

Zadanie 1 (Mateusz Wojaczek)

Niech X będzie zmienną losową o rozkładzie normalnym $N(\mu, \sigma^2)$. Wyznaczyć wszystkie momenty centralne zmiennej X , tj.

$$\mathbb{E}[(X - \mu)^k] \quad \text{dla } k \in \mathbb{N}.$$

Rozwiązanie

Niech

$$Y \sim N(0, 1).$$

Wtedy

$$X = \mu + \sigma Y.$$

Zatem

$$\mathbb{E}[(X - \mu)^k] = \mathbb{E}[(\sigma Y)^k] = \sigma^k \mathbb{E}[Y^k].$$

Obliczamy więc momenty zmiennej Y .

Obliczenie $\mathbb{E}(Y^k)$

Z definicji:

$$\mathbb{E}[Y^k] = \frac{1}{\sqrt{2\pi}} \int_0^\infty y^k e^{-\frac{y^2}{2}} dy.$$

Oznaczmy:

$$I_k = \int_0^\infty y^k e^{-\frac{y^2}{2}} dy.$$

Obliczamy całkę I_k przez całkowanie przez części:

$$\begin{aligned} I_k &= \int_0^\infty y^{k-1} \cdot y e^{-\frac{y^2}{2}} dy \\ &= \left[y^{k-1} \left(-e^{-\frac{y^2}{2}} \right) \right]_0^\infty + \int_0^\infty (k-1) y^{k-2} e^{-\frac{y^2}{2}} dy. \end{aligned}$$

Wyraz brzegowy znika, zatem otrzymujemy rekurencję:

$$I_k = (k-1) I_{k-2}.$$

Wartości początkowe

$$I_0 = \int_0^\infty e^{-\frac{y^2}{2}} dy = \sqrt{2\pi},$$

$$I_1 = 0,$$

ponieważ jest to wartość oczekiwana zmiennej $Y \sim N(0, 1)$.

Postać ogólna

Dla k parzystego:

$$I_{2k} = \prod_{i=1}^k (2i-1) \sqrt{2\pi}.$$

Dla k nieparzystego:

$$I_{2k+1} = 0.$$

Stąd:

$$\mathbb{E}[Y^k] = \begin{cases} 0, & \text{gdy } k \text{ jest nieparzyste,} \\ \prod_{i=1}^{k-1} (k-2i), & \text{gdy } k \text{ jest parzyste.} \end{cases}$$

Wniosek końcowy

$$\mathbb{E}[(X - \mu)^k] = \sigma^k \mathbb{E}[Y^k],$$

czyli:

$$\mathbb{E}[(X - \mu)^k] = \begin{cases} 0, & \text{gdy } k \text{ jest nieparzyste,} \\ \sigma^k \prod_{i=1}^{\frac{k}{2}-1} (k-2i), & \text{gdy } k \text{ jest parzyste.} \end{cases}$$

Odpowiedź

$$\mathbb{E}[(X - \mu)^k] = \begin{cases} 0, & \text{gdy } k \text{ jest nieparzyste,} \\ \sigma^k \prod_{i=1}^{\frac{k}{2}-1} (k-2i), & \text{gdy } k \text{ jest parzyste.} \end{cases}$$

Zadanie 2 (Hubert Jastrzębski)

Niech X, Y, Z będą niezależnymi zmiennymi losowymi ze standardowym rozkładem normalnym. Która z wartości $P((X + Y + Z)/3 < 10)$ i $P(X < 10)$ jest większa? Odpowiedź uzasadnij.

Rozwiązanie

$$1. P(X < 10) = \Phi(10)$$

$$X + Y + Z \sim N(0, 3) \text{ (lemat 1).}$$

$$2. P((X + Y + Z)/3 < 10) = P((X + Y + Z) < 30) = \Phi\left(\frac{30}{\sqrt{3}}\right) = \Phi(10\sqrt{3}) \text{ (lemat 2).}$$

$$\Phi(10\sqrt{3}) > \Phi(10), \text{ więc } P((X + Y + Z)/3 < 10) > P(X < 10).$$

Nie pamiętam czy to było na wykładzie (pewnie tak, ale **ktoś** go nie spisał, a mi się nie chce robić notatek), ale mogę przy okazji udowodnić dwie rzeczy z których korzystam, bo samo zadanie jest dość proste:

Lemat 1: Jeżeli $X \sim N(\mu_1, \sigma_1^2)$ i $Y \sim N(\mu_2, \sigma_2^2)$ są niezależne, to $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

Dowód:

Wiemy, że gęstości X i Y to:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}$$

Gęstość sumy dwóch niezależnych zmiennych losowych to spłot ich gęstości:

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx$$

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2} e^{-\left(\frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{(z-x-\mu_2)^2}{2\sigma_2^2}\right)} dx$$

$$\text{Możemy wykazać, że: } \frac{(x-A)^2}{a} + \frac{(x-B)^2}{b} = \frac{a+b}{ab} \left(x - \frac{Ab+Ba}{a+b}\right)^2 + \frac{(A-B)^2}{a+b}$$

$$\text{Podstawiając odpowiednio dostajemy: } \frac{(x-\mu_1)^2}{2\sigma_1^2} + \frac{(z-x-\mu_2)^2}{2\sigma_2^2} = \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2} (x - \tilde{\mu})^2 + \frac{(z-(\mu_1+\mu_2))^2}{\sigma_1^2+\sigma_2^2}$$

$$f_{X+Y}(z) = \underbrace{\frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{(z-(\mu_1+\mu_2))^2}{2(\sigma_1^2+\sigma_2^2)}}}_{\text{stała względem } x} \cdot \int_{-\infty}^{\infty} e^{-\frac{1}{2} \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\sigma_2^2} (x-\tilde{\mu})^2} dx$$

$$\text{Mamy jakiś taki wzór (oczywiście każdy go zna z analizy): } \int_{-\infty}^{\infty} e^{-a(x-b)^2} dx = \sqrt{\frac{\pi}{a}}$$

Korzystając z niego dostajemy:

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2} \frac{\sigma_1^2 + \sigma_2^2}{\sigma_1^2 \sigma_2^2} (x - \tilde{\mu})^2} dx = \sqrt{\frac{2\pi\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}}$$

Ostatecznie więc:

$$f_{X+Y}(z) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} e^{-\frac{(z - (\mu_1 + \mu_2))^2}{2(\sigma_1^2 + \sigma_2^2)}}$$

Lemat 2: Jeżeli $X \sim N(\mu, \sigma^2)$, to $P(X < x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

Dowód:

Wiemy, że:

$$\Phi(x) = \int_{-\infty}^x \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}}_{\text{gęstość } N(0,1)} dt \implies \Phi\left(\frac{x-\mu}{\sigma}\right) = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$P(X < x) = \int_{-\infty}^x \underbrace{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}}_{\text{gęstość } N(\mu, \sigma^2)} dt$$

Podstawmy w drugim całkowaniu $u = \frac{t-\mu}{\sigma} \implies t = \sigma u + \mu \implies dt = \sigma du$.

