

Ostatni raz wygenerowano: 13 czerwca 2023. Kontakt: thasiow(at)onet.pl

1 Podstawowe pojęcia

1. Kwantyl

Kwantylem rzędu p , gdzie $0 \leq p \leq 1$, w rozkładzie empirycznym P_X zmiennej losowej X nazywamy taką wartość zmiennej losowej x_p , dla której spełnione są nierówności

$$P_X((-\infty, x_p]) \geq p$$

oraz

$$P_X([x_p, \infty)) \geq 1 - p.$$

W szczególności, kwantylem rzędu p jest taka wartość x_p zmiennej losowej, że wartości mniejsze lub równe od x_p są przyjmowane z prawdopodobieństwem co najmniej p , zaś wartości większe lub równe od x_p są przyjmowane z prawdopodobieństwem co najmniej $1 - p$.

2. Statystyki pozycyjne

Niech $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ będzie ciągiem zmiennych losowych powstałych z X_1, \dots, X_n , po ich uporządkowaniu w ciąg niemalejący. Zmienną $X_{k:n}$, $k = 1, \dots, n$, nazywamy k -tą statystyką pozycyjną. W szczególności

$$X_{1:n} = \min(X_1, \dots, X_n)$$

$$X_{n:n} = \max(X_1, \dots, X_n)$$

k -ta statystyka pozycyjna ma dystrybuantę:

$$F_{k:n}(x) = P(X_{k:n} \leq x) = \sum_{i=k}^n \binom{n}{i} F(x)^i (1 - F(x))^{n-i}$$

3. Transformacja zmiennych losowych

Niech X - zmienna o gęstości f . Żeby obliczyć gęstość zmiennej Y zdefiniowanej jako:

$$Y = g(x)$$

używamy następującego wzoru:

$$g(y) = f(h(y)) |h'(y)|$$

gdzie h - funkcja odwrotna do g (tzn. $h(g(t)) = t$).

Dla przykładu jeżeli $Y = aX + b$ to

$$g(y) = f\left(\frac{y-b}{a}\right) \frac{1}{|a|}$$

4. Momenty i współczynniki

Moment zwykły rzędu k :

$$m_k = E(X^k) = \int_{-\infty}^{\infty} x^k f(x) dx$$

Moment centralny rzędu k

$$\mu_k = E((X - (EX))^k)$$

Współczynnik asymetrii:

$$A = \frac{M_3}{\sigma^3}$$

gdzie M_3 - 3 moment centralny, σ - odchylenie standardowe

$$A = \frac{M_4}{\sigma^4} - 3$$

Własności:

(1) jeżeli X i Y niezależne to 3 moment centralny ich sumy równa się sumie trzecich momentów centralnych. Czyli $E((X + Y - E(X + Y))^3) = E((X - EX)^3) + E((Y - EY)^3)$

5. Funkcja Gamma

$$\Gamma(n) = (n - 1)!$$

$$\Gamma(1/2) = \sqrt{\pi}$$

$$\Gamma(1) = 1$$

$$\Gamma(n + 1/2) = \frac{(2n - 1)!!}{2^n} \sqrt{\pi}$$

Silnia podwójna z krokiem dwa np. $9!! = 9 \cdot 7 \cdot 5 \cdot 3 \cdot 1$

6. Błąd średniokwadratowy estymatora

$$\text{MSE}(\hat{\theta}) = E((\hat{\theta} - \theta)^2)$$

7. Twierdzenia graniczne

SUMA DUŻEJ LICZBY ZMIENNYCH LOSOWYCH Z JEDNAKOWEGO ROZKŁADU MA ROZKŁAD NORMALNY

$$\lim_{n \rightarrow \infty} \left(1 + \frac{k}{n}\right)^n = e^k$$

2 Rozkłady

Rozkład Poissona

$$f(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$E(X) = \lambda$$

$$Var(X) = \lambda$$

Rozkład jednostajny ciągły na przedziale $[a, b]$

$$f(x) = \frac{1}{b-a}$$

$$F(x) = \frac{x-a}{b-a}$$

$$E(X) = \frac{a+b}{2}$$

$$Var(X) = \frac{(b-a)^2}{12}$$

Dla rozkładu jednostajnego na przedziale $[0, \theta]$ estymator największej wiarygodności to: $\hat{\theta} = \max\{x_1, \dots, x_n\}$

