# Metody probabilistyczne

## 13. Elementy statystki matematycznej II

Wojciech Kotłowski

Instytut Informatyki PP
http://www.cs.put.poznan.pl/wkotlowski/

16.01.2018

# Parametryczna rodzina rozkładów

Założymy teraz, że znany jest typ rozkładu cechy $X$ w populacji, ale nie są znane jego parametry.

# Parametryczna rodzina rozkładów

Załóżymy teraz, że znany jest typ rozkładu cechy $X$ w populacji, ale nie są znane jego parametry.

Przykłady:

- $X \sim B(p)$, ale nie znamy wartości parametru $p$ (jak w wyborach z dwoma kandydatami)
- $X \sim \mathrm{Exp}(\lambda)$, ale nie znamy wartości $\lambda$ (np. czas oczekiwania na zdarzenie)
- $X \sim N(\mu, \sigma^2)$, ale nie znamy wartości $\mu$ i $\sigma^2$ (np. rozkład wzrostu w populacji)

Mamy więc rodzinę rozkładów $\{P_\theta \colon \theta \in \Theta\}$ różniących się tylko wartością parametru $\theta$, gdzie $\Theta$ jest zbiorem możliwych wartości parametru.

# Funkcja wiarogodności

Niech cecha $X$ ma rozkład $P_\theta$ należący do parametrycznej rodziny rozkładów.

Załóżmy, że w próbie zaobserwowaliśmy dane $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

# Funkcja wiarogodności

Niech cecha $X$ ma rozkład $P_\theta$ należący do parametrycznej rodziny rozkładów.

Załóżmy, że w próbie zaobserwowaliśmy dane $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

Funkcją wiar**o**godności $L(\boldsymbol{x}; \theta)$ nazywamy:

- Dla dyskretnej zmiennej $X$: łączne prawdopodobieństwo zaobserwowanych danych $\boldsymbol{x}$:

$$L(\boldsymbol{x}; \theta) \;=\; P_\theta(\boldsymbol{X} = \boldsymbol{x}) \;=\; p_\theta(x_1) p_\theta(x_2) \cdot \ldots \cdot p_\theta(x_n),$$

gdzie używamy skrótowo $p_\theta(x_i) = P_\theta(X_i = x_i)$

# Funkcja wiarogodności

Niech cecha $X$ ma rozkład $P_\theta$ należący do parametrycznej rodziny rozkładów.

Załóżmy, że w próbie zaobserwowaliśmy dane $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

Funkcją wiar**o**godności $L(\boldsymbol{x}; \theta)$ nazywamy:

- Dla dyskretnej zmiennej $X$: łączne prawdopodobieństwo zaobserwowanych danych $\boldsymbol{x}$:

  $$L(\boldsymbol{x}; \theta) \; = \; P_\theta(\boldsymbol{X} = \boldsymbol{x}) \; = \; p_\theta(x_1)p_\theta(x_2) \cdot \ldots \cdot p_\theta(x_n),$$

  gdzie używamy skrótowo $p_\theta(x_i) = P_\theta(X_i = x_i)$
- Dla ciągłej zmiennej losowej $X$: łączną gęstość prawdopodobieństwa zaobserwowanych danych $\boldsymbol{x}$:

  $$L(\boldsymbol{x}; \theta) \; = \; f_{\boldsymbol{X};\theta}(\boldsymbol{x}) \; = \; f_\theta(x_1)f_\theta(x_2) \cdot \ldots \cdot f_\theta(x_n)$$

# Funkcja wiarogodności

Niech cecha $X$ ma rozkład $P_\theta$ należący do parametrycznej rodziny rozkładów.

