Wykład 12

1 Estymacja przedziałowa.

DEFINICJA. *Estymatorem przedziałowym parametru rzeczywistego θ nazywamy parę funkcji $L(x_1, x_2, ..., x_n)$ i $U(x_1, x_2, ..., x_n)$ próby $x_1, x_2, ..., x_n$ spełniającą dla wszystkich $x = (x_1, x_2, ..., x_n)$ nierówność*

$$L(x) \leqslant U(x).$$

Dla x, realizacji próby losowej $(X_1, ..., X_n)$, oszacowanie przedziałowe parametru θ mówi, że

$$L(x) \leqslant \theta \leqslant U(x).$$

Przedział losowy $[L(X), U(X)]$ nazywamy estymatorem przedziałowym parametru θ.

PRZYKŁAD. Dla próby losowej $(X_1, ..., X_4)$ z rozkładu $N(\mu, 1)$ przedział $[\bar{X} - 1, \bar{X} + 1]$ jest estymatorem przedziałowym parametru μ.

PRZYKŁAD CD. Gdy estymujemy parametr μ przy pomocy średniej próbkowej $\bar{X}$ to

$$P(\bar{X} = \mu) = 0.$$

Dla estymatrora przedziałowego $[\bar{X} - 1, \bar{X} + 1]$ mamy natomiast

$$\begin{aligned} P\Big(\mu \in [\bar{X} - 1, \bar{X} + 1]\Big) &= P\Big(\bar{X} - 1 \leqslant \mu \leqslant \bar{X} + 1\Big) &(1)\\ &= P\Big(-1 \leqslant \mu - \bar{X} \leqslant 1\Big) &(2)\\ &= P\Big(-1 \leqslant \bar{X} - \mu \leqslant 1\Big) &(3)\\ &= P\left(-2 \leqslant \frac{\bar{X} - \mu}{\sqrt{\frac{1}{4}}} \leqslant 2\right) &(4)\\ &= P\left(-2 \leqslant Z \leqslant 2\right) = 0,9544... &(5) \end{aligned}$$

Skorzystaliśmy z faktu, że $\mu - \bar{X}$ ma taki sam rozkład jak $\mu - \bar{X}$ oraz, że zmienna losowa

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{1}{4}}}$$

jest zmienną losową o standardowym rozkładzie normalnym $N(0,1)$.
Mamy więc 95% szans, że nasz przedział pokryje nieznaną wartość parametru θ.

DEFINICJA. *Dla estymatora przedziałowego $[L(X),\, U(X)]$ parametru θ, prawdopodobieństwem pokrycia nazywamy*

$$P(\theta \in [L(X),\, U(X)]).$$

DEFINICJA. *Dla estymatora przedziałowego $[L(X),\, U(X)]$ parametru θ, poziomem ufności nazywamy*

$$\inf_{\theta} P_\theta \left(\theta \in [L(X),\, U(X)]\right).$$

Zwykle gdy poziom ufności jest równy ustalonej wartości $1-\alpha$ to mówimy o przedziale ufności na poziomie $1-\alpha$.

1.1 Przedział ufności dla wartości średniej w rozkładzie normalnym.

Niech $(X_1, ..., X_n)$ będzie próba losowa z rozkładu $N(\mu, \sigma^2)$ o znanej wartości parametru σ^2. Wyznaczymy przedział ufności dla nieznanej wartości parametru μ.

Średnia w próbie

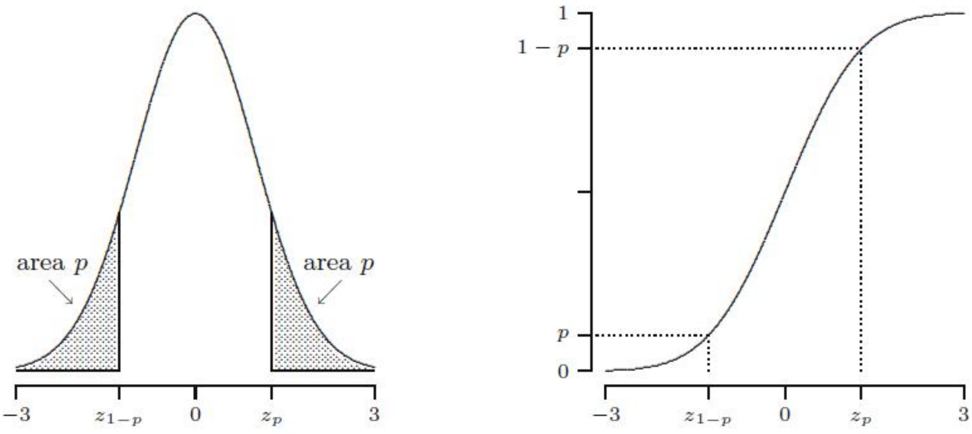
$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

