jó közelítéssel megegyezik a kék színű tartomány területével, vagyis

$$P(25 \leq \xi \leq 35) \approx \int_{25}^{35} f_{\eta}(x) dx = P(25 \leq \eta \leq 35).$$



de Moivre–Laplace-tétel (1738)

Legyen

- ξ binomiális eloszlású változó n és p paraméterekkel,
- η normális eloszlású változó $\mu = np$ illetve $\sigma = \sqrt{np(1-p)}$ paraméterekkel!

Ekkor tetszőleges $a \leq b$ valós számok esetén

$$P(a \leq \xi \leq b) \approx P(a \leq \eta \leq b).$$

A közelítés annál pontosabb, minél nagyobb az n paraméter értéke.

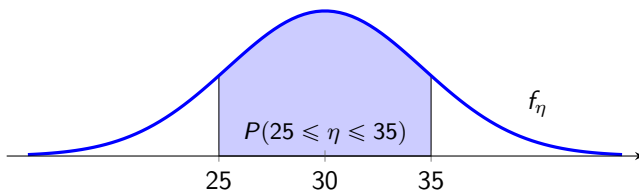
Megjegyzések:

- Azt a normális eloszlást keressük, amelyik a legjobban hasonlít a binomiálisra. A tételben: $E(\eta) = E(\xi)$, $D(\eta) = D(\xi)$.
- A tételt $n \geq 50$ esetén már szokták alkalmazni, de a tapasztalatok szerint $n \geq 100$ esetén válik igazán pontossá.

Feladat. Mennyi annak az esélye, hogy a megkérdezett 100 ember közül legalább 25 és legfeljebb 35 dohányzik?

Legyen η normális eloszlású $\mu = 30$ és $\sigma = 4,6$ paraméterekkel!
Ekkor a de Moivre–Laplace-tétel alkalmazásával:

$$\begin{aligned} P(25 \leq \xi \leq 35) &\approx P(25 \leq \eta \leq 35) = P\left(\frac{25 - 30}{4,6} \leq \frac{\eta - \mu}{\sigma} \leq \frac{35 - 30}{4,6}\right) \\ &= P(-1,09 \leq \eta_{0,1} \leq 1,09) = \Phi(1,09) - \Phi(-1,09) \\ &= \Phi(1,09) - [1 - \Phi(1,09)] = 0,86 - [1 - 0,86] = 0,72 = 72\% \end{aligned}$$



Feladat. Milyen a valós számra teljesül, hogy $P(\xi \geq a) = 95\%$?

Most $P(\xi < a) = 5\%$, tehát a ξ változó 5%-os kvantilisét keressük.

Ötlet: határozzuk meg az η változó 5%-os kvantilisét!

$$0,05 = P(\eta < q_{5\%}) = P\left(\frac{\eta - \mu}{\sigma} < \frac{q_{5\%} - 30}{4,6}\right) = \Phi\left(\frac{q_{5\%} - 30}{4,6}\right)$$

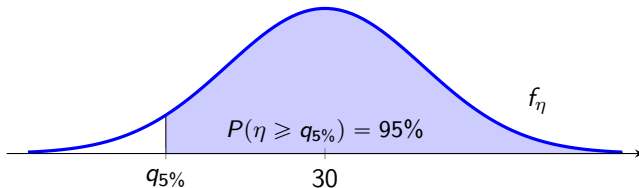
Tehát:

$$\Phi\left(-\frac{q_{5\%} - 30}{4,6}\right) = 1 - 0,05 = 0,95 = \Phi(1,64).$$

Ebből következik: $-(q_{5\%} - 30)/4,6 = 1,64$, vagyis $q_{5\%} = 22,5$.

A de Moivre–Laplace-tétel ismételt alkalmazásával:

$$P(\xi \geq 22,5) \approx P(\eta \geq 22,5) = 95\%.$$



További nevezetes diszkrét eloszlások

Hogyan dolgoztunk eddig diszkrét valószínűségi változókkal:

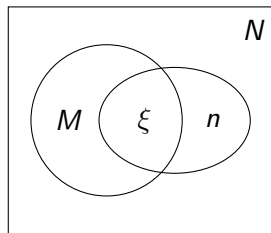
- 1 Meghatároztuk az értékkészletet.
- 2 Felírtuk a valószínűségeloszlást. (Ez a nehéz!)
- 3 Kiszámoltuk a várható értéket és a szórást.

Nevezetes eloszlások: a feladat szövege alapján ráismerünk arra, hogy melyik eloszlásról van szó, majd kipuskázzuk a formulákat.

Mintavételezési feladat: Adott egy halmaz, benne néhány elem meg van különböztetve valamilyen módon. Véletlenszerűen kiveszünk néhány elemet a halmazból (minta). Hány megkülönböztetett elem van a mintában?

Jelölések:

- N = halmaz elemszáma,
- M = megkülönböztetett elemek száma,
- n = ennyi elemet veszünk ki,
- ξ = megkülönböztetett elemek a mintában.



A mintavételezés típusai:

- **Visszatevéses:** Számít a sorrend; ha azonos elemeket húzunk ki, de más sorrendben, akkor az egy másik minta.
- **Visszatevés nélküli:** A húzás sorrendjét nem vesszük figyelembe.

A binomiális eloszlás előállítása

Legyen adva egy Bernoulli-kísérlet, ahol p a kedvező kimenetel esélye! Ha ezt a kísérletet egymástól függetlenül n alkalommal megismételjük, akkor az összesen kapott kedvező kimenetek száma binomiális eloszlást követ n és p paraméterrel.

A visszatevéses mintavételezés esete

Visszatevéses mintavételezés esetén a ξ valószínűségi változó binomiális eloszlást követ n és $p = M/N$ paraméterekkel.

Hipergeometrikus eloszlás

Egy ξ valószínűségi változó hipergeometrikus eloszlást követ N , M és n pozitív egész paraméterekkel, ha

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

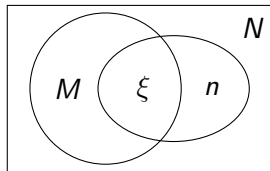
A hipergeometrikus eloszlás várható értéke $E(\xi) = nM/N$.

A hipergeometrikus eloszlás előállítása

Visszatevés nélküli mintavételezéskor a megkülönböztetett elemek száma a mintában hipergeometrikus eloszlást követ N , M és n paraméterekkel.

Bizonyítás.

$$P(\xi = k) = \frac{\text{kedvező}}{\text{összes}} = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}.$$



Feladat. Egy szelvénnel játszunk az ötöslottón. Határozzuk meg a találatok számának az eloszlását és várható értékét!

Legyen ξ a találatok száma, $R_\xi = \{0, 1, 2, 3, 4, 5\}$.

ξ hipergeometrikus eloszlású, $N = 90$, $M = 5$, $n = 5$,

$$P(\xi = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} = \frac{\binom{5}{k} \binom{85}{5-k}}{\binom{90}{5}}, \quad k = 0, 1, \dots, 5.$$

Továbbá: $E(\xi) = nM/N = 5 \cdot 5/90 \approx 0,28$.

Feladat. Hogyan változna a megoldás, ha a lottószámokat visszatevéssel húznák ki?

Legyen ξ a találatok száma, $R_\xi = \{0, 1, 2, 3, 4, 5\}$.

ξ binomiális eloszlású, $n = 5$, $p = M/N = 5/90$,

$$P(\xi = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{5}{k} \left(\frac{5}{90}\right)^k \left(\frac{85}{90}\right)^{5-k}, \quad k = 0, 1, \dots, 5.$$

Továbbá: $E(\xi) = np = 5 \cdot 5/90 \approx 0,28$.

