

1. Alapfeladatok

1. Egy 32 lapos magyar kártyapakliból a 4 ász és a 4 király van a kezünkben. Ebből választunk 2 lapot visszatevés nélkül. Jelölje X a kapott piros színű, Y pedig a kapott zöld színű lapok számát.

(a) Adjuk meg X és Y együttes eloszlását.

(b) Számoljuk ki X és Y várható értékét, szórásnégyzetét, valamint X és Y korrelációs együtthatóját.

Megoldás. Az együttes eloszlást az alábbi táblázattal adhatjuk meg:

X/Y	0	1	2	összesen
0	$\frac{6}{\binom{8}{2}}$	$\frac{8}{\binom{8}{2}}$	$\frac{1}{\binom{8}{2}}$	$\frac{15}{\binom{8}{2}}$
1	$\frac{8}{\binom{8}{2}}$	$\frac{4}{\binom{8}{2}}$	0	$\frac{12}{\binom{8}{2}}$
2	$\frac{1}{\binom{8}{2}}$	0	0	$\frac{1}{\binom{8}{2}}$
összesen	$\frac{15}{\binom{8}{2}}$	$\frac{12}{\binom{8}{2}}$	$\frac{1}{\binom{8}{2}}$	1

Az X hipergeometrikus eloszlású, $N = 8, M = 2, n = 2$ paraméterekkel, várható értéke és szórása ebből is kiszámítható. Nézzük meg definíció alapján is:

$$\mathbb{E}(X) = \sum_{k=0}^2 k \mathbb{P}(X = k) = \frac{12}{\binom{8}{2}} + 2 \cdot \frac{1}{\binom{8}{2}} = \frac{1}{2} = 0,5 = \frac{2 \cdot 2}{8} = n \frac{M}{N}.$$

Illetve

$$\mathbb{E}(X^2) = \sum_{k=0}^2 k^2 \mathbb{P}(X = k) = \frac{12}{\binom{8}{2}} + 2^2 \cdot \frac{1}{\binom{8}{2}} = \frac{16}{28} = \frac{4}{7}.$$

Ebből:

$$D^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{4}{7} - \frac{1}{4} = \frac{9}{28} = 0,312.$$

A zöld és piros színek szimmetriája miatt Y várható értéke és szórásnégyzete ugyanannyi.

Most számítsuk ki a korrelációs együtthatót.

$$\mathbb{E}(XY) = \sum_{k=0}^2 \sum_{l=0}^2 k l \mathbb{P}(X = k, Y = l) = 1 \cdot \frac{1}{\binom{8}{2}} = \frac{1}{28}.$$

Ebből a korrelációs együttható:

$$R(X, Y) = \frac{\text{cov}(X, Y)}{D(X)D(Y)} = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{D(X)D(Y)} = \frac{\frac{1}{28} - \frac{1}{4}}{\frac{9}{28}} = -\frac{6}{9} = -0,66.$$

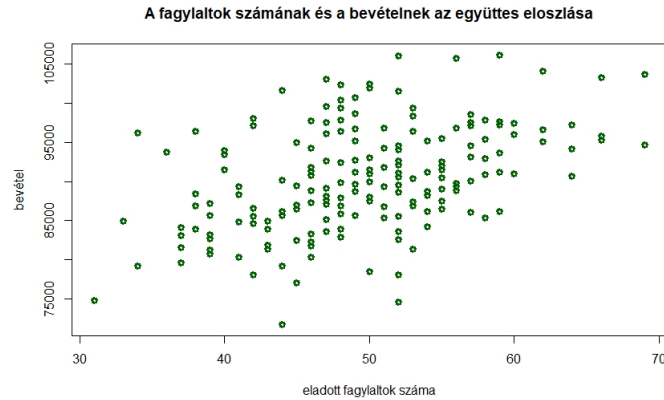
A korrelációs együttható közepes (gyenge közepes) lineáris jellegű összefüggést mutat, ami negatív: minél nagyobb X , annál kisebb Y (ami a feladat jellegéből is természetesen adódik).

2. Egy cukrászdában kétféle terméket árulnak. Tegyük fel, hogy a fagylaltot kérők száma (ez legyen X), Poisson-eloszlású 50 paraméterrel, a süteményt kérők száma Poisson-eloszlású 150 paraméterrel (ez legyen Y), és hogy X és Y függetlenek. A fagylalt ára 300 forint, a süteményé 500.

(a) Mennyi a napi bevétel várható értéke, illetve szórása?

(b) Számítsuk ki X -nek és napi bevételnek a korrelációs együtthatóját.

