Rachunek prawdopodobieństwa i statystyka

Wykład 6

Funkcja generująca moment

Definicja: Funkcją generującą moment (fgm) dla zmiennej losowej X nazywamy funkcję zmiennej rzeczywistej t postaci:

$$m_X(t) = \mathcal{E}\left[e^{tX}\right]$$

Uwaga: mgf istnieje zawsze dla $t = 0$. Jeśli X jest dodatnią zmienną losową oraz istnieje fgm dla określonej wartości t_0 to wtedy fgm istnieje także dla $t \leq t_0$

Przykład: fgm dla rozkładu dwumianowego

$$m_X(t) = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} = \left(pe^t + 1 - p\right)^n$$

Przykład: fgm dla rozkładu wykładniczego

$$m_X(t) = \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^\infty e^{-(\lambda - t)x} dx = \frac{\lambda}{\lambda - t}$$

Twierdzenie: Jeśli spełniony jest warunek $\mathcal{E}\left[|X|^n\right] < \infty$ to wówczas dla k-ty moment $(k = 1, 2, \ldots, n)$ zmiennej losowej X dany jest przez:

$$m_k = \left.\frac{d^k}{dt^k} m_X(t)\right|_{t=0}$$

Funkcja generująca moment

Przykład: fgm dla rozkładu dwumianowego

$$m_X(t) = \left(pe^t + 1 - p\right)^n$$

$$m_X'(t) = n\left(pe^t + 1 - p\right)^{n-1} pe^t \quad \Rightarrow \quad \mathcal{E}[X] = m_X'(0) = np$$

$$m_X''(t) = n(n-1)\left(pe^t + 1 - p\right)^{n-2}\left(pe^t\right)^2 + n\left(e^t + 1 - p\right)^{n-1} pe^t \quad \Rightarrow$$

$$\Rightarrow \quad \mathcal{E}\left[X^2\right] = m_X''(0) = n(n-1)p^2 + np$$

Twierdzenie: Jeśli $m_X(t)$ jest fgm zmiennej X, to fgm zmiennej $Y = \alpha X + \beta$ dane jest przez

$$m_Y(t) = e^{\beta t} m_X(\alpha t)$$

D: $m_Y(t) = \mathcal{E}\left[e^{tY}\right] = \mathcal{E}\left[e^{t(\alpha X + \beta)}\right] = \mathcal{E}\left[e^{t\alpha X} e^{\beta t}\right] = e^{\beta t} m_X(\alpha t)$

Twierdzenie: Jeśli X i Y są statystycznie niezależnymi zmiennymi losowymi o fgm odpowiednio $m_X(t)$ i $m_Y(t)$ to fgm zmiennej $X + Y$ dana jest przez

$$m_{X+Y}(t) = m_X(t)\, m_Y(t)$$

Funkcja charakterystyczna

<u>Definicja:</u> Funkcją charakterystyczną zmiennej losowej X nazywamy funkcję zmiennej rzeczywistej t postaci:

$$\varphi_{\mathrm{X}}(t) = \mathcal{E}\left[e^{it\mathrm{X}}\right]$$

<u>Uwaga:</u> Funkcja charakterystyczna jest funkcją rzeczywistą wtedy i tylko wtedy gdy rozkład zmiennej losowej X jest symetryczny względem $x = 0$

<u>Przykład:</u> Funkcja charakterystyczna dla rozkładu dwumianowego

$$\varphi_X(t) = \sum_{k=0}^{n} e^{itk} \binom{n}{k} p^k (1-p)^{n-k} = \left(pe^{it} + 1 - p\right)^n$$

<u>Przykład:</u> Funkcja charakterystyczna dla rozkładu wykładniczego

$$\varphi_X(t) = \int_0^{\infty} e^{itx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda - it)x} dx = \frac{\lambda}{\lambda - it}$$

<u>Twierdzenie:</u> Jeśli spełniony jest warunek $\int_{-\infty}^{+\infty} |\varphi(t)| dt < \infty$ to wówczas mamy:

$$f(x) = F'(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi(t) dt \qquad p_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itx} \varphi(t) dt$$