Załóżmy, że w próbie zaobserwowaliśmy dane $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

Funkcją wiar**o**godności $L(\boldsymbol{x}; \theta)$ nazywamy:

- Dla dyskretnej zmiennej $X$: łączne prawdopodobieństwo zaobserwowanych danych $\boldsymbol{x}$:

$$L(\boldsymbol{x}; \theta) \; = \; P_\theta(\boldsymbol{X} = \boldsymbol{x}) \; = \; p_\theta(x_1) p_\theta(x_2) \cdot \ldots \cdot p_\theta(x_n),$$

  gdzie używamy skrótowo $p_\theta(x_i) = P_\theta(X_i = x_i)$

- Dla ciągłej zmiennej losowej $X$: łączną gęstość prawdopodobieństwa zaobserwowanych danych $\boldsymbol{x}$:

$$L(\boldsymbol{x}; \theta) \; = \; f_{\boldsymbol{X};\theta}(\boldsymbol{x}) \; = \; f_\theta(x_1) f_\theta(x_2) \cdot \ldots \cdot f_\theta(x_n)$$

Uwaga: funkcja wiarogodności jest funkcją danych i parametru $\theta$

# Estymacja metodą największej wiarogodności

Estymatorem największej wiarogodności (NW) parametru $\theta$ nazywamy estymator równy parametrowi, który maksymalizuje funkcję wiarogodności:

$$\widehat{\theta}(\boldsymbol{x}) = \operatorname*{argmax}_{\theta} L(\boldsymbol{x}; \theta)$$

Estymator NW szacuje nieznany parametr poprzez wybór wartości parametru, przy którym zaobserwowane dane są najbardziej prawdopodobne!

# Estymacja metodą największej wiarogodności

Estymatorem największej wiarogodności (NW) parametru $\theta$ nazywamy estymator równy parametrowi, który maksymalizuje funkcję wiarogodności:

$$\widehat{\theta}(\boldsymbol{x}) = \operatorname*{argmax}_{\theta} L(\boldsymbol{x};\theta)$$

Estymator NW szacuje nieznany parametr poprzez wybór wartości parametru, przy którym zaobserwowane dane są najbardziej prawdopodobne!

Uwaga: zwykle znacznie wygodniejsze bywa działanie na ujemnym logarytmie wiarogodności $-\ln L(\boldsymbol{x};\theta)$. Ponieważ $L$ ma maksimum wtedy i tylko wtedy gdy $-\ln L$ ma minimum, możemy alternatywnie zdefiniować:

$$\widehat{\theta}(\boldsymbol{x}) = \operatorname*{argmin}_{\theta} -\ln L(\boldsymbol{x};\theta)$$

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować nieznaną wartość oczekiwaną $\mu$.
- Funkcja wiarogodności:

$$L(\boldsymbol{x}; \mu) = \prod_{i=1}^{n} f_\mu(x_i)$$

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować nieznaną wartość oczekiwaną $\mu$.
- Funkcja wiarogodności:

$$L(\boldsymbol{x}; \mu) = \prod_{i=1}^{n} f_\mu(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować nieznaną wartość oczekiwaną $\mu$.
- Funkcja wiarogodności:

$$
\begin{aligned}
L(\boldsymbol{x};\mu) &= \prod_{i=1}^{n} f_\mu(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\
-\ln L(\boldsymbol{x};\mu) &= -n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}
\end{aligned}
$$

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować nieznaną wartość oczekiwaną $\mu$.
- Funkcja wiarogodności:

$$\begin{aligned} L(\boldsymbol{x};\mu) &= \prod_{i=1}^{n} f_\mu(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\ -\ln L(\boldsymbol{x};\mu) &= -n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2} \end{aligned}$$

- Aby znaleźć minimum funkcji $-\ln L(\boldsymbol{x};\mu)$ liczymy pochodną po $\mu$ i przyrównujemy ją do zera:

$$\frac{\partial}{\partial\mu} - \ln L(\boldsymbol{x};\mu) = \sum_{i=1}^{n} \frac{\mu - x_i}{\sigma^2}$$

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować nieznaną wartość oczekiwaną $\mu$.
- Funkcja wiarogodności:

$$\begin{aligned} L(\boldsymbol{x};\mu) &= \prod_{i=1}^{n} f_\mu(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\ -\ln L(\boldsymbol{x};\mu) &= -n\ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2} \end{aligned}$$