Geometriai eloszlás

Egy ξ valószínűségi változó **geometriai eloszlást** követ $p \in (0, 1)$ paraméterrel, ha

$$P(\xi = k) = (1 - p)^{k-1}p, \quad k = 1, 2, \dots$$

Az eloszlás várható értéke és szórása: $E(\xi) = 1/p$, $D(\xi) = \sqrt{1-p}/p$.

A geometriai eloszlás előállítása

Legyen adva egy Benoulli-kísérlet p sikervalószínűséggel. Amennyiben ezt a kísérletet egymástól függetlenül addig ismételjük, míg kedvező kimenetelt nem kapunk, akkor az ismételések száma geometriai eloszlást követ.

Bizonyítás.

$$\begin{aligned} P(\xi = k) &= P(\text{sorban } k-1 \text{ kedvezőtlen, majd egy kedvező kimenetel}) \\ &= \underbrace{(1-p) \cdot \dots \cdot (1-p)}_{(k-1) \text{ darab}} \cdot p = (1-p)^{k-1}p, \quad k = 1, 2, \dots \end{aligned}$$

Feladat. Piri néni szeretne venni egy szép pálmát a közeli faiskolából, de sajnos nem mindig talál neki tetszőt. A faiskolába minden héten érkezik új áru, de ebben a korábbi hetektől függetlenül csak 40% valószínűséggel van olyan csemete, amivel Piri néni elégedett. Emiatt Piri néni minden héten elmegy a faiskolába egészen addig, míg nem talál egy neki tetsző pálmát. Jelölje ξ azt, hogy hányszor kell elmennie a faiskolába. Határozzuk meg a ξ változó eloszlását, várható értékét és szórását!

Legyen ξ az utazások száma, $R_\xi = \{1, 2, \dots\}$

Bernoulli-kísérlet: egy utazás, kedvező kimenetel: szép pálma, $p = 0,4$.

ξ = kísérletek száma az első kedvező kimenetelig, geometriai eloszlású,

$$P(\xi = k) = (1 - p)^{k-1}p = 0,6^{k-1}0,4,$$

$$E(\xi) = \frac{1}{p} = \frac{1}{0,4} = 2,5, \quad D(\xi) = \frac{\sqrt{1-p}}{p} = \frac{\sqrt{0,6}}{0,4} \approx 1,94.$$

Feladat. Mennyi annak a valószínűsége, hogy Piri néni legfeljebb 5 hét alatt talál neki tetsző pálmát?

Annak az esélye, hogy 5 nap alatt nem tud pálmát venni:

$$P(\xi > 5) = P(5 \text{ napon át csak rossz pálma}) = 0,6^5 = 0,078$$

Tehát: $P(\xi \leq 5) = 1 - P(\xi > 5) = 0,922 = 92,2\%$

Feladat. Milyen k értékre teljesül, hogy k nap alatt Piri néni legalább 95% százalék eséllyel talál neki tetsző pálmát?

Az előző feladat mintájára:

$$P(\xi > k) = P(k \text{ napon át csak rossz pálma}) = 0,6^k,$$

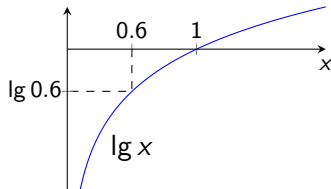
$$P(\xi \leq k) = 1 - P(\xi > k) = 1 - 0,6^k \geq 0,95,$$

$$0,6^k \leq 0,05$$

$$k \lg 0,6 = \lg 0,6^k \leq \lg 0,05$$

$$k \geq \lg 0,05 / \lg 0,6 = 5,86$$

$$k \geq 6$$

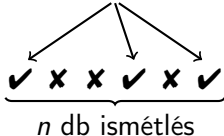


Gyakori hiba, hogy keverik a binomiális és a geometriai eloszlást, ugyanis mindkét esetben egy Bernoulli-kísérletet ismételünk meg több alkalommal. Pedig van különbség:

- Binomiális eloszlás: a végrehajtások száma (n) előre rögzített, és az a kérdés, hogy hány kedvező kimenetelt kapunk.
- Geometriai eloszlás: az van megadva, hogy hány kedvező kimenetelt akarunk (1), és a szükséges végrehajtások száma a véletlen.

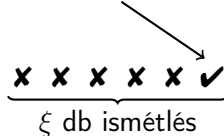
Binomiális eloszlás:

ξ db kedvező kimenetel



Geometriai eloszlás:

1 db kedvező kimenetel



A várható érték és a szórás tulajdonságai

Konstansváltozó (degenerált változó)

Egy ξ valószínűségi változó **konstansváltozó**, ha létezik olyan a valós szám, hogy $P(\xi = a) = 1$.

A várható érték és a szórás tulajdonságai

Legyen ξ tetszőleges valószínűségi változó, és legyen a valós szám.

- 1 Ha $P(\xi = a) = 1$, akkor $E(\xi) = a$ és $D(\xi) = 0$.
- 2 Egy változónak pontosan akkor 0 a szórása, ha a változó konstans.
- 3 $E(a\xi) = aE(\xi)$, $D^2(a\xi) = a^2D^2(\xi)$.

Bizonyítás.

- (1) Diszkrét valószínűségi változó várható értéke és szórása, házi feladat.
- (2) Ha a változó nem konstans, akkor $P(\xi = E(\xi)) < 1$, tehát

$$\text{Var}(\xi) = E\left([\xi - E(\xi)]^2\right) = [\xi - E(\xi)]^2 \text{ átlagos értéke} > 0$$

Vagyis ha $\text{Var}(\xi) = 0$, akkor a változó konstans.

(3) A várható értékre vonatkozó azonosság:

Független megfigyelések a ξ változóra: $\xi_1, \xi_2, \dots$

Független megfigyelések az $a\xi$ változóra: $a\xi_1, a\xi_2, \dots$

Nagy számok Kolmogorov-féle törvénye:

$$E(a\xi) \leftarrow \frac{a\xi_1 + \dots + a\xi_n}{n} = a \frac{\xi_1 + \dots + \xi_n}{n} \rightarrow aE(\xi).$$

Tehát: $E(a\xi) = aE(\xi)$.

A szórásra vonatkozó azonosság:

$$\begin{aligned} D^2(a\xi) &= E\left([a\xi - E(a\xi)]^2\right) = E\left([a\xi - aE(\xi)]^2\right) \\ &= E\left(a^2[\xi - E(\xi)]^2\right) = a^2 E\left([\xi - E(\xi)]^2\right) = a^2 D^2(\xi). \end{aligned}$$

Összeg várható értéke és varianciája

Legyenek $\xi^{(1)}, \dots, \xi^{(m)}$ tetszőleges valószínűségi változók, és tekintsünk $a_1, \dots, a_m$ valós számokat. Ekkor teljesülnek az alábbi azonosságok.

❶ A változók összegének a várható értéke:

$$E(\xi^{(1)} + \dots + \xi^{(m)}) = E(\xi^{(1)}) + \dots + E(\xi^{(m)}).$$

Ebből az is következik, hogy

$$E(a_1\xi^{(1)} + \dots + a_m\xi^{(m)}) = a_1E(\xi^{(1)}) + \dots + a_mE(\xi^{(m)}).$$

❷ Az összegvarianciája: ha $\xi^{(1)}, \dots, \xi^{(m)}$ függetlenek, akkor

$$D^2(\xi^{(1)} + \dots + \xi^{(m)}) = D^2(\xi^{(1)}) + \dots + D^2(\xi^{(m)}).$$

Ebből az is következik, hogy

$$D^2(a_1\xi^{(1)} + \dots + a_m\xi^{(m)}) = a_1^2D^2(\xi^{(1)}) + \dots + a_m^2D^2(\xi^{(m)}).$$

Bizonyítás.

