

Estymacja przedziałowa

Przykład wstępny. Model gaussowski ze znaną wariancją

$$(\mathbb{R}, \{N(\mu, \sigma^2), \mu \in \mathbb{R}\})$$

Zadanie polega na oszacowaniu wartości średniej μ .

Jakie wartości obserwacji X są „wysoce prawdopodobne” dla różnych μ ?

Niech „wysoce prawdopodobne” wynosi 0.9. Szukamy takich $a(\mu)$ oraz $b(\mu)$, że

$$P_\mu\{X \in (a(\mu), b(\mu))\} = 0.9, \forall \mu \in \mathbb{R}$$

Rozwiązanie

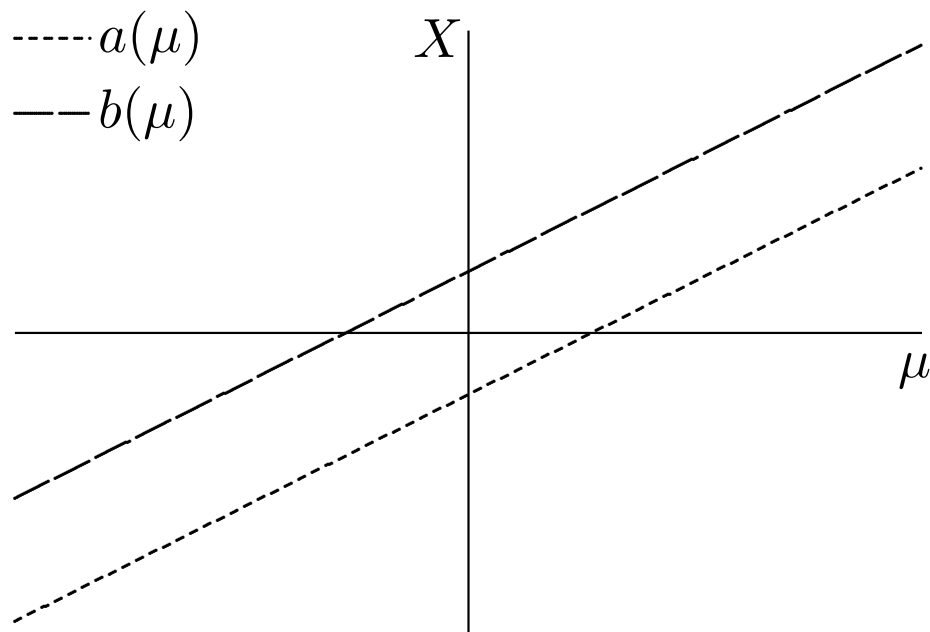
$$P_\mu\{X \in (a(\mu), b(\mu))\} = \Phi\left(\frac{b(\mu) - \mu}{\sigma}\right) - \Phi\left(\frac{a(\mu) - \mu}{\sigma}\right)$$

Niech z_{α_1} oraz z_{α_2} będą takie, że

$$\Phi(z_{\alpha_2}) - \Phi(z_{\alpha_1}) = 0.9$$

$$b(\mu) = \mu + \sigma z_{\alpha_2}, \quad a(\mu) = \mu + \sigma z_{\alpha_1}$$

$$b(\mu) - a(\mu) = \min! \Rightarrow z_{\alpha_2} \stackrel{?}{=} -z_{\alpha_1} = z_{0.95}$$



$$X \in (\mu + \sigma z_{\alpha_1}, \mu + \sigma z_{\alpha_2})$$

⇓

$$\mu \in (X - \sigma z_{\alpha_2}, X - \sigma z_{\alpha_1})$$

Zatem

$$P_{\mu}\{X \in (a(\mu), b(\mu))\} = 0.9, \forall \mu \in \mathbb{R}$$

⇓

$$P_{\mu}\{\mu \in (X - \sigma z_{\alpha_2}, X - \sigma z_{\alpha_1})\} = 0.9, \forall \mu \in \mathbb{R}$$

.....

Definicja 9.1 Przedział $(\underline{\theta}(X), \bar{\theta}(X))$ taki, że

$$P_{\theta}\{\theta \in (\underline{\theta}(X), \bar{\theta}(X))\} = 1 - \alpha, \quad \forall \theta \in \Theta$$

nazywamy **przedziałem ufności** na **poziomie ufności** $1 - \alpha$. \square

Definicja 9.2 Zbiór $S_X \subset \Theta$ taki, że

$$P\{\theta \in S_X\} = 1 - \alpha, \quad \forall \theta \in \Theta$$

nazywamy **obszarem ufności** na **poziomie ufności** $1 - \alpha$. \square

Konstrukcja obszaru ufności

1. W przestrzeni $\mathcal{X} \times \Theta$ wyznaczyć taki zbiór S , że dla każdego θ

$$P_{\theta}\{(\mathcal{X} \times \{\theta\}) \cap S\} = 1 - \alpha$$

Wówczas dla danego $x \in \mathcal{X}$ zbiór

$$(\{x\} \times \Theta) \cap S$$

jest zbiorem ufności na poziomie ufności $1 - \alpha$

2. Skonstruować taką funkcję $t(x, \theta)$, że przy wartości parametru θ jej rozkład nie zależy od θ oraz (jeżeli $\Theta \subset \mathbb{R}$) jest ona monotoniczną funkcją θ . Taka funkcja nazywa się **funkcją centralną**. Wówczas można wyznaczyć takie dwie liczby, że

$$P_{\theta}\{t_1 < t(x, \theta) < t_2\} = 1 - \alpha$$

Ze względu na monotoniczność funkcji centralnej

$$t_1 < t(x, \theta) < t_2 \Leftrightarrow \underline{\theta}(X) < \theta < \bar{\theta}(X)$$

