

Задача 1. Докажете, че събитията A и B са независими, ако индикаторите $1_A, 1_B$ са независими случайни величини.

Доказателство:

$$\mathbb{P}(A \cap B) = \mathbb{P}(1_{A \cap B} = 1) = \underbrace{\mathbb{P}(1_A 1_B = 1)}_{\substack{\text{само когато} \\ \text{и двата индикатора} \\ \text{са 1-ца}}} = \mathbb{P}(1_A = 1 \cap 1_B = 1) \stackrel{\substack{\text{незав.} \\ 1_A \perp 1_B}}{=} \mathbb{P}(1_A = 1) \mathbb{P}(1_B = 1)$$

$$= \mathbb{P}(1_A = 1) \mathbb{P}(1_B = 1) = \mathbb{P}(A) \mathbb{P}(B), \text{ което искахме да докажем.}$$

□

Задача 2. (независимост дискретни сл. вел.)

1. Кога наричаме две събития независими? Дефинирайте кога наричаме две дискретни случайни величини независими и кога некорелирани.
2. Нека хвърляме $n > 3$ пъти монета с вероятност за ези p и дефинираме събитията $A =$ "третото хвърляне е ези" и $B =$ "общо са се паднали 3 езита". При какви условия A и B са независими?

Решение:

2.1. Две събития A и B се наричат независими и бележим $A \perp B$, тогава и само тогава когато $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$ (Ако $\mathbb{P}(A) > 0 \Rightarrow \mathbb{P}(B|A) = \mathbb{P}(B)$, т.е. независимостта означава, че случването на събитието A не ни носи никаква информация за B).

Две дискретни случайни величини X, Y във вероятностно пространство V се наричат независими и бележим $X \perp Y$, тогава и само тогава когато

$$\mathbb{P}(X = x_j; Y = y_k) = \mathbb{P}(X = x_j \cap Y = y_k) \stackrel{\text{def.}}{=} \mathbb{P}(X = x_j) \mathbb{P}(Y = y_k), \forall j, k.$$

Целта на корелацията е да мери някаква степен на линейност между X и Y . Ако $DX < \infty$ и $DY < \infty$, тогава $\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{DX} \sqrt{DY}}$ се нарича коефициент на корелация между X и Y .

Това реално е нормираната ковариация на X и Y , където $\text{cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]$ се нарича ковариация. За да бъдат двете случайни величини X и Y некорелирани е необходимо и достатъчно

$$\begin{aligned} \text{cov}(X, Y) &= \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)] = \mathbb{E}(XY - X\mathbb{E}Y - Y\mathbb{E}X + \mathbb{E}X\mathbb{E}Y) \stackrel{\substack{\text{лин.} \\ \text{функц.}}}{=} \\ &= \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y = 0, \text{ т.е. } \mathbb{E}XY = \mathbb{E}X\mathbb{E}Y. \text{ За дискретния случай това е:} \end{aligned}$$

$$\sum_i \sum_j x_i x_j \mathbb{P}(X = x_i \cap Y = y_j) = \sum_i x_i \mathbb{P}(X = x_i) \sum_j y_j \mathbb{P}(Y = y_j)$$

2.2 Дефинираме си с успех падането на "ези" от хвърлянето на дадена монета.

$$(1) \quad \mathbb{P}(A) = p \text{ (интересуваме се само от третия успех, който е с вероятност } p)$$

B е конкретен изход от биномно разпределена случайна величина $\text{Bin}(n, p)$ за $k = 3$ успеха.

$$(2) \quad \mathbb{P}(B) = \mathbb{P}(\text{общо 3 успеха}) = \binom{n}{3} p^3 (1-p)^{n-3}$$

Нека $X_i = \{\text{на } i\text{-тата позиция се пада "ези"}\}$, $1 \leq i \leq n$. X_i са независими в съвкупност (бернулиеви експерименти).

$$\begin{aligned} (3) \quad \mathbb{P}(A \cap B) &= \mathbb{P}(X_3 \cap \{\text{на останалите } n-1 \text{ позиции разпределяме 2 успеха}\}) = \\ &= \mathbb{P}(X_3) \times \mathbb{P}(\{\text{на останалите } n-1 \text{ позиции разпределяме 2 успеха}\}) = \\ &= p \times \binom{n-1}{2} p^2 (1-p)^{n-3}. \end{aligned}$$

За да са независими събитията A и B е необходимо и достатъчно да е изпълнено условието: $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, това се случва когато:

$$\binom{n-1}{2} p^3 (1-p)^{n-3} = \binom{n}{3} p^4 (1-p)^{n-3}, \text{ т.е. когато } p = \frac{\binom{n-1}{2}}{\binom{n}{3}} = \frac{\frac{(n-1)!}{(n-3)!2!}}{\frac{n!}{(n-3)!3!}} = \frac{3}{n}.$$

Следователно A и B са независими когато $np = 3$ □

Задача 3. Нека X е непрекъсната случайна величина с функция на разпределение F , която е строго монотонно растяща върху реалната права. Покажете, че $Y = F(X) \in \mathcal{U}(0,1)$.