- Aby znaleźć minimum funkcji $-\ln L(\boldsymbol{x};\mu)$ liczymy pochodną po $\mu$ i przyrównujemy ją do zera:

$$\frac{\partial}{\partial\mu} - \ln L(\boldsymbol{x};\mu) = \sum_{i=1}^{n} \frac{\mu - x_i}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(\mu - x_i)$$

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować nieznaną wartość oczekiwaną $\mu$.
- Funkcja wiarogodności:

$$
\begin{aligned}
L(\boldsymbol{x}; \mu) &= \prod_{i=1}^{n} f_\mu(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\
-\ln L(\boldsymbol{x}; \mu) &= -n \ln \left( \frac{1}{\sqrt{2\pi\sigma^2}} \right) + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}
\end{aligned}
$$

- Aby znaleźć minimum funkcji $-\ln L(\boldsymbol{x}; \mu)$ liczymy pochodną po $\mu$ i przyrównujemy ją do zera:

$$
\begin{aligned}
\frac{\partial}{\partial \mu} - \ln L(\boldsymbol{x}; \mu) &= \sum_{i=1}^{n} \frac{\mu - x_i}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (\mu - x_i) \\
\frac{1}{\sigma^2} \sum_{i=1}^{n} (\mu - x_i) &= 0 \quad \Longleftrightarrow \quad \mu = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}_n
\end{aligned}
$$

## Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować nieznaną wartość oczekiwaną $\mu$.
- Funkcja wiarogodności:

$$\begin{aligned} L(\boldsymbol{x};\mu) &= \prod_{i=1}^{n} f_\mu(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \\ -\ln L(\boldsymbol{x};\mu) &= -n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2} \end{aligned}$$

- Aby znaleźć minimum funkcji $-\ln L(\boldsymbol{x};\mu)$ liczymy pochodną po $\mu$ i przyrównujemy ją do zera:

$$\begin{aligned} \frac{\partial}{\partial\mu} - \ln L(\boldsymbol{x};\mu) &= \sum_{i=1}^{n} \frac{\mu - x_i}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (\mu - x_i) \\ \frac{1}{\sigma^2} \sum_{i=1}^{n} (\mu - x_i) &= 0 \quad \Longleftrightarrow \quad \mu = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}_n \end{aligned}$$

- Wniosek: estymatorem NW parametru $\mu$ w rozkładzie normalnym jest średnia arytmetyczna z próby $\widehat{\mu} = \overline{X}_n$.

# Przykład: rozkład wykładniczy

- $X \sim \mathrm{Exp}(\lambda)$ i chcemy estymować nieznany parametr $\lambda$.
- Funkcja wiarogodności:

$$L(\boldsymbol{x}; \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

# Przykład: rozkład wykładniczy

- $X \sim \mathrm{Exp}(\lambda)$ i chcemy estymować nieznany parametr $\lambda$.
- Funkcja wiarogodności:

$$L(\boldsymbol{x};\lambda) \;=\; \prod_{i=1}^{n} \lambda e^{-\lambda x_i} \;=\; \lambda^n e^{-\lambda(x_1+\ldots+x_n)}$$

# Przykład: rozkład wykładniczy

- $X \sim \mathrm{Exp}(\lambda)$ i chcemy estymować nieznany parametr $\lambda$.
- Funkcja wiarogodności:

$$L(\boldsymbol{x};\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda(x_1+\ldots+x_n)} = \lambda^n e^{-n\lambda\overline{x}_n}$$

# Przykład: rozkład wykładniczy

- $X \sim \mathrm{Exp}(\lambda)$ i chcemy estymować nieznany parametr $\lambda$.
- Funkcja wiarogodności:

$$
\begin{aligned}
L(\boldsymbol{x};\lambda) &= \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda(x_1+\ldots+x_n)} = \lambda^n e^{-n\lambda\overline{x}_n} \\
-\ln L(\boldsymbol{x};\lambda) &= -n\ln\lambda + n\lambda\overline{x}_n
\end{aligned}
$$