(1) Az állítást csak $m = 2$ esetén bizonyítjuk. Cél:

$$E(\xi + \eta) = E(\xi) + E(\eta).$$

Független megfigyelések a ξ változóra: $\xi_1, \xi_2, \dots$

Független megfigyelések az η változóra: $\eta_1, \eta_2, \dots$

Független megfigyelések a $\xi + \eta$ összegre: $\xi_1 + \eta_1, \xi_2 + \eta_2, \dots$

Nagy számok Kolmogorov-féle törvénye:

$$\frac{(\xi_1 + \eta_1) + \dots + (\xi_n + \eta_n)}{n} = \frac{\xi_1 + \dots + \xi_n}{n} + \frac{\eta_1 + \dots + \eta_n}{n}$$

$$\downarrow \qquad \qquad \qquad \downarrow \qquad \qquad \qquad \downarrow$$

$$E(\xi + \eta) \qquad \qquad \qquad E(\xi) \qquad \qquad \qquad E(\eta)$$

Tehát: $E(\xi + \eta) = E(\xi) + E(\eta)$.

Ebből következik, hogy

$$E(a_1\xi + a_2\eta) = E(a_1\xi) + E(a_2\eta) = a_1E(\xi) + a_2E(\eta).$$

Most bebizonyítunk egy olyan állítást, ami korábban már kimondtunk.

A variancia meghatározása

Ha ξ olyan valószínűségi változó, melynek véges a várható értéke, akkor

$$\text{Var}(\xi) = E(\xi^2) - (E(\xi))^2.$$

Bizonyítás.

$$\begin{aligned}\text{Var}(\xi) &= E\left([\xi - E(\xi)]^2\right) = E\left(\xi^2 - 2E(\xi)\xi + (E(\xi))^2\right) \\ &= E(\xi^2) + E(-2E(\xi) \cdot \xi) + E([E(\xi)]^2) \\ &= E(\xi^2) - 2E(\xi) \cdot E(\xi) + [E(\xi)]^2 \\ &= E(\xi^2) - (E(\xi))^2\end{aligned}$$

Feladat. Egy cég úgy próbálja meg népszerűsíteni az általa forgalmazott chipset, hogy a zacskókba a Micimackó című mese figuráit rejti el: Micimackót, Malackát, Tigrist és Fülest. Minden zacskóban pontosan egy figura található, és minden figurának azonos a gyakorisága. Mi addig vásároljuk a terméket, míg meg nem kapjuk a teljes kollekció, tehát míg végül mindegyik figurából lesz legalább egy darab. Várhatóan hány zacskó chipset kell megvennünk? Mennyi a vásárolt chipsek számának a szórása?

Legyen: ξ = a vásárolt chipsek száma, $R_\xi = \{4, 5, \dots\}$.

Amit tudunk:

$$P(\xi = 4) = 0,09, \quad P(6 \text{ chips elég}) = P(\xi \leq 6) = 0,38.$$

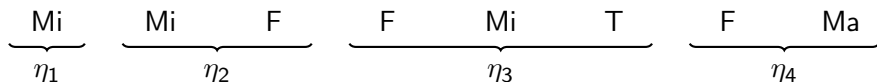
Ha minden k -ra sikerülne meghatározni a $P(\xi \leq k)$ valószínűséget, akkor ebből az eloszlás:

$$P(\xi = k) = P(\xi \leq k) - P(\xi \leq k - 1), \quad k = 5, 6, \dots$$

De ez egy nehéz út lenne...

Vezessük be az $\eta_1, \eta_2, \eta_3, \eta_4$ változókat:

η_i = ha már van $i - 1$ különböző figuránk, akkor hány további chipset kell vásárolnunk ahhoz, hogy találjunk egy új figurát



Észrevételek:

- az η_1 változó értéke mindig 1;
- η_2 geometriai eloszlású $p = 3/4$ paraméterrel;
- η_3 geometriai eloszlású $p = 2/4$ paraméterrel;
- η_4 geometriai eloszlású $p = 1/4$ paraméterrel;
- $\xi = \eta_1 + \eta_2 + \eta_3 + \eta_4$;
- $\eta_1, \eta_2, \eta_3, \eta_4$ függetlenek.

A várható érték és a szórás:

$$E(\xi) = E(\eta_1) + E(\eta_2) + E(\eta_3) + E(\eta_4) = 1 + \frac{1}{3/4} + \frac{1}{1/2} + \frac{1}{1/4} \approx 8,33,$$

$$\begin{aligned} D^2(\xi) &= D^2(\eta_1) + D^2(\eta_2) + D^2(\eta_3) + D^2(\eta_4) \\ &= 0 + \frac{1 - 3/4}{(3/4)^2} + \frac{1 - 1/2}{(1/2)^2} + \frac{1 - 1/4}{(1/4)^4} \approx 14,44, \end{aligned}$$

$$D(\xi) = \sqrt{\text{Var}(\xi)} \approx 3,8.$$

Feladat. Ha egy zacskó chips ára 300 forint, akkor mennyi az összesen elköltött pénzünk várható értéke és szórása?

Most ξ chipset vásárolunk meg, aminek 300ξ forint az ára. Az elköltött pént várható értéke és szórása:

$$E(300\xi) = 300E(\xi) = 2500,$$

$$D^2(300\xi) = 300^2 D^2(\xi) = 1.299.600,$$

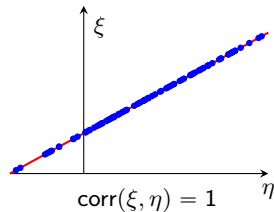
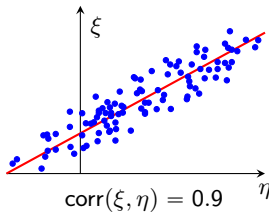
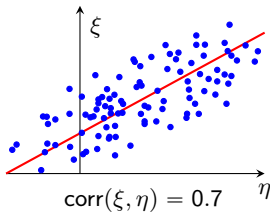
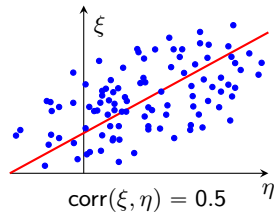
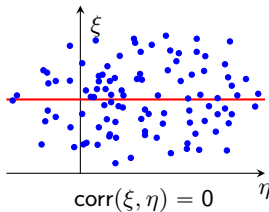
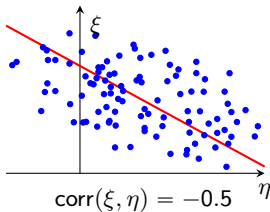
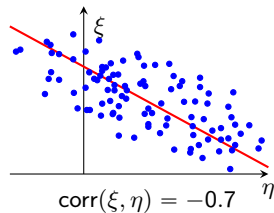
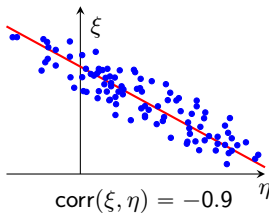
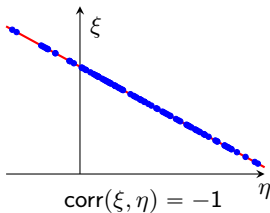
$$D(300\xi) = \sqrt{1.299.600} = 1140.$$

Legyenek ξ és η valószínűségi változók! A **korrelációs együttható** a két változó közötti kapcsolat irányát és erősségét számszerűsíti.

Jelölés: $\text{corr}(\xi, \eta)$.