Коментар: Всъщност условията върху F могат да се облекчат, но идеята е, че ако можем да симулираме равномерно разпределение с компютър и знаем F^{-1} , то $F^{-1}(Y)$ ще ни е симулация за X .

Решение:

$$F_Y(y) \stackrel{\text{def.}}{=} \mathbb{P}(Y < y) = \mathbb{P}(F(X) < y) \stackrel{F \uparrow \text{мон.}}{=} \mathbb{P}(X < F^{-1}(y)) = F(F^{-1}(y)) = y.$$

Следователно $Y \in \mathcal{U}(0,1)$, тъй като $F_{\mathcal{U}}(t) = t = F_Y(t)$, $\forall t \in [0,1]$ (функциите на разпределение на равномерното и Y са равни). □

Задача 4. (независимост непрекъснати сл. вел.)

1. Дефинирайте кога наричаме две непрекъснати случайни величини независими и кога некорелирани.
2. Дефинирайте функция пораждаща моментите $M_X(t)$ на случайната величина X . Нека $X \sim \mathcal{N}(0,1)$. Пресметнете $M_X(t)$. На колко са равни $\mathbb{E}X$, $\mathbb{E}X^2$, $\mathbb{E}X^3$?
3. Нека $X \sim \mathcal{N}(0,1)$. Потърсете случайна величина, която е полином на X и е некорелирана, но не е независима с X .

Решение:

4.1. Нека $X = (X_1, X_2)$. Тогава $X_1 \perp\!\!\!\perp X_2$, когато $F_X(x) = F_{X_1}(x_1)F_{X_2}(x_2)$ или

$\mathbb{P}(X_1 < x_1, X_2 < x_2) = \mathbb{P}(X_1 < x_1)\mathbb{P}(X_2 < x_2)$ за всяко $x = (x_1, x_2) \in \mathbb{R}^2$. Ако X е вектор от НСВ, то независимостта е еквивалентна на $f_X(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$, $\forall x = (x_1, x_2) \in \mathbb{R}^2$.

Аналогично както при дискретния вариант имаме, че X и Y са некорелирани когато $\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y$:

$$\int_{D_X} \int_{D_Y} xy \, dF_Y(y) dF_X(x) = \int_{D_X} x \, dF_X(x) \int_{D_Y} y \, dF_Y(y), \text{ където } F_Z \text{ е функцията на}$$

разпределение на НСВ Z .

4.2. Нека X е случайна величина. Ако $\mathbb{E}e^{tX}$ съществува за $t \in (-\varepsilon, \varepsilon)$ и някое $\varepsilon > 0$, то $M_X(t) = \mathbb{E}e^{tX}$ за $t \in (-\varepsilon, \varepsilon)$ се нарича функция на моментите.

Нека $X \sim \mathcal{N}(0,1)$. Тогава,

$$f_X(x) = f_{\mathcal{N}(0,1)} = f_{\mathcal{N}(\mu,\sigma^2)}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \Big|_{\substack{\mu=0 \\ \sigma^2=1}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx.$$

$$\begin{aligned} M_X(t) &= \mathbb{E}e^{tX} = \int_{-\infty}^{\infty} e^{tx} f_X(x) \, dx = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2 - 2tx + t^2) + \frac{t^2}{2}} \, dx = \\ &= e^{\frac{t^2}{2}} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-t)^2}{2}} \, dx}_{f_{\mathcal{N}(t,1)}(x)} = e^{\frac{t^2}{2}}. \end{aligned}$$

$$\mathbb{E}X^k = \frac{\partial}{\partial t^k} M_X(t) \Big|_{t=0}, \text{ за } k \geq 1. \text{ Следователно:}$$

$$\mathbb{E}X = \frac{\partial}{\partial t} M_X(t) \Big|_{t=0} = \frac{\partial}{\partial t} e^{\frac{t^2}{2}} \Big|_{t=0} = \frac{2t}{2} e^{\frac{t^2}{2}} \Big|_{t=0} = 0$$

$$\mathbb{E}e^{tX} = \mathbb{E}e^{t\mathcal{N}(0,1)} = \int_0^1 e^{tx} \, dx = \frac{1}{t}(e^t - 1) = f(t), \quad \lim_{n \rightarrow \infty} \frac{e^t - 1}{t} \stackrel{\text{Лопитал}}{=} \lim_{t \rightarrow 0} \frac{e^t}{1} = 1$$

$$\mathbb{E}X^2 = \frac{\partial}{\partial t^2} M_X(t) \Big|_{t=0} = \frac{\partial}{\partial t} (te^{\frac{t^2}{2}}) \Big|_{t=0} = e^{\frac{t^2}{2}} + t \times \frac{2t}{2} \times e^{\frac{t^2}{2}} \Big|_{t=0} = e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}} \Big|_{t=0} = 1$$

$$\mathbb{E}X^3 = \frac{\partial}{\partial t^3} M_X(t) \Big|_{t=0} = \frac{\partial}{\partial t} (e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}}) \Big|_{t=0} = te^{\frac{t^2}{2}} + 2te^{\frac{t^2}{2}} + t^3 e^{\frac{t^2}{2}} \Big|_{t=0} = 3te^{\frac{t^2}{2}} + t^3 e^{\frac{t^2}{2}} \Big|_{t=0} = 0$$

4.3.