# Przykład: rozkład wykładniczy

- $X \sim \mathrm{Exp}(\lambda)$ i chcemy estymować nieznany parametr $\lambda$.
- Funkcja wiarogodności:

$$\begin{aligned} L(\boldsymbol{x};\lambda) &= \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda(x_1+\ldots+x_n)} = \lambda^n e^{-n\lambda\overline{x}_n} \\ -\ln L(\boldsymbol{x};\lambda) &= -n\ln\lambda + n\lambda\overline{x}_n \end{aligned}$$

- Przyrównanie pochodnej do zera:

$$\frac{\partial}{\partial\lambda} - \ln L(\boldsymbol{x};\lambda) = -\frac{n}{\lambda} + n\overline{x}_n$$

# Przykład: rozkład wykładniczy

- $X \sim \mathrm{Exp}(\lambda)$ i chcemy estymować nieznany parametr $\lambda$.
- Funkcja wiarogodności:

$$\begin{aligned} L(\boldsymbol{x};\lambda) &= \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda(x_1+\ldots+x_n)} = \lambda^n e^{-n\lambda\overline{x}_n} \\ -\ln L(\boldsymbol{x};\lambda) &= -n\ln\lambda + n\lambda\overline{x}_n \end{aligned}$$

- Przyrównanie pochodnej do zera:

$$\begin{aligned} \frac{\partial}{\partial\lambda} - \ln L(\boldsymbol{x};\lambda) &= -\frac{n}{\lambda} + n\overline{x}_n \\ \frac{\partial}{\partial\lambda} - \ln L(\boldsymbol{x};\lambda) &= 0 \quad \Longleftrightarrow \quad \lambda = \frac{1}{\overline{x}_n} \end{aligned}$$

# Przykład: rozkład wykładniczy

- $X \sim \mathrm{Exp}(\lambda)$ i chcemy estymować nieznany parametr $\lambda$.
- Funkcja wiarogodności:

$$
\begin{aligned}
L(\boldsymbol{x};\lambda) &= \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda(x_1+\ldots+x_n)} = \lambda^n e^{-n\lambda\overline{x}_n} \\
-\ln L(\boldsymbol{x};\lambda) &= -n\ln\lambda + n\lambda\overline{x}_n
\end{aligned}
$$

- Przyrównanie pochodnej do zera:

$$
\begin{aligned}
\frac{\partial}{\partial\lambda} - \ln L(\boldsymbol{x};\lambda) &= -\frac{n}{\lambda} + n\overline{x}_n \\
\frac{\partial}{\partial\lambda} - \ln L(\boldsymbol{x};\lambda) &= 0 \quad \Longleftrightarrow \quad \lambda = \frac{1}{\overline{x}_n}
\end{aligned}
$$

- Wniosek: estymator NW parametru $\lambda$ w rozkładzie wykładniczym:

$$
\widehat{\lambda} = \frac{1}{\overline{X}_n}
$$

# Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n - k$ zer

# Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n - k$ zer

$$L(\boldsymbol{x}; p) = \prod_{i=1}^{n} p(x_i)$$

# Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n-k$ zer

$$L(\boldsymbol{x};p) = \prod_{i=1}^{n} p(x_i) = p(1)^k p(0)^{n-k}$$

# Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n - k$ zer

$$L(\boldsymbol{x}; p) = \prod_{i=1}^{n} p(x_i) = p(1)^k p(0)^{n-k} = p^k (1-p)^{n-k}$$

# Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n-k$ zer

$$\begin{aligned} L(\boldsymbol{x};p) &= \prod_{i=1}^{n} p(x_i) = p(1)^k p(0)^{n-k} = p^k(1-p)^{n-k} \\ -\ln L(\boldsymbol{x};p) &= -k\ln p - (n-k)\ln(1-p) \end{aligned}$$

## Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n-k$ zer

$$
\begin{aligned}
L(\boldsymbol{x};p) &= \prod_{i=1}^{n} p(x_i) = p(1)^k p(0)^{n-k} = p^k(1-p)^{n-k} \\
-\ln L(\boldsymbol{x};p) &= -k\ln p - (n-k)\ln(1-p)
\end{aligned}
$$

- Przyrównanie pochodnej do zera:

$$
\frac{\partial}{\partial p} - \ln L(\boldsymbol{x};p) = -\frac{k}{p} + \frac{n-k}{1-p}
$$

# Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n-k$ zer

$$\begin{aligned} L(\boldsymbol{x}; p) &= \prod_{i=1}^{n} p(x_i) = p(1)^k p(0)^{n-k} = p^k (1-p)^{n-k} \\ -\ln L(\boldsymbol{x}; p) &= -k \ln p - (n-k) \ln(1-p) \end{aligned}$$

- Przyrównanie pochodnej do zera:

$$\begin{aligned} \frac{\partial}{\partial p} - \ln L(\boldsymbol{x}; p) &= -\frac{k}{p} + \frac{n-k}{1-p} \\ \frac{\partial}{\partial p} - \ln L(\boldsymbol{x}; p) &= 0 \iff \frac{k}{p} = \frac{n-k}{1-p} \end{aligned}$$

# Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n-k$ zer

$$\begin{aligned} L(\boldsymbol{x}; p) &= \prod_{i=1}^{n} p(x_i) = p(1)^k p(0)^{n-k} = p^k (1-p)^{n-k} \\ -\ln L(\boldsymbol{x}; p) &= -k \ln p - (n-k) \ln(1-p) \end{aligned}$$

- Przyrównanie pochodnej do zera:

$$\begin{aligned} \frac{\partial}{\partial p} - \ln L(\boldsymbol{x}; p) &= -\frac{k}{p} + \frac{n-k}{1-p} \\ \frac{\partial}{\partial p} - \ln L(\boldsymbol{x}; p) &= 0 \iff \frac{k}{p} = \frac{n-k}{1-p} \iff p = \frac{k}{n} \end{aligned}$$

# Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n-k$ zer

$$\begin{aligned} L(\boldsymbol{x}; p) &= \prod_{i=1}^{n} p(x_i) = p(1)^k p(0)^{n-k} = p^k (1-p)^{n-k} \\ -\ln L(\boldsymbol{x}; p) &= -k \ln p - (n-k) \ln(1-p) \end{aligned}$$

- Przyrównanie pochodnej do zera:

$$\begin{aligned} \frac{\partial}{\partial p} - \ln L(\boldsymbol{x}; p) &= -\frac{k}{p} + \frac{n-k}{1-p} \\ \frac{\partial}{\partial p} - \ln L(\boldsymbol{x}; p) &= 0 \iff \frac{k}{p} = \frac{n-k}{1-p} \iff p = \frac{k}{n} \end{aligned}$$

- Wniosek: estymator NW parametru $p$ w rozkładzie dwupunktowym jest częstość empiryczna

$$\widehat{p} = \frac{|\{i : X_i = 1\}|}{n}$$

# Przykład: rozkład dwupunktowy

- $X \sim B(p)$ i chcemy estymować nieznany parametr $p$.
- Załóżmy że wśród $n$ obserwacji jest $k$ jedynek i $n-k$ zer

$$\begin{aligned} L(\boldsymbol{x};p) &= \prod_{i=1}^{n} p(x_i) = p(1)^k p(0)^{n-k} = p^k (1-p)^{n-k} \\ -\ln L(\boldsymbol{x};p) &= -k \ln p - (n-k)\ln(1-p) \end{aligned}$$

- Przyrównanie pochodnej do zera:

$$\begin{aligned} \frac{\partial}{\partial p} - \ln L(\boldsymbol{x};p) &= -\frac{k}{p} + \frac{n-k}{1-p} \\ \frac{\partial}{\partial p} - \ln L(\boldsymbol{x};p) &= 0 \iff \frac{k}{p} = \frac{n-k}{1-p} \iff p = \frac{k}{n} \end{aligned}$$