ma rozkład normalny $N\left(\mu, \frac{\sigma^2}{n}\right)$. Stąd zmienna losowa

$$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

ma standardowy rozkład normalny $N(0,1)$. Teraz z łatwością możemy wyznaczyć przedział ufności na poziomie $1-\alpha$, gdzie α jest zadaną liczbą z przedziału $(0,1)$. W tym celu korzystamy z kwantyli standardowego rozkładu normalnego $N(0,1)$:

- $z_{\frac{\alpha}{2}}$ – kwantyl rzędu $\frac{\alpha}{2}$,
- $z_{1-\frac{\alpha}{2}}$ – kwantyl rzędu $1-\frac{\alpha}{2}$.



Rysunek 1: Kwantyle rozkładu normalnego.

Dla zmiennej losowej Z, o standardowym rozkładzie normalnym

$$P\left(z_{\frac{\alpha}{2}} \leqslant Z \leqslant z_{1-\frac{\alpha}{2}}\right) = 1-\alpha. \tag{6}$$

Ze względu na symetrię gęstości standardowego rozkładu normalnego mamy równość

$$z_{\frac{\alpha}{2}} = -z_{1-\frac{\alpha}{2}}$$

w konsekwencji czego równanie (6) przyjmuje postać

$$P\left(-z_{1-\frac{\alpha}{2}} \leqslant Z \leqslant z_{1-\frac{\alpha}{2}}\right) = 1-\alpha.$$

Stąd otrzymujemy

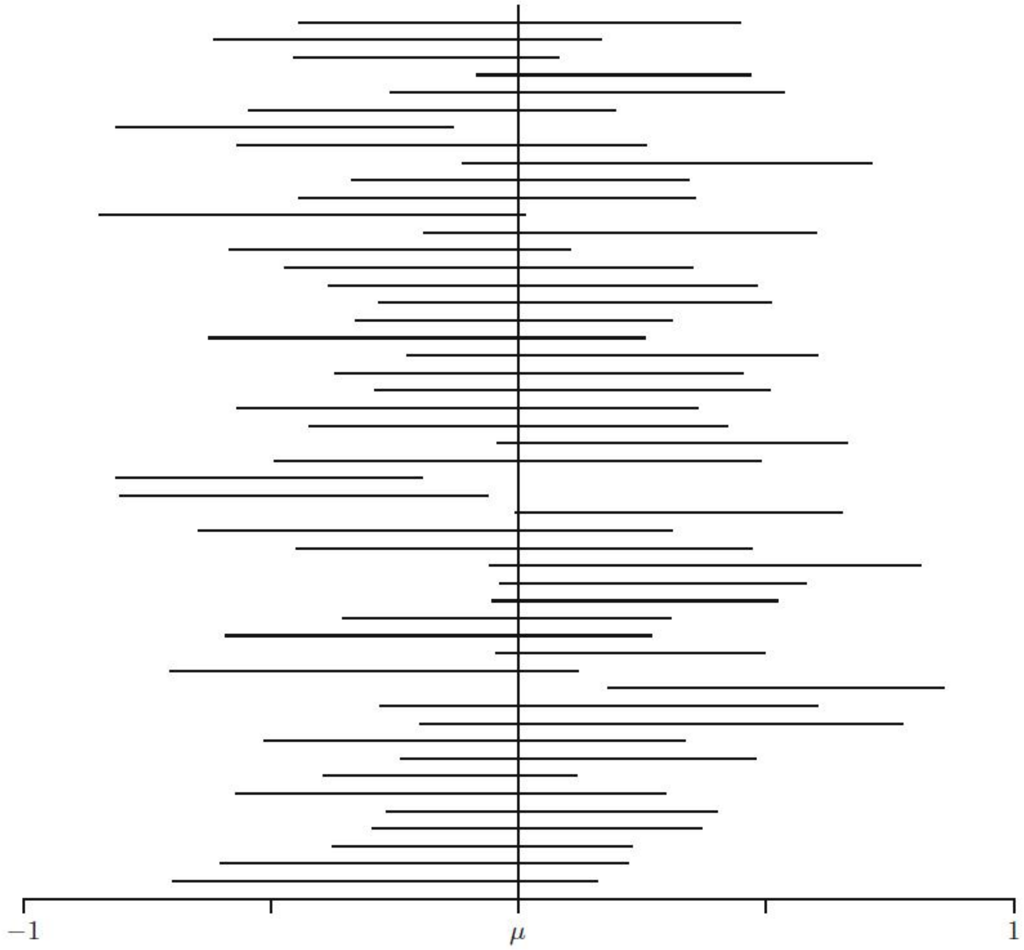
$$P\left(-z_{1-\frac{\alpha}{2}} \leqslant \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \leqslant z_{1-\frac{\alpha}{2}}\right) = 1-\alpha,$$