Rozkład Gamma

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad x > 0$$

$$E(X) = \frac{\alpha}{\beta}$$

$$Var(X) = \frac{\alpha}{\beta^2}$$

(1) gdy $X \sim Gamma(a_1, b)$, $Y \sim Gamma(a_2, b)$ to $X+Y \sim Gamma(a_1+a_2, b)$

(2) gdy $\alpha = 1$ to mamy rozkład wykładniczy $Exp(\beta)$

(3) gdy $\alpha = \frac{n}{2}$ oraz $\beta = \frac{1}{2}$ to mamy rozkład ch-kwadrat z n stopniami swobody

Rozkład χ^2 to szczególny przypadek rozkładu Gamma zachodzi:

$$\chi^2(n) + \chi^2(k) = \chi^2(n+k)$$

(4) gdy $X \sim Gamma(a, b) \rightarrow 2 \cdot b \cdot X \sim \chi^2(n)$, $n = 2 \cdot a$

(5) gdy $X \sim Gamma(a, b) \rightarrow kX \sim Gamma(a, \frac{b}{k})$

Dla przypomnienia $F(n) = (n-1)!$, $F(n+1) = n!$

Rozkład wykładniczy

$$f(x) = \beta e^{-\beta x}$$

$$E(X) = \frac{1}{\beta}$$

$$VAR(X) = \frac{1}{\beta^2}$$

Własności:

(1) gdy $X \sim \text{Exp}(\beta)$ to $kX \sim \text{Exp}(\beta/k)$.

(2) gdy $X \sim \text{Exp}(\beta)$ i $Y \sim \text{Exp}(\beta)$ to $X + Y \sim G(1+1, \beta)$

(3) jeżeli $X_i \sim \text{Exp}(\beta)$ to $\min(X_1, \dots, X_n) \sim \text{Exp}(\beta_1 + \dots + \beta_n)$

Rozkład ujemny dwumianowy

$$P(X = k) = \binom{k+r-1}{k} \cdot (1-p)^r p^k,$$

$$EX = \frac{pr}{1-p}$$

$$Var X = \frac{pr}{(1-p)^2}$$

Rozkład ujemny dwumianowy 2 (W. Wołyński załącznik)

$$p_k = \frac{\Gamma(r+k)}{\Gamma(r)k!} p^r (1-p)^k$$

$$E(X) = \frac{r(1-p)}{p}$$

$$VAR(X) = \frac{r(1-p)}{p^2}$$

gdzie

$$\frac{\Gamma(r+k)}{\Gamma(r)k!} = \frac{(r+k-1)!}{(r-1)!k!} = \binom{r+k-1}{k}$$

Rozkład dwumianowy

$$p_k = \binom{n}{k} p^k (1-p)^{n-k}$$

gdzie $k = 0, 1, \dots, n$

$$E(X) = np$$

$$Var(X) = np(1 - p)$$

Rozkład normalny

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Rozkład log-normalny

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}$$

$$E(X) = e^{\mu + \frac{\sigma^2}{2}}$$

$$Var(X) = (e^{\sigma^2} - 1) e^{2\mu + \sigma^2}$$

Rozkład geometryczny

$$P(X = k) = (1 - p)^{k-1} \cdot p$$

dla $k = 1, 2, 3, \dots$ Interpretowany jako p-p pierwszego sukcesu w k-tej próbie

$$E(X) = \frac{1}{p}$$

$$Var(X) = \frac{1 - p}{p^2}$$

Rozkład geometryczny 2

$$P(X = k) = (1 - p)^k \cdot p$$

dla $k = 0, 1, 2, 3, \dots$ Interpretowany jako liczba porażek z rzędu przed pojawieniem się sukcesu

$$E(X) = \frac{1 - p}{p}$$

$$Var(X) = \frac{1 - p}{p^2}$$

Rozkład F Snedecora

Jeżeli X i Y są niezależne oraz $X \sim \chi^2(n_1)$ i $Y \sim \chi^2(n_2)$, to:

$$\frac{\frac{X}{n_1}}{\frac{Y}{n_2}} \sim F(n_1, n_2)$$

Własności rozkładów

1. Jeżeli Z i Y są niezależnymi zmiennymi losowymi przy czym $Z \sim N(0, 1)$ i $Y \sim \chi^2(n)$ to

$$T = \frac{Z}{\sqrt{\frac{Y}{n}}} \sim T(n)$$

czyli zmienna T ma rozkład t-studenta

2.

$$\frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim T(n-1)$$

gdzie $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$

3.

