Rachunek prawdopodobieństwa i statystyka

Komentarz do wykładu z 2. kwietnia

Definicja 1. *Funkcją tworzącą momenty zmiennej losowej X nazywamy*

$$M_X(t) = \mathrm{E}\,(\mathrm{e}^{tX}). \tag{1}$$

W wypadku dyskretnym

$$M_X(t) = \sum_k \exp\,(t x_k)\cdot p_k,$$

natomiast w wypadku ciągłym

$$M_X(t) = \int_{\mathbb{R}} \mathrm{e}^{tx}\cdot f(x)\,dx.$$

Uzasadnienie nazwy: Ponieważ $\mathrm{e}^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^k}{k!} + \ldots$. Jest zatem (przynajmniej formalnie): $\mathrm{e}^{tX} = 1 + tX + \frac{t^2X^2}{2!} + \ldots + \frac{t^kX^k}{k!} + \ldots$. *Moment zwykły* zmiennej losowej X to $m_k = \mathrm{E}(X^k)$. Stąd $M_X(t) = 1 + m_1\,t + m_2\,\frac{t^2}{2} + \ldots$. Można sprawdzić – dla przykładu – że $m_1 = \mathrm{E}X = M_X'(0)$ oraz $m_2 = \mathrm{E}(X^2) = M_X''(0)$. Stąd już $\mathrm{V}X = M_X''(0) - (M_X'(0))^2$.

Twierdzenie 1. *Załóżmy, że zmienne X, Y są niezależne. Niech $V = aX + b$, $Z = X + Y$, gdzie $a \neq 0$. Jest wówczas $M_V(t) = \mathrm{e}^{\,tb}\cdot M_X(at)$ oraz $M_Z(t) = M_X(t)\cdot M_Y(t)$.*

Dowód.

$$\begin{aligned} M_V(t) &= \mathrm{E}\left(\mathrm{e}^{t(aX+b)}\right) = \mathrm{E}\left(\mathrm{e}^{tb}\,\mathrm{e}^{(at)X}\right) = \mathrm{e}^{tb}\cdot M_X(at),\\ M_Z(t) &= \mathrm{E}\left(\mathrm{e}^{t(X+Y)}\right) = \mathrm{E}\left(\mathrm{e}^{tX}\,\mathrm{e}^{tY}\right) \overset{(a)}{=} \mathrm{E}\left(\mathrm{e}^{tX}\right)\cdot \mathrm{E}\left(\mathrm{e}^{tY}\right) = M_X(t)\cdot M_Y(t). \end{aligned}$$

(a) Ponieważ zmienne X, Y są niezależne, to niezależne są też zmienne e^{tX}, e^{tY}. □

Kilka przykładów "MGF-ów"

Przykłady:

(i) Niech $X \sim B(n,p)$. Prawdopodobieństwa p_k są dane wzorem $p_k = \binom{n}{k} p^k(1-p)^{n-k}$. Stąd

$$M_X(t) = \sum_{k=0}^{n} \mathrm{e}^{tk}\,p_k = \sum_{k=0}^{n} \mathrm{e}^{tk}\binom{n}{k} p^k(1-p)^{n-k} = \sum_{k=0}^{n}\binom{n}{k}\left(p\cdot \mathrm{e}^t\right)^k\,(1-p)^{n-k} \overset{(b)}{=} \left(p\cdot\mathrm{e}^t + q\right)^n.$$

(b) Wzór dwumianowy: $(a+b)^n = \sum_{k=0}^{n}\binom{n}{k} a^k\,b^{n-k}$.

(ii) Niech $X \sim \mathrm{Poisson}(\lambda)$. Teraz $p_k = \mathrm{e}^{-\lambda}\cdot\frac{\lambda^k}{k!}$. Ponieważ w rozwinięciu e^x w szereg nieskończony (c): $\mathrm{e}^x = \sum_{k=0}^{\infty}\frac{x^k}{k!}$, więc wynika stąd, że $M_X(t) = \mathrm{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\,\mathrm{e}^{tk} = \mathrm{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{\left(\lambda\cdot\mathrm{e}^t\right)^k}{k!} \overset{(c)}{=} \mathrm{e}^{\lambda(\mathrm{e}^t-1)}$.

(iii) Rozkład B(n,p) jako suma rozkładów "0-1". Przeprowadzamy n niezależnych prób z ppb sukcesu p w każdej próbie. Jeżeli X_k oznacza zmienną losową zliczającą sukcesy w k-tej próbie, to rozkład zmiennej X_k nazywamy rozkładem "0-1".

