Estymacja parametrów

Krzysztof Topolski

Wykład 1

Wrocław, 08 lutego 2012

Rozkład zmiennej losowej

Definicja.

Dystrybuantą rozkładu zmiennej losowej X nazywamy funkcję $F_X(x)$ zdefiniowaną dla wszystkich x jako

$$F_X(x) = P(\omega : X(\omega) \leq x).$$

Rozkład zmiennej losowej

Przykład.

Rzucamy monetą aż do momentu wyrzucenia orła. Niech p będzie prawdopodobieństwem wyrzucenia orła w jednym rzucie. Niech X będzie zmienną losową opisującą liczbę rzutów do momentu wyrzucenia orła.

Wtedy

$$P(X = x) = (1 - p)^{x-1} p,$$

Ponieważ zakładamy, że rzuty są niezależne, a aby orzeł pojawił się w rzucie o numerze x musimy najpierw wyrzucić $x - 1$ razy reszki.

Rozkład zmiennej losowej

Przykład.

Rzucamy monetą aż do momentu wyrzucenia orła. Niech p będzie prawdopodobieństwem wyrzucenia orła w jednym rzucie. Niech X będzie zmienną losową opisującą liczbę rzutów do momentu wyrzucenia orła.

Wtedy

$$P(X = x) = (1 - p)^{x-1} p,$$

Ponieważ zakładamy, że rzuty są niezależne, a aby orzeł pojawił się w rzucie o numerze x musimy najpierw wyrzucić $x - 1$ razy reszki.

Przykład.

Rzucamy monetą aż do momentu wyrzucenia orła. Niech p będzie prawdopodobieństwem wyrzucenia orła w jednym rzucie. Niech X będzie zmienną losową opisującą liczbę rzutów do momentu wyrzucenia orła.

Wtedy
$$P(X = x) = (1 - p)^{x-1} p,$$

Ponieważ zakładamy, że rzuty są niezależne, a aby orzeł pojawił się w rzucie o numerze x musimy najpierw wyrzucić $x - 1$ razy reszki.

Przykład cd.

Dla dowolnej liczby naturalnej x,

$$P(X \leq x) = \sum_{i=1}^{x} P(X = i) = \sum_{i=1}^{x} (1-p)^{i-1} p.$$

$$\sum_{k=1}^{n} q^{k-1} = \frac{1-q^n}{1-q}, \quad t \neq 1.$$

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= \frac{1-(1-t)^x}{1-(1-p)} p \\ &= 1-(1-t)^x, \quad x = 1, 2, \ldots . \end{aligned}$$

Przykład cd.

Dla dowolnej liczby naturalnej x,

$$P(X \leq x) = \sum_{i=1}^{x} P(X = i) = \sum_{i=1}^{x} (1-p)^{i-1} p.$$

$$\sum_{k=1}^{n} q^{k-1} = \frac{1-q^n}{1-q}, \quad t \neq 1.$$

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= \frac{1-(1-t)^x}{1-(1-p)} p \\ &= 1-(1-t)^x, \quad x = 1, 2, \ldots . \end{aligned}$$

Przykład cd.

Dla dowolnej liczby naturalnej x,

$$P(X \leq x) = \sum_{i=1}^{x} P(X = i) = \sum_{i=1}^{x} (1-p)^{i-1} p.$$

$$\sum_{k=1}^{n} q^{k-1} = \frac{1-q^n}{1-q}, \quad t \neq 1.$$

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= \frac{1-(1-t)^x}{1-(1-p)} p \\ &= 1-(1-t)^x, \quad x = 1, 2, \ldots . \end{aligned}$$

Twierdzenie.

Następujące warunki są równoważne

- Zmienne losowe X i Y mają ten sam rozkład.
- dla każdego x, zachodzi równość $F_X(x) = F_Y(x)$.

Funkcja gęstości

Definicja.

Funkcją gęstości rozkładu dyskretnej zmiennej losowej X nazywamy funkcję $f_X(x)$ zdefiniowaną dla wszystkich x jako

$$f_X(x) = P(\omega : X(\omega) = x).$$

Przykład

Dla zmiennej losowej o rozkładzie geometrycznym funkcja gęstości ma postać

$$f_X(x) = P(X = x) = \begin{cases} (1-p)^{x-1}p, & \text{dla} \quad x = 1, 2, \ldots \\ 0, & \text{dla pozostałych } x. \end{cases}$$

Funkcja gęstości

Definicja.

