

Rachunek Prawdopodobieństwa

Anna Janicka

wykład IX, 27.11.2018

**ROZKŁADY WIELOWYMIAROWE – cd.
NIEZALEŻNOŚĆ ZMIENNYCH LOSOWYCH**

Plan na dzisiaj

16. Rozkłady wielowymiarowych zmiennych losowych – cd.

- wartości oczekiwane funkcji zm. los.
- kowariancja, korelacja
- wartość oczekiwana wektora losowego
- liniowe przekształcenia wektorów losowych

17. Niezależność zmiennych losowych

- definicja
- charakterystyki nzal. zm. los



16. Rozkład łączny zmiennych losowych – cd

1. Twierdzenie 21

(i) *Jeśli (X, Y) jest dyskretną zmienną losową skoncentrowaną na zbiorze S i $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ jest funkcją borelowską, to*
$$\mathbb{E}\phi(X, Y) = \sum_{(x, y) \in S} \phi(x, y) \mathbb{P}((X, Y) = (x, y))$$
(o ile wartość oczekiwana istnieje).

(ii) *Jeśli (X, Y) jest zmienną losową o rozkładzie ciągłym (z gęstością g) i $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ jest funkcją borelowską, to*
$$\mathbb{E}\phi(X, Y) = \iint_{\mathbb{R}^2} \phi(x, y) g(x, y) dx dy$$
(o ile wartość oczekiwana istnieje).

2. Przykłady



16. Rozkład łączny zmiennych losowych – cd(2)

3. Kowariancja, korelacja

Niech (X, Y) będzie dwuwymiarową zmienną losową, taką że X, Y są całkowalnymi zmiennymi losowymi, i taką, że $\mathbb{E}|XY| < \infty$. Kowariancją zmiennych X, Y nazywamy liczbę

$$\text{Cov}(X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y).$$

Jeśli dodatkowo $\text{Var}X, \text{Var}Y > 0$, definiujemy współczynnik korelacji liniowej zmiennych X, Y , jako

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}X \cdot \text{Var}Y}} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}.$$



16. Rozkład łączny zmiennych losowych – cd(3)

4. Własności kowariancji i korelacji:

- stałe względem przesunięć
- dwuliniowość kowariancji
- przydatny wzór

$$\text{Cov}(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}X \cdot \mathbb{E}Y.$$

- wariancja jako kowariancja
- „łapie” liniową zależność
- kowariancja wrażliwa na skalę,
współczynnik korelacji nie:



16. Rozkład łączny zmiennych losowych – cd(4)

5. Ograniczenie współczynnika korelacji:

Nierówność Schwarza:

Niech $X, Y : \Omega \rightarrow \mathbb{R}$ będą zmiennymi losowymi takimi, że $\mathbb{E}X^2 < \infty$, $\mathbb{E}Y^2 < \infty$. Wówczas
$$|\mathbb{E}XY| \leq (\mathbb{E}X^2)^{1/2}(\mathbb{E}Y^2)^{1/2}.$$

Co więcej, równość w powyższej nierówności zachodzi wtedy i tylko wtedy, gdy istnieją takie dwie liczby a, b (z których co najmniej jedna jest różna od zera), że
$$\mathbb{P}(aX = bY) = 1.$$

Wniosek:

Jeśli $X, Y : \Omega \rightarrow \mathbb{R}$ są zmiennymi losowymi o skończonej i niezerowej wariancji, to $|\rho(X, Y)| \leq 1$.

