

# Rachunek prawdopodobieństwa i statystyka

## Komentarz do wykładów z 7. i 12. kwietnia

Założmy, że zmienne losowe  $X_1, \dots, X_n$  są niezależne i podlegają rozkładowi  $N(\mu, \sigma^2)$  każda. Określmy dwie nowe zmienne losowe  $\bar{X}, S^2$  jako

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n X_k, \quad S^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X})^2. \quad (1)$$

Używając funkcji generujących momentów (MGFs-ów), stwierdzamy iż

**Twierdzenie 1.**  $\bar{X} \sim N(\mu, \sigma^2/n)$ .

*Dowód.* Z notatki 5, tw. 1 wynika, że:

$$M_{\bar{X}}(t) = \prod_{k=1}^n M_{X_k}\left(\frac{t}{n}\right) = \exp\left(n\mu u + \frac{n\sigma^2 u^2}{2}\right)\Big|_{u=t/n} = \exp\left(t\mu + \frac{\sigma^2 t^2}{2n}\right) \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

□

Zanim zajmujemy się zmienną  $S^2$ , wprowadźmy nową zmienną losową  $S_\mu^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \mu)^2$ . Po niewielkim przekształceniu, znajdujemy iż  $\frac{nS_\mu^2}{\sigma^2} = \sum_{k=1}^n \left(\frac{X_k - \mu}{\sigma}\right)^2 \equiv \chi^2(n)$ <sup>1</sup>. Znajdziemy teraz związek pomiędzy  $S_\mu^2$  a  $S^2$ .

**Twierdzenie 2.**

$$S_\mu^2 = S^2 + (\bar{X} - \mu)^2. \quad (2)$$

*Dowód.*

$$\begin{aligned} n \cdot S_\mu^2 &= \sum_{k=1}^n (X_k - \mu)^2 = \sum_{k=1}^n (X_k - \bar{X} + \bar{X} - \mu)^2 = \\ &= \sum_{k=1}^n (X_k - \bar{X})^2 + n \cdot (\bar{X} - \mu)^2 + 2 \cdot \sum_{k=1}^n (\bar{X} - \mu)(X_k - \bar{X}). \end{aligned}$$

Sprawdzamy, że

$$\sum_{k=1}^n (\bar{X} - \mu)(X_k - \bar{X}) = (\bar{X} - \mu) \cdot \sum_{k=1}^n (X_k - \bar{X}) = (\bar{X} - \mu) \cdot \left(\sum_{k=1}^n X_k - n \cdot \bar{X}\right) = 0,$$

co kończy dowód twierdzenia. □

UWAGA: Nieformalnie możemy zapisać **twierdzenie 1** w postaci  $\chi^2(n) \equiv (\text{pewien\_rozkład}) + \chi^2(1)$ . Gdyby **pewien\\_rozkład** i  $\chi^2(1)$  były niezależne, to moglibyśmy powiedzieć coś o **pewnym rozkładzie** – czyli o rozkładzie  $S^2$  (Twierdzenie 1. z notatki 5.) Twierdzenie 1. z bieżącej notatki można bowiem przepisać w postaci

$$\frac{nS_\mu^2}{\sigma^2} = \frac{nS^2}{\sigma^2} + \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2. \quad (3)$$

Po lewej stronie powyższego równania rozpoznajemy rozkład  $\chi^2(n)$ , drugi składnik prawej strony podlega rozkładowi  $\chi^2(1)$ .

---

<sup>1</sup>Lista 5., zadanie 6.

## Niezależność $\bar{X}$ oraz $S^2$

Wprowadzamy nowe zmienne  $Y_k$  następująco:

$$Y_1 = \bar{X}, \quad Y_k = X_k - \bar{X}, \quad (k = 2, \dots, n). \quad (4)$$

Ponieważ zmienne  $X_1, \dots, X_n$  są niezależne więc  $n$ -wymiarowa gęstość  $f_{X_1, \dots, X_n}(x_1, \dots, x_n)$  wyraża się wzorem

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = (2\pi)^{-n/2} \sigma^{-n} \prod_{k=1}^n \exp \left\{ -\frac{1}{2} \left( \frac{x_k - \mu}{\sigma} \right)^2 \right\} = (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2} \sum_{k=1}^n \left( \frac{x_k - \mu}{\sigma} \right)^2 \right\}. \quad (5)$$