Funkcją gęstości rozkładu dyskretnej zmiennej losowej X nazywamy funkcję $f_X(x)$ zdefiniowaną dla wszystkich x jako

$$f_X(x) = P(\omega : X(\omega) = x).$$

Przykład

Dla zmiennej losowej o rozkładzie geometrycznym funkcja gęstości ma postać

$$f_X(x) = P(X = x) = \begin{cases} (1-p)^{x-1}p, & \text{dla} \quad x = 1, 2, \ldots \\ 0, & \text{dla pozostałych}\, x. \end{cases}$$

Funkcja gęstości

Definicja.

Funkcją gęstości rozkładu ciągłej zmiennej losowej X nazywamy funkcję $f_X(x)$, dla której dla wszystkich x

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt.$$

Uwaga.

$$\frac{d}{dx} F_X(x) = f_X(x).$$

Funkcja gęstości

Definicja.

Funkcją gęstości rozkładu ciągłej zmiennej losowej X nazywamy funkcję $f_X(x)$, dla której dla wszystkich x

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt.$$

Uwaga.

$$\frac{d}{dx} F_X(x) = f_X(x).$$

Funkcja gęstości

Przykład

Dla zmiennej losowej o rozkładzie logistycznym

$$F_X(x) = \frac{1}{1 + e^{-x}},$$

funkcja gęstości ma postać

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{e^{-x}}{(1 + e^{-x})^2}.$$

Transformacje zmiennych losowych

Twierdzenie.

Niech X będzie zmienną losową o rozkładzie z gęstością $f_X(x)$ oraz niech $Y = g(X)$, gdzie g jest funkcją ściśle monotoniczną. Załóżmy, że $f_X(x)$ jest funkcją ciągłą oraz, że $g^{-1}(y)$ jest funkcją z ciągłą pochodną. Wtedy gęstość rozkładu zmiennej losowej Y jest postaci

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|.$$

Transformacje zmiennych losowych

Przykład.

Dla ciągłej zmiennej losowej X o dystrybuancie $F_X(x)$ wyznaczmy rozkład zmiennej losowej $Y = X^2$.

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}).$$

Ponieważ zmienna losowa ma ciągły rozkład

$$\begin{aligned} F_Y(y) &= P(-\sqrt{y} < X \leq \sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

Transformacje zmiennych losowych

Przykład.

Dla ciągłej zmiennej losowej X o dystrybuancie $F_X(x)$ wyznaczmy rozkład zmiennej losowej $Y = X^2$.

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}).$$

Ponieważ zmienna losowa ma ciągły rozkład

$$\begin{aligned} F_Y(y) &= P(-\sqrt{y} < X \leq \sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

Przykład.

Dla ciągłej zmiennej losowej X o dystrybuancie $F_X(x)$ wyznaczmy rozkład zmiennej losowej $Y = X^2$.

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}).$$

Ponieważ zmienna losowa ma ciągły rozkład

$$\begin{aligned} F_Y(y) &= P(-\sqrt{y} < X \leq \sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

Przykład cd.

Gęstość rozkładu zmiennej losowej Y otrzymamy różniczkując jej dystrybuantę

$$F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) \\ &= \frac{d}{dy}[F_X(\sqrt{y}) - F_X(-\sqrt{y})] \\ &= \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}). \end{aligned}$$

Funkcja gęstości ma postać

$$f_Y(y) = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})].$$

Przykład cd.

Gęstość rozkładu zmiennej losowej Y otrzymamy różniczkując jej dystrybuantę

$$F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) \\ &= \frac{d}{dy}[F_X(\sqrt{y}) - F_X(-\sqrt{y})] \\ &= \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}). \end{aligned}$$

Funkcja gęstości ma postać

$$f_Y(y) = \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})].$$

Wartość oczekiwana

Definicja.