- Wniosek: estymator NW parametru $p$ w rozkładzie dwupunktowym jest częstość empiryczna

$$\widehat{p} = \frac{|\{i \colon X_i = 1\}|}{n} = \overline{X}_n$$

## Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować zarówno $\mu$ jak i $\sigma^2$.
- Funkcję wiarogodności policzyliśmy już wcześniej:

$$-\ln L(\boldsymbol{x}; \mu, \sigma^2) \quad = \quad -n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2}$$

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować zarówno $\mu$ jak i $\sigma^2$.
- Funkcję wiarogodności policzyliśmy już wcześniej:

$$\begin{aligned} -\ln L(\boldsymbol{x}; \mu, \sigma^2) &= -n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2} \\ &= \frac{1}{2} n \ln \sigma^2 \; + \; n \ln \sqrt{2\pi} \; + \; \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2} \end{aligned}$$

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować zarówno $\mu$ jak i $\sigma^2$.
- Funkcję wiarogodności policzyliśmy już wcześniej:

$$
\begin{aligned}
-\ln L(\boldsymbol{x}; \mu, \sigma^2) &= -n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2} \\
&= \frac{1}{2} n \ln \sigma^2 \; + \; n \ln \sqrt{2\pi} \; + \; \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2}
\end{aligned}
$$

- Aby znaleźć minimum funkcji $-\ln L(\boldsymbol{x}; \mu, \sigma^2)$ przyrównujemy pochodne po $\mu$ i $\sigma^2$ do zera.

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować zarówno $\mu$ jak i $\sigma^2$.
- Funkcję wiarogodności policzyliśmy już wcześniej:

$$\begin{aligned} -\ln L(\boldsymbol{x}; \mu, \sigma^2) &= -n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2} \\ &= \frac{1}{2} n \ln \sigma^2 \; + \; n \ln \sqrt{2\pi} \; + \; \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2} \end{aligned}$$

- Aby znaleźć minimum funkcji $-\ln L(\boldsymbol{x}; \mu, \sigma^2)$ przyrównujemy pochodne po $\mu$ i $\sigma^2$ do zera.
  Pochodną po $\mu$ policzyliśmy już wcześniej:

$$\frac{\partial}{\partial \mu} - \ln L(\boldsymbol{x}; \mu, \sigma^2) = 0 \quad \Longleftrightarrow \quad \mu = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}_n$$

# Przykład: rozkład normalny

- $X \sim N(\mu, \sigma^2)$ i chcemy estymować zarówno $\mu$ jak i $\sigma^2$.
- Funkcję wiarogodności policzyliśmy już wcześniej:

$$
\begin{aligned}
-\ln L(\boldsymbol{x}; \mu, \sigma^2) &= -n \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) + \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2} \\
&= \frac{1}{2} n \ln \sigma^2 \; + \; n \ln \sqrt{2\pi} \; + \; \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2}
\end{aligned}
$$

- Aby znaleźć minimum funkcji $-\ln L(\boldsymbol{x}; \mu, \sigma^2)$ przyrównujemy pochodne po $\mu$ i $\sigma^2$ do zera.
  Pochodną po $\mu$ policzyliśmy już wcześniej:

$$
\frac{\partial}{\partial \mu} - \ln L(\boldsymbol{x}; \mu, \sigma^2) = 0 \quad \Longleftrightarrow \quad \mu = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{x}_n
$$

  Pozostaje nam policzyć pochodną po $\sigma^2$.