Megoldás.



1. ábra. Az X és $300X+500Y$ együttes eloszlása 200 kísérletből

- (a) Használjuk fel a várható érték linearitását, illetve azt, hogy a Poisson eloszlás paramétere a várható értékkel egyezik meg. $E(300 \cdot X + 500 \cdot Y) = 300 \cdot E(X) + 500 \cdot E(Y) = 300 \cdot 50 + 500 \cdot 150 = 90000$

Mivel X és Y függetlenek, és $D^2(X) = 50$, $D^2(Y) = 150$, ezért $D^2(300 \cdot X + 500 \cdot Y) = 300^2 D^2(X) + 500^2 D^2(Y) = 42000000 \Rightarrow D(300 \cdot X + 500 \cdot Y) = \sqrt{42000000} \approx 6480,74$

- (b) A kovariancia tulajdonságai, illetve X és Y függetlensége miatt $\text{cov}(X, 300 \cdot X + 500 \cdot Y) = 300 \cdot \text{cov}(X, X) + 500 \cdot \text{cov}(X, Y) = 300 \cdot D^2(X) + 0 = 300 \cdot 50$. Tehát a korrelációs együttható:

$$R(X, 300 \cdot X + 500 \cdot Y) = \frac{\text{cov}(X, 300 \cdot X + 500 \cdot Y)}{D(X) \cdot D(300 \cdot X + 500 \cdot Y)} = \frac{300 \cdot 50}{\sqrt{50} \cdot 6480,74} \approx 0,327$$

```
fagyi=rpois(200, lambda=50)
```

```
suti=rpois(200, lambda=150)
```

```
bevetel=300*fagyi+500*suti
```

```
eladas=fagyi+suti
```

```
plot(bevetel~fagyi, lwd="3", col="darkgreen", main="A fagyaltok számának és a bevételek az együttes eloszlása", xlab="eladott fagyaltok száma", ylab="bevétel")
```

3. Mihez és hogyan konvergál független szabályos kockadobások mértani közepe?

Megoldás. Emlékeztető: ha $X_1, X_2, \dots$ független, azonos eloszlású, véges várható értékű valószínűségi változók, akkor

$$\frac{X_1 + X_2 + \dots + X_n}{n} \rightarrow \mathbb{E}(X_1)$$

teljesül 1 valószínűséggel, ha $n \rightarrow \infty$. Ez a **nagy számok Kolmogorov-féle erős törvénye**.

A mértani közép, ha $X_1, \dots, X_n, \dots$ a dobások:

$$\sqrt[n]{X_1 X_2 \dots X_n}$$

Legyen X_i : i -edik kockadobás és $Y_n = \sqrt[n]{X_1 \dots X_n}$. Ekkor

$$Y_n = e^{\ln(Y_n)} = e^{\frac{1}{n}(\ln(X_1) + \dots + \ln(X_n))} \rightarrow e^{\mathbb{E}(\ln(X_1))}$$

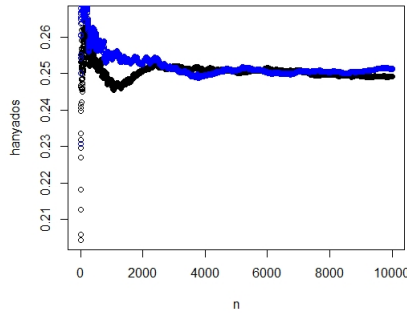
1 valószínűséggel.

Itt $\ln X_1, \ln X_2, \dots$ független azonos eloszlású valószínűségi változók, a várható értékük véges, erre alkalmazhatjuk a nagy számok erős törvényét.

$$\mathbb{E}(\ln(X_1)) = \frac{1}{6}(\ln(1) + \dots + \ln(6)) = \frac{\ln 6!}{6}$$

Ezután azt használtuk, hogy az exponenciális függvény folytonos.

Mindebből $Y_n \rightarrow \sqrt[6]{6!}$.



2. ábra. A $(\xi_1 + \dots + \xi_n)/(\xi_1^2 + \dots + \xi_n^2)$ hányados n függvényében két független sorsolásból

4. ξ_j -k független 3 paraméterű Poisson eloszlású változók. Mihez és hogyan konvergál

$$\frac{\xi_1 + \dots + \xi_n}{\xi_1^2 + \dots + \xi_n^2}?$$

Megoldás.

$$\frac{\xi_1 + \dots + \xi_n}{\xi_1^2 + \dots + \xi_n^2} = \frac{(\xi_1 + \dots + \xi_n)/n}{(\xi_1^2 + \dots + \xi_n^2)/n} \rightarrow \frac{E\xi_1}{E\xi_1^2} = \frac{3}{3 + 3^2} = \frac{1}{4}$$

A számlálóban a nagy számok erős törvényét alkalmazva (teljesülnek a feltételek) a limesz 1 valószínűséggel 3.