Granice całkowania: kiedy $t \rightarrow -\infty$, to $u \rightarrow -\infty$; kiedy $t = x$, to $u = \frac{x-\mu}{\sigma}$.

$$P(X < x) = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \sigma du = \int_{-\infty}^{\frac{x-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Dodatek na koniec:

Jeśli wiemy, że $(X + Y + Z)/3 \sim N(0, 1/3)$, to można łatwo zrozumieć czemu

$$P((X + Y + Z)/3 < 10) > P(X < 10)$$

Rozkład $N(0, 1/3)$ jest bardziej skupiony wokół średniej (mniejsza wariancja) niż rozkład $N(0, 1)$. Zatem prawdopodobieństwo, że zmienna losowa z rozkładu $N(0, 1/3)$ przyjmie wartość mniejszą niż 10 jest większa niż prawdopodobieństwo, że zmienna losowa z rozkładu $N(0, 1)$ przyjmie wartość mniejszą niż 10.

Odpowiedź

$$\boxed{P((X + Y + Z)/3 < 10) > P(X < 10)}$$

Zadanie 3 (Dominik)

Niech X będzie zmienną losową z rozkładu normalnego $N(0, \sigma^2)$, zaś Z zmienną przyjmującą wartości ze zbioru $\{-1, +1\}$ z rozkładem jednostajnym. Niech $Y = XZ$.

- (i) Pokaż, że zmienna Y ma rozkład $N(0, \sigma^2)$.
- (ii) Wyjaśnij, dlaczego zmienne X, Y nie są niezależne.
- (iii) Pokaż, że zmienne X, Y są nieskorelowane (czyli $Cov(X, Y) = 0$).

Rozwiązanie podpunktu (i)

Dystrybuanta zmiennej X to z własności rozkładu normalnego

$$F_X(x) = P(X \leq x) = P\left(\frac{X}{\sigma} \leq \frac{x}{\sigma}\right) = \Phi\left(\frac{x}{\sigma}\right).$$

$\frac{X}{\sigma}$ jest zmienną o standardowym rozkładzie normalnym, a Φ oznacza dystrybucję standardowego rozkładu normalnego.

Pokażmy, że XZ ma taką samą dystrybucję, a więc też taki sam rozkład, co X .

$$\begin{aligned} F_{XZ}(x) &= P(XZ \leq x) \\ &= P\left((Z = 1 \wedge X \leq x) \vee (Z = -1 \wedge -X \leq x)\right) \end{aligned}$$

Z rozłączności zdarzeń oraz niezależności X i Z :

$$\begin{aligned} F_{XZ}(x) &= P(Z = 1)P(X \leq x) + P(Z = -1)P(-X \leq x) \\ &= \frac{1}{2}P(X \leq x) + \frac{1}{2}P(-X \leq x) \end{aligned}$$

$P(X \leq x)$ przekształciliśmy już powyżej, zróbmy to samo z $P(-X \leq x)$. Ostatnią równość mamy z wykładu:

$$P(-X \leq x) = P\left(\frac{X}{\sigma} \geq \frac{-x}{\sigma}\right) = 1 - P\left(\frac{X}{\sigma} \leq \frac{-x}{\sigma}\right) = 1 - \Phi\left(\frac{-x}{\sigma}\right) = \Phi\left(\frac{x}{\sigma}\right)$$

Zatem:

$$F_{XZ}(x) = \frac{1}{2}\Phi\left(\frac{x}{\sigma}\right) + \frac{1}{2}\Phi\left(\frac{x}{\sigma}\right) = \Phi\left(\frac{x}{\sigma}\right).$$

X i $Y = XZ$ mają taką samą dystrybucję, więc też taki sam rozkład.

Rozwiązanie podpunktu (ii)

Założmy **nie wprost**, że X i Y są niezależne. W takim razie:

$$P(X \in [-1, 1] \wedge Y \in [-1, 1]) = P(X \in [-1, 1]) \cdot P(Y \in [-1, 1]).$$

Ale $X \in [-1, 1]$ jest tym samym zdarzeniem, co $Y \in [-1, 1]$, zatem

$$P(X \in [-1, 1] \wedge Y \in [-1, 1]) = P(X \in [-1, 1]) = P(Y \in [-1, 1]).$$

Z poprzedniego równania otrzymujemy oczywiście fałszywy wniosek, że $P(X \in [-1, 1]) = 1$, zatem założenie jest sprzeczne.

Rozwiązanie podpunktu (iii)

Rozpiszmy $Cov(X, Y)$ z definicji:

$$\begin{aligned} Cov(X, Y) &= E[XY] - \underbrace{E[X]}_{=0} \cdot E[Y] \\ &= E[XY] \\ &= P(Z = -1)E[XY | Z = -1] + P(Z = +1)E[XY | Z = +1] \end{aligned}$$

Rozpiszmy wiedząc, że $Y = ZX$ i korzystając z niezależności X i Z :

- $E[XY | Z = -1] = E[X(-X) | Z = -1] = E[-X^2] = -E[X^2]$
- $E[XY | Z = +1] = E[X \cdot X | Z = +1] = E[X^2]$

Zatem

$$Cov(X, Y) = -\frac{1}{2}E[X^2] + \frac{1}{2}E[X^2] = 0.$$

Zadanie 4 (Pavlo Tsikalysyn)

Załóżmy, że mamy do dyspozycji generator liczb z rozkładu jednostajnego z przedziału $[-1, 1]$. Naszym celem będzie wygenerowanie liczb z rozkładu normalnego $N(0, 1)$. Niech U, V będą wygenerowanymi niezależnie wartościami na $[-1, 1]$. Procedura wygląda następująco. Jeśli $U^2 + V^2 > 1$, to odrzucamy wynik i próbujemy jeszcze raz. W przeciwnym przypadku niech $S = U^2 + V^2$ i niech

$$X = U\sqrt{\frac{-2\ln S}{S}}, Y = V\sqrt{\frac{-2\ln S}{S}}.$$

Udowodnij, że X, Y są niezależnymi zmiennymi losowymi z rozkładu $N(0, 1)$. Pamiętaj o wstępnym warunkowaniu! Wskazówka: przeczytaj sekcję 9.5 w podręczniku.

Rozwiązanie

Warunek $U^2 + V^2 \leq 1$, oznacza że losujemy punktu z koła z równym prawdopodobieństwem. Wtedy niech $\frac{U}{\sqrt{S}} = \cos\alpha$, $\frac{V}{\sqrt{S}} = \sin\alpha$, z tego mamy:

$$X = \cos\alpha\sqrt{-2\ln S}, Y = \sin\alpha\sqrt{-2\ln S}$$

Z tego że losujemy punkty z koła, mamy że α jest niezależna od S oraz ma rozkład jednostajny w $[0, 2\pi]$.

$$P(S < s) = P(U^2 + V^2 < s) = \frac{\pi s}{\pi} = s$$

Z tego wynika że S ma rozkład jednostajny w $[0, 1]$.