Funkcje charakterystyczne

Przykład: Znajdź gęstość p-twa zmiennej losowej X, której funkcja charakterystyczna dana jest przez

$$\varphi(t) = \exp(2it - 3|t|)$$

Sprawdzamy czy funkcja charakterystyczna jest całkowalna:

$$\int_{-\infty}^{+\infty} |\varphi(t)| dt = \int_{-\infty}^{+\infty} e^{-3|t|} dt = 2\int_{0}^{\infty} e^{-3t} dt = \frac{2}{3} < \infty$$

Znajdujemy funkcję gęstości p-twa: $f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(-itx + 2it - 3|t|) = \frac{1}{\pi} \frac{3}{(x-2)^2 + 9}$

Przykład: Wyznacz rozkład p-twa, którego funkcja charakterystyczna ma postać

$$\varphi(t) = \exp(2it)$$

Sprawdzamy czy funkcja charakterystyczna jest całkowalna: $\int_{-\infty}^{+\infty} |\varphi(t)| dt = \int_{-\infty}^{+\infty} dt = \infty$

Ponieważ funkcja charakterystyczna jest okresowa (2π) więc mamy do czynienia ze zmienną dyskretną:

$$p_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itk + 2it} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\cos(2-k)t + i\sin(2-k)t \right) dt = \frac{1}{\pi} \frac{\sin(2-k)t}{2-k} \Bigg|_0^{\pi} = \begin{cases} 0 & \text{dla} \quad k \neq 2 \\ 1 & \text{dla} \quad k = 2 \end{cases}$$

Funkcje charakterystyczne

Twierdzenie: Jeśli istnieje k-ty moment zmiennej losowej $\mathbf{X}$ to jej funkcja charakterystyczna jest k-krotnie różniczkowalna i zachodzi związek:

$$m_k = \frac{1}{i^k} \frac{d^k}{dt^k} \varphi_{\mathrm{X}}(t) \Big|_{t=0}$$

Twierdzenie: Jeśli $\varphi_{\mathrm{X}}(t)$ jest f. ch. zmiennej $\mathbf{X}$, to f.ch. zmiennej $\mathbf{Y} = \alpha \mathbf{X} + \beta$ dane jest przez

$$\varphi_{\mathrm{Y}}(t) = e^{i\beta t} \varphi_{\mathrm{X}}(\alpha t)$$

Twierdzenie: Jeśli $\mathbf{X}$ i $\mathbf{Y}$ są statystycznie niezależnymi zmiennymi losowymi o f. ch. odpowiednio $\varphi_{\mathrm{X}}(t)$ i $\varphi_{\mathrm{Y}}(t)$ to f. ch. zmiennej $\mathbf{X}+\mathbf{Y}$ dana jest przez

$$\varphi_{\mathrm{X+Y}}(t) = \varphi_{\mathrm{X}}(t)\, \varphi_{\mathrm{Y}}(t)$$

Przykład: Znajdź rozkład sumy dwóch zmiennych pochodzących z rozkładów Poissona o parametrach μ_1 i μ_2.

$$\varphi(t) = \exp\left(\mu\left(e^{it} - 1\right)\right)$$

$$\Rightarrow \quad \varphi_k(t)\varphi_l(t) = \exp\left(\mu_1\left(e^{it}-1\right)\right)\exp\left(\mu_2\left(e^{it}-1\right)\right) = \exp\left(\left(\mu_1+\mu_2\right)\left(e^{it}-1\right)\right)$$

Warunkowe wartości oczekiwane

Wartością oczekiwaną dyskretnej zmiennej losowej k pod warunkiem, że zmienna losowa m przyjmuje wartość *m* nazywamy wielkość:

$$\mathcal{E}\left[\mathrm{k} \mid \mathrm{m}=m\right] \equiv \mu_{\mathrm{k}}(m) \equiv \sum_k k\, \mathrm{P}\left(\mathrm{k}=k \mid \mathrm{m}=m\right) = \frac{1}{\mathrm{P}_{\bullet,\mathrm{m}}} \sum_k k\, \mathrm{P}_{\mathrm{km}}$$