A $\xi_1^2, \xi_2^2, \dots$ is függetlenek. Az is látható (mivel ξ és ξ^2 között most egy bijekció van), hogy azonos eloszlásúak. Véges várható értékűek, hiszen a Poisson-eloszlás szórása is véges. A nevezőben is alkalmazhatjuk a nagy számok erős törvényét. A limesz 1 valószínűségi értelemben

$$\mathbb{E}(\xi_1^2) = D^2(\xi_1) + \mathbb{E}(\xi_1)^2 = \lambda + \lambda^2 = 3 + 9 = 12.$$

Azt használjuk, hogy két 0 valószínűségű esemény uniója is 0 valószínűségű (a szitaformula felső becslést ad), ezért annak valószínűsége is 1, hogy a számláló és nevező is a megadott értékekhez konvergál.

5. Adjunk példát olyan (X_n) valószínűségi változókból álló sorozatra, mely

- (a) 1 valószínűséggel konvergens, de L^p -ben nem konvergens;
- (b) L^p -ben konvergens, de 1 valószínűséggel nem konvergens.

Megoldás. A valószínűségi mező mindkét esetben legyen a $[0, 1]$ intervallum a Borel-halmazokkal és rajta a Lebesgue-mértékkel.

(a) $X_n : [0, 1] \rightarrow \mathbb{R}$, $X_n(\omega) = 2^n \cdot \mathbb{I}_{\{\omega < 1/n\}}$.

(b) A valószínűségi változókat, melyek tehát $X_n : [0, 1] \rightarrow \mathbb{R}$ függvények, csoportokban definiáljuk, az egyes csoportok mérete $2, 4, 8, 16, \dots$, azaz a 2 -hatványok. A k . csoporthoz bontsuk fel a $[0, 1]$ intervallumot 2^k egyforma hosszú intervallumra (minden pontot pontosan egyszer fedve le), és a k . csoporton belül a j . valószínűségi változó legyen a j . kis intervallum indikátora (1 az intervallumon belül, 0 azon kívül).

Minden $\omega \in [0, 1]$ szám (elemi esemény) minden k -ra valamelyik kis intervallumba esik, így minden k -ra az adott 2^k nagyságú csoport tagjai közül egy esetben 1, a többi esetben 0 lesz az érték, vagyis $X_n(\omega)$. Ebből következik, hogy $X_n(\omega)$ végtelen sokszor veszi fel az 1 és a 0 értékeket is, tehát nem konvergens. Így a sorozat 0 valószínűséggel konvergens. Viszont a különbség L^p normában 0-hoz tart.

6. Tegyük fel, hogy $X_n \rightarrow X$ eloszlásban. Következik-e ebből, hogy $X_n - X \rightarrow 0$ eloszlásban? Itt X , illetve (X_n) ugyanazon a valószínűségi mezőn értelmezett valószínűségi változók sorozata.

Megoldás.

Nem. Két független kockadobás (Y , és Z) azonos eloszlású, de a különbségük nem konstans nulla. Legyen tehát $X_n = Y$ minden n -re, és $X = Z$, ez ellenpélda.

7. Mutassunk arra példát, hogy X_n, X azonos valószínűségi mezőn értelmezett valószínűségi változók és $X_n \rightarrow X$ eloszlásban, de nem sztochasztikusan.

Megoldás.

Ugyanaz a példa, mint az előző feladatban. A különbség konstans pozitív valószínűséggel nagyobb, mint 1.

2. Gondolkodtató feladatok

1. Legyenek $X_1, X_2, \dots$ független valószínűségi változók úgy, hogy minden n -re $\mathbb{P}(X_n = n^2 - 1) = 1/n^2$ és $\mathbb{P}(X_n = -1) = 1 - 1/n^2$. Bizonyítsuk be, hogy $\mathbb{E}(X_n) = 0$ minden n -re, és

$$\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n} = -1$$

teljesül 1 valószínűséggel.