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$$

$$\frac{1}{\sigma^2} \sum (X_i - \bar{X})^2 \sim \chi^2(n-1)$$

3 Wartość oczekiwana

$$E(g(X_1, \dots, X_n)) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(X_1, \dots, X_n) \cdot f(X_1, \dots, X_n) dx_1 \dots dx_n$$

4 Kowariancja

$$COV(S, T) = E(S \cdot T) - ES \cdot ET$$

5 Warunkowa wartość oczekiwana

Gęstość warunkowa:

Dla wektora losowego (X, Y) gęstość warunkową zmiennej losowej X przy warunku $Y = y$ nazywamy funkcję

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

Warunkowa wartość oczekiwana:

$$E(X|Y = y) = \int_{-\infty}^{\infty} x f(x|y) dx = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} x f(x, y) dx = \frac{\int_{-\infty}^{\infty} x f(x, y) dx}{\int_{-\infty}^{\infty} f(x, y) dx}$$

<https://www.youtube.com/watch?v=Ki2HpTCPwhM>

Iteracyjność oraz Law of total expectation:

Iteracyjność (law of iterated expectation)

$$E(E(Y|X)) = EY$$

Gdy

$$F_1 \subset F_2 \subset M \rightarrow E(Y|F_1) = E(E(Y|F_2)|F_1) = E(E(Y|F_1)|F_2)$$

Uogólniona iteracyjność:

$$E(Y|X) = E(E(Y|Z)|X)$$

One special case states that if A_1, A_2, \dots, A_n is a partition of the whole outcome space, i.e. these events are mutually exclusive and exhaustive, then

$$E(X) = \sum_{i=1}^n E(X | A_i) P(A_i).$$

Others

$$E(X + Y|X) = X + EY$$

$$E(XY|X) = X \cdot EY$$

The following formulation of the law of iterated expectations plays an important role in many economic and finance models:

$$E(X | I_1) = E(E(X | I_2) | I_1),$$

6 Testowanie hipotez statystycznych

Testowanie hipotez statystycznych - na podstawie WYKŁADY ZE STATYSTYKI MATEMATYCZNEJ (II rok WNE), Agata Boratyńska (s.76).

Niech:

K - zbiór krytyczny, zbiór wyników obserwacji przy których odrzucamy H_0

A - zbiór afirmacji, zbiór wyników, przy których nie odrzucamy H_0

Jeżeli mamy K to mamy test statystyczny:

$$K = \{T(x) > c\},$$

,co oznacza: odrzucić H_0 , gdy obliczona wartość funkcji $T(x)$ jest większa niż c .
 T to statystyka testowa, c to wartość krytyczna.

Błąd I rodzaju: **odrzućenie H_0 , gdy jest prawdziwa**

Błąd II rodzaju: **nie odrzucenie H_0 , gdy jest fałszywa**

Uwaga: zmniejszenie prawdopodobieństwa błędu I rodzaju powoduje wzrost prawdopodobieństwa błędu II rodzaju.

Prawdopodobieństwo błędu I rodzaju:

$$P_\theta(K), \quad \theta \in \Theta_0$$

gdzie $H_0 : \theta \in \Theta_0, H_1 : \theta \in \Theta_1$.

Prawdopodobieństwo błędu II rodzaju:

$$P_\theta(A) = 1 - P_\theta(K) \quad \theta \in \Theta_1$$

Test jest na poziomie α , jeśli:

$$\forall_{\theta \in \Theta_0} P_\theta(K) \leq \alpha$$

Wielkość $P_\theta(K)$ nazywamy **mocą testu** przy alternatywie $\theta \in \Theta_1$.