Wartością oczekiwaną ciągłej zmiennej losowej x pod warunkiem, że zmienna losowa y przyjmuje wartość *y* nazywamy wielkość:

$$\mathcal{E}\left[\mathrm{x} \mid \mathrm{y}=y\right] \equiv \mu_{\mathrm{x}}(y) \equiv \int_{-\infty}^{+\infty} x \cdot f\left(x \mid y\right) dx = \frac{1}{f_2(y)} \int_{-\infty}^{+\infty} x \cdot f\left(x, y\right) dx$$

Podobnie określamy wartość oczekiwaną dyskretnej zmiennej losowej m pod warunkiem, że zmienna losowa k przyjmuje wartość *k*:

$$\mathcal{E}\left[\mathrm{m} \mid \mathrm{k}=k\right] \equiv \mu_{\mathrm{m}}(k) \equiv \sum_m m\, \mathrm{P}\left(\mathrm{m}=m \mid \mathrm{k}=k\right) = \frac{1}{\mathrm{P}_{k,\bullet}} \sum_m m\, \mathrm{P}_{\mathrm{km}}$$

oraz wartość oczekiwaną ciągłej zmiennej losowej y pod warunkiem, że zmienna losowa x przyjmuje wartość *x*:

$$\mathcal{E}\left[\mathrm{y} \mid \mathrm{x}=x\right] \equiv \mu_{\mathrm{y}}(x) \equiv \int_{-\infty}^{+\infty} y \cdot f\left(y \mid x\right) dy = \frac{1}{f_1(x)} \int_{-\infty}^{+\infty} y \cdot f\left(x, y\right) dy$$

Linie regresji I-go rodzaju

Linią regresji I-go rodzaju zmiennej losowej $\mathbf{k}$ ($\mathbf{x}$) względem zmiennej losowej $\mathbf{m}$ ($\mathbf{y}$) nazywamy zbiór punktów o współrzędnych (k,m) w przypadku zmiennych dyskretnych i (x,y) w przypadku zmiennych ciągłych, spełniających równania:

$$k = \mu_{\mathrm{k}}(m) \qquad x = \mu_{\mathrm{x}}(y)$$

Analogicznie, **linią regresji I-go rodzaju** zmiennej losowej $\mathbf{m}$ ($\mathbf{y}$) względem zmiennej $\mathbf{k}$ ($\mathbf{x}$) nazywamy, odpowiednio dla zmiennych dyskretnych i ciągłych, zbiory punktów (k,m) lub (x,y), spełniające równania:

$$m = \mu_{\mathrm{m}}(k) \qquad y = \mu_{\mathrm{y}}(x)$$

Własności:

Średnie odchylenie kwadratowe zmiennej losowej $\mathbf{x}$ od pewnej funkcji $g(y)$ jest najmniejsze, gdy ta funkcja z p-twem 1 jest równa $\mu_x(y)$:

$$\mathcal{E}\left[\left(\mathbf{x}-\mu_{\mathrm{x}}(y)\right)^2\right] = \min_g \mathcal{E}\left[\left(\mathbf{x}-g(y)\right)^2\right]$$

Średnie odchylenie kwadratowe zmiennej losowej $\mathbf{y}$ od pewnej funkcji $f(x)$ jest najmniejsze, gdy ta funkcja z p-twem 1 jest równa $\mu_y(x)$:

$$\mathcal{E}\left[\left(\mathbf{y}-\mu_{\mathrm{y}}(x)\right)^2\right] = \min_g \mathcal{E}\left[\left(\mathbf{y}-f(x)\right)^2\right]$$

Linie regresji II-go rodzaju

Linią regresji II-go rodzaju nazywamy zadaną krzywą $y = h(x; a,b,\dots)$ gdy spełnia ona warunek:

$$\mathcal{E}\left[\left(\mathrm{y}-h\left(x;a,b,\dots\right)\right)^2\right]=\min\left(a,b,\dots\right)$$