Fontosabb tulajdonságai:

- $\text{corr}(\xi, \eta) \in [-1, +1]$.
- Ha ξ és η függetlenek, akkor $\text{corr}(\xi, \eta) = 0$.
- Ha $\text{corr}(\xi, \eta) > 0$, akkor a két változó között pozitív irányú kapcsolat van.
- Ha $\text{corr}(\xi, \eta) < 0$, akkor a két változó között negatív irányú kapcsolat van.
- Minél közelebb van a korrelációs együttható $+1$ -hez vagy -1 -hez, annál erősebb a két változó közötti kapcsolat.
- Ha $\text{corr}(\xi, \eta) = \pm 1$, akkor a két változó tökéletesen meghatározza egymást, továbbá $\eta = a\xi + b$ valamilyen a, b valós számokkal.



Valószínűségi változók összegének a varianciája

Legyenek ξ és η tetszőleges valószínűségi változók, és legyenek a és b valós számokat. Ekkor

$$D^2(a\xi + b\eta) = a^2 D^2(\xi) + b^2 D^2(\eta) + 2ab D(\xi) D(\eta) \operatorname{corr}(\xi, \eta).$$

Ebből $a = b = 1$ választással az is következik, hogy

$$D^2(\xi + \eta) = D^2(\xi) + D^2(\eta) + 2D(\xi)D(\eta) \operatorname{corr}(\xi, \eta).$$

Megjegyzések:

- A fenti formula kettőnél több változó összegére is felírható, de azt nem tanuljuk.
- Ha ξ és η függetlenek, akkor $\operatorname{corr}(\xi, \eta) = 0$, vagyis

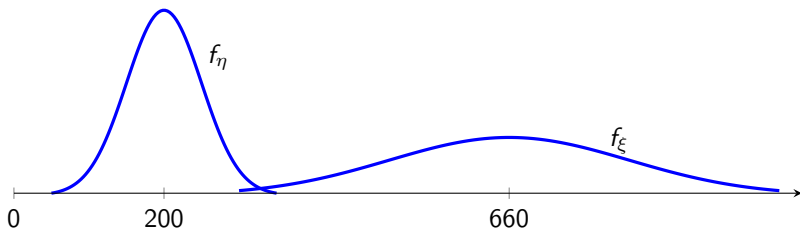
$$D^2(a\xi + b\eta) = a^2 D^2(\xi) + b^2 D^2(\eta).$$

Feladat. A Tisza és a Maros vízhozama egy számunkra ismeretlen eloszlást követ. Azt tudjuk, hogy a Maros torkolata felett a Tisza vízhozamának a várható értéke $660 \text{ m}^3/\text{s}$, szórása $160 \text{ m}^3/\text{s}$, míg a Maros vízhozamának a várható értéke $200 \text{ m}^3/\text{s}$, szórása $50 \text{ m}^3/\text{s}$. Határozzuk meg a Tisza belvárosi hídnál mért vízhozamának a várható értékét és szórását, ha

- a két vízhozam független;
- a két vízhozam korrelációs együtthatója $+0,8$;
- a két vízhozam korrelációs együtthatója $-0,8$;

Legyen ξ = a Tisza vízhozama, η = a Maros vízhozama!

Ekkor $E(\xi) = 660$, $D(\xi) = 160$, $E(\eta) = 200$, $D(\eta) = 50$.



A Tisza belvárosi hídnál mért vízhozam várható értéke és szórása:

$$E(\xi + \eta) = E(\xi) + E(\eta) = 660 + 200 = 860,$$

$$\begin{aligned} D^2(\xi + \eta) &= D^2(\xi) + D^2(\eta) + 2D(\xi)D(\eta) \operatorname{corr}(\xi, \eta) \\ &= 160^2 + 50^2 + 2 \cdot 160 \cdot 50 \cdot \operatorname{corr}(\xi, \eta) = 28.100 + 16.000 \operatorname{corr}(\xi, \eta). \end{aligned}$$

- Ha a két vízhozam független, akkor $\operatorname{corr}(\xi, \eta) = 0$, tehát

$$D^2(\xi + \eta) = 28.100 + 16.000 \cdot 0 = 28.100, \quad D(\xi + \eta) = 168.$$

- Ha $\operatorname{corr}(\xi, \eta) = +0,8$, akkor

$$D^2(\xi + \eta) = 28.100 + 16.000 \cdot (+0,8) = 40.900, \quad D(\xi + \eta) = 202.$$

- Ha $\operatorname{corr}(\xi, \eta) = -0,8$, akkor

$$D^2(\xi + \eta) = 28.100 + 16.000 \cdot (-0,8) = 15.300, \quad D(\xi + \eta) = 124.$$

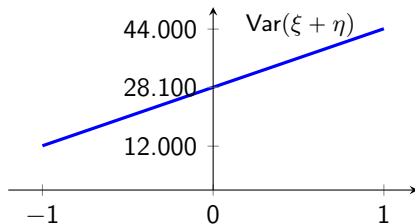
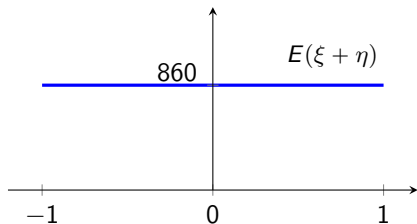
Feladat. Vajon a valóságban milyen előjelű a két vízhozam közötti korrelációs együttható? Az árvízi védekezés szempontjából melyik lenne jobb, a negatív vagy a pozitív korreláció?

A valóságban a két vízhozam közötti korreláció pozitív. Árvízvédelmi szempontból a negatív korreláció lenne előnyösebb.

Feladat. Írjuk fel és ábrázoljuk a $\xi + \eta$ változó várható értékét és varianciáját a korrelációs együttható függvényében!

Várható érték: $E(\xi + \eta) = 860$

Variancia: $\text{Var}(\xi + \eta) = 28.100 + 16.000 \text{corr}(\xi, \eta)$



Feladat. Adott egy kötvény és egy részvény, mindkettőnek 1000 forint a jelenlegi ára. A kötvény 10 százalékos éves hozamot fizet. A részvény értéke egy év múlva egy valószínűségi változó, mely a 900, az 1200 és az 1500 értéket veheti fel rendre 0,3, 0,4 illetve 0,3 valószínűséggel. Ha 1 millió forintból állíthatunk össze egy portfóliót, akkor várhatóan mennyit fog érni ez a portfólió egy év múlva? Mennyi a portfólió értékének a szórása?

Az árak egy év múlva: ξ = részvény ára, η = kötvény ára.

A pénzügyi matematikában a kötvény kockázatmentes:

$$P(\eta = 1100) = 1, \quad E(\eta) = 1100, \quad D(\eta) = 0.$$

A részvény árának eloszlása:

$$P(\xi = 900) = 0,3, \quad P(\xi = 1200) = 0,4, \quad P(\xi = 1500) = 0,3,$$

Várható érték és szórás:

$$E(\xi) = 900 \cdot 0,3 + 1200 \cdot 0,4 + 1500 \cdot 0,3 = 1200,$$

$$E(\xi^2) = 900^2 \cdot 0,3 + 1200^2 \cdot 0,4 + 1500^2 \cdot 0,3 = 1.494.000,$$

$$\text{Var}(\xi) = E(\xi^2) - (E(\xi))^2 = 54.000, \quad D(\xi) = \sqrt{\text{Var}(\xi)} \approx 232.$$

Ha a részvényt és b kötvényt veszünk, akkor a portfólió értéke ma:

$$a \cdot 1000 + b \cdot 1000 = 1.000.000, \quad \text{amiből} \quad a + b = 1000.$$

A portfólió értéke egy év múlva: $a\xi + b\eta$.

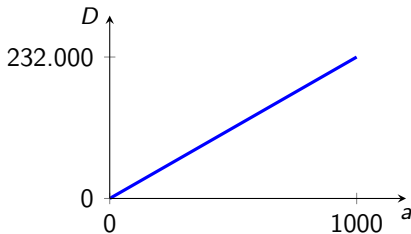
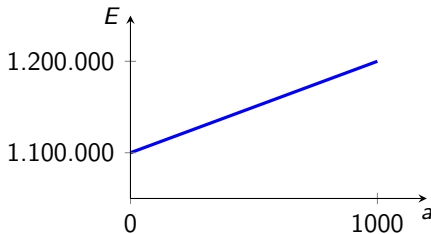
A kötvény ára nem függ a részvény árától, tehát ξ és η függetlenek.