# Przykład: rozkład normalny – c.d.

$$-\ln L(\boldsymbol{x};\mu,\sigma^2) \quad = \quad \frac{1}{2} n \ln \sigma^2 \; + \; n \ln \sqrt{2\pi} \; + \; \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2}$$

# Przykład: rozkład normalny – c.d.

$$
\begin{aligned}
-\ln L(\boldsymbol{x};\mu,\sigma^2) &= \frac{1}{2} n \ln \sigma^2 + n \ln \sqrt{2\pi} + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2} \\
\frac{\partial}{\partial \sigma^2} - \ln L(\boldsymbol{x};\mu,\sigma^2) &= \frac{1}{2} \frac{n}{\sigma^2} - \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i-\mu)^2
\end{aligned}
$$

# Przykład: rozkład normalny – c.d.

$$
\begin{aligned}
-\ln L(\boldsymbol{x};\mu,\sigma^2) &= \frac{1}{2} n \ln \sigma^2 + n \ln \sqrt{2\pi} + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2} \\
\frac{\partial}{\partial \sigma^2} - \ln L(\boldsymbol{x};\mu,\sigma^2) &= \frac{1}{2}\frac{n}{\sigma^2} - \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2
\end{aligned}
$$

Przyrównując pochodną do zera dostajemy:

$$
\frac{1}{2}\frac{n}{\sigma^2} = \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2
$$

# Przykład: rozkład normalny – c.d.

$$
\begin{aligned}
-\ln L(\boldsymbol{x};\mu,\sigma^2) &= \frac{1}{2} n \ln \sigma^2 + n \ln \sqrt{2\pi} + \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2} \\
\frac{\partial}{\partial \sigma^2} - \ln L(\boldsymbol{x};\mu,\sigma^2) &= \frac{1}{2}\frac{n}{\sigma^2} - \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2
\end{aligned}
$$

Przyrównując pochodną do zera dostajemy:

$$
\frac{1}{2}\frac{n}{\sigma^2} = \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 \iff \sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2
$$

# Przykład: rozkład normalny – c.d.

$$-\ln L(\boldsymbol{x};\mu,\sigma^2) \;=\; \frac{1}{2}n\ln\sigma^2 \;+\; n\ln\sqrt{2\pi} \;+\; \sum_{i=1}^{n}\frac{(x_i-\mu)^2}{2\sigma^2}$$

$$\frac{\partial}{\partial\sigma^2}-\ln L(\boldsymbol{x};\mu,\sigma^2) \;=\; \frac{1}{2}\frac{n}{\sigma^2} \;-\; \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

Przyrównując pochodną do zera dostajemy:

$$\frac{1}{2}\frac{n}{\sigma^2} \;=\; \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 \iff \sigma^2 \;=\; \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$$

Wniosek: estymatory NW parametrów $\mu$ i $\sigma^2$ w rozkładzie normalnym:

$$\widehat{\mu} \;=\; \overline{X}_n, \qquad \widehat{\sigma}^2 \;=\; \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2$$

Estymator wariancji jest obciążony

# Zadanie

**Zadanie 1**

Wyznacz estymator największej wiarogodności parametru $\lambda$ dla rozkład Poissona

# Własności estymatora największej wiarogodności

Estymatory największej wiarogodności mogą być czasem kłopotliwe do wyznaczenia (wymagają rozwiązania problemu optymalizacji), mają jednak pożądane własności asymptotyczne (tzn. dla $n \to \infty$):

- Zgodność
- Asymptotyczna nieobciążoność
- Asymptotyczna efektywność

# Estymacja przedziałowa

Dotychczas rozważaliśmy tzw. estymatory punktowe, które szacują wartość parametru $\theta$ za pomocą jednej liczby $\widehat{\theta}$.

# Estymacja przedziałowa

Dotychczas rozważaliśmy tzw. estymatory punktowe, które szacują wartość parametru $\theta$ za pomocą jednej liczby $\widehat{\theta}$.

Często chcemy mieć nie tylko oszacowanie parametru $\theta$, ale również przedział, w którym (z dużym prawdopodobieństwem) $\theta$ się znajdzie:

$$[\widehat{\theta}_L,\ \widehat{\theta}_P]$$

Taki przedział nazywamy przedziałem ufności, a zadanie jego wyznaczenia nazywamy estymacją przedziałową.

# Estymacja przedziałowa

Dotychczas rozważaliśmy tzw. estymatory punktowe, które szacują wartość parametru $\theta$ za pomocą jednej liczby $\widehat{\theta}$.