Dodatkowe kryteria

1. Jak najmniejszy zbiór ufności. Jeżeli $\Theta \subset \mathbb{R}$ oraz $t(x, \theta)$ jest funkcją centralną, to

$$\bar{\theta}(X) - \underline{\theta}(X) = \min!$$

2. Zbiór ufności o zadanej wielkości. Jeżeli $\Theta \subset \mathbb{R}$, $t(x, \theta)$ jest funkcją centralną oraz $d > 0$, to

$$\bar{\theta}(X) - \underline{\theta}(X) < d$$

Przykład (Model dwumianowy).

$$(\{0, 1, \dots, n\}, \{B(n, \theta), \theta \in [0, 1]\})$$

Zaobserwowano X sukcesów.

Skonstruować przedział ufności $(\theta_1(X), \theta_2(X))$ dla θ na poziomie ufności 0.95

$$P_\theta\{\theta_1(X) < \theta < \theta_2(X)\} = 0.95, \quad \forall \theta \in [0, 1]$$

$\theta_1(X), \theta_2(X)$ są takie, że

$$\binom{n}{X} \theta_1^X (1 - \theta_1)^{n-X} = \gamma_1 (= 0.025)$$

$$\binom{n}{X} \theta_2^X (1 - \theta_2)^{n-X} = \gamma_2 (= 0.975)$$

$$\gamma_2 - \gamma_1 = 0.95$$

Rozwiązanie numeryczne!

.....

Przykład (Model gaussowski).

Przedział ufności dla μ

Model dla próby X_1, X_2, \dots, X_n :

$$(\mathbb{R}^n, \{N(\mu \mathbf{1}_n, \sigma^2 \mathbf{I}_n), \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\})$$

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

Znane fakty:

1. $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$
2. $\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi_{n-1}^2$
3. \bar{X} oraz $\sum_{i=1}^n (X_i - \bar{X})^2$ są niezależne
4. Jeżeli $\xi \sim N(0, 1)$, $\eta \sim \chi_v^2$, ξ i η są niezależne, to $t = \xi / \sqrt{\eta/v} \sim t_v$.

Funkcja

$$t_{n-1} = \frac{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 / (n-1)}} = \frac{\bar{X} - \mu}{S} \sqrt{n}$$

jest funkcją centralną.

Niech t_1, t_2 będą takimi liczbami, że

$$P\{t_1 < t_{n-1} < t_2\} = 1 - \alpha$$

Przedział ufności dla μ

$$\left(\bar{X} - t_2 \frac{S}{\sqrt{n}}; \bar{X} - t_1 \frac{S}{\sqrt{n}} \right)$$

Przedział jest najkrótszy, jeżeli $t_1 = -t_2$, czyli wybieramy taką liczbę $t(\alpha; n - 1)$, że

$$P\{|t_{n-1}| < t(\alpha; n - 1)\} = 1 - \alpha$$

Przedział ufności dla μ

$$\left(\bar{X} - t(\alpha; n - 1) \frac{S}{\sqrt{n}}; \bar{X} + t(\alpha; n - 1) \frac{S}{\sqrt{n}} \right)$$

.....

Przykład (Model gaussowski).

Przedział ufności dla σ^2

Model dla próby X_1, X_2, \dots, X_n :

$$(\mathbb{R}^n, \{N(\mu \mathbf{1}_n, \sigma^2 \mathbf{I}_n), \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\})$$

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

Funkcja

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

jest funkcją centralną.

Niech c_1, c_2 będą takimi liczbami, że

$$P\{c_1 < \chi_{n-1}^2 < c_2\} = 1 - \alpha$$

Przedział ufności dla σ^2

$$\left(\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{c_2}, \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{c_1} \right)$$

1. Taki wybór granic, że

$$P\{\chi_{n-1}^2 < c_1\} = \alpha/2 \text{ oraz } P\{\chi_{n-1}^2 > c_2\} = \alpha/2$$

Oznaczenia

$$c_1 = \chi^2 \left(1 - \frac{\alpha}{2}; n - 1\right); \quad c_2 = \chi^2 \left(\frac{\alpha}{2}; n - 1\right)$$

2. Przedział jest najkrótszy, tzn. $\frac{1}{c_1} - \frac{1}{c_2} = \min!$
($f_{n-1}(x)$ oznacza gęstość rozkładu)

$$\int_{c_1}^{c_2} f_{n-1}(x) dx = 1 - \alpha$$

$$\frac{1}{c_1} - \frac{1}{c_2} = \min!$$

Rozwiązanie

$$\int_{c_1}^{c_2} f_{n-1}(x) dx = 1 - \alpha$$

$$c_2^2 f_{n-1}(c_2) = c_1^2 f_{n-1}(c_1)$$

.....