Várható érték és szórás:

$$\begin{aligned} E(a\xi + b\eta) &= aE(\xi) + bE(\eta) = a \cdot 1200 + b \cdot 1100 \\ &= 1200a + 1100(1000 - a) = 100a + 1.100.000, \end{aligned}$$

$$D^2(a\xi + b\eta) = a^2 D^2(\xi) + b^2 D^2(\eta) = a^2 \cdot 54.000 + b \cdot 0 = 54.000a^2,$$

$$D(a\xi + b\eta) = \sqrt{54.000a^2} \approx 232a.$$



Feladat. Hogyan fektessük be a pénzünket, ha az a célunk, hogy várható értékben 300 ezer forint hozamot realizáljunk? Mekkora ebben az esetben a portfóliói szórása?

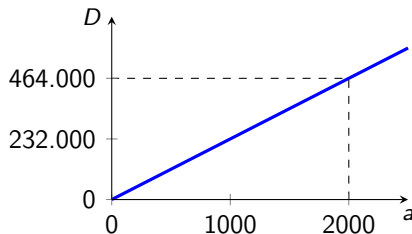
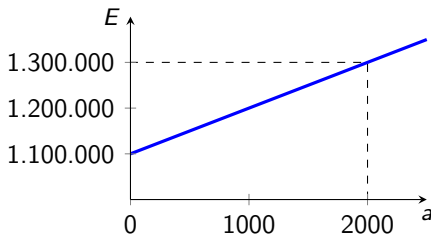
A 300 ezer forint várható hozam esetén:

$$1.300.000 = E(a\xi + b\eta) = 100a + 1.100.000.$$

A megoldás: $a = 2000$, $b = -1000$.

El kell adnunk 1000 kötvényt, azaz fel kell vennünk 1 millió forint hitelt, majd a rendelkezésre álló 2 millió forintból részvényt kell vennünk.

A szórás: $D(a\xi + b\eta) = 232a = 464.000$.



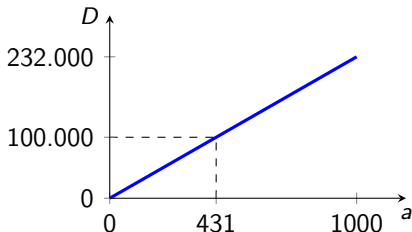
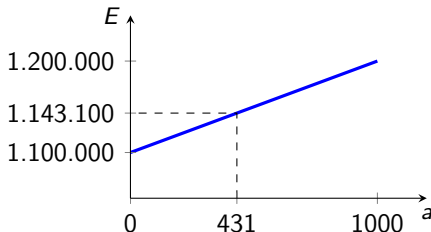
Feladat. Úgy akarjuk befektetni a pénzünket, hogy a portfólió értékének a szórása legfeljebb 100 ezer forint legyen. Mennyi a maximálisan elérhető várható hozam ebben az esetben?

A portfólió jövőbeli értékének a szórása

$$100.000 \geq D(a\xi + b\eta) = 232a, \quad \text{amiből} \quad a \leq 431.$$

A hozam várható értéke annál nagyobb, minél több részvényt veszünk. A maximális várható hozamot akkor érjük el, ha $a = 431$ és $b = 569$.

A várható érték: $E(a\xi + b\eta) = 100a + 1.100.000 = 1.143.100$.



A centrális határeloszlás-tétel

Feladat. Egy bankban egy adott napon 200 ember akar készpénzt felvenni egymástól függetlenül. Egy-egy felvétel során a felvett összeg nagysága egy olyan valószínűségi változó, melynek várható értéke 50 ezer forint, szórása 30 ezer forint. Adjuk meg a teljes kifizetés várható értékét és szórását! Mennyi annak az esélye, hogy a teljes kifizetés legalább 10,5 millió forint?

Legyen ξ egy készpénzfelvétel nagysága ezer forintban kifejezve!

Tudjuk: $E(\xi) = 50$, $D(\xi) = 30$.

Független megfigyelések a ξ változóra: $\xi_1, \dots, \xi_{200}$.

Teljes kifizetés: $\xi_1 + \dots + \xi_{200}$.

$$E(\xi_1 + \dots + \xi_{200}) = E(\xi_1) + \dots + E(\xi_{200}) = 200E(\xi) = 10.000,$$

$$D^2(\xi_1 + \dots + \xi_{200}) = D^2(\xi_1) + \dots + D^2(\xi_{200}) = 200D^2(\xi) = 180.000,$$

$$D(\xi_1 + \dots + \xi_{200}) = \sqrt{200}D(\xi) = 424.$$

Második kérdés: nem ismertjük az összegváltozó eloszlását!

Centrális határeloszlás-tétel, CHT (Lévy, 1919)

Legyenek

- $\xi_1, \dots, \xi_n$ egy véges szórású ξ változó független megfigyelései;
- η normális eloszlású változó, $\mu = nE(\xi)$ és $\sigma = \sqrt{nD(\xi)}$.

Ekkor tetszőleges $a \leq b$ valós számok esetén

$$P(a \leq \xi_1 + \dots + \xi_n \leq b) \approx P(a \leq \eta \leq b),$$

A közelítés annál pontosabb, minél nagyobb az n paraméter értéke.

Megjegyzések:

- Az η normális eloszlású változó paramétereit úgy választjuk meg, hogy a lehető legjobban hasonlítson a $\xi_1 + \dots + \xi_n$ összeghez:

$$E(\xi_1 + \dots + \xi_n) = nE(\xi) = \mu = E(\eta);$$

$$D^2(\xi_1 + \dots + \xi_n) = nD^2(\xi) = \sigma^2 = D^2(\eta).$$

- A mintaméret legyen $n \geq 50$, de inkább $n \geq 100$.

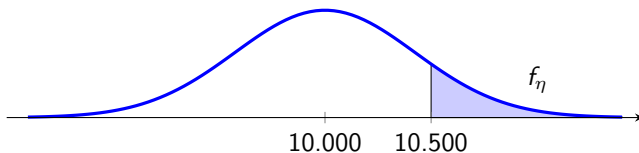
Feladat. A bankos feladatban mennyi az esélye, hogy a teljes kifizetés legalább 10,5 millió forint?

Legyen η normális eloszlású változó a következő paraméterekkel:

$$\mu = nE(\xi) = 200 \cdot 50 = 10.000 \quad \text{és} \quad \sigma = \sqrt{nD(\xi)} = \sqrt{200} \cdot 30 = 424.$$

A centrális határeloszlás-tétel alkalmazásával:

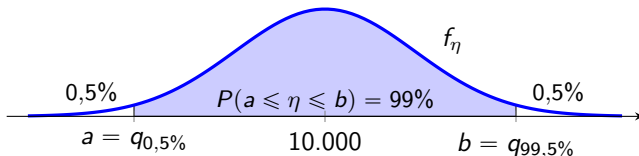
$$\begin{aligned} P(\xi_1 + \dots + \xi_{200} \geq 10.500) &\approx P(\eta \geq 10.500) \\ &= P\left(\frac{\eta - \mu}{\sigma} \geq \frac{10.500 - 10.000}{424}\right) = P(\eta_{0,1} \geq 1,18) \\ &= 1 - P(\eta_{0,1} < 1,18) = 1 - \Phi(1,18) = 1 - 0,88 = 0,12. \end{aligned}$$



Feladat. Adjunk meg egy olyan $[a, b]$ intervallumot, mely 99% valószínűséggel tartalmazza a teljes kifizetés nagyságát!