i dalej

$$P\left(\bar{X} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leqslant \mu \leqslant \bar{X} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

Przedział ufności, na poziomie ufności $1-\alpha$, dla średniej μ, gdy znana jest wariancja σ^2, ma postać

$$\left[\bar{X} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$



Rysunek 2: Pięćdziesiąt 90% przedziałów ufności dla $\mu = 0$.

UWAGA. Przedziały jednostronne

$$\mu \in \left(-\infty,\ \bar{X} + z_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right],$$

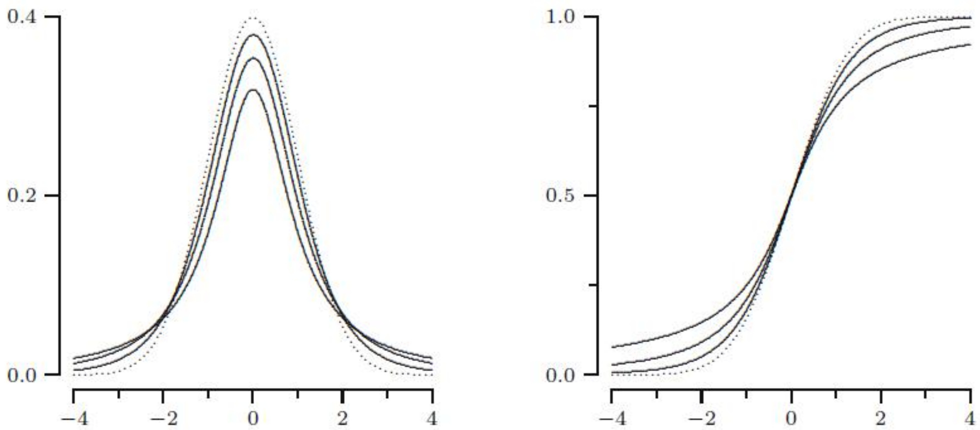
$$\mu \in \left[\bar{X} - z_{1-\alpha}\frac{\sigma}{\sqrt{n}},\ \infty\right),$$

1.2 Przedział ufności dla wartości średniej w rozkładzie normalnym o nieznanej wariancji.

W celu konstrukcji przedziału ufności skorzystamy ze statystyki

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}},$$

gdzie S jest estymatorem wariancji. Statystyka T ma rozkład $t-$Studenta z $n-1$ stopniami swobody.



Rysunek 3: Krzywa punktowa to wykres gęstości standardowego rozkładu normalnego. Pozostałe krzywe to wykresy gęstości rozkładów t-Studenta odpowiednio z 1, 2 i 5 stopniami swobody. Wraz ze wzrostem stopni swobody krzywe coraz lepiej przybliżają krzywą normalną.

Przediał ufności na poziomie ufnosci $1-\alpha$ ma postać

$$\left[\bar{X} - t_{(1-\frac{\alpha}{2},n-1)}\frac{S}{\sqrt{n}},\ \bar{X} + t_{(1-\frac{\alpha}{2},n-1)}\frac{S}{\sqrt{n}}\right],$$

gdzie $t_{(1-\frac{\alpha}{2},n-1)}$ oznacza kwantyl rzędu $1-\frac{\alpha}{2}$, rozkładu $t-$Studenta z $n-1$ stopniami swobody.