W szczególności prostą regresji II-go rodzaju nazywamy linię prostą $y = ax+b$ spełniającą warunek:

$$\mathcal{E}\left[\left(\mathrm{y}-a\,\mathrm{x}-b\right)^2\right]=\mathcal{E}\left[\left(\left(\mathrm{y}-\mu_y\right)-a\left(\mathrm{x}-\mu_x\right)+\mu_y-a\mu_x-b\right)^2\right]=$$

$$=\sigma_y^2+a^2\sigma_x^2-2a\,\mathrm{cov}[\mathrm{x},\mathrm{y}]+\left(\mu_y-a\mu_x-b\right)^2=\sigma_y^2+a^2\sigma_x^2-2a\sigma_x\sigma_y\rho+\left(\mu_y-a\mu_x-b\right)^2=\min\left(a,b\right)$$

Szukamy minimum ze względu na parametry a i b:

$$\begin{cases}\dfrac{\partial}{\partial a}\mathcal{E}\left[\left(\mathrm{y}-a\,\mathrm{x}-b\right)^2\right]=2a\sigma_x^2-2\sigma_x\sigma_y\rho-2\mu_x\left(\mu_y-a\mu_x-b\right)=0\\[2ex]\dfrac{\partial}{\partial b}\mathcal{E}\left[\left(\mathrm{y}-a\,\mathrm{x}-b\right)^2\right]=-2\left(\mu_y-a\mu_x-b\right)=0\end{cases}\quad\Rightarrow\quad\begin{cases}a=\rho\dfrac{\sigma_y}{\sigma_x}\\[2ex]b=\mu_y-a\mu_x\end{cases}$$

Prosta regresji II-go rodzaju zmiennej y względem zmiennej x ma postać:

$$y=\rho\frac{\sigma_y}{\sigma_x}x-\rho\frac{\sigma_y}{\sigma_x}\mu_x+\mu_y$$

Podobnie znajdujemy prostą regresji II-go rodzaju zmiennej x względem y:

$$x=\rho\frac{\sigma_x}{\sigma_y}y-\rho\frac{\sigma_x}{\sigma_y}\mu_y+\mu_x\qquad\Rightarrow\qquad y=\frac{1}{\rho}\frac{\sigma_y}{\sigma_x}x-\frac{1}{\rho}\frac{\sigma_y}{\sigma_x}\mu_x+\mu_y$$

Linie regresji I-go rodzaju

<u>Przykład:</u> Linie regresji I-go rodzaju dla dwuwymiarowego rozkładu normalnego.

$$\mathcal{N}\left(x,y;\mu_x,\mu_y,\sigma_x,\sigma_y\right)=$$

$$=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2+\left(\frac{y-\mu_y}{\sigma_y}\right)^2-2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right)\right]\right\}$$

Rozkłady brzegowe to jednowymiarowe rozkłady normalne:

$$\mathcal{N}_1\left(x;\mu_x,\sigma_x\right)=\frac{1}{\sqrt{2\pi}\sigma_x}\exp\left(-\frac{\left(x-\mu_x\right)^2}{2\sigma_x^2}\right) \qquad \mathcal{E}[\mathrm{x}]=\mu_x \qquad \mathcal{V}[\mathrm{x}]=\sigma_x^2$$

$$\mathcal{N}_2\left(y;\mu_y,\sigma_y\right)=\frac{1}{\sqrt{2\pi}\sigma_y}\exp\left(-\frac{\left(y-\mu_y\right)^2}{2\sigma_y^2}\right) \qquad \mathcal{E}[\mathrm{y}]=\mu_y \qquad \mathcal{V}[\mathrm{y}]=\sigma_y^2$$