Először az η változóra adunk meg egy ilyen intervallumot. Legyenek $a = q_{0,5\%}$ és $b = q_{99,5\%}$ az η változó két kvantilise! Ekkor:

$$P(q_{0,5\%} \leq \eta \leq q_{99,5\%}) = P(\eta < q_{99,5\%}) - P(\eta < q_{0,5\%}) = 99,5\% - 0,5\% = 99\%.$$



Ez jó megoldás lesz a feladatra, ugyanis a CHT alkalmazásával:

$$P(a \leq \xi_1 + \dots + \xi_{200} \leq b) \approx P(a \leq \eta \leq b) = 99\%.$$

A 99,5%-os kvantilis, $b = q_{99,5\%}$:

$$0,995 = P(\eta < b) = P\left(\frac{\eta - \mu}{\sigma} < \frac{b - 10.000}{424}\right) = \Phi\left(\frac{b - 10.000}{424}\right)$$

Az eloszlástáblázatban: $\Phi(2,58) = 0,995$. Kapjuk:

$$\frac{b - 10.000}{424} = 2,58, \quad \text{tehát} \quad b = 11.094 = 11 \text{ millió } 94 \text{ ezer.}$$

A 0,5%-os kvantilis, $a = q_{0,5\%}$:

$$0,005 = P(\eta < a) = P\left(\frac{\eta - \mu}{\sigma} < \frac{a - 10.000}{424}\right) = \Phi\left(\frac{a - 10.000}{424}\right).$$

Vagyis:

$$\Phi\left(-\frac{a - 10000}{424}\right) = 1 - 0,005 = 0,995 = \Phi(2,58).$$

Kapjuk:

$$-\frac{a - 10.000}{424} = 2,58, \quad \text{tehát} \quad a = 8.906 = 8 \text{ millió } 906 \text{ ezer.}$$

A kérdéses intervallum: $[a, b] = [8 \text{ millió } 906 \text{ ezer}, 11 \text{ millió } 94 \text{ ezer}]$.

Együttes eloszlás, kovariancia, korreláció

Diszkrét valószínűségi változók együttes eloszlása

Legyenek ξ és η diszkrét valószínűségi változók!

- A két változó **együttes eloszlása**:

$$P(\xi = x, \eta = y), \quad x \in R_\xi, y \in R_\eta.$$

- **Marginális eloszlások, peremeloszlások**: a két változó külön-külön vett eloszlása.
- Akkor mondjuk, hogy a két diszkrét valószínűségi változó **független egymástól**, ha tetszőleges $x \in R_\xi$ és $y \in R_\eta$ értékek esetén:

$$P(\xi = x, \eta = y) = P(\xi = x)P(\eta = y).$$

Szemléletesen: a $\xi = x$ és az $\eta = y$ esemény független egymástól.

Feladat. Adott két szabályos pénzérme, melyek egyik oldalára a 0, a másik oldalára az 1 értéket írjuk. Feldobjuk a két érmét, és legyen ξ a dobott számok összege, η pedig a dobott számok szorzata.

ξ = dobott számok összege, $R_\xi = \{0, 1, 2\}$

η = dobott számok szorzata, $R_\eta = \{0, 1\}$

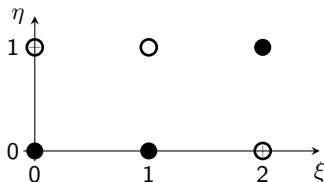
Együttes eloszlás:

$$P(\xi = 0, \eta = 0) = P(\text{két 0-t dobunk}) = 1/4,$$

$$P(\xi = 1, \eta = 0) = P(\text{egy 0-t és egy 1-est dobunk}) = 1/2,$$

$$P(\xi = 2, \eta = 1) = P(\text{két 1-est dobunk}) = 1/4,$$

$$P(\xi = 0, \eta = 1) = P(\xi = 1, \eta = 1) = P(\xi = 2, \eta = 0) = 0.$$



	$\xi = 0$	$\xi = 1$	$\xi = 2$
$\eta = 1$	0	0	1/4
$\eta = 0$	1/4	1/2	0

Feladat. Adjuk meg a marginális eloszlásokat!

A ξ változó eloszlása:

$$P(\xi = 0) = P(\xi = 0, \eta = 0) + P(\xi = 0, \eta = 1) = 1/4 + 0 = 1/4,$$

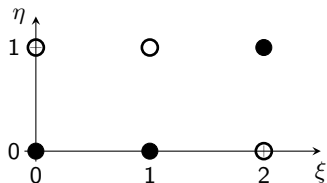
$$P(\xi = 1) = P(\xi = 1, \eta = 0) + P(\xi = 1, \eta = 1) = 1/2 + 0 = 1/2,$$

$$P(\xi = 2) = P(\xi = 2, \eta = 0) + P(\xi = 2, \eta = 1) = 0 + 1/4 = 1/4.$$

Az η változó eloszlása:

$$P(\eta = 0) = P(\xi = 0, \eta = 0) + P(\xi = 1, \eta = 0) + P(\xi = 2, \eta = 0) = 3/4,$$

$$P(\eta = 1) = P(\xi = 0, \eta = 1) + P(\xi = 1, \eta = 1) + P(\xi = 2, \eta = 1) = 1/4.$$



	$\xi = 0$	$\xi = 1$	$\xi = 2$	össz.
$\eta = 1$	0	0	1/4	1/4
$\eta = 0$	1/4	1/2	0	3/4
össz.	1/4	1/2	1/4	1

Feladat. Határozzuk meg a ξ és az η várható értékét és szórását!

	$\xi = 0$	$\xi = 1$	$\xi = 2$	össz.
$\eta = 1$	0	0	1/4	1/4
$\eta = 0$	1/4	1/2	0	3/4
össz.	1/4	1/2	1/4	1

$$E(\xi) = 0 \cdot 1/4 + 1 \cdot 1/2 + 2 \cdot 1/4 = 1,$$

$$E(\xi^2) = 0^2 \cdot 1/4 + 1^2 \cdot 1/2 + 2^2 \cdot 1/4 = 2,5,$$

$$D^2(\xi) = E(\xi^2) - (E(\xi))^2 = 1,5, \quad D(\xi) = 1,225.$$

$$E(\eta) = 0 \cdot 3/4 + 1 \cdot 1/4 = 0,25,$$

$$E(\eta^2) = 0^2 \cdot 3/4 + 1^2 \cdot 1/4 = 0,25,$$

$$D^2(\eta) = E(\eta^2) - (E(\eta))^2 = 0,1875, \quad D(\eta) = 0,433.$$

Kovariancia és korreláció

Legyenek ξ és η véges szórású valószínűségi változók!

- A két változó **kovarianciája**:

$$\text{Cov}(\xi, \eta) = E\left([\xi - E(\xi)][\eta - E(\eta)]\right)$$

- A két változó **korrelációja** avagy **korrelációs együtthatója**:

$$\text{corr}(\xi, \eta) = \frac{\text{Cov}(\xi, \eta)}{D(\xi)D(\eta)}$$

- A két változó akkor **korrelálatlan**, ha $\text{corr}(\xi, \eta) = 0$.

Megjegyzések:

- A kovariancia és a korreláció szimmetrikus a két változóban:

$$\text{Cov}(\xi, \eta) = \text{Cov}(\eta, \xi) \quad \text{és} \quad \text{corr}(\xi, \eta) = \text{corr}(\eta, \xi).$$

- $\text{Cov}(\xi, \xi) = \text{Var}(\xi)$.

A kovariancia és a korreláció fontosabb tulajdonságai

Legyenek ξ és η véges szórású valószínűségi változók!