От 4.2. Доказахме, че за $X \in \mathcal{N}(0,1) \rightarrow \mathbb{E}X = 0$.

Търсим полином $Y = g(X)$, за който е изпълнено $\text{cov}(X, Y) = 0$ и $Y \not\perp X$.

Тъй като $Y = g(X)$ е функция на X , то ще следва че $Y \not\perp X$. За да бъдат некорелирани е

$$\text{необходимо } 0 = \text{cov}(X, Y) = \mathbb{E}XY - \underbrace{\mathbb{E}X}_{=0} \mathbb{E}Y = \mathbb{E}XY = \mathbb{E}Xg(X) = \int_{-\infty}^{\infty} xg(x)f_X(x) \, dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} xg(x)e^{-\frac{x^2}{2}} \, dx. \text{ Последното ще се нулира когато функцията под интеграла е}$$

нечетна, което е еквивалентно на това $xg(x)$ да е нечетна или $g(x)$ да е четна. Т.е.

$Y = X^2, Y = X^4, \dots, Y = \cos(X), Y = e^{x^2} + e^{-x^2}$ са все валидни примери. □

Задача 5. Нека X е случайна величина с плътност $3(1-x)^2$ за $x \in (0,1)$. Намерете първите два цели момента и изчислете функцията на моментите.

Решение:

Нека първо проверим дали плътността X е добре дефинирана:

$$1 \stackrel{?}{=} \int_{D_X} f_X(x) dx = \int_0^1 3(1-x)^2 dx = 3 \int_0^1 1 - 2x + x^2 dx = 3 \left(x - \frac{2x}{2} + \frac{x^3}{3} \right) \Big|_0^1 = \\ = 3 \left(1 - 1 + \frac{1}{3} \right) = 1.$$

$$M_X(t) = \mathbb{E}e^{tX} = \int_0^1 e^{tx} 3(1-x)^2 dx = 3 \underbrace{\int_0^1 e^{tx} dx}_{I_1} - 6 \underbrace{\int_0^1 x e^{tx} dx}_{I_2} + 3 \underbrace{\int_0^1 x^2 e^{tx} dx}_{I_3}$$

$$I_1 := \int_0^1 e^{tx} dx = \frac{1}{t} \int_0^1 e^{tx} dtx = \frac{1}{t} e^{tx} \Big|_0^1 = \frac{e^t}{t} - \frac{1}{t}$$

$$I_2 := \int_0^1 x e^{tx} dx = \frac{1}{t} \int_0^1 x d e^{tx} \stackrel{\text{и.ч.}}{=} \frac{1}{t} \left(x e^{tx} \Big|_0^1 - \underbrace{\int_0^1 e^{tx} dx}_{I_1} \right) = \frac{x e^{tx}}{t} \Big|_0^1 - \frac{1}{t} \left(\frac{e^t}{t} - \frac{1}{t} \right) =$$

$$= \frac{e^t}{t} - 0 - \frac{e^t - 1}{t^2} = \frac{e^t}{t} - \frac{e^t}{t^2} + \frac{1}{t^2}$$

$$I_3 := \int_0^1 x^2 e^{tx} dx = \frac{1}{t} \int_0^1 x^2 d e^{tx} \stackrel{\text{и.ч.}}{=} \frac{1}{t} x^2 e^{tx} \Big|_0^1 - \frac{1}{t} \int_0^1 e^{tx} d x^2 = \frac{e^t}{t} - \frac{2}{t} \underbrace{\int_0^1 x e^{tx} dx}_{I_2} =$$

$$= \frac{e^t}{t} - \frac{2}{t} \left(\frac{e^t}{t} - \frac{e^t}{t^2} + \frac{1}{t^2} \right) = \frac{e^t}{t} - \frac{2e^t}{t^2} + \frac{2e^t - 2}{t^3}. \text{ Следователно,}$$

$$M_X(t) = 3I_1 - 6I_2 + I_3 = \cancel{\frac{3e^t}{t}} - \frac{3}{t} - \cancel{\frac{6e^t}{t}} + \cancel{\frac{6e^t}{t^2}} - \frac{6}{t^2} + \cancel{\frac{3e^t}{t}} - \cancel{\frac{6e^t}{t^2}} + \frac{6e^t}{t^3} - \frac{6}{t^3} = \\ = -\frac{3}{t} - \frac{6}{t^2} + \frac{6e^t}{t^3} - \frac{6}{t^3}.$$