Rozkłady warunkowe i warunkowe wartości oczekiwane:

$$g\left(y\,|\,x\right)=\frac{\mathcal{N}\left(x,y;\mu_x,\mu_y,\sigma_x,\sigma_y\right)}{\mathcal{N}_1\left(x;\mu_x,\sigma_x\right)}=\frac{1}{\sqrt{2\pi}\sigma_y\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2\sigma_y^2\left(1-\rho^2\right)}\left[y-(\mu_y+\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x))\right]^2\right)$$

$$g\left(x\,|\,y\right)=\frac{\mathcal{N}\left(x,y;\mu_x,\mu_y,\sigma_x,\sigma_y\right)}{\mathcal{N}_2\left(y;\mu_y,\sigma_y\right)}=\frac{1}{\sqrt{2\pi}\sigma_x\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2\sigma_x^2\left(1-\rho^2\right)}\left[x-(\mu_x+\rho\frac{\sigma_x}{\sigma_y}(y-\mu_y))\right]^2\right)$$

Linie regresji I-go rodzaju

Linie regresji I-go rodzaju to proste o równaniach:

$$y = \mathcal{E}[y \mid x] = \mu_y + \rho \frac{\sigma_y}{\sigma_x}(x - \mu_x)$$

$$x = \mathcal{E}[x \mid y] = \mu_x + \rho \frac{\sigma_x}{\sigma_y}(y - \mu_y) \quad \Rightarrow \quad y = \mu_y + \frac{1}{\rho}\frac{\sigma_y}{\sigma_x}(x - \mu_x)$$

Uwaga: W przypadku dwuwymiarowego rozkładu normalnego, linie regresji I-go rodzaju są jednocześnie prostymi regresji II-go rodzaju.

Linie regresji I-go rodzaju

<u>Przykład:</u> Chcemy umieć przewidzieć wzrost pewnego gatunku żyta na podstawie koncentracji fosforu w glebie. W tym celu badany cztery koncentracje 2, 4, 8, 16 ppm fosforu, i obserwujemy wzrost roślin przy każdej z nich począwszy od nasiona aż do momentu kwitnienia. Następnie dokonujemy pomiaru suchej masy każdej z roślin. Wyniki pomiarów przedstawia tabelka.

Roślina	Fosfor [ppm]	Masa [g]
1	2	4.1
2	2	3.8
3	2	4.0
4	2	3.9
5	4	5.2
6	4	4.9
7	4	5.0
8	4	4.8
9	8	5.7
10	8	5.9
11	8	6.0
12	8	6.2
13	16	11.7
14	16	8.9
15	16	10.1
16	16	10.3

Zmienne losowe: x - koncentrację fosforu, y - masa roślin.

$$\mu_x \equiv \mathcal{E}[x] = \frac{1}{16}(4\times 2 + 4\times 4 + 4\times 8 + 4\times 16) = 7.5$$

$$\mu_y \equiv \mathcal{E}[y] = \frac{1}{16}(4.1+3.8+4.0+3.9+5.2+4.9+5.0+4.8+5.7$$

$$+5.9+6.0+6.2+11.7+8.9+10.1+10.3) = \frac{100.5}{16} \cong 6.28$$

$$\mathcal{E}[x^2] = \frac{1360}{16} = 85 \qquad \mathcal{E}[y^2] = \frac{727.49}{16} = 45.468 \qquad \mathcal{E}[xy] = \frac{957.6}{16} = 59.85$$

$$\mathrm{cov}[x,y] = \mathcal{E}[xy] - \mathcal{E}[x]\mathcal{E}[y] = 59.85 - 7.5\times 6.28 = 12.75$$

$$\mathcal{D}[x] = \sqrt{\mathcal{E}[x^2] - (\mathcal{E}[x])^2} = \sqrt{85 - 7.5^2} = 5.36$$

$$\mathcal{D}[y] = \sqrt{\mathcal{E}[y^2] - (\mathcal{E}[y])^2} = \sqrt{45.468 - 6.28^2} = 2.46$$

$$\rho = \frac{\mathrm{cov}[x,y]}{\mathcal{D}[x]\mathcal{D}[y]} = 0.967$$

$$y = 0.44x + 2.96$$