Ponadto, jeśli $|\rho(X, Y)| = 1$, to istnieją takie liczby
$$a, b \in \mathbb{R}, \text{ że } Y = aX + b.$$

16. Rozkład łączny zmiennych losowych – cd(5)

6. Wartość oczekiwana, macierz kowariancji

Jeśli (X, Y) jest dwuwymiarową zmienną losową, to:

a) jeśli X oraz Y są całkowalne, to wartością oczekiwaną $\mathbb{E}(X, Y)$ zmiennej (X, Y) nazywamy wektor $(\mathbb{E}X, \mathbb{E}Y)$.

b) jeśli X oraz Y są całkowalne z kwadratem, to macierzą kowariancji zmiennej (X, Y) nazywamy macierz

$$\begin{bmatrix} \text{Var}X & \text{Cov}(X, Y) \\ \text{Cov}(X, Y) & \text{Var}Y \end{bmatrix}$$

Analogiczne wzory zachodzą dla zmiennych losowych o wartościach w $\mathbb{R}^d$, $d \geq 3$: wartość oczekiwaną definiujemy jako wektor $(\mathbb{E}X_1, \mathbb{E}X_2, \dots, \mathbb{E}X_d)$, a macierz kowariancji - jako $(\text{Cov}(X_i, X_j))_{1 \leq i, j \leq d}$.



16. Rozkład łączny zmiennych losowych – cd(6)

7. Liniowe przekształcenia wektorów los.

Założmy, że $X = (X_1, X_2, \dots, X_n)$ jest zmienną losową, a A - macierzą $m \times n$.

(i) Jeśli X posiada skończoną wartość oczekiwaną, to zmienna AX też ma skończoną wartość oczekiwaną i $\mathbb{E}(AX) = A\mathbb{E}X$.

(ii) Jeśli istnieje macierz kowariancji Q_X zmiennej X , to istnieje też macierz kowariancji zmiennej AX i wynosi ona $Q_{AX} = AQ_X A^t$.

Założmy, że X jest zmienną losową o rozkładzie ciągłym w $\mathbb{R}^n$, m jest pewnym wektorem w $\mathbb{R}^n$, a T jest pewnym nieosobliwym przekształceniem $\mathbb{R}^n$ w siebie (tzn. $\det T \neq 0$, gdzie T utożsamiamy z jego macierzą). Wówczas zmienna $TX + m$ ma gęstość $g_{TX+m}(x) = \frac{1}{|\det T|} g_X(T^{-1}(x - m))$.



17. Niezależne zmienne losowe

1. Definicja niezależności

Zmienne losowe $X_1, \dots, X_n: \Omega \rightarrow \mathbb{R}$ nazywamy **niezależnymi**, jeśli dla dowolnego ciągu $B_1, B_2, \dots, B_n$ zbiorów borelowskich zachodzi

$$\begin{aligned} \mathbb{P}(X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n) \\ = \mathbb{P}(X_1 \in B_1) \mathbb{P}(X_2 \in B_2) \dots \mathbb{P}(X_n \in B_n). \end{aligned}$$

- 2. Niezależność zmiennych losowych a niezależność zdarzeń
 - 3. Niezależność parami a niezależność łączna
-



17. Niezależne zmienne losowe – cd.

4. Badanie niezależności dyskretnych zm.

Niech $X_1, X_2, \dots, X_n$ będą dyskretnymi zmiennymi losowymi, zaś S_{X_i} takimi zbiorami przeliczalnymi, że $\mathbb{P}(X_i \in S_{X_i}) = 1$. Wówczas zmienne $X_1, X_2, \dots, X_n$ są niezależne wtedy i tylko wtedy, gdy dla każdego ciągu $x_1, x_2, \dots, x_n$ takiego, że $x_i \in S_{X_i}$, $i = 1, 2, \dots, n$, zachodzi równość

$$\begin{aligned} \mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ = \mathbb{P}(X_1 = x_1) \mathbb{P}(X_2 = x_2) \cdot \dots \cdot \mathbb{P}(X_n = x_n). \end{aligned}$$

5. Przykład



17. Niezależne zmienne losowe – cd(2)

6. Badanie niezależności ciągłych zm.

Niech $X_1, X_2, \dots, X_n: \Omega \rightarrow \mathbb{R}$ będą ciągłymi zmiennymi losowymi o gęstościach $g_1, g_2, \dots, g_n$, odpowiednio.