Niech X', Y' - będą niezależnymi zmiennymi z rozkładem $N(0, 1)$, wtedy:

$$F(x', y') = \int_{-\infty}^{y'} \int_{-\infty}^{x'} \frac{1}{2\pi} e^{-(x^2+y^2)/2} dx dy$$

Zauważymy że wybieramy punkt na płaszczyźnie, to przejdziemy na współrzędne biegunowe z $R^2 = X'^2 + Y'^2$, a $\theta = \tan^{-1} \frac{Y'}{X'}$. Wtedy:

$$X' = R \cos \theta, Y' = R \sin \theta$$

. Za pomocą Jacobiana otrzymujemy:

$$F(r', \theta') = \int_0^{\theta'} \int_0^{r'} \frac{1}{2\pi} e^{-r'^2/2} r dr d\theta = \frac{\theta'}{2\pi} (1 - e^{-r'^2/2})$$

Zauważymy że $F(r', \theta') = G(r')H(\theta')$, gdzie $G(r') = 1 - e^{-r'^2/2}$, $H(\theta') = \frac{\theta'}{2\pi}$, to oznacza że R oraz θ są niezależne.

$$P(R \geq r') = e^{-r'^2/2}, P(S \geq s) = 1 - s$$

Podstawimy $\theta = \alpha$, $R = \sqrt{-2\ln S}$, wtedy:

$$X' = \cos \alpha \sqrt{-2\ln S}, Y' = \sin \alpha \sqrt{-2\ln S} \implies X' = X, Y' = Y$$

Zadanie 5 (Kacper Orszulak)

Niech X, Y będą niezależnymi zmiennymi o takich samych rozkładach i skończonej wartości oczekiwanej oraz wariancji. Niech $U = X + Y$ i $V = X - Y$.

- (i) Wykaż, że U, V są nieskorelowane.
- (ii) Podaj przykład X, Y takich, że U, V nie są niezależne.
- (iii) Udowodnij, że jeśli X, Y mają rozkład $N(0, 1)$, to U, V są niezależne.

Rozwiązanie

(i)

Z definicji zmienne U, V są skorelowane wtw. $Cov(U, V) = 0$. Bardziej użyteczny będzie natomiast równoważny warunek: $\mathbb{E}[UV] = \mathbb{E}[U]\mathbb{E}[V]$. Sprawdźmy go zatem:

$$\mathbb{E}[UV] = \mathbb{E}[(X + Y)(X - Y)] = \mathbb{E}[X^2 - Y^2] = \mathbb{E}[X^2] - \mathbb{E}[Y^2]$$

$$\mathbb{E}[U]\mathbb{E}[V] = \mathbb{E}[X + Y]\mathbb{E}[X - Y] = (\mathbb{E}[X] + \mathbb{E}[Y])(\mathbb{E}[X] - \mathbb{E}[Y]) = \mathbb{E}[X]^2 - \mathbb{E}[Y]^2$$

Ostatnie porównanie:

$$\begin{aligned} \mathbb{E}[X^2] - \mathbb{E}[Y^2] &= \mathbb{E}[X]^2 - \mathbb{E}[Y]^2 \\ \mathbb{E}[X^2] - \mathbb{E}[X]^2 &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \end{aligned}$$

Oczywiście jest to prawda, ponieważ obie zmienne mają ten sam rozkład. □

(ii)

Niech X, Y to niezależne zmienne losowe o tym samym rozkładzie t. że

$$\mathcal{P}(X = -1) = \mathcal{P}(X = +1) = \frac{1}{2}$$

Wtedy oczywiście

$$\begin{aligned}\mathcal{P}(X + Y = -2) &= \mathcal{P}(X - Y = -2) = \frac{1}{4} \\ \mathcal{P}(X + Y = -2 \wedge X - Y = -2) &= 0 \neq \frac{1}{4}\end{aligned}$$

Znaleźliśmy więc szukany kontrprzykład.

(iii)

Niech $X, Y \sim N(0, 1)$ będą niezależne. Dowodzimy, że U, V są niezależne.

Oczywiście z własności rozkładu normalnego $-Y \sim N(0, 1)$. Dla przypomnienia $\begin{smallmatrix} U:=X+Y \\ V:=X-Y \end{smallmatrix}$. Z twierdzenia z wykładu wiemy zatem, że $U, V \sim N(0, 2)$.

Szukamy równości: $\mathcal{P}(U \leq u_0 \wedge V \leq v_0) = \mathcal{P}(U \leq u_0)\mathcal{P}(V \leq v_0)$

Prawa strona równości, to po prostu:

$$\mathcal{P}(U \leq u_0 \wedge V \leq v_0) = \left(\int_{-\infty}^{u_0} f_U(u) du \right) \left(\int_{-\infty}^{v_0} f_V(v) dv \right).$$

Rozpisujemy lewą stronę równości. Aby to zrobić posłużymy się zamianą współrzędnych znaną z analizy. Przypomnijmy ogólną formułę:

$$\int_A f = |\Phi'| \int_{\Phi^{-1}(A)} (f \circ \Phi)$$

Jak to rozumieć? Pomysł jest taki, że chcemy policzyć całkę funkcji f na jakimś “brzydkim” zbiorze A . Okazuje się jednak, że znamy przekształcenie Φ (intuicyjnie wystarczy by to było coś podobnego do bijekcji - w szczególności bijekcja jest dobra), które zamienia pewien prostokąt (po którym oczywiście łatwo całkować) na tę “brzydką” przestrzeń. W tym zapisie $\Phi^{-1}(A)$ oznacza zatem przeciwobraz “brzydkiego” zbioru, a więc nasz docelowy prostokąt.

Niech

$$\Phi : \mathbb{R}^2 \ni (u, v) \rightarrow (x, y) = \left(\frac{u+v}{2}, \frac{u-v}{2} \right) \in \mathbb{R}^2.$$

Jacobian takiego przekształcenia liniowego to

$$\begin{aligned}J_\Phi &= \Phi' = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \\ \det \Phi' &= -\frac{1}{4} - \frac{1}{4} = -\frac{1}{2} \\ |\Phi'| &= |\det \Phi'| = \frac{1}{2}.\end{aligned}$$

Definiujemy A jako taki zbiór punktów (x, y) , które spełniają nierówność lewej strony:

$$A := \left\{ (x, y) : \begin{cases} x + y \leq u_0 \\ x - y \leq v_0 \end{cases} \right\}$$

Oczywiście takiemu zbiorowi odpowiada prostokąt

$$\Phi^{-1}(A) = \left\{ (u, v) : \begin{cases} u \leq u_0 \\ v \leq v_0 \end{cases} \right\}.$$

Możemy wreszcie przejść do właściwego rozpisywania lewej strony równości:

$$\begin{aligned} \mathcal{P}(U \leq u_0 \wedge V \leq v_0) &= \mathcal{P}((x, y) \in A) \\ &= \iint_A f_{X,Y}(x, y) dy dx \\ &= |\Phi'| \iint_{\Phi^{-1}(A)} (f_{X,Y} \circ \Phi)(u, v) dv du \\ &= \frac{1}{2} \int_{-\infty}^{u_0} \int_{-\infty}^{v_0} f_{X,Y} \left(\frac{u+v}{2}, \frac{u-v}{2} \right) dv du \end{aligned}$$