Moc testu statystycznego rozumiana jest jako prawdopodobieństwo nieodrzućenia H_1 przy jej prawdziwości (odrzućenie H_0 przy jej fałszywości).

Test o obszarze krytycznym K_1 jest mocniejszy niż test o obszarze krytycznym K_2 gdy

$$\forall_{\theta \in \Theta_0} P_\theta(K_1) \leq \alpha \text{ i } P_\theta(K_2) \leq \alpha$$

$$\forall_{\theta \in \Theta_1} P_\theta(K_1) \geq P_\theta(K_2)$$

co oznacza, że jeżeli H_0 jest prawdziwe to większe prawdopodobieństwo odrzucenia gdy jest fałszywa.

$$\exists_{\theta_1 \in \Theta_1} P_{\theta_1}(K_1) > P_{\theta_1}(K_2)$$

Test jednostajnie najmocniejszy

$$\forall_{K \subset X} P_\theta(K) \leq \alpha \quad \text{gdy } \theta \in \Theta_0$$

zachodzi:

$$\forall_{\theta \in \Theta_1} P_\theta(K^*) \geq P_\theta(K)$$

Test jednostajnie najmocniejszy liczymy przez badanie ilorazu funkcji wiarygodności (bez pochodnych, przekształcamy do skutku). Test oparty na ilorazie wiarygodności potrzebuje pochodnych.

7 Metoda największej wiarygodności

Funkcja wiarygodności:

$$L(x_1, \dots, x_n; \theta_1, \dots, \theta_r) = \prod_{i=1}^n f(x_i; \theta_1, \dots, \theta_r)$$

Przykład wyznaczania parametrów metodą największej wiarygodności:
Niech

$$f(x; \mu; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Wyznaczamy L :

$$L(x_1, \dots, x_N; \mu, \sigma^2) = \prod_{i=1}^N f(x_i, \mu, \sigma)$$

Następnie wyznaczmy $\ln L$ (funkcje $\ln L(\theta)$ i $L(\theta)$ osiągają maksimum dla tej samej wartości, a często zamiast $L(\theta)$ wygodniej jest używać logarytmu funkcji wiarygodności).

$$\ln L(x_1, \dots, x_N; \mu, \sigma^2) = \sum_{i=1}^N \ln f(x_i; \mu, \sigma^2)$$

$$\ln f(x; \mu; \sigma^2) = -\ln \sigma\sqrt{2\pi} - \frac{(x-\mu)^2}{2\sigma^2}$$

Maksymalizujemy L ze względu na μ i mamy:

$$0 = \frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^N \frac{\partial \ln f(x, \mu, \sigma^2)}{\partial \mu} = \sum_{i=1}^N \frac{x_i - \mu}{\sigma^2} = \frac{1}{\sigma^2} \left(\left(\sum_{i=1}^N x_i \right) - N \cdot \mu \right)$$

Stąd estymatorem μ jest:

$$m = \frac{1}{N} \sum_{i=1}^N x_i$$

Postępując w analogiczny sposób dla σ^2 otrzymujemy estymator dla wariancji:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2$$

8 Przedziały ufności

Niech cecha X ma rozkład w populacji z nieznanym parametrem θ . Z populacji wybieramy próbę losową (X_1, X_2, \dots, X_n) . Przedziałem ufności o współczynniku ufności $1 - \alpha$ nazywamy taki przedział (θ_1, θ_2) , który spełnia warunek:

$$P(\theta_1 < \theta < \theta_2) = 1 - \alpha$$

gdzie θ_1 i θ_2 są funkcjami wyznaczonymi na podstawie próby losowej.

Przedział ufności dla wartości oczekiwanej

Jeśli X jest próbą prostą z rozkładu normalnego z nieznanym parametrem μ i znanym σ , to przedział ufności dla μ na poziomie $1 - \alpha$ ma postać:

$$\bar{X}_n \pm u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

czyli

$$P\left(\bar{X}_n - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X}_n + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

gdzie u_p oznacza kwantyl rzędu p w rozkładzie normalnym $N(0, 1)$.