Wówczas zmienne $X_1, X_2, \dots, X_n$ są niezależne wtedy i tylko wtedy, gdy funkcja $g: \mathbb{R}^n \rightarrow [0, \infty)$ dana wzorem $g(x_1, x_2, \dots, x_n) = g_1(x_1)g_2(x_2) \cdot \dots \cdot g_n(x_n)$, jest gęstością rozkładu $\mu_{(X_1, X_2, \dots, X_n)}$.

7. Przykłady

- rozkład jednostajny na kwadracie $\{\pm 1, \pm 1\}$
 - rozkład jednostajny na kole
 - inny rozkład
-



17. Niezależne zmienne losowe – cd. (3)

8. Niezależność funkcji zmiennych niezal.

*Rozważmy zmienne losowe $X_{1,1}, X_{1,2}, \dots, X_{1,k_1},$
 $X_{2,1}, X_{2,2}, \dots, X_{2,k_2}, \dots, X_{n,1}, X_{n,2}, \dots, X_{n,k_n}$
oraz funkcje borelowskie $\varphi_i: \mathbb{R}^{k_i} \rightarrow \mathbb{R}, i = 1, 2, \dots, n.$
Jeśli powyższe zmienne są niezależne, to również zmienne*

$$Y_1 = \varphi_1(X_{1,1}, X_{1,2}, \dots, X_{1,k_1}),$$

$$Y_2 = \varphi_2(X_{2,1}, X_{2,2}, \dots, X_{2,k_2}),$$

$\dots$

$$Y_n = \varphi_n(X_{n,1}, X_{n,2}, \dots, X_{n,k_n})$$

są niezależne.



17. Niezależne zmienne losowe – cd. (4)

9. Charakterystyki niezależnych zm. los.

Niech $X_1, X_2, \dots, X_n$ będą niezależnymi zmiennymi losowymi, posiadającymi wartość oczekiwaną. Wówczas zmienna $X_1 \cdot X_2 \cdot \dots \cdot X_n$ również ma wartość oczekiwaną i zachodzi równość

$$\mathbb{E}(X_1 X_2 \cdot \dots \cdot X_n) = \mathbb{E}X_1 \cdot \mathbb{E}X_2 \cdot \dots \cdot \mathbb{E}X_n.$$

10. Przykład

11. Kowariancja niezależnych zm. los.

Jeśli X, Y są niezależnymi zmiennymi losowymi takimi, że zachodzi $\mathbb{E}|XY| < \infty$, to $\text{Cov}(X, Y) = 0$. Jeśli dodatkowo zmienne X, Y mają skończoną, niezerową wariancję, to zeruje się także ich współczynnik korelacji, $\rho(X, Y) = 0$.

17. Niezależne zmienne losowe – cd. (5)

12. Nieskorelowane zmienne losowe

Zmienne losowe X, Y nazwiemy nieskorelowanymi, jeżeli $\rho(X, Y) = 0$.

13. Niezależność $\Rightarrow$ nieskorelowanie, ale

$\Leftarrow$ **NIE ZACHODZI!**

14. Przykład – rozkład jednostajny na kole



17. Niezależne zmienne losowe – cd. (6)

15. Wariancja sumy nzał zm. los.

Niech $X_1, X_2, \dots, X_n$ będą ~~niezależnymi~~ zmiennymi losowymi o skończonej wariancji. Wówczas zmienna $X_1 + X_2 + \dots + X_n$ także ma skończoną wariancję oraz

$$\text{Var}(X_1 + X_2 + \dots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n) + 2 \sum_{i < j} \text{Cov}(X_i, X_j).$$