Megoldás. A várható érték kiszámítása:

$$\mathbb{E}(X_n) = (n^2 - 1) \cdot \frac{1}{n^2} + (-1) \cdot \left(1 - \frac{1}{n^2}\right) = 0.$$

Másrészt $\sum_{n=1}^{\infty} \mathbb{P}(X_n = n^2 - 1) < \infty$, így a Borel–Cantelli-lemma alapján 1 valószínűséggel ezek közül az események közül csak véges sok teljesül. Tehát 1 valószínűséggel van egy n_0 (ami függ ω -tól), hogy minden $n \geq n_0$ -ra $X_n = -1$. Ha van ilyen n_0 , akkor világos, hogy az átlag -1 -hez tart. Tehát az átlag 1 valószínűséggel -1 -hez tart.

2. Adjunk példát olyan (X_n) valószínűségi változókból álló sorozatra, mely sztochasztikusan konvergens, de 1 valószínűséggel nem konvergens (ebből következik, hogy nem 1 valószínűséggel konvergens).

Megoldás. Ugyanaz, mint az 5/b résznél. Minden $0 < \varepsilon < 1$ -ra $\mathbb{P}(|X_n| > \varepsilon) \rightarrow 0$, hiszen ez a valószínűség éppen annak az intervallumnak a hossza, amelynek X_n az indikátora, ez pedig a k . csoportban $1/2^k$ volt, így 0-hoz tart.

Tehát a megadott sorozat sztochasztikusan tart a 0-hoz (és L^1 -ben is), de nem konvergál 1 valószínűséggel.

Másik példa: független, nullához tartó, de végtelen összegű valószínűségű események indikátorainak a sorozata (a Borel–Cantelli-lemma használható arra, hogy belássuk, hogy ezek közül 1 valószínűséggel végtelen sok bekövetkezik).

3. Bizonyítsuk be, hogy az (X_n) valószínűségi változók sorozata pontosan akkor konvergál sztochasztikusan az X valószínűségi változóhoz, ha minden (X_{n_k}) részsorozatnak van olyan részsorozata, mely 1 valószínűséggel konvergál X -hez.

Megoldás. Tegyük fel, hogy minden részsorozatnak van 1 valószínűséggel konvergens részsorozata. Tegyük fel, hogy nem igaz a sztochasztikus konvergencia, vagyis valamely $\varepsilon > 0$ -ra

$$\limsup_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = \delta > 0.$$

Ez azt jelenti, hogy van egy olyan részsorozat, ami mentén

$$\lim_{k \rightarrow \infty} \mathbb{P}(|X_{n_k} - X| > \varepsilon) = \delta > 0.$$

Ebből a részsorozatból válasszunk ki egy 1 valószínűséggel konvergens részsorozatot. Mivel az 1 valószínűségű konvergenciából következik a sztochasztikus konvergencia, ez a részsorozat sztochasztikusan is konvergál, ez viszont ellentmond ennek a legutóbbi feltételnek. Ellentmondásra jutottunk, vagyis a sorozat sztochasztikusan konvergál.

Tegyük fel, hogy $X_n \rightarrow X$ sztochasztikusan, és (X_{n_k}) egy részsorozat. Kellene: ennek egy 1 valószínűséggel konvergens részsorozata.

Tudjuk: minden $\varepsilon > 0$ -ra

$$\mathbb{P}(|X_{n_k} - X| > \varepsilon) \rightarrow 0.$$

Az (X_{n_k}) is sztochasztikusan konvergál. Ezért találhatunk olyan részsorozatot, hogy minden j -re

$$\mathbb{P}\left(|X_j^* - X| > \frac{1}{j}\right) \leq \frac{1}{2^j}.$$

Ezeknek az összege véges, a Borel–Cantelli-lemma szerint 0 valószínűséggel következik be végtelen sok. Ez elég, mert ha ebből csak véges sok következik be, akkor egy küszöbtől kezdve ez már nem igaz egyetlen j -re sem. Ha minden elég nagy j -re legfeljebb $1/j$ a különbség, az viszont már elég.

Vagy másképpen: elég nagy n -től kezdve már az unió valószínűsége is kisebb ε -nál.