- 1 $\text{Cov}(\xi, \eta) \in \mathbb{R}$
- 2 $\text{corr}(\xi, \eta) \in [-1, 1]$
- 3 $\text{Cov}(\xi, \eta) = E(\xi\eta) - E(\xi)E(\eta)$
- 4 Ha változók függetlenek, akkor korrelálatlanok is, azaz $\text{corr}(\xi, \eta) = 0$. Ennek megfordítása nem igaz, a korrelálatlanságból nem következik a függetlenség.

Bizonyítás. (3)

$$\begin{aligned}\text{Cov}(\xi, \eta) &= E([\xi - E(\xi)][\eta - E(\eta)]) \\ &= E(\xi\eta - E(\xi)\eta - E(\eta)\xi + E(\xi)E(\eta)) \\ &= E(\xi\eta) - E(E(\xi)\eta) - E(E(\eta)\xi) + E(E(\xi)E(\eta)) \\ &= E(\xi\eta) - E(\xi)E(\eta) - E(\eta)E(\xi) + E(\xi)E(\eta) \\ &= E(\xi\eta) - E(\xi)E(\eta)\end{aligned}$$

Feladat. Adjuk meg a két változó kovarianciáját és korrelációját!

	$\xi = 0$	$\xi = 1$	$\xi = 2$	össz.
$\eta = 1$	0	0	1/4	1/4
$\eta = 0$	1/4	1/2	0	3/4
össz.	1/4	1/2	1/4	1

A szorzat várható értéke:

$$E(\xi\eta) = \sum_{x \in R_\xi, y \in R_\eta} xyP(\xi = x, \eta = y) = (0 \cdot 0) \cdot 1/4 + (1 \cdot 0) \cdot 1/2 + (2 \cdot 1) \cdot 1/4 = 0,5$$

Kovariancia:

$$\text{Cov}(\xi, \eta) = E(\xi\eta) - E(\xi)E(\eta) = 0,5 - 1 \cdot 0,25 = 0,25$$

Korrelációs együttható:

$$\text{corr}(\xi, \eta) = \frac{\text{Cov}(\xi, \eta)}{D(\xi)D(\eta)} = \frac{0,25}{1,225 \cdot 0,433} = 0,47$$

Feladat. A pénzérmés feladatban a két változó független egymástól?

1. Megoldás. Nem, ugyanis $\text{corr}(\xi, \eta) \neq 0$.

2. Megoldás. Nem, ugyanis $x = 0$ és $y = 1$ választással:

$$P(\xi = 0, \eta = 1) = 0 \neq 1/4 \cdot 1/4 = P(\xi = 0)P(\eta = 1).$$

Feladat. Adjuk meg ξ és η együttes eloszlását olyan módon, hogy a marginális eloszlások ne változzanak, de a két változó független legyen!

A két változó pontosan akkor független, ha tetszőleges x és y esetén:

$$P(\xi = x, \eta = y) = P(\xi = x)P(\eta = y).$$

	$\xi = 0$	$\xi = 1$	$\xi = 2$	össz.
$\eta = 1$	1/16	1/8	1/16	1/4
$\eta = 0$	3/16	3/8	3/16	3/4
össz.	1/4	1/2	1/4	1

Feladat. Adjunk példát olyan változókra, melyek korrelálatlanok, de nem függetlenek!

Tekintsük az alábbi együttes eloszlással definiált változókat:

	$\xi = -1$	$\xi = 0$	$\xi = +1$	össz.
$\eta = +1$	0	1/4	0	1/4
$\eta = 0$	1/4	0	1/4	1/2
$\eta = -1$	0	1/4	0	1/4
össz.	1/4	1/2	1/4	1

Ekkor teljesülnek az alábbiak:

- Standard számolással levezethető, hogy $\text{corr}(\xi, \eta) = 0$.
- Viszont a változók nem függetlenek:

$$P(\xi = 0, \eta = 0) = 0 \neq 1/2 \cdot 1/2 = P(\xi = 0)P(\eta = 0).$$

A nagy számok törvénye

A nagy számok Kolmogorov-féle törvénye (1933)

Legyen $\xi_1, \xi_2, \dots$ a ξ valószínűségi változó egymástól független megfigyelései! Ekkor

$$\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow E(\xi), \quad n \rightarrow \infty.$$

Bizonyítás. A jelenleg tudásunk kevés hozzá.

A nagy számok Borel-féle törvénye

Tekintsünk egy véletlen kísérletet, és legyen A egy tetszőleges esemény! Jelölje $k_n(A)$ az A esemény bekövetkezési gyakoriságát a kísérlet n alkalommal való független megismétlése során! Ekkor:

$$\frac{k_n(A)}{n} \rightarrow P(A), \quad n \rightarrow \infty.$$

Bizonyítás. Vezessük be a következő valószínűségi változót:

$$\zeta = \begin{cases} 1, & \text{ha az } A \text{ esemény bekövetkezik,} \\ 0, & \text{ha az } A \text{ esemény nem következik be.} \end{cases}$$

Valószínűségeloszlása, várható értéke és szórásnégyzete:

$$P(\zeta = 1) = P(A \text{ bekövetkezik}) = P(A),$$

$$P(\zeta = 0) = P(A \text{ nem következik be}) = 1 - P(A).$$

$$E(\zeta) = 1 \cdot P(\zeta = 1) + 0 \cdot P(\zeta = 0) = P(A),$$

$$E(\zeta^2) = 1^2 \cdot P(\zeta = 1) + 0^2 \cdot P(\zeta = 0) = P(A),$$

$$D^2(\zeta) = \text{Var}(\zeta) = E(\zeta^2) - (E(\zeta))^2 = P(A) - (P(A))^2 = P(A)[1 - P(A)].$$

Újra és újra megismételjük a kísérletet.

Független megfigyelések a ζ változóra: $\zeta_1, \zeta_2, \dots$

Az esemény bekövetkezési gyakorisága: $k_n(A) = \zeta_1 + \dots + \zeta_n$

A nagy számok Kolmogorov-féle törvényét alkalmazva:

$$\frac{k_n(A)}{n} = \frac{\zeta_1 + \dots + \zeta_n}{n} \rightarrow E(\zeta) = P(A)$$

Szerencsejáték: Egy profitorientált társaság szervezésében olyan játékot játszunk, melynél a nyeremény véletlen nagyságú.

- A játékban két fél vesz részt, a kaszinó és a játékos.
- c = a játék ára, előre meghatározott.
- ξ = a játékos nyereménye, véletlen nagyságú.

A kaszinó szempontjai:

- A kaszinó mindig sok játékosal áll kapcsolatban.
- Az a cél, hogy sok játék után termeljen profitot.

Kérdések:

- Milyen c ár mellett lesz a kaszinó nyereséges?
- Milyen ár lesz a racionálisan gondolkodó játékos számára elfogadható?
- Létezik-e **igazságos ár**, ami mindkét félnek megfelel?

Tegyük fel, hogy a játékos többször is játszik!

Nyeremények az egyes játékokban: $\xi_1, \xi_2, \dots$

Az egy játékra jutó átlagos nyeremény:

$$\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow E(\xi).$$

Három eset lehetséges:

- Ha $c < E(\xi)$, akkor a játék a játékosnak kedvez.
Az egy játékra jutó átlagos nettó nyeremény: $E(\xi) - c > 0$.
- Ha $c > E(\xi)$, akkor a játék a kaszinónak kedvez.
Az egy játékra jutó átlagos nettó bevétel: $c - E(\xi) > 0$.
- Ha $c = E(\xi)$, akkor várható értékben egyik fél sem nyer.
Ezt nevezzük a matematikában **igazságos árnak**.

Profitorientált társaságoknál az árazás első számú szabálya: $c > E(\xi)$.