Odwrócenie wzorów (??) daje

$$\begin{aligned} X_k &= \bar{X} + Y_k = Y_1 + Y_k \quad (k = 2, \dots, n), \\ nY_1 &= X_1 + \dots + X_n = X_1 + (Y_1 + Y_2) + \dots + (Y_1 + Y_{n-1}) + (Y_1 + Y_n), \text{ czyli} \\ X_1 &= Y_1 - Y_2 - \dots - Y_n. \end{aligned} \quad (6)$$

Wyznacznik Jacobianu odwrócenia (??) daje wartość  $n \cdot 2$ . Wzór (??) można zatem przepisać, – w nowym układzie współrzędnych – w postaci

$$g_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = n \cdot (2\pi)^{-n/2} \sigma^{-n} \exp \left\{ -\frac{1}{2} \sum_{k=1}^n \left( \frac{x_k - \mu}{\sigma} \right)^2 \right\}, \quad (7)$$

gdzie w miejsce  $x_k$  należy podstawić wzory (??). Tymczasowo, nie dokonujemy tego podstawienia, wzory byłyby nazbyt złożone. Zajmiemy się tylko wykładnikiem wzoru (??), bez czynnika  $-1/2$ . Uwzględniając wzór (??) możemy wspomniany wykładnik przedstawić w postaci

$$\frac{nS_\mu^2}{\sigma^2} = \sum_{k=1}^n \left( \frac{X_k - \mu}{\sigma} \right)^2 = \sum_{k=1}^n \left( \frac{X_k - \bar{X}}{\sigma} \right)^2 + \left( \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \right)^2 = \frac{nS^2}{\sigma^2} + \chi^2(1). \quad (8)$$

Drugi składnik po prawej stronie jest zapisany nieformalnie.<sup>3</sup>

**$S^2$  jest funkcją  $Y_2, \dots, Y_n$ .**

Rozpatrzmy  $nS^2 = \sum_{k=1}^n (X_k - \bar{X})^2 = (X_1 - \bar{X})^2 + \sum_{k=2}^n (X_k - \bar{X})^2$ . Uwzględniając wzory (??) można je (to równanie) przepisać w postaci

$$nS^2 = \sum_{k=1}^n (X_k - \bar{X})^2 = \left( \sum_{k=2}^n Y_k \right)^2 + \sum_{k=2}^n Y_k^2. \quad (9)$$

**$S^2$  oraz  $\bar{X}$  są niezależne.**

Pamiętając o tym, że  $Y_1 \equiv \bar{X}$  i podstawiając równość (??) do równania (??) otrzymujemy

$$\begin{aligned} K &= n \cdot (2\pi)^{-n/2} \sigma^{-n}, \\ g_{Y_1, \dots, Y_n}(y_1, \dots, y_n) &= K \exp \left\{ -\frac{1}{2} \sum_{k=1}^n \left( \frac{x_k - \mu}{\sigma} \right)^2 \right\} \stackrel{(8)}{=} \\ &\stackrel{(8)}{=} K \exp \left\{ -\frac{1}{2} \left( \frac{nS_\mu^2}{\sigma^2} \right) \right\} = K \exp \left\{ -\frac{1}{2} \left( \frac{nS^2}{\sigma^2} \right) \right\} \cdot \exp \left\{ -\frac{1}{2} \left( \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \right)^2 \right\}. \end{aligned} \quad (10)$$

To dowodzi następującego twierdzenia

<sup>2</sup>Por. zadanie 5. z listy 1.

<sup>3</sup>Por. twierdzenie ??.

**Twierdzenie 3.** Zmienne losowe  $\bar{X}$  oraz  $S^2$  są niezależne. Inaczej: zmienne losowe  $Y_1$  oraz  $(Y_2, \dots, Y_n)$  są niezależne.

Równanie (??) można przepisać – w języku funkcji tworzących momenty – następująco:

$$(1 - 2t)^{-n/2} = M_{\frac{nS^2}{\sigma^2}}(t) = M_{\frac{nS^2}{\sigma^2}}(t) \cdot (1 - 2t)^{-1/2}.$$

Stąd wynika

**Twierdzenie 4.**  $\frac{nS^2}{\sigma^2} \sim \chi^2(n-1) \equiv \text{Gamma}\left(\frac{1}{2}, \frac{n-1}{2}\right)$ .

Oczywiście, twierdzenia ?? oraz ?? są prawdziwe przy założeniu, że zmienne losowe  $X_1, \dots, X_n$  są niezależne i podlegają rozkładowi  $N(\mu, \sigma^2)$  każda.