Skorzystamy teraz z niezależności X, Y oraz z następujących wzorów:

$$\begin{aligned} \phi(x) &= \frac{1}{\sqrt{2\pi}\sigma} \exp \left(-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right) \\ f_X(x) = f_Y(x) = \phi(x) &= \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x^2}{2} \right) \\ f_U(u) = f_V(u) = \phi \left(\frac{u}{\sqrt{2}} \right) &= \frac{1}{2\sqrt{\pi}} \exp \left(-\frac{u^2}{4} \right) \end{aligned}$$

Rozpisujemy:

$$\begin{aligned} f_{X,Y} \left(\frac{u+v}{2}, \frac{u-v}{2} \right) &= f_X \left(\frac{u+v}{2} \right) f_Y \left(\frac{u-v}{2} \right) \\ &= \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{u+v}{2} \right)^2 \right) \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{u-v}{2} \right)^2 \right) \\ &= \frac{1}{2\pi} \exp \left(-\frac{1}{2} \left(\left(\frac{u+v}{2} \right)^2 + \left(\frac{u-v}{2} \right)^2 \right) \right) \\ &= \frac{1}{2\pi} \exp \left(-\frac{u^2 + v^2}{4} \right) \\ &= 2 \frac{1}{2\sqrt{\pi}} \frac{1}{2\sqrt{\pi}} \exp \left(-\frac{u^2}{4} \right) \exp \left(-\frac{v^2}{4} \right) \\ &= 2 f_U(u) f_V(v) \end{aligned}$$

Możemy teraz dopiąć naszą równość:

$$\begin{aligned}\mathcal{P}(U \leq u_0 \wedge V \leq v_0) &= \dots = \frac{1}{2} \int_{-\infty}^{u_0} \int_{-\infty}^{v_0} f_{X,Y} \left(\frac{u+v}{2}, \frac{u-v}{2} \right) dv du \\ &= \frac{1}{2} \int_{-\infty}^{u_0} \int_{-\infty}^{v_0} 2f_U(u)f_V(v) dv du \\ &= \int_{-\infty}^{u_0} f_U(u) \left(\int_{-\infty}^{v_0} f_V(v) dv \right) du \\ &= \left(\int_{-\infty}^{u_0} f_U(u) du \right) \left(\int_{-\infty}^{v_0} f_V(v) dv \right) \\ &= \mathcal{P}(U \leq u_0 \wedge V \leq v_0)\end{aligned}$$

□

Zadanie 6 (Wiktoria K)

Niech X, Y będą niezależnymi zmiennymi z rozkładu $N(0, 1)$. Oblicz

$$E(X + Y | X > 0, Y > 0).$$

Rozwiązanie

Z liniowości wartości oczekiwanej i niezależności X, Y :

$$E(X + Y | X > 0, Y > 0) = E(X | X > 0, Y > 0) + E(Y | X > 0, Y > 0) = E(X | X > 0) + E(Y | Y > 0)$$

X i Y mają ten sam rozkład zatem:

$$E(X + Y | X > 0, Y > 0) = 2E(X | X > 0)$$

$$E(X | X > 0) = \frac{1}{P(X > 0)} \int_0^{\infty} x f(x) dx$$

Obliczymy teraz całkę

$$\begin{aligned}\int \frac{1}{\sqrt{2\pi}} x e^{-\frac{1}{2}x^2} dx &= -\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \\ \int_0^{\infty} \frac{1}{\sqrt{2\pi}} x e^{-\frac{1}{2}x^2} dx &= \frac{1}{\sqrt{2\pi}}\end{aligned}$$

Z symetrii rozkładu normalnego $P(X > 0) = \frac{1}{2}$

Podsumowując:

$$E(X + Y | X > 0, Y > 0) = 2 \cdot \frac{1}{\frac{1}{2}} \cdot \frac{1}{\sqrt{2\pi}} = 2\sqrt{\frac{2}{\pi}}$$

Odpowiedź

$$2\sqrt{\frac{2}{\pi}}$$

Zadanie 7 (Wiktor M)

Niech X, Y będą niezależnymi zmiennymi z rozkładu $N(0, 1)$. Wyznacz rozkład $X + Y$ pod warunkiem, że $X > 0$ i $Y > 0$. To znaczy, dla każdego z oblicz:

$$P(X + Y < z | X > 0, Y > 0).$$

Wskazówka: Wynik podaj w terminach standardowego rozkładu normalnego.

Rozwiązanie

$$P(X + Y < z | X > 0, Y > 0) = \frac{P(X + Y < z, X > 0, Y > 0)}{P(X > 0, Y > 0)}$$

Mianownik tego wyrażenia wynosi $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ (z niezależności X, Y).

Aby łatwiej obliczyć licznik, wprowadźmy zmienne $U = \frac{X+Y}{\sqrt{2}}, V = \frac{X-Y}{\sqrt{2}}$. Z Zadania 5 (podpunkt *iii*) wiemy, że $X + Y$ oraz $X - Y$ są niezależne, więc U oraz V są niezależne. Oprócz tego $X + Y$ ma rozkład $N(0, 2)$, a po 'przeskalowaniu' przez $\sqrt{2}$ dostajemy, że U, V mają rozkład standardowy normalny.

Obliczmy więc licznik, po drodze korzystając z faktu, że $\phi(u)$ to pochodna $\Phi(u)$, więc $\int \phi(u)\Phi(u)du = \frac{1}{2}\Phi(u)^2$, co z kolei implikuje $\int_0^t \phi(u)\Phi(u)du = \frac{1}{2}(\Phi(t)^2 - \Phi(0)^2) = \frac{1}{2}\Phi(t)^2 - \frac{1}{8}$.

$$\begin{aligned} P(X + Y < z, X > 0, Y > 0) &= P(0 < U < \frac{z}{\sqrt{2}}, -U < V < U) = \int_0^{\frac{z}{\sqrt{2}}} \phi(u) \int_{-u}^u \phi(v) dv du = \\ &= \int_0^{\frac{z}{\sqrt{2}}} \phi(u)(2\Phi(u) - 1) du = 2 \int_0^{\frac{z}{\sqrt{2}}} \phi(u)\Phi(u) du - \int_0^{\frac{z}{\sqrt{2}}} \phi(u) du = \Phi(\frac{z}{\sqrt{2}})^2 - \Phi(0)^2 - \Phi(\frac{z}{\sqrt{2}}) + \frac{1}{2} = \\ &= \Phi(\frac{z}{\sqrt{2}})^2 - \Phi(\frac{z}{\sqrt{2}}) + \frac{1}{4} = (\Phi(\frac{z}{\sqrt{2}}) - \frac{1}{2})^2 \end{aligned}$$

Łącząc otrzymane wyniki, otrzymujemy wynik $4(\Phi(\frac{z}{\sqrt{2}}) - \frac{1}{2})^2$

Odpowiedź

$$4(\Phi(\frac{z}{\sqrt{2}}) - \frac{1}{2})^2 \text{ dla } z > 0, \text{ wpp. } 0$$

Zadanie 8 (Krzysztof Peszko)

Rozważmy następujący eksperyment. Stojąc na przeciwko bardzo długiej ściany wprawiają latarkę w ruch obrotowy i obserwujemy losowy punkt na ścianie na którym zatrzyma się światło latarki. Dla ustalenia uwagi, powiedzmy, że stoimy w punkcie $(0, 1)$ a ścianą jest oś OX . Niech X będzie współrzędną w punkcie na osi OX w który zwrócone jest światło latarki. Jeśli światło nie jest zwrócone na ścianę, to możemy powtórzyć eksperyment do skutku.