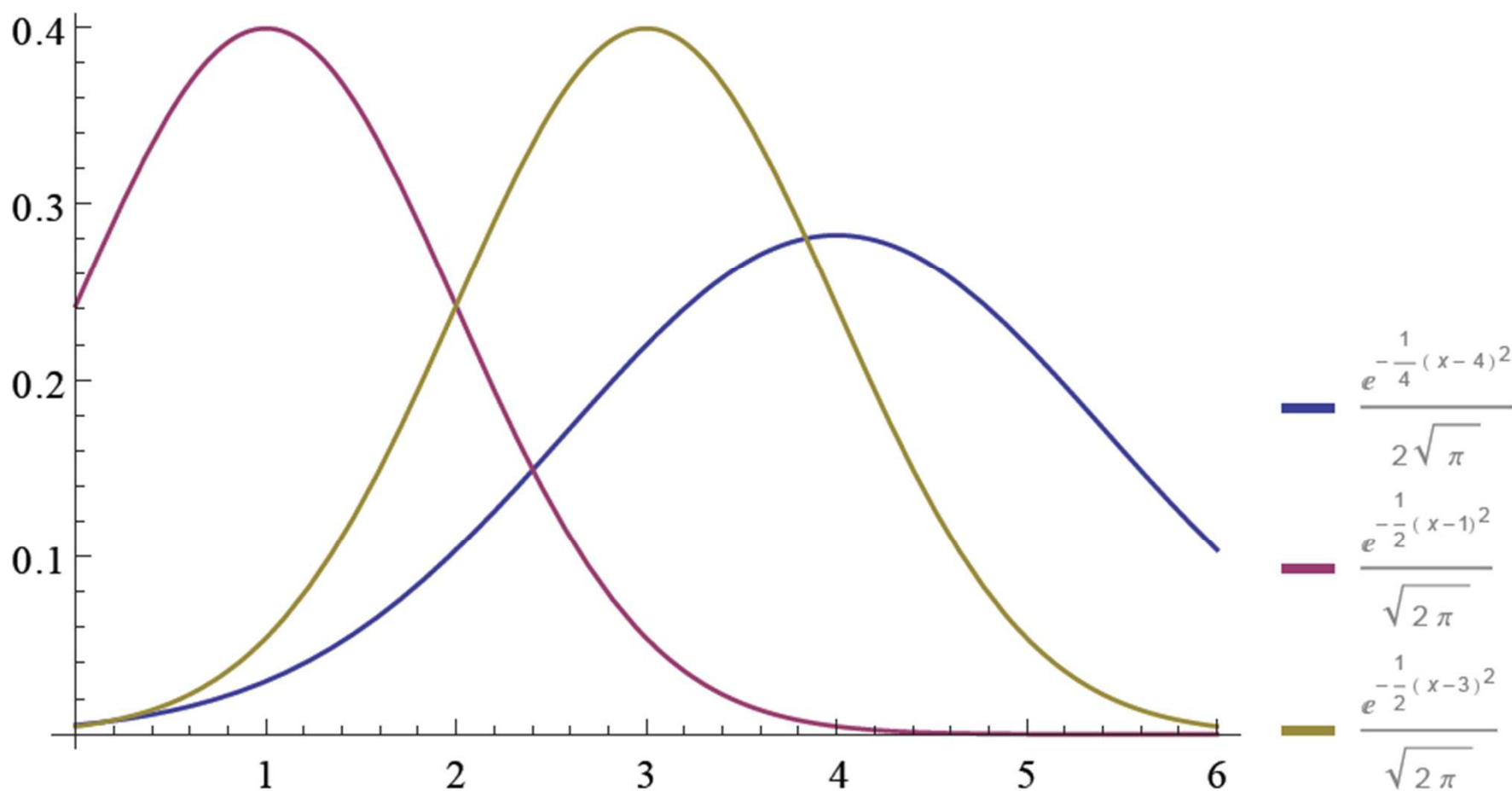
16. Przykład – suma oczek w n rzutach

17. Gęstość sumy – spłot funkcji gęstości

Założmy, że X, Y są niezależnymi zmiennymi losowymi o rozkładach ciągłych z gęstościami g_X, g_Y . Wówczas zmienna $X + Y$ ma rozkład z gęstością będącą spłotem funkcji g_X i g_Y , tzn.

$$g_{X+Y}(x) = g_X * g_Y(x) = \int_{\mathbb{R}} g_X(y) g_Y(x - y) dy = \int_{\mathbb{R}} g_X(x - y) g_Y(y) dy.$$


Suma niezależnych zmiennych losowych





WARSAW UNIVERSITY
Faculty of Economic Sciences