## Wielowymiarowy rozkład normalny.

**Definicja 1.** Mówimy, że zmienna losowa  $(X_1, \dots, X_n)^T$  podlega  $n$ -wymiarowemu rozkładowi normalnemu – w skrócie  $X \sim N(\mu, \Sigma)$  – wtedy i tylko wtedy gdy

$$f_X(x) = f_{X_1, \dots, X_n}(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(X - \mu)^T \Sigma^{-1}(X - \mu)\right]. \quad (11)$$

UWAGI:

1. Wektor  $\mu = [\mu_1, \dots, \mu_n]^T \in \mathbb{R}^n$ . Liczbę  $\mu_k$  nazywamy wartością oczekiwaną jednowymiarowej zmiennej brzegowej  $X_k$ .
2. Symetryczną, dodatnio określoną macierz  $\Sigma = (s_{ij})_{i,j=1}^n \in \mathbb{R}^{n \times n}$  nazywamy macierzą wariancji-kowariancji. Liczbę  $s_{ii}$  nazywamy wariancją zmiennej  $X_i$ , liczbę  $s_{ij}$  – kowariancją zmiennych  $X_i, X_j$ .
3. Symbol  $|\Sigma|$  oznacza wyznacznik macierzy  $\Sigma$ .
4. Dla  $n = 1$  wzór (??) upraszcza się do postaci

$$f_X(x) = f_{X_1(x_1)} = \frac{1}{(\sqrt{2\pi}) \sigma_1} \exp\left[-\frac{1}{2}(X_1 - \mu_1)^T (\sigma_1^2)^{-1} (X_1 - \mu_1)\right].$$

5. Wzór (3) z notatki 6. ilustruje wypadek  $n = 2$ .
6. Symetryczna macierz  $M \in \mathbb{R}^{n \times n}$  jest nazywana macierzą dodatnio określoną jedynie wtedy gdy  $\forall x \in \mathbb{R}^n \ x^T M x \geq 0$ , natomiast równość zachodzi jedynie dla  $x = \mathbb{O}$ . Macierz dodatnio określona ma  $n$  wartości własnych  $\lambda_k > 0$ , w szczególności jest to macierz odwracalna.

Niech macierz  $A \in \mathbb{R}^n$  będzie macierzą odwracalną, a zmienna losowa  $X$  niech ma rozkład  $N(\mu, \Sigma)$ . Rozważmy zmienną losową  $Y = AX$ , to znaczy  $Y$  jest obrazem zmiennej losowej  $X$  względem przekształcenia liniowego  $A$ . Przy tych założeniach prawdziwe jest poniższe

**Twierdzenie 5.**

$$Y \sim N(A\mu, A\Sigma A^T).$$

*Dowód.* Zajmijmy się wykładnikiem wzoru (??), bez czynnika  $-1/2$ . Ponieważ macierz jest odwracalna, więc  $X = A^{-1}Y$ , co podstawieniu do wykładnika daje

$$(X - \mu)^T \Sigma^{-1} (X - \mu) = (A^{-1}Y - \mu)^T \Sigma^{-1} (A^{-1}Y - \mu) = (A^{-1}(Y - A\mu))^T \Sigma^{-1} (A^{-1}(Y - A\mu)) = (*)$$

Korzystając z zależności  $(SU)^T = U^T S^T$  oraz  $(SU)^{-1} = U^{-1} S^{-1}$  mamy

$$(*) = (Y - A\mu)^T (A^{-1})^T \Sigma^{-1} A^{-1} (Y - A\mu) = (Y - A\mu)^T ((A\Sigma A^T))^{-1} (Y - A\mu).$$

Zauważamy podobieństwo otrzymanego wzoru do wykładnika wzoru (??). Obydwa wzory mają postać  $z^T M^{-1} z$ .

Przejdźmy teraz do Jacobianu przekształcenia  $X = A^{-1}Y$ . Jeżeli elementy macierzy  $A^{-1}$  oznaczymy symbolem  $b_{ij}$ , to  $X_i = \sum_{k=1}^n b_{ik} Y_k$ . Stąd  $\frac{\partial X_i}{\partial Y_j} = b_{ij}$ , czyli  $J = |A^{-1}| = \frac{1}{|A|}$ . Przejście do zmiennej  $Y$  we wzorze (??) daje

$$f_Y(y) = f_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2} |A|} \exp \left[ -\frac{1}{2} (Y - A\mu)^T ((A\Sigma A^T))^{-1} (Y - A\mu) \right].$$

Uwzględniając równość  $|A\Sigma A^T| = |A| |\Sigma| |A^T| = |\Sigma| |A|^2$  otrzymujemy tezę twierdzenia.

□

←

Z poważaniem,  
Witold Karczewski