(I) Udowodnij, że dystrybuanta i gęstość zmiennej X zadane są przez:

$$P(X < x) = \frac{1}{2} + \frac{1}{\pi} \arctg(x), f(x) = \frac{1}{\pi(1+x^2)}$$

(II) Jaka jest $E(X)$? A jaka jest $E(X|X > 0)$?

(III) Jeśli zmienna X jest zdefiniowana w analogiczny sposób jak powyżesz, ale zamiast punktu $(0, 1)$, bierzemy punkt (x_0, γ) to mówimy, że X ma rozkład Cauchy'ego z parametrem (x_0, γ) . Rozkład Cauchy'ego z parametrem $(0, 1)$ nazywamy standardowym rozkładem Cauchy'ego. Wyznacz dystrybuantę, gęstość, wartość oczekiwaną i wariancję dla rozkładu Cauchy'ego.

Rozwiązanie

(I) Podzielimy sobie problem, na dwa podproblemy. Jeśli $x > 0$ i $x < 0$. Jeśli $x > 0$. To z prawdopodobieństwem $\frac{1}{2}$, $X < 0$, Ciekawsza jest dla fragmentu między $0, ax$. Zdefiniujemy, trójkąt o wierzchołkach w $(0, 0)$, $(0, 1)$ i $(x, 0)$. Zdefiniujemy kąt α jako kąt w trójkącie przy wierzchołku $(0, 1)$. Bazując na podstawowej trygonometrii, wiemy, że $\tg(\alpha) = x/1$, czyli $\alpha = \arctg(x)$. Można również, zauważyć, że prawdopodobieństwo wejście między 0 , a x , jest równe $\frac{1}{2} * \frac{\alpha}{\frac{\pi}{2}} = \frac{\alpha}{\pi} = \frac{\arctg(x)}{\pi}$. Także suma, $P(X < x) = \frac{1}{2} + \frac{\arctg(x)}{\pi}$, czyli dokładnie to co potrzebowaliśmy.

Jeśli natomiast mamy $x < 0$, to robimy podobny patent, tylko $\arctg$ trzeba odjąć od $\frac{1}{2}$ oraz zmienić znak x . Po przeliczeniu jeszcze raz wychodzi: $\frac{1}{2} - \frac{\arctg(-x)}{\pi}$, z analizy możemy, wyciągnąć minusa i otrzymujemy $\frac{1}{2} + \frac{\arctg(x)}{\pi}$, co ostatecznie udowadnia dystrybuantę.

By pokazać gęstość, wystarczy zrobić pochodną dystrybuanty, czyli $\frac{d}{dx} \frac{1}{2} + \frac{\arctg(x)}{\pi} = \frac{1}{\pi \cdot (1+x^2)}$, co należało pokazać.

(II)

$$E(X) = \int_{-\infty}^{\infty} x * f(X) = \int_{-\infty}^{\infty} x * f(X) = \int_{-\infty}^{\infty} x * \frac{1}{\pi \cdot (1+x^2)} = [t = 1+x^2, dt = 2x] = \int_1^{\infty} \frac{1}{2\pi \cdot t} + \int_{-\infty}^1 \frac{1}{2\pi \cdot t} = \int_1^{\infty} \frac{1}{2\pi \cdot t} - \int_1^{\infty} \frac{1}{2\pi \cdot t} = 0$$

Haha, mały błąd $\infty - \infty$, jest znakiem niezdefiniowanym. Niestety nie jestem w stanie udowodnić, że nie da się tego zdefiniować, ale według zewnętrznych źródeł tak jest. Przykro mi za niedogodność.

$$E(X|X > 0) = \dots$$

$$P(X < x | X > 0) = \frac{2}{\pi} \operatorname{arctg}(x)$$

$$f(X | X > 0) = \frac{2}{\pi \cdot (1 + x^2)}$$

$$E(X | X > 0) = \int_0^\infty x \cdot f(X | X > 0) = \int_0^\infty x \cdot \frac{2}{\pi \cdot (1 + x^2)} = [t = x^2 + 1, dt = 2x dx] = \int_1^\infty \frac{1}{\pi \cdot t} = \frac{1}{\pi} \cdot \int_1^\infty \frac{1}{t} = \infty$$

(III) Dystrybuanta: Z prostych przesunięć geometrycznych (jak popatrzeć się na rysunku), to można zapisać jako:

$$P(X < x) = \frac{1}{2} + \frac{1}{\pi} \operatorname{arctg}\left(\frac{x - x_0}{\gamma}\right)$$

Gęstość: (pochodna dystrybuanty)

$$f(x) = \frac{1}{\pi \cdot (1 + (\frac{x - x_0}{\gamma})^2)}$$

$E(X)$, podobnie jak $Var(X)$, z przyczyn z (II) też są niezdefiniowane.

Zadanie 9 (Michał)

Mając do dyspozycji generator liczb losowych z rozkłady jednostajnego na odcinku $[0, 1]$ zaprojektuj generator liczb z rozkładu Cauchy'ego

Rozwiązanie

W tym zadaniu skorzystamy z metody odwrotnej dystrybuanty.

Standardowy rozkład Cauchy'ego ma funkcję gęstości prawdopodobieństwa:

$$f(x) = \frac{1}{\pi(1 + x^2)}$$

Znajdujemy dystrybuantę:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{1}{\pi(1 + t^2)} dt$$

$$F(x) = \frac{1}{\pi} \arctan x + \frac{1}{2}, \quad x \in \mathbb{R}$$

Jest to funkcja ściśle rosnąca, a zatem posiada funkcję odwrotną:

$$F^{-1}(u) = \tan\left(\pi\left(u - \frac{1}{2}\right)\right), \quad u \in [0, 1]$$

Niech U będzie zmienną losową o rozkładzie jednostajnym na przedziale $[0, 1]$. Definiujemy zmienną losową

$$X = F^{-1}(U) = \tan\left(\pi\left(U - \frac{1}{2}\right)\right).$$

Sprawdźmy czy genertor działa, rzeczywiście dla dowolnego $x \in \mathbb{R}$ zachodzi

$$\begin{aligned}\mathbb{P}(X \leq x) &= \mathbb{P}\left(\tan\left(\pi\left(U - \frac{1}{2}\right)\right) \leq x\right) \\ &= \mathbb{P}\left(U \leq \frac{1}{\pi} \arctan x + \frac{1}{2}\right) \\ &= \frac{1}{\pi} \arctan x + \frac{1}{2}\end{aligned}$$

co oznacza, że zmienna losowa X ma standardowy rozkład Cauchy'ego.