X_k	0	1
p_k	$1-p$	p

Dla funkcji tworzącej momenty zmiennej X_k mamy proste wyrażenie

$$M_{X_k}(t) = \mathrm{E}\left(\mathrm{e}^{tX_k}\right) = (1-p)\cdot \mathrm{e}^0 + p\cdot \mathrm{e}^t = p\cdot \mathrm{e}^t + q.$$

Ponieważ $X = X_1 + \ldots + X_n$, a zmienne X_k są niezależne, więc[a]

$$M_X(t) = \prod_{k=1}^{n} M_{X_k}(t) = \left(p\cdot \mathrm{e}^t + q\right)^n.$$

(iv) Rozkład Gamma(b,p). Gęstość $f(x) = \dfrac{b^p}{\Gamma(p)}\, x^{p-1}\, \mathrm{e}^{-bx}$, dla $x \in (0,\infty)$, parametry $b, p \in \mathbb{R}$.

$$M_X(t) = \frac{b^p}{\Gamma(p)} \int_0^\infty \mathrm{e}^{tx}\, x^{p-1}\, \mathrm{e}^{-bx}\, dx = \left\downarrow \begin{array}{ccc} u & = & (b-t)x \\ du & = & (b-t)\, dx \end{array} \right\uparrow = (*).$$

Dla $t \in (-\infty, b)$ jest dalej (granice całkowania)

$$(*) = \frac{b^p}{\Gamma(p)} \cdot \frac{1}{(b-t)^p} \int_0^\infty \mathrm{e}^{-u}\, u^{p-1}\, du = \left(\frac{b}{b-t}\right)^p = \left(1 - \frac{t}{b}\right)^{-p}.$$

(v) Znajdziemy teraz funkcję tworzącą momenty dla rozkładu $N(0,1)$. Gęstość tego rozkładu to $f(x) = \dfrac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-x^2/2}$, gdzie $x \in \mathbb{R}$. Funkcja tworząca momenty to:

$$\begin{aligned} M_X(t) \quad &= \mathrm{E}\left(\mathrm{e}^{tX}\right) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{tx}\, \mathrm{e}^{-x^2/2}\, dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left[-\frac{x^2 - 2xt}{2}\right] dx = \\ &= \exp\left(\frac{t^2}{2}\right) \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left[-\frac{(x-t)^2}{2}\right] dx = \left| \begin{array}{l} u = x - t \\ du = dx \end{array} \right| = \\ &= \exp\left(\frac{t^2}{2}\right) \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-u^2/2}\, du = \mathrm{e}^{t^2/2}. \end{aligned}$$

[a] MGF sumy niezależnych zmiennych to iloczyn ich MGFs-ów.

Rozkład normalny

Definicja 2. *Mówimy, że zmienna losowa X podlega rozkładowi normalnemu z parametrami μ, σ^2 ($X \sim N(\mu, \sigma^2)$) jedynie wtedy gdy gęstość jest określona wzorem*

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \text{ dla } x \in \mathbb{R}. \tag{2}$$

Twierdzenie 2. *Niech $X \sim N(\mu, \sigma^2)$ i niech $Y = \dfrac{X-\mu}{\sigma}$, gdzie $\sigma > 0$. Wówczas $Y \sim N(0,1)$.*

Dowód. Udowodnimy twierdzenie klasycznym sposobem: od dystrybuanty do gęstości.

$$F_Y(t) = P(Y \leq t) = P\left(\frac{X-\mu}{\sigma} \leq t\right) = P(X \leq \sigma t + \mu) = F_X(\sigma t + \mu).$$

Różniczkując powyższe równanie obustronnie względem zmiennej t otrzymujemy

$$f_Y(t) = f_X(\sigma t + \mu)\cdot \sigma = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) \sim \mathrm{N}(0,1).$$

$\square$

Definicja 3. *Niech zmienne losowe* $X_1, \ldots, X_n$ *będą niezależne i podlegają rozkładowi* $N(0,1)$. *Rozkład zmiennej losowej* $Z = \sum_{k=1}^{n} X_k^2$ *nazywamy rozkładem* chi-kwadrat z n stopniami swobody *i oznaczamy symbolem* $Z \sim \chi^2(n)$.