$$\mathbb{E}X = 3 \int_0^1 x(1-x)^2 dx = 3 \int_0^1 x - 2x^2 + x^3 dx = 3 \left(\frac{x^2}{2} - \frac{2x^3}{3} + \frac{x^4}{4} \right) \Big|_0^1 = \\ = 3 \left(\frac{1}{2} - \frac{2}{3} + \frac{1}{4} \right) = 3 \left(\frac{6 - 8 + 3}{12} \right) = \frac{1}{4}$$

$$\begin{aligned} \mathbb{E}X^2 &= 3 \int_0^1 x^2(1-x)^2 dx = 3 \int_0^1 x^2 - 2x^3 + x^4 dx = 3 \left(\frac{x^3}{3} - \frac{2x^4}{4} + \frac{x^5}{5} \right) \Big|_0^1 = \\ &= 3 \left(\frac{1}{3} - \frac{1}{2} + \frac{1}{5} \right) = \frac{1}{10} (10 - 15 + 6) = \frac{1}{10}. \end{aligned}$$

□

Задача 6. Нека X, Y и Z са случайни величини със стойности в \mathbb{N} и $f : \mathbb{N} \rightarrow \mathbb{N}$. Кога наричаме X и Y еднакво разпределени? Да предположим, че последното е изпълнено. Вярно ли е, че $f(X)$ и $f(Y)$ са еднакво разпределени? А $X + Z$ и $Y + Z$? Докажете или дайте контрапримери. Вярно ли е, че ако X и Z са независими, то стига $\mathbb{E}(f(X)) < \infty$ и $\mathbb{E}(g(Z)) < \infty$, където $g : \mathbb{N} \rightarrow \mathbb{N}$, то $\mathbb{E}(f(X)g(Z)) = \mathbb{E}(f(X))\mathbb{E}(g(Z))$.

Решение:

X и Y са еднакво разпределени $\Leftrightarrow F_X = F_Y$. Ако X, Y са НСВ, то $\Rightarrow f_X = f_Y$. Вярно ли е че $f(X)$ и $f(Y)$ са еднакво разпределени, ако $X \stackrel{d}{=} Y$.

Трябва да покажем, че за всяко $n \in \mathbb{N}$ е вярно, че $\mathbb{P}(f(X) = n) = \mathbb{P}(f(Y) = n)$. Но $\{f(X) = n\} = \bigcup_{k \in \mathbb{N}: f(k)=n} \{X = k\}$ или събитието $\{f(X) = n\}$ е обединението на независимите събития $\{X = k\}$ за тези k , за които $f(k) = n$. Пример, ако $f(l) = l^2$, то $\{f(X) = n\}$ е празното множество, ако n не е квадрат и $\{f(X) = n\} = \{x = \sqrt{n}\}$ иначе. Тогава

$$\begin{aligned} \mathbb{P}(f(X) = n) &= \mathbb{P}\left(\bigcup_{k \in \mathbb{N}: f(k)=n} \{X = k\}\right) = \sum_{k \in \mathbb{N}: f(k)=n} \mathbb{P}(X = k) = \sum_{k \in \mathbb{N}: f(k)=n} \mathbb{P}(Y = k) = \\ &= \mathbb{P}\left(\bigcup_{k \in \mathbb{N}: f(k)=n} \{Y = k\}\right) = \mathbb{P}(f(Y) = n). \end{aligned}$$

$X + Z$ и $Y + Z$?

$X \perp\!\!\!\perp Y$ и X и Y са разпределени така:

$$\mathbb{P}(X = 1) = \mathbb{P}(X = 2) = \mathbb{P}(Y = 1) = \mathbb{P}(Y = 2) = \frac{1}{2}. \text{ Нека}$$

$$Z = -X \Rightarrow \mathbb{P}(Z = -1) = \mathbb{P}(Z = -2) = \frac{1}{2}. \text{ Тогава}$$

$$\mathbb{P}(X + Z = n) = \mathbb{P}(n = 0) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

$$\mathbb{P}(Y + Z = n) = \mathbb{P}(Y - X = n). \text{ Нека } n = 1 \Rightarrow \mathbb{P}(X + Z = 1) = 0$$

$$\mathbb{P}(Y + Z = 1) = \mathbb{P}(Y - X = 1) = \mathbb{P}(Y = 2; X = 1) \stackrel{X \perp\!\!\!\perp Y}{=} \mathbb{P}(Y = 2)\mathbb{P}(X = 1) = \frac{1}{4}$$

\Rightarrow при $n = 1 : \mathbb{P}(X + Z = n) \neq \mathbb{P}(Y + Z = n) \Rightarrow X + Z$ и $Y + Z$ не са еднакво разпределени.

Вярно ли е, че ако X и Z са независими, то стига $\mathbb{E}(f(X)) < \infty$ и $\mathbb{E}(g(Z)) < \infty$, където $g : \mathbb{N} \rightarrow \mathbb{N}$, то $\mathbb{E}(f(X)g(Z)) = \mathbb{E}(f(X))\mathbb{E}(g(Z))$. Стига да докажем, че $f(X)$ и $g(Z)$ са независими.