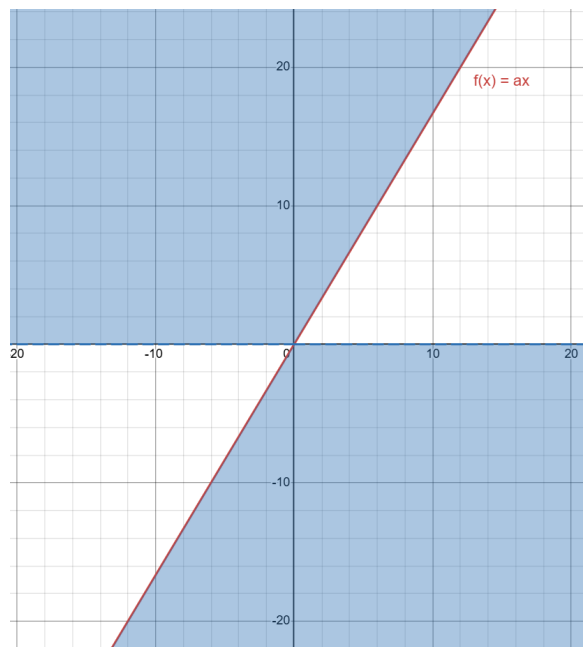
Odpowiedź

$$X = \tan\left(\pi\left(U - \frac{1}{2}\right)\right).$$

Zadanie 10 (Kamil)

Niech X, Y będą niezależnymi zmiennymi losowymi o rozkładzie $N(0, 1)$. Wyznacz rozkład zmiennej losowej $Z = \frac{X}{Y}$.

Rozwiązanie



Policzymy dystrybuantę, całka musi być liczona po obszarze, który wyznacza wyrażenie w prawdopodobieństwie (obszar zaznaczony na rysunku):

$$P\left(\frac{X}{Y} \leq a\right)$$

Nierówność

$$\frac{X}{Y} \leq a$$

jest równoważna

$$X \leq aY \quad \text{dla } Y > 0, \quad X \geq aY \quad \text{dla } Y < 0.$$

Ponieważ zmienne losowe X i Y są niezależne oraz mają rozkład normalny $N(0, 1)$, ich wspólna gęstość to iloczyn gęstości i ma postać:

$$f_{X,Y}(x, y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}.$$

Dystrybuanta zmiennej $\frac{X}{Y}$ wynosi więc

$$P\left(\frac{X}{Y} \leq a\right) = \int_0^\infty \int_{-\infty}^{ay} \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dx dy + \int_{-\infty}^0 \int_{ay}^\infty \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} dx dy.$$

Patrząc na rysunek można dojść do wniosku, że łatwiej przejść po obszarze za pomocą kąta i promienia, a nie współrzędnych kartezjańskich. Wprowadzamy więc współrzędne biegunowe:

$$x = r \cos \theta, \quad y = r \sin \theta,$$

Wówczas

$$x^2 + y^2 = r^2, \quad dx dy = r dr d\theta,$$

a gęstość wspólna przyjmuje postać

$$\frac{1}{2\pi} e^{-\frac{r^2}{2}} r dr d\theta.$$

$$P\left(\frac{X}{Y} \leq a\right) = \int_0^\infty \int_{\arctan \frac{1}{a}}^\pi \frac{1}{2\pi} e^{-\frac{r^2}{2}} r d\theta dr + \int_0^\infty \int_{\pi+\arctan \frac{1}{a}}^{2\pi} \frac{1}{2\pi} e^{-\frac{r^2}{2}} r d\theta dr.$$

Wyciągając stałe przed całki otrzymujemy

$$P\left(\frac{X}{Y} \leq a\right) = \frac{1}{2\pi} \left(\int_0^\infty e^{-\frac{r^2}{2}} r dr \right) \left(\int_{\arctan \frac{1}{a}}^\pi d\theta + \int_{\pi+\arctan \frac{1}{a}}^{2\pi} d\theta \right).$$

Najpierw liczymy całkę po zmiennej r :

$$\int_0^\infty e^{-\frac{r^2}{2}} r dr = \left[-e^{-\frac{r^2}{2}} \right]_0^\infty = 1.$$

Teraz liczymy całki po kącie:

$$\begin{aligned} \int_{\arctan \frac{1}{a}}^\pi d\theta &= \pi - \arctan \frac{1}{a}, \\ \int_{\pi+\arctan \frac{1}{a}}^{2\pi} d\theta &= 2\pi - \left(\pi + \arctan \frac{1}{a} \right) = \pi - \arctan \frac{1}{a}. \end{aligned}$$

Zatem

$$\begin{aligned} P\left(\frac{X}{Y} \leq a\right) &= \frac{1}{2\pi} \left(2\pi - 2 \arctan \frac{1}{a} \right) = 1 - \frac{1}{\pi} \arctan \frac{1}{a}. \\ P\left(\frac{X}{Y} \leq a\right) &= 1 - \frac{1}{\pi} \arctan \frac{1}{a} = 1 - \frac{1}{\pi} \left(\frac{\pi}{2} - \arctan a \right) = \frac{1}{2} + \frac{1}{\pi} \arctan a. \end{aligned}$$

Jest to dystrybuanta rozkładu Cauchy'ego

Odpowiedź

$$\frac{X}{Y} \sim \text{Cauchy}(0, 1)$$

Zadanie 11 (Kuba Sułkowski)

Niech (X, Y) będzie punktem wylosowanym jednostajnie ze zbioru $\{(x, y) \mid x^2 + y^2 \leq 1, x \geq 0\}$ (czyli jednostkowego półkola). Jaki rozkład ma zmienna Y/X ?

Rozwiązanie

Poprawmy treść zadania i zamieńmy warunek $x \geq 0$ na $x > 0$, aby uniknąć dzielenia przez zero ($x = 0$ i tak jest zdarzeniem miary zero).

Zauważmy, że wylosowanie punktu jednostajnie z półkola jest równoważne z wylosowaniem jednostajnie kąta i (już niejednostajnie) odległości od $(0, 0)$.

Nazwijmy nasze półkole $S = \{(x, y) \mid x^2 + y^2 \leq 1, x > 0\}$

$$F(z) = \int_S [Y/X \leq z] ds = \int_{-\pi/2}^{\pi/2} \int_0^1 2r [Y/X \leq z] dr d\alpha$$

Zdarzenie $[Y/X \leq z]$ jest niezależne od r

$$\begin{aligned} F(z) &= \int_{-\pi/2}^{\pi/2} [Y/X \leq z] \int_0^1 \frac{2r}{\pi} dr d\alpha = \int_{-\pi/2}^{\pi/2} [Y/X \leq z] \frac{1}{\pi} d\alpha = \int_{-\pi/2}^{\pi/2} [\tan \alpha \leq z] \frac{1}{\pi} d\alpha \\ &= \int_{-\pi/2}^{\tan^{-1} z} \frac{d\alpha}{\pi} = \frac{1}{\pi} \left(\tan^{-1}(z) + \frac{\pi}{2} \right) \end{aligned}$$

Jest to dystrybucja rozkładu Cauchy'ego

Odpowiedź

$$\frac{Y}{X} \sim \text{Cauchy}(0, 1)$$

Zadanie 12 (Olek Wieczorek)

Czas obsługi i -tego klienta dany jest zmienną losową o wartości oczekiwanej wynoszącej 2 minuty oraz wariancji równej 1 minucie. Zmienne opisujące każdego z klientów są niezależne. Niech Y oznacza sumaryczny czas obsługi 50 klientów. Wyznacz $P(90 \leq Y \leq 110)$.