Wartość oczekiwaną zmiennej losowej X, oznaczoną przez EX, i gdy X jest dyskretną zmienną losową definiujemy jako

$$EX = \sum_x x\, P(X = x) = \sum_x x\, f_X(x),$$

natomiast gdy X jest ciągłą zmienną losową definiujemy jako

$$EX = \int_{-\infty}^{\infty} x\, f_X(x)\, dx.$$

Wartość oczekiwana

Przykład.

Niech X będzie zmienną losową o rozkładzie wykładniczym z gęstością rozkładu

$$f_X(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}}, \quad 0 \leq x < \infty, \; \lambda > 0.$$

Przykład cd.

Wartość oczekiwana EX dla zmiennej losowej o rozkładzie wykładniczym jest równa

$$EX = \int_{-\infty}^{\infty} x\, f_X(x) = \int_0^{\infty} x \frac{1}{\lambda} e^{-\frac{x}{\lambda}}\, dx$$

(całkując przez części)

$$= -xe^{-\frac{x}{\lambda}} \Big|_0^{\infty} + \int_0^{\infty} e^{-\frac{1}{\lambda}x}\, dx$$

$$= \int_0^{\infty} e^{-\frac{1}{\lambda}x}\, dx = \lambda$$

Przykład cd.

Wartość oczekiwana EX dla zmiennej losowej o rozkładzie wykładniczym jest równa

$$\begin{aligned} EX = \int_{-\infty}^{\infty} x\, f_X(x) &= \int_0^{\infty} x \frac{1}{\lambda} e^{-\frac{x}{\lambda}}\, dx \\ \text{(całkując przez części)} \\ &= -xe^{-\frac{x}{\lambda}} \Big|_0^{\infty} + \int_0^{\infty} e^{-\frac{1}{\lambda}x}\, dx \\ &= \int_0^{\infty} e^{-\frac{1}{\lambda}x}\, dx = \lambda \end{aligned}$$

Fakt.

Dla funkcji $g(x)$, dla której istnieje $Eg(X)$

$$Eg(X) = \sum_i g(i)P(X = i) = \sum_i g(i)f_X(i),$$

gdy X jest dyskretną zmienną losową

$$Eg(X) = \int_{-\infty}^{\infty} g(x)f_X(x),$$

gdy X jest ciągłą zmienną losową.

Własności wartości oczekiwanej

Twierdzenie.

Niech X będzie zmienną losową natomiast a, b i c stałymi. Wtedy dla dowolnych funkcji $g(x)$ i $f(x)$, dla których istnieją wartości oczekiwane

$$E(a\,g(X) + b\,f(X) + c) = a\,Eg(X) + b\,Ef(X) + c.$$

Jeśli dla wszystkich x, $g(x) \geq 0$ to $Eg(X) \geq 0$.

Jeśli dla wszystkich x, $a \leq g(X) \leq b$ to $a \leq Eg(X) \leq b$.

Definicja.

Momentem rzędu n rozkładu zmiennej losowej X nazywamy

$$\mu_n^{'} = EX^n.$$

Momentem centralnym rzędu n zmiennej losowej X nazywamy

$$\mu_n = E(X - EX)^n.$$

Wariancja

Definicja.

Moment centralny rzędu 2 rozkładu zmiennej losowej X nazywamy wariancją

$$Var\,X = E(X - EX)^2,$$

lub co jest równoważne

$$Var\,X = EX^2 - [EX]^2.$$

Wariancja

Twierdzenie.

Jeśli zmienna losowa X ma skończoną wariancję to dla dowolnych stałych a i b,

$$Var(a\,X + b) = a^2\,Var\,X.$$

Jeśli zmienne losowe X i Y są niezależne i mają skończone wariancje to

$$Var(X + Y) = Var\,X + Var\,Y.$$

$$Var(a\,X + b\,Y + c) = a^2\,Var\,X + b^2\,Var\,Y.$$

Wariancja

Twierdzenie.

Jeśli zmienna losowa X ma skończoną wariancję to dla dowolnych stałych a i b,

$$Var(a\,X + b) = a^2\, Var\, X.$$

Jeśli zmienne losowe X i Y są niezależne i mają skończone wariancje to

$$Var(X + Y) = Var\, X + Var\, Y.$$

$$Var(a\,X + b\,Y + c) = a^2\, Var\, X + b^2\, Var\, Y.$$