$X \perp\!\!\!\perp Z \Leftrightarrow \mathbb{P}(X = x_j, Z = Z_k) = \mathbb{P}(X = x_j)\mathbb{P}(Z = z_k), \forall j, k$. Трябва да покажем, че за всяко $m, n \in \mathbb{N}$ е вярно, че $\mathbb{P}(f(X) = m, g(Z) = n) = \mathbb{P}(f(X) = m)\mathbb{P}(g(Z) = n)$.

$$\begin{aligned}
\mathbb{P}(f(X) = m, g(Z) = n) &= \mathbb{P}\left(\bigcup_{k \in \mathbb{N}: f(k)=m} \{X = k\}; \bigcup_{l \in \mathbb{N}: g(l)=n} \{Z = l\}\right) = \\
&= \sum_{k \in \mathbb{N}: f(k)=m} \sum_{l \in \mathbb{N}: g(l)=n} \mathbb{P}(X = k; Z = l) = \sum_{k \in \mathbb{N}: f(k)=m} \sum_{l \in \mathbb{N}: g(l)=n} \mathbb{P}(X = k) \mathbb{P}(Z = l) = \\
&= \sum_{k \in \mathbb{N}: f(k)=m} \mathbb{P}(X = k) \sum_{l \in \mathbb{N}: g(l)=n} \mathbb{P}(Z = l) = \mathbb{P}\left(\bigcup_{k \in \mathbb{N}: f(k)=m} \{X = k\}\right) \mathbb{P}\left(\bigcup_{l \in \mathbb{N}: g(l)=n} \{Z = l\}\right) = \\
&= \mathbb{P}(f(X) = m) \mathbb{P}(g(Z) = n)
\end{aligned}$$

□

Задача 7. (контрапример ЗГЧ)

1. Разполагаме със зар с 2 червени и 4 черни страни и със зар с 4 червени и 2 черни страни. Вероятността да се падне, която и да е от страните е $1/6$.

Избираме с вероятност $1/2$ един от двата зара и го хвърляме безкраен брой пъти. Да дефинираме за $n \geq 1$

$$X_n = \begin{cases} 1, & \text{ако на } n\text{-тото хвърляне се е паднала черна страна,} \\ 0, & \text{иначе.} \end{cases}$$

Докажете, че дефинираните по-горе случайни величини са еднаков разпределени и пресметнете очакването им. Независими ли са?

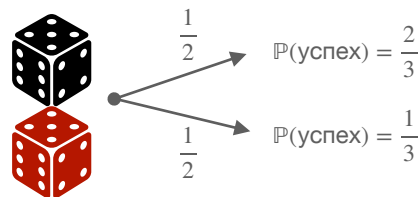
2. Формулирайте слабия ЗГЧ. Докажете, че той е в сила/не е в сила за редицата $(X_n)_n$.

Решение:

Дефинираме събитието "пада се черна страна при хвърляне" с успех. Нека Y е случайната величина {избираме един от двата зара}

Номираме заровете с числата 1 и 2 (1 е с повече черни страни)

$$\mathbb{P}(Y = 1) = \frac{1}{2}, \mathbb{P}(Y = 2) = \frac{1}{2}.$$



$$\frac{\sum_{i=1}^n X_i}{N} \xrightarrow[n \rightarrow \infty]{\text{п.с.}} \underbrace{V}_{\text{сл.вел.}}, \text{ където } V = \begin{cases} \frac{2}{3}, & \text{когато } Y = 1 \\ \frac{1}{3}, & \text{когато } Y = 2 \end{cases}, \text{ но това не е ЗГЧ, т.к. } V \text{ е случайна}$$

величина, а не константа.

$$X_i = \frac{2}{3} \times 1_{\{Y=1\}} + \frac{1}{3} \times 1_{\{Y=2\}}$$

$$\mathbb{E}V = \frac{2}{3}\mathbb{E}1_{\{Y=1\}} + \frac{1}{3}\mathbb{E}1_{\{Y=2\}} = \frac{2}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{1}{2}. \text{ Казваме, че за } X \text{ е изпълнен ЗГЧ}$$

$$(\text{слаб}), \text{ ако } \frac{\sum_{i=1}^n (X_i - \mathbb{E}X_i)}{n} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 0 (*)$$

$X = (X_i)_{i=1}^\infty$, наричаме редица от независими в съвкупност и еднакво разпределение сл.

вел., ако $X_i \stackrel{d}{=} X_1, \forall i \geq 1$ и са независими. Нека $X = (X_i)_{i=1}^\infty$ от нез. едн. раз. сл. вел.

Нека в доп. $\mathbb{E}|X_1| = \mu < \infty$, тогава е изпълнено (*). □

Задача 8. Докажете, че вероятността броят на шестниците при хвърляне на стандартен зар 900 пъти да е между 120 и 180 е поне 31/36.

Доказателство:

Неравенство на Чебишев: Нека A е множеството от стойности, за които $|X - \mathbb{E}X| > a$.