Rozwiązanie

W tym zadaniu trzeba skorzystać z CTG. Ogranicza się to do tego, że po prostu musimy tak zmodyfikować daną nierówność, żeby wyszedł nam wzór z CTG. Wtedy aplikujemy twierdzenie i mamy wynik.

Policzmy potrzebne rzeczy:

Y_i – czas obsługi i -tego klienta

$$Y = \sum_{i=1}^{50} Y_i \quad \bar{Y} = \frac{Y}{50}$$

$$E(\bar{Y}) = \frac{1}{50} E(Y) = \frac{1}{50} \cdot \sum_{i=1}^{50} E(Y_i) = \frac{1}{50} \cdot 50 \cdot 2 = 2$$

$$Var(\bar{Y}) = \frac{1}{50^2} \cdot \sum_{i=1}^{50} Var(Y_i) = \frac{1}{50^2} \cdot 50 \cdot 1 = \frac{1}{50}$$

$$\frac{1}{\sqrt{Var(\bar{Y})}} = \sqrt{50} = 5\sqrt{2}$$

I teraz wystarczy odpowiednio przekształcić nierówność z treści:

$$\begin{aligned} P(90 \leq Y \leq 110) &= P\left(\frac{9}{5} \leq \bar{Y} \leq \frac{11}{5}\right) = P\left(-\frac{1}{5} \leq \bar{Y} - E(\bar{Y}) \leq \frac{1}{5}\right) \\ &= P\left(-\sqrt{2} \leq \frac{\bar{Y} - E(\bar{Y})}{\sqrt{Var(\bar{Y})}} \leq \sqrt{2}\right) = \Phi(\sqrt{2}) - \Phi(-\sqrt{2}) = 2 \cdot \Phi(\sqrt{2}) - 1 \end{aligned}$$

Przedostatnia równość to CTG, a ostatnia wynika z symetrii rozkładu normalnego ($\Phi(-x) = 1 - \Phi(x)$).

Teraz możemy jedynie odczytać dystrybuantę z tabelki.

$$2 \cdot \Phi(\sqrt{2}) - 1 \approx 2 \cdot \Phi(1.41) - 1 \approx 0.84$$

Odpowiedź

$$\Phi(\sqrt{2}) - \Phi(-\sqrt{2}) \approx 0.84$$

Zadanie 13 (OSKIBOSKI123)

Niech Y będzie zmienną z rozkładu dwumianowego z parametrami $n = 20$ oraz $p = \frac{1}{2}$. Wyznacz $P(8 \leq Y \leq 10)$ dokładnie. Następnie rozpisz Y jako sumę indykatorów i wyznacz to samo prawdopodobieństwo za pomocą Centralnego Twierdzenia Granicznego. Różnica pomiędzy tymi dwoma wynikami wynika z faktu, iż przybliżamy rozkład dyskretny z małym n za pomocą rozkładu ciągłego. Sytuację można poprawić zauważając, że dla Y zachodzi $P(8 \leq Y \leq 10) = P(7.5 \leq Y \leq 10.5)$. Wyznacz szukane prawdopodobieństwo po zastosowaniu tej poprawki oraz porównaj otrzymane wyniki numerycznie.

Rozwiązanie

Niech $Y \sim \text{Bin}(n = 20, p = \frac{1}{2})$. Wyznaczamy $P(8 \leq Y \leq 10)$.

1. Obliczenie dokładne

$$P(Y = k) = \binom{20}{k} \left(\frac{1}{2}\right)^{20}, \quad k = 0, \dots, 20$$

$$P(8 \leq Y \leq 10) = \sum_{k=8}^{10} \binom{20}{k} \cdot \frac{1}{2^{20}}$$

$$\binom{20}{8} = 125970, \quad \binom{20}{9} = 167960, \quad \binom{20}{10} = 184756$$

$$P_{\text{exact}} = \frac{125970 + 167960 + 184756}{2^{20}} = \frac{478686}{1048576} \approx 0.4565$$

2. Przybliżenie CTG bez poprawki ciągłości

Niech $X_1, \dots, X_{20}$ i.i.d., $X_i \sim \text{Bernoulli}(1/2)$. Wtedy:

$$\mathbb{E}[X_i] = \mu = \frac{1}{2}, \quad \text{Var}(X_i) = \sigma^2 = \frac{1}{4}, \quad \sigma = \frac{1}{2}.$$

$$\bar{X}_{20} = \frac{1}{20} \sum_{i=1}^{20} X_i = \frac{Y}{20}$$

Z Twierdzenia 15.3.1:

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq b\right) = \Phi(b) - \Phi(a)$$

Dla $n = 20$ stosujemy przybliżenie.

Zdarzenie $\{8 \leq Y \leq 10\}$ równoważne jest $\{0.4 \leq \bar{X}_{20} \leq 0.5\}$.

Podstawiamy do wzoru CTG:

$$a = \frac{0.4 - \mu}{\sigma/\sqrt{20}} = \frac{0.4 - 0.5}{0.5/\sqrt{20}}$$

$$b = \frac{0.5 - \mu}{\sigma/\sqrt{20}} = \frac{0.5 - 0.5}{0.5/\sqrt{20}}$$

Obliczamy liczbowo:

$$\sigma/\sqrt{20} = 0.5/\sqrt{20} \approx 0.5/4.472136 = 0.1118034$$

$$a = \frac{-0.1}{0.1118034} = -0.8944272 \approx -0.8944$$

$$b = 0$$

Zatem:

$$\begin{aligned} P(8 \leq Y \leq 10) &\approx P(-0.8944 \leq Z \leq 0), \quad Z \sim N(0, 1) \\ &= \Phi(0) - \Phi(-0.8944) \approx 0.5 - 0.1856 = 0.3144 \end{aligned}$$

3. CTG z poprawką ciągłości

Poprawka: $P(8 \leq Y \leq 10) = P(7.5 \leq Y \leq 10.5)$, czyli:

$$\left\{ \frac{7.5}{20} \leq \bar{X}_{20} \leq \frac{10.5}{20} \right\} = \{0.375 \leq \bar{X}_{20} \leq 0.525\}$$

Podstawiamy do wzoru CTG:

$$a' = \frac{0.375 - 0.5}{0.5/\sqrt{20}} = \frac{-0.125}{0.1118034} = -1.118034 \approx -1.1180$$

$$b' = \frac{0.525 - 0.5}{0.5/\sqrt{20}} = \frac{0.025}{0.1118034} = 0.2236068 \approx 0.2236$$

Zatem:

$$\begin{aligned} P(8 \leq Y \leq 10) &\approx P(-1.1180 \leq Z \leq 0.2236) \\ &= \Phi(0.2236) - \Phi(-1.1180) \approx 0.5885 - 0.1318 = 0.4567 \end{aligned}$$

5. Porównanie

Dokładna wartość: 0.4564. CTG bez poprawki: 0.3144 (błąd ≈ 0.1420). CTG z poprawką: 0.4567 (błąd ≈ 0.0003).