A sztochasztikus konvergencia miatt minden $j \geq 1$ -re van olyan $N(j)$, hogy $n \geq N(j)$ esetén

$$\mathbb{P}\left(|X_n - X| > \frac{1}{j^2}\right) \leq \frac{1}{j^2},$$

ugyanis $\varepsilon = \frac{1}{j^2}$ esetén a bal oldalon álló valószínűség 0-hoz tart. Az (X_{n_k}) -ből válasszunk egy olyan részsorozatot, hogy a j . tagra $n_{k_j} \geq N(j)$ legyen, amiből következik, hogy

$$\mathbb{P}\left(|X_{n_{k_j}} - X| > \frac{1}{j^2}\right) \leq \frac{1}{j^2}.$$

Ezeknek az eseményeknek a valószínűségeinek véges az összege, így a Borel–Cantelli-lemma szerint közülük 1 valószínűséggel csak véges sok következik be. Ha viszont ezek közül csak véges sok következik be, akkor minden $\varepsilon > 0$ -ra csak véges sok olyan j lesz, amire $|X_{n_{k_j}} - X| > \varepsilon$, vagyis ebből következik, hogy $X_{n_{k_j}} \rightarrow X$. Ezzel megkaptuk az 1 valószínűségű konvergenciát.

3. Gyakorló feladatok

1. Egy osztályba 16 fiú és 20 lány jár. Tegyük fel, hogy minden tanítási napon egymástól függetlenül a fiúk 0,04, a lányok 0,05 valószínűséggel hiányoznak. Legyen X a jövő hétfőn hiányzó fiúk, Y pedig a jövő hétfőn hiányzó lányok száma.

(a) Számítsuk ki az összes jövő hétfői hiányzó, vagyis $X + Y$ várható értékét.

(b) Számítsuk ki X , Y és $X + Y$ szórását.

(c) Mennyi X és Y kovarianciája?

(d) Mennyi X és $X + Y$ kovarianciája?

(e) Mennyi X és $X + Y$ korrelációs együtthatója?

Megoldás.

(a) X és Y is binomiális eloszlású, így

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y) = 16 \cdot 0,04 + 20 \cdot 0,05 = 1,64.$$

Itt még nem is használtuk, hogy X és Y függetlenek.

(b) Binomiális eloszlásnál a szórásnégyzet $np(1-p)$, így a függetlenség miatt

$$D(X) = \sqrt{16 \cdot 0,04 \cdot 0,96} = 0,784; \quad D(Y) = \sqrt{20 \cdot 0,05 \cdot 0,95} = 0,975;$$

$$D(X + Y) = \sqrt{D^2(X) + D^2(Y)} = 1,251.$$

(c) Mivel X és Y függetlenek, kovarianciájuk 0.

(d) A kovariancia bilineáris tulajdonsága és X, Y függetlensége alapján

$$\text{cov}(X, X + Y) = \text{cov}(X, X) + \text{cov}(X, Y) = D^2(X) + 0 = D^2(X) = 0,615.$$

$$\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

(e) Az előzőek és a korrelációs együttható definíciója alapján

$$R(X, X + Y) = \frac{\text{cov}(X, X + Y)}{D(X)D(X + Y)} = \frac{0,614}{0,784 \cdot 1,251} = 0,626.$$

A korrelációs együttható mindig a $[-1, 1]$ intervallumba esik. Ez egy közepes erősségű, pozitív irányú (pozitív az előjel, minél nagyobb X , tipikusan annál nagyobb $X + Y$ is), lineáris összefüggésre utal.

2. Legyenek X és Y független Poisson-eloszlású valószínűségi változók, X várható értéke 4, Y várható értéke pedig 10.

(a) Mennyi $X + Y$ várható értéke?

(b) Mennyi $X + Y$ szórása?

(c) Mennyi X és $X + Y$ korrelációs együtthatója?

(d) Mennyi $2X + 3Y$ és $X - Y$ kovarianciája?

(e) Mennyi $2X - Y$ szórása?

Megoldás.

(a) A várható érték additivitás alapján $\mathbb{E}(X) + \mathbb{E}(Y) = 14$.

(b) Itt $X + Y$ is Poisson-eloszlású, paramétere a két paraméter összege, azaz 14. Ezért $D(X + Y) = \sqrt{14} = 3,74$.

(c) Az előző feladathoz hasonlóan

$$R(X, X + Y) = \frac{\text{cov}(X, X + Y)}{D(X)D(X + Y)} = \frac{D^2(X)}{D(X)D(X + Y)} = \frac{4}{2 \cdot \sqrt{14}} = 0,53.$$

(d) A kovariancia bilineáris tulajdonsága alapján és X, Y függetlensége miatt

$$\text{cov}(2X + 3Y, X - Y) = \text{cov}(2X, X) + \text{cov}(2X, -Y) + \text{cov}(3Y, X) + \text{cov}(3Y, -Y) = 2D^2(X) - 3D^2(Y) = -22.$$

(e)

$$D(2X - Y) = \sqrt{4D^2(X) + D^2(-Y)} = \sqrt{4D^2(X) + D^2(Y)} = \sqrt{26} = 5,1.$$