Тогава $A = \{|X - \mathbb{E}X| > a\} = \{(X - \mathbb{E}X)^2 > a^2\}$

Искаме да ограничим вероятността за случването на събитието A отгоре.

$$\begin{aligned} DX &\stackrel{\text{def.}}{=} \mathbb{E}(X - \mathbb{E}X)^2 \times 1 = \mathbb{E}(X - \mathbb{E}X)^2 \times 1_{\{A\}} + \underbrace{\mathbb{E}(X - \mathbb{E}X)^2 \times 1_{\{\bar{A}\}}}_{\geq 0} \geq \\ &\geq \mathbb{E}(X - \mathbb{E}X)^2 \times 1_A \geq a^2 \mathbb{E}1_{\{A\}} = a^2 \mathbb{P}(A). \text{ Следователно } \mathbb{P}(A) \leq \frac{DX}{a^2}. \end{aligned}$$

$$\text{Т.е. } \mathbb{P}(|X - \mathbb{E}X| > a) = \mathbb{P}(\mathbb{E}X - a < X < \mathbb{E}X + a) \leq \frac{DX}{a^2}.$$

Сега обратно към задачата.

Нека $X_i = \{\text{пада се } 6\text{-ца при хвърляне на } i\text{-ТИЯ зар}\}, i = \overline{1, n}$ и $S_n = \sum_{i=1}^n X_i$. Очевидно

$X_i \in \text{Ber}\left(p = \frac{1}{6}\right)$ и $S_{900} \in \text{Bin}\left(n = 900, p = \frac{1}{6}\right)$, тъй като броеви успехите в

бернулиево разпределени случайни величини. Тъй като бернулиевите експерименти са независими и еднакво разпределени със средно $\mu = \mathbb{E}X_1 = p = \frac{1}{6}$ и

$$\sigma^2 = \mathbb{D}X_1 = p(1-p) = \frac{1}{6} \times \frac{5}{6} = \frac{5}{36}, \text{ то от ЦГТ може да направим следното приближение}$$

(считаме, че $n = 900$ е достатъчно голямо):

$$\frac{\sum_{i=1}^{900} X_i - n\mu}{\sqrt{n\sigma^2}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0,1), \text{ т.е. } \frac{S_{900} - 900 \times \frac{1}{6}}{\sqrt{900 \times \frac{5}{36}}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0,1)$$

$$\text{Следователно, } \mathbb{P}(120 \leq S_{900} \leq 180) = \mathbb{P}(120 - 150 \leq S_{900} - 150 \leq 180 - 150) =$$

$$= \mathbb{P}\left(-\frac{30 \times 6}{30 \times \sqrt{5}} \leq \mathcal{N}(0,1) \leq \frac{30 \times 6}{30 \times \sqrt{5}}\right) = \mathbb{P}\left(|Z - \underbrace{\mathbb{E}Z}_{=0}| \leq \frac{6}{\sqrt{5}}\right) =$$

$$= \mathbb{P} \left(|Z| \leq \frac{6}{\sqrt{5}} \right) = 1 - \mathbb{P} \left(|Z| > \frac{6}{\sqrt{5}} \right) \stackrel{\substack{\text{неравенство} \\ \text{на Чебишов}}}{\geq} 1 - \frac{DZ}{\left(\frac{6}{\sqrt{5}} \right)^2} = 1 - \frac{5}{36} \times 1 = \frac{31}{36}.$$

□

Задача 9.

Хвърляте монета 1000 пъти и получавате 800 езита. Това ви усъмнява, че монетата е честна. Нека θ е вероятността за ези.

- Пресметнете каква е вероятността да наблюдавате 800 езита при допускане, че монетата е честна;
- Използвайте ЦГТ, за да конструирате доверителен интервал с ниво на доверие за точковата оценка на θ . Най-вероятно няма да можете да използвате понятието централна статистика, но се опитайте чрез увеличаване на доверителния интервал, което е резултат от оценка на дисперсията (зависеща от θ)
- (**) Ако приемете, че вероятността за честна монета е 0.99 и с вероятност 0.01 е точковата оценка, която получавате от тези 1000 хвърляния, т.е. 4/5. Как бихте преизчислили вероятността за честност при настъпването на тези данни?

Решение:

Според лектора може би задачата е грешна, но все пак каква е идеята:

$Ber(\theta)$

$Bin(1000, \theta)$

$$T = \frac{S_n - n\theta}{\sqrt{n\theta(1-\theta)}} \in \mathcal{N}(0,1)$$

$$\mathbb{P}(T > a) \stackrel{\theta=\frac{1}{2}}{=} \mathbb{P} \left(\frac{S_n - n \times \frac{1}{2}}{\sqrt{n \times \frac{1}{4}}} > \frac{a - n \times \frac{1}{2}}{\sqrt{n \times \frac{1}{4}}} \right) \stackrel[n=1000]{a=800}{=} \mathbb{P} \left(\mathcal{N}(0,1) > \frac{800 - 1000 \times \frac{1}{2}}{\sqrt{1000 \times \frac{1}{4}}} \right) =$$