Przybliżenie CTG z poprawką ciągłości daje wynik bardzo bliski dokładnemu.

Zadanie 14 (Stanisław Macura)

Zaprosiłeś na przyjęcie 64 gości. Zakładając, że każdy z nich będzie chciał zjeść niezależnie od pozostałych 0, 1 albo 2 kanapki z prawdopodobieństwami $\frac{1}{4}$, $\frac{1}{2}$ oraz $\frac{1}{4}$ odpowiednio, ile kanapek musisz przygotować, aby mieć 95% pewności, że ich nie zabraknie?

Rozwiązanie

Niech X_i oznacza liczbę kanapek zjedzonych przez i -tego gościa.

$$P(X_i = 0) = \frac{1}{4}, \quad P(X_i = 1) = \frac{1}{2}, \quad P(X_i = 2) = \frac{1}{4}$$

$$\mathbb{E}[X_i] = 1, \quad \text{Var}(X_i) = \frac{1}{2}$$

Niech $S = \sum_{i=1}^{64} X_i$. Wówczas

$$\mathbb{E}[S] = 64, \quad \text{Var}(S) = 32, \quad \sigma_S = \sqrt{32}$$

Z centralnego twierdzenia granicznego:

$$S \approx \mathcal{N}(64, 32)$$

Szukamy s takiego, że $P(S \leq s) = 0.95$. Mamy

$$P(S \leq s) = P\left(\frac{S - 64}{\sqrt{32}} \leq \frac{s - 64}{\sqrt{32}}\right)$$

Niech $Z = \frac{S-64}{\sqrt{32}} \sim \mathcal{N}(0, 1)$. Z tablic:

$$P(Z \leq 1.645) = 0.95$$

Stąd

$$\frac{s - 64}{\sqrt{32}} = 1.645 \quad \Rightarrow \quad s = 64 + 1.645\sqrt{32} \approx 73.3$$

Zatem należy przygotować co najmniej 74 kanapki.

Zadanie 15 (Filip Manijak)

Niech $X_1, \dots, X_n$ będzie ciągiem niezależnych zmiennych losowych z rozkładu wykładniczego z parametrem $\lambda = 1$. Niech $\bar{X}$ będzie średnią arytmetyczną tego ciągu. Jak duże musi być n , aby

$$P(0.9 \leq \bar{X} \leq 1.1) \geq 0.95 ?$$

Rozwiązanie

Zmienne X_i mają $E[X_i] = 1, Var[X_i] = 1$. Wtedy $S = \sum_{i=1}^n X_i$. Skoro n jest wystarczająco duże, to podstawowe CTG daje nam:

$$P\left(a \leq \frac{S-1}{\frac{1}{\sqrt{n}}} \leq b\right) = \Phi(b) - \Phi(a)$$

Ponieważ średnia ma $E[S] = 1, Var[S] = 1/n$. Przekształcając:

$$P\left(a \leq \frac{S-1}{\frac{1}{\sqrt{n}}} \leq b\right) = P\left(\frac{a}{\frac{1}{\sqrt{n}}} + 1 \leq S \leq \frac{b}{\frac{1}{\sqrt{n}}} + 1\right)$$

Czyli $a \geq -0.1\sqrt{n}$ oraz $b \leq 0.1\sqrt{n}$. Minimalne n otrzymamy oczywiście dla najmniejszego a i największego b , czyli $-a = b = 0.1\sqrt{n}$. Czyli chcemy znaleźć minimalne n że:

$$\Phi(b) - \Phi(a) = 2 * \Phi(b) - 1 = 2 * \Phi(0.1\sqrt{n}) - 1 \geq 0.95$$

Niestety nie da się tego ładnie policzyć, więc sprawdzamy w tablicach. Wychodzi nam, że $\Phi(x) \geq 0.975$ dla $x \geq 1.959964$. Czyli

$$0.1\sqrt{n} \geq 1.959964 \implies n \geq 385$$

Odpowiedź

385

Zadanie 16 (Kuba Sułkowski)

Wskaż, jak można wykorzystać eksperyment, w którym rzucamy n razy igłą długości 1 na płaszczyznę z równoległymi liniami narysowanymi o odstępach jeden, do oszacowania wartości $\frac{1}{\pi}$. Jak duże powinno być n , aby z prawdopodobieństwem 0.95 nasze oszacowanie było z dokładnością do 0.01? Skorzystaj z CTG.

Rozwiązanie

Nasze zdarzenie to standardowa igła Buffona, czyli każdy rzut jest próbą Bernoulliego o prawdopodobieństwie sukcesu $\frac{2}{\pi}$.

Wartość oczekiwana sumy tych indykatorów po n próbach jest $E(\sum_{i=1}^n X_i) = \frac{2}{\pi}n$, a wariancja wyniesie $Var(\sum_{i=1}^n X_i) = n \cdot \frac{2}{\pi} \cdot \left(1 - \frac{2}{\pi}\right)$

Dla średniej mamy: $E(S) = \frac{2}{\pi}$, $Var(S) = \frac{1}{n} \cdot \frac{2}{\pi} \cdot \left(1 - \frac{2}{\pi}\right)$

Zamiast obliczać $\frac{1}{\pi}$ z dokładnością 0.01 obliczymy $\frac{2}{\pi}$ z dokładnością 0.02.

Przyjmijmy $p = \frac{2}{\pi}$, $\varepsilon = 0.02$

$$P(p - \varepsilon \leq S \leq p + \varepsilon) = P(-\varepsilon \leq S - p \leq \varepsilon)$$

$$G = P\left(-\varepsilon \cdot \sqrt{\frac{\pi^2 n}{2(\pi - 2)}} \leq (S - p) \cdot \sqrt{\frac{\pi^2 n}{2(\pi - 2)}} \leq \varepsilon \cdot \sqrt{\frac{\pi^2 n}{2(\pi - 2)}}\right)$$

Kopiuując z zadania wyżej (lub sprawdzając w tablicach), widzimy że $\Phi(x) \geq 0.975$ dla $x \geq 1.959964$. Czyli

$$0.02 \cdot \sqrt{\frac{\pi^2 n}{2(\pi - 2)}} \geq 1.959964 \implies n \geq 2222$$

Moglibyśmy założyć, że mamy jakąś ograniczoną wiedzę na temat π (gdybyśmy znali dokładną wartość π , obliczenie $\frac{1}{\pi}$ mogłoby być wykonane prościej).

Wiedząc jedynie, że $\pi > 2$, dostalibyśmy $n \geq 2401$ (pesymistycznym przypadkiem jest tutaj $\pi = 4$)

Odpowiedź

2222