Feladat. Egy kaszinó egy olyan játékot indít, melyben 5% eséllyel 10.000 dollárt lehet nyerni, míg a többi esetben 0 a nyeremény. Határozzuk meg a játék igazságos árát. Milyen áron fogja a kaszinó eladni ezt a játékot?

ξ = nyeremény nagysága, $R_\xi = \{0, 10.000\}$,

$P(\xi = 0) = 0,95$, $P(\xi = 10.000) = 0,05$,

$E(\xi) = 0 \cdot P(\xi = 0) + 10.000 \cdot P(\xi = 10.000) = 500$.

Igazságos ár = egy játékra jutó átlagos nyeremény = 500.

A kaszinó által kiszabott ár: $c > 500$.

Kérdés. Miért játszanak az emberek szerencsejátékot? Lehetséges okok:

- Rosszul mérik fel a valószínűségeket, nincsen tisztában vele, hogy a szabályok a játékos ellen dolgoznak.
- Reményt vásárolnak: a nyeremény megváltoztatná az életüket.
- Nem csak a nyereményért játszanak, hanem az élményért is.
- Szerencsejáték-függőség: az egyén nem racionális.

Biztosításmatematika:

- Szereplők: a biztosító és a biztosított.
- A biztosított befizet előre meghatározott c díjat.
- Ezért cserébe a biztosító megtéríti a biztosítottat érő károkat.
- Kérdés: mennyi az igazságos biztosítási díj?

Veszélyközösség: a hasonló kockázatnak kitett biztosítottak csoportja.

- ξ = a veszélyközösségre jellemző kárnagyság (valószínűségi változó).
- $\xi_1, \xi_2, \dots$ = a veszélyközösségen belül az egyes alanyok kárnagyságai, ezek független megfigyelések a ξ változóra.
- Átlagos kárérték a veszélyközösségen belül:

$$\frac{\xi_1 + \dots + \xi_n}{n} \rightarrow E(\xi).$$

- Törvényi előírás: a veszélyközösségen belül a biztosítottak befizetései fedezzék a veszélyközösséget érő károkat.

A biztosítás igazságos ára: $c = E(\xi)$

Profitorientált biztosítóknál az árazás első számú szabálya: $c > E(\xi)$.

Feladat. A munkámhoz szükségem van egy 10.000 dolláros berendezésre, ami az elkövetkező évben 5% eséllyel javíthatatlanul tönkremegy. Szeretnék biztosítást kötni erre az eseményre. Milyen biztosítási díj mellett fog egy biztosító szerződést kötni velem?

ξ = kárérték, $R_\xi = \{0, 10.000\}$,

$P(\xi = 0) = 0,95$, $P(\xi = 10.000) = 0,05$,

$E(\xi) = 0 \cdot P(\xi = 0) + 10.000 \cdot P(\xi = 10.000) = 500$.

Tehát a biztosító 500 dollárnál nagyobb összeget fog majd kérni.

Kérdés. Ez az ár megéri nekem, mint biztosítottnak?

- A káresemény átlagosan 20 évente átlagosan egyszer következik be. Ha minden évben félreteszek 500 dollárt, akkor 20 év alatt összejön 10.000 dollár. Miért fizetnék 500 dollárnál többet a biztosítónak?
- De mi történik akkor, ha a berendezés már az első évben tönkremegy?
- A biztosítottak döntéseit a rövid távú kockázatkerülés határozza meg.

Sztochasztikus hasznosságelmélet

A korábbiakban megvizsgáltuk, hogy mennyi a szerencsejátékok és a biztosítások „igazságos ára”:

- A kaszinó/biztosító sok ügyféllel áll kapcsolatban.
- Nem az a cél, hogy a kaszinó/biztosító minden ügyfélen keressenek.
- igazságos ár = átlagos kifizetés = kifizetés várható értéke

Kérdés: ez az ár elfogadható-e a játékosok/biztosítottak számára.

- A játékosok potenciálisan csak egy-két játékot játszanak a kaszinóban.
- A biztosítottakat a rövidtávú kockázatkerülés jellemzi.

Sztochasztikus hasznosságelmélet: A klasszikus hasznosságelmélet kiterjesztése véletlen nagyságú jószágkosarakra. A véletlen jószágkosarakat tipikusan az elért hasznosság várható értékével szoktuk jellemezni.

Feladat. Egy kaszinó egy olyan játékot indít, melyben 5% eséllyel 10.000 dollárt lehet nyerni, míg a többi esetben 0 a nyeremény. Tegyük fel, hogy a játékos az $u(x) = \sqrt{x}$ hasznossági függvényen keresztül értékeli az anyagi helyzetét. Mi a játékos számára elfogadható ár?

A játékosnak most két jószágkosár közül kell választania:

- 1. kosár: Nem játszik, tehát megtartja a játék c árát.
Elért hasznosság: $u(c) = \sqrt{c}$.
- 2. kosár: Játszik, tehát feláldozza a c összeget a ξ nagyságú nyereményért cserébe. Az elért hasznosság várható értéke:

$$\begin{aligned} E(u(\xi)) &= u(0) \cdot P(\xi = 0) + u(10.000) \cdot P(\xi = 10.000) \\ &= 0 \cdot 0,95 + 100 \cdot 0,05 = 5. \end{aligned}$$

A játékos számára a két kosár akkor közömbös, ha

$$u(c) = E(u(\xi)), \quad \text{tehát} \quad c = 25.$$

Lehetséges esetek:

- Ha $c < 25$, akkor $u(c) < 5 = E(u(\xi))$. Ekkor megéri játszani.
- Ha $c > 25$, akkor $u(c) > 5 = E(u(\xi))$. Nem éri meg játszani.

Feladat. A munkámhoz szükségem van egy 10.000 dolláros berendezésre, ami az elkövetkező évben 5% eséllyel javíthatatlanul tönkremegy. Az anyagi helyzetemet az $u(x) = \sqrt{x}$ hasznossági függvénnyel értékelem. Mi a számomra elfogadható biztosítási díj?

Legyen c a biztosítás díja! A kezdeti vagyonom: $10.000 + c$.
Két jószágkosár közül kell választanom:

- 1. kosár: Megkötöm a biztosítást. Év végén a vagyonom: 10.000 .
Az elért hasznosság: $u(10.000) = 100$.
- 2. kosár: Nem kötök biztosítást. Év végén a vagyonom:

$$\eta = \begin{cases} 10.000 + c, & \text{ha nem ér kár,} \\ c, & \text{ha kár ér.} \end{cases}$$

Az elért hasznosság várható értéke:

$$E(u(\eta)) = u(10.000 + c) \cdot 0,95 + u(c) \cdot 0,05 = 0,95\sqrt{10.000 + c} + 0,05\sqrt{c}.$$

Racionálisan gondolkodó egyénként azt az opciót fogom preferálni, amelyiknek nagyobb a hasznossága. A két kosár akkor közömbös, ha:

$$E(u(\eta)) = 100$$

$$0,95\sqrt{10.000 + c} + 0,05\sqrt{c} = 100$$

$$0,95^2(10.000 + c) = (100 - 0,05\sqrt{c})^2$$

$$0,9c + 10\sqrt{c} - 975 = 0$$

Ebből $d = \sqrt{c}$ helyettesítéssel: $0,9d^2 + 10d - 975 = 0$.

A másodfokú egyenlet megoldóképletével:

- $d = 27,8$, amiből következik: $c = d^2 = 773$;
- $d = -38,9$, nem ad értelmezhető megoldást.

A feladat megoldása:

- $c = 773$: a két kosár közömbös.
- $c < 773$: az 1. kosár hasznossága nagyobb, megéri biztosítást kötni.
- $c > 773$: a 2. kosár hasznossága nagyobb, nem kötünk biztosítást.