$$\mathbb{P} \left(\mathcal{N}(0,1) > \frac{300 \times 2}{100} \right) = \mathbb{P} (\mathcal{N}(0,1) > 6) = 1 - \mathbb{P} (\mathcal{N}(0,1) \leq 6) \approx 1 - 1 = 0.$$

$$T = \frac{S_n - n\theta}{\sqrt{n\theta(1-\theta)}} \in \mathcal{N}(0,1)$$

$$\mathbb{P}(q_1 \leq T \leq q_2) = \mathbb{P} \left(q_1 \sqrt{n\theta(1-\theta)} + n\theta \leq S_n \leq q_2 \sqrt{n\theta(1-\theta)} + n\theta \right)$$

$$\gamma = \mathbb{P} \left(\frac{S_n}{n} - q_1 \sqrt{\frac{\theta(1-\theta)}{n}} \geq \theta \geq \frac{S_n}{n} - q_2 \sqrt{\frac{\theta(1-\theta)}{n}} \right)$$

Имаме, че $\left(\theta - \frac{1}{2}\right)^2 \geq 0 \Rightarrow \theta(1 - \theta) \leq \frac{1}{4}$

$$1 - \gamma = \mathbb{P}\left(\theta < \frac{S_n}{n} - q_2 \sqrt{\frac{\theta(1 - \theta)}{n}}\right) + \mathbb{P}\left(\theta > \frac{S_n}{n} - q_1 \sqrt{\frac{\theta(1 - \theta)}{n}}\right) \geq$$

$$\stackrel{\theta(1 - \theta) \leq \frac{1}{4}}{\geq} \mathbb{P}\left(\theta < \frac{S_n}{n} - q_2 \frac{\sqrt{\frac{1}{4}}}{\sqrt{n}}\right)$$

$$\mathbb{P}(T > a) \stackrel{\theta = \frac{1}{2}}{=} \mathbb{P}\left(\frac{S_n - n \times \frac{1}{2}}{n \sqrt{n \times \frac{1}{4}}} > a\right) = \mathbb{P}\left(\frac{S_n - \frac{n}{2}}{\frac{\sqrt{n}}{2}} > a\right) = \mathbb{P}\left(\frac{2S_n - n}{\sqrt{n}} > a\right)$$

Задача 10. Нека X е случайна величина с разпределение

$f_X(x; \theta) = C(\theta)e^{-\theta x^2}$, $x > 0$, $\theta > 0$. Намерете максимално правдоподобна оценка за θ от n наблюдения. Можете да използвате, че $C(\theta) = K\theta^{1/2}$, където K не зависи от θ . Вярно ли е, че оценката е състоятелна?

Решение:

$$f_X(x; \theta) = C(\theta)e^{-\theta x^2}, x > 0, \theta > 0$$

$$L_X(x, \theta) = C^n(\theta) \prod_{j=1}^n e^{-\theta X_j^2} = K^n \theta^{\frac{n}{2}} e^{-\theta \sum_{j=1}^n X_j^2}$$

$$\ln L_X(x; \theta) = n \ln K + \frac{n}{2} \ln \theta - \theta \sum_{j=1}^n X_j^2$$

$$\frac{\partial}{\partial \theta} \ln L_X(x; \theta) = \frac{n}{2\theta} - \sum_{j=1}^n X_j^2 = 0 \Rightarrow \hat{\theta} = \frac{n}{2 \sum_{j=1}^n X_j^2}$$

Състоятелна ли е оценката? Ако искаме да докажем, че е състоятелна, трябва проверим следното (*):

$$\hat{\theta} = \frac{n}{2 \sum_{j=1}^n X_j^2}, \text{ нека н.е.р. } Y_j = X_j^2. \text{ Тогава } \hat{\theta}(\vec{X}) = \frac{1}{2} \frac{n}{\sum_{j=1}^n X_j^2} = \frac{1}{2} \frac{n}{\sum_{j=1}^n Y_j}.$$

$$\frac{\sum_{j=1}^n Y_j}{n} \xrightarrow[n \rightarrow \infty]{\text{п.с.}} \mathbb{E} Y_1 = \mathbb{E} X_1^2.$$

$$\hat{\theta}(\vec{X}) \xrightarrow[n \rightarrow \infty]{\text{п.с.}} \frac{1}{2 \mathbb{E} X_1^2} \stackrel{?}{=} \theta \text{ (ако е състоятелна)}$$

$$\mathbb{E}X_1^2 = K\theta^{\frac{1}{2}} \int_0^\infty x^2 e^{-\theta x^2} dx = \frac{K\theta^{\frac{1}{2}}}{\theta^{\frac{3}{2}}} \int_0^\infty \theta x^2 e^{-\theta x^2} d\sqrt{\theta}x =$$

$$\stackrel{\sqrt{\theta}x=y}{=} \frac{K}{\theta} \int_0^\infty \underbrace{y^2 e^{-y^2}}_{\text{нечетна}} dy \stackrel{\text{нечетност}}{=} \frac{K}{2\theta} \frac{\sqrt{2\pi}^{\frac{1}{2}}}{\sqrt{2\pi}^{\frac{1}{2}}} \int_{-\infty}^\infty y^2 e^{-y^2} dy = \frac{K\sqrt{\pi}}{2\theta} \times \frac{1}{2}, \text{ т.е. оценката}$$

е състоятелна (схожда се добре към функция на тита, от която може да изразим тита с обратната функция).