Wnioski (część z nich będzie treścią zadań na ćwiczeniach)

1. "Odwrócenie" twierdzenia jest następujące: Jeżeli $S \sim \mathrm{N}(0,1)$ i $T = \sigma S + \mu$, to $T \sim \mathrm{N}(\mu, \sigma^2)$.
2. Wiemy, że dla rozkładu $U \sim \mathrm{N}(0,1)$ jest $M_U(t) = \exp\left(\frac{t^2}{2}\right)$. Korzystając z tw. 1 (pierwszy wzór) otrzymujemy MGF dla rozkładu $\mathrm{N}(\mu, \sigma^2)$
$$M_U(t) = \exp\left(\mu\, t + \frac{\sigma^2 t^2}{2}\right).$$
3. Jeżeli $U \sim \mathrm{N}(0,1)$, to $U^2 \sim \mathrm{Gamma}(1/2, 1/2) \equiv \chi^2(1)$.
4. Niech $U_1, \ldots, U_n$ będą niezależnymi zmiennymi losowymi o rozkładzie $\mathrm{N}(0,1)$ każda. Niech, oprócz tego, $Z = \sum_{k=1}^{n} U_k^2$. Wówczas $Z \sim \mathrm{Gamma}(1/2, n/2) \equiv \chi^2(n)$.
5. Niech $U_1, \ldots, U_n$ będą niezależnymi zmiennymi losowymi o rozkładach (odpowiednio) $\mathrm{N}(\mu_k, \sigma_k^2)$. Niech, dodatkowo, $Z = \sum_{k=1}^{n} \alpha_k U_k$. Wówczas $Z \sim \mathrm{N}\left(\sum_{k=1}^{n} \alpha_k \mu_k, \sum_{k=1}^{n} \alpha_k^2 \sigma_k^2\right)$.
6. Jeżeli $U_1, \ldots, U_n$ są niezależnymi zmiennymi losowymi o rozkładzie $\mathrm{N}(\mu, \sigma^2)$ każda, oraz $S^2 = \frac{1}{n} \cdot \sum_{k=1}^{n} (U_k - \mu)^2$, to $\frac{nS^2}{\sigma^2} \sim \chi^2(n)$.
7. $\mapsto \infty$. Pozostałe uwagi…

Uwaga (o rozkładzie normalnym 2-wymiarowym)

Poniżej, nieformalne określenie 2-wymiarowego rozkładu normalnego. Nieformalne, ponieważ można podać zwarty wzór na gęstość n-wymiarowego rozkładu normalnego (co nastąpi w niedalekiej przyszłości).

Dany jest wektor $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ oraz macierz $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$. O macierzy Σ zakładamy też, że jest nieosobliwa (odwracalna). Funkcją gęstości $f(x,y)$ zmiennej losowej (X,Y) nazywamy

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(y-\mu_2)^2}{\sigma_2^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}\right]\right). \tag{3}$$

Z powyższego wzoru można wyliczyć, że zmienne brzegowe X, Y mają rozkłady $X \sim \mathrm{N}(\mu_1, \sigma_1^2)$, $Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$. Oprócz tego współczynnik kowariancji zmiennych X, Y to $\rho\sigma_1\sigma_2$.

Trójwymiarowa zmienna o rozkładzie normalnym: wzór zajmuje 3–4 wiersze. W podsumowaniu: jeżeli na muralu zobaczymy wzór (3) to można przejść obok napisu bez żadnych emocji, jest to po prostu **normalny** rozkład.

Rozpatrzmy 2-wymiarową zmienną $(X,Y) \sim \mathrm{N}(\mu, \Sigma)$. Jeżeli kowariancja $\mathrm{Cov}(X,Y) = 0$, to wynika stąd, że $\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$, czyli $\rho = 0$. Wzór (3) można przepisać jako

$$f(x,y) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}\right) \cdot \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{1}{2}\frac{(y-\mu_2)^2}{\sigma_2^2}\right). \tag{4}$$

Wnioskujemy stąd, że zmienne brzegowe X, Y są niezależne; dodatkowo $X \sim \mathrm{N}(\mu_1, \sigma_1^2)$, a także $Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$. Mamy zatem szczególny wypadek, kiedy tw. 2 z notatki 4. można odwrócić, a mianowicie

Twierdzenie 3. *Dana jest zmienna losowa* $(X,Y) \sim N(\mu, \Sigma^2)$. *Jeżeli* $\mathrm{Cov}(X,Y) = 0$, *to zmienne brzegowe* X, Y *są niezależne*

$\leftarrow$

Z poważaniem,
Witold Karczewski