Gdy σ jest nieznane, do oszacowania μ używamy przedziału

$$\bar{X}_n \pm t_{1-\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}$$

gdzie $t_p(n-1)$ jest kwantylem rzędu p w rozkładzie t-Studenta z $n-1$ stopniami swobody i $S_n = \sqrt{S_n^2}$, gdzie S_n^2 jest dane wzorem:

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

9 Estymatory

9.1 Estymatory uzyskane metodą największych kwadratów

$$a = \frac{SS_{XY}}{SS_X} = \frac{Cov(X, Y)}{Var(X)}$$

$$b = \bar{y} - a \cdot \bar{x}$$

$$SS_{XY} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$SS_X = \sum_{i=1}^n (x_i - \bar{x})^2$$

$$SS_Y = \sum_{i=1}^n (y_i - \bar{y})^2$$

9.2 Nieobciążoność

Estymator jest nieobciążony, jeśli wartość oczekiwana rozkładu estymatora jest równa wartości szacowanego parametru:

$$E(\hat{\theta}) = \theta$$

Jeśli różnica pomiędzy wartością oczekiwaną rozkładu estymatora a wartością szacowanego parametru jest zależna funkcyjnie od estymatora:

$$E(\hat{\theta}) - \theta = b(\hat{\theta})$$

to estymator nazywamy obciążonym, zaś samą różnicę nazywamy obciążeniem estymatora.

9.3 Asymptotyczna nieobciążoność

Estymator nazywamy asymptotycznie nieobciążonym, jeśli obciążenie estymatora dąży do zera przy rosnącej liczbie próby:

$$\lim_{n \rightarrow \infty} b(\hat{\theta}) = 0$$

Każdy estymator nieobciążony jest estymatorem asymptotycznie nieobciążonym.

9.4 Zgodność

Estymator nazywamy zgodnym, jeśli jest stochastycznie zbieżny do szacowanego parametru:

$$\bigwedge_{\epsilon > 0} \lim_{n \rightarrow \infty} P\{|\hat{\theta} - \theta| < \epsilon\} = 1$$

Oznacza to, że jeśli rośnie liczba próby, rośnie też prawdopodobieństwo, że oszacowanie przy pomocy estymatora będzie przyjmować wartości coraz bliższe wartości szacowanego parametru. Inaczej: zwiększając liczbę próby, zmniejszamy ryzyko popełnienia błędu większego niż pewna ustalona wielkość.

9.5 Efektywność

Spośród zbioru wszystkich nieobciążonych estymatorów $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r$ najefektywniejszym nazywamy estymator o najmniejszej wariancji. Definicja ta jest bardzo niewygodna, ponieważ do wyznaczenia najefektywniejszego estymatora potrzebna jest znajomość wariancji wszystkich estymatorów nieobciążonych danego parametru rozkładu. W praktyce o wiele łatwiej jest skorzystać z nierówności Rao-Cramera.

9.6 Asymptotyczna efektywność

Estymator $\hat{\theta}$ jest asymptotycznie najefektywniejszy, jeśli przy wzrastającej liczbie próby wariancja estymatora $\hat{\theta}$ dąży do wariancji estymatora najefektywniejszego $\hat{\theta}^*$:

$$\lim_{n \rightarrow \infty} \frac{D^2(\hat{\theta})}{D^2(\hat{\theta}^*)} = 1$$

gdzie $D^2(\hat{\theta})$ oznacza wariancję estymatora.

9.7 Dostateczność

Estymator $\hat{\theta}$ jest dostateczny, jeśli można, ze względu na niego dokonać faktoryzacji (rozłożenia na iloczyn) łącznej funkcji gęstości $f(x)$ wektora wyników próby $X = (X_1, X_2, \dots, X_n)$

$$f(x, \theta) = f(\hat{\theta}, \theta)g(x, \hat{\theta})$$

gdzie $g(x, \hat{\theta})$ jest funkcją niezależną od parametru θ .

9.8 Funkcja wiarygodności

$$L(x_1, \dots, x_n; \theta_1, \dots, \theta_r) = \prod_{i=1}^n f(x_i; \theta_1, \dots, \theta_r)$$