11. Нека $X \in \mathcal{N}(\mu, 4)$. Постройте 90 % доверителен интервал за μ , ако наблюденията над X са: 1,3,4,4.

Решение:

Знаем, че $X \in \mathcal{N}(\mu, \sigma^2)$, като σ^2 е известно ($\sigma = 2$), т.е. $\mu = \theta$.

$$\text{Имаме, че } \hat{\mu} = \bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j = \frac{1+3+4+4}{4} = 3.$$

$\underbrace{\hspace{10em}}_{\in \mathcal{N}(\mu, \frac{\sigma^2}{n})}$

Пояснения защо $\frac{1}{n} \sum_{j=1}^n X_j \in \mathcal{N}(\mu, \frac{\sigma^2}{n})$:

$$\sum_{j=1}^n X_j \stackrel{\substack{\text{линейност} \\ \text{на } \mathcal{N} \\ \text{и незав.}}}{\in} \mathcal{N}(n\mu, n\sigma^2) \Rightarrow \frac{1}{n} \sum_{j=1}^n X_j \in \mathcal{N}\left(\frac{n\mu}{n}, \frac{n\sigma^2}{n^2}\right) = \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

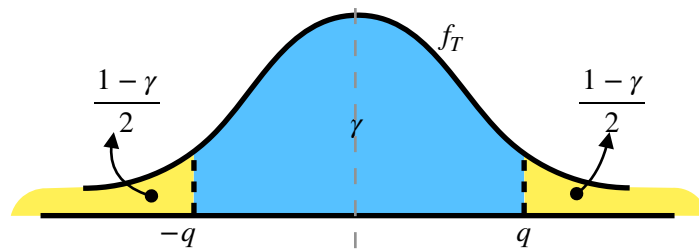
Използвахме свойството $\mathcal{N}(\mu_1, \sigma_1^2) + \mathcal{N}(\mu_2, \sigma_2^2) = \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ за независими нормално разпределени сл. вел. (какви то в случая са нашите опити \vec{X} , тъй като ще са прототипи на X - от там ги взимаме).

$$\text{Тогава } T(\vec{X}, \mu) = \underbrace{\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}}_{\text{нормираме}} \in \mathcal{N}(0,1), \sigma \text{ е известно число. Следователно}$$

получаваме, че T е намаляваща функция по μ и $\mathbb{P}(T \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{y^2}{2}} dy$, т.е.

не зависи от $\mu \stackrel{\text{def.}}{\Rightarrow} T$ е централна статистика за μ (тя е монотонна и намаляваща по μ и нейното разпределение съвпада с $\mathcal{N}(0,1)$, т.е. не зависи от θ)

Тогава, $\gamma = \mathbb{P}(-q < T < q)$, тъй като $\mathcal{N}(0,1)$ е симетрично:



Това ни гарантира, че $(-q, q)$ ще е най-тесния интервал, тъй като в $(-\infty, -q)$ и (q, ∞) е сбита най-малко маса (навсякъде извън тези интервали е сбита повече маса)

$q = q_{\frac{1}{2} + \frac{\gamma}{2}}$, който квантил го има в таблицата за нормалното стандартно разпределение.

$$\gamma = \mathbb{P}(-q < T < q) = \mathbb{P}\left(-q_{\frac{1}{2} + \frac{\gamma}{2}} < \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} < q_{\frac{1}{2} + \frac{\gamma}{2}}\right) \Rightarrow$$

$$\Rightarrow \mathbb{P}\left(\mu \in \underbrace{\left(\bar{X}_n - \frac{\sigma}{\sqrt{n}} \times q_{\frac{1}{2} + \frac{\gamma}{2}}, \bar{X}_n + \frac{\sigma}{\sqrt{n}} \times q_{\frac{1}{2} + \frac{\gamma}{2}}\right)}_{I_1}\right) = \gamma \Rightarrow$$

\Rightarrow

(Даденото е $\gamma = 0.90$, $\sigma = 2$, $n = 4$)

$$I_1 = \bar{X}_n - \frac{\sigma}{\sqrt{n}} \times q_{\frac{1}{2} + \frac{\gamma}{2}} = 3 - \frac{2}{\sqrt{4}} \times q_{0.95} \approx 3 + 1.645 = 4.645$$

$$I_2 = \bar{X}_n + \frac{\sigma}{\sqrt{n}} \times q_{\frac{1}{2} + \frac{\gamma}{2}} = 3 + \frac{2}{\sqrt{4}} \times q_{0.95} \approx 3 - 1.645 = 1.355$$

□