Często chcemy mieć nie tylko oszacowanie parametru $\theta$, ale również przedział, w którym (z dużym prawdopodobieństwem) $\theta$ się znajdzie:

$$[\widehat{\theta}_L,\ \widehat{\theta}_P]$$

Taki przedział nazywamy przedziałem ufności, a zadanie jego wyznaczenia nazywamy estymacją przedziałową.

Zaletą przedziału ufności jest wskazanie stopnia dokładności oszacowania parametru.

# Przedział ufności

Przedziałem ufności dla parametru $\theta$ na poziomie ufności $1-\alpha$ nazywamy przedział $[\widehat{\theta}_L, \widehat{\theta}_P]$ taki, że $\widehat{\theta}_L$ i $\widehat{\theta}_P$ są statystykami wyznaczonymi z próby, oraz:

$$P\left(\widehat{\theta}_L \leqslant \theta \leqslant \widehat{\theta}_P\right) = 1 - \alpha$$

# Przedział ufności

Przedziałem ufności dla parametru $\theta$ na poziomie ufności $1-\alpha$ nazywamy przedział $[\widehat{\theta}_L, \widehat{\theta}_P]$ taki, że $\widehat{\theta}_L$ i $\widehat{\theta}_P$ są statystykami wyznaczonymi z próby, oraz:

$$P\left(\widehat{\theta}_L \leqslant \theta \leqslant \widehat{\theta}_P\right) = 1 - \alpha$$

- Z prawdopodobieństwem $1-\alpha$ prawdziwa wartość parametru znajduje się w przedziale

# Przedział ufności

Przedziałem ufności dla parametru $\theta$ na poziomie ufności $1-\alpha$ nazywamy przedział $[\widehat{\theta}_L, \widehat{\theta}_P]$ taki, że $\widehat{\theta}_L$ i $\widehat{\theta}_P$ są statystykami wyznaczonymi z próby, oraz:

$$P\left(\widehat{\theta}_L \leqslant \theta \leqslant \widehat{\theta}_P\right) = 1 - \alpha$$

- Z prawdopodobieństwem $1-\alpha$ prawdziwa wartość parametru znajduje się w przedziale
- Uwaga: to przedział jest losowy, a nie wartość parametru!

# Przedział ufności

Przedziałem ufności dla parametru $\theta$ na poziomie ufności $1-\alpha$ nazywamy przedział $[\widehat{\theta}_L, \widehat{\theta}_P]$ taki, że $\widehat{\theta}_L$ i $\widehat{\theta}_P$ są statystykami wyznaczonymi z próby, oraz:

$$P\left(\widehat{\theta}_L \leqslant \theta \leqslant \widehat{\theta}_P\right) = 1 - \alpha$$

- Z prawdopodobieństwem $1 - \alpha$ prawdziwa wartość parametru znajduje się w przedziale
- Uwaga: to przedział jest losowy, a nie wartość parametru!
- Czasem $\alpha$ (zamiast $1 - \alpha$) nazywa się poziomem ufności

# Przedział ufności

Przedziałem ufności dla parametru $\theta$ na poziomie ufności $1-\alpha$ nazywamy przedział $[\widehat{\theta}_L, \widehat{\theta}_P]$ taki, że $\widehat{\theta}_L$ i $\widehat{\theta}_P$ są statystykami wyznaczonymi z próby, oraz:

$$P\left(\widehat{\theta}_L \leqslant \theta \leqslant \widehat{\theta}_P\right) = 1 - \alpha$$

- Z prawdopodobieństwem $1 - \alpha$ prawdziwa wartość parametru znajduje się w przedziale
- Uwaga: to przedział jest losowy, a nie wartość parametru!
- Czasem $\alpha$ (zamiast $1 - \alpha$) nazywa się poziomem ufności
- Czasem używam się przedziału otwartego $(\widehat{\theta}_L, \widehat{\theta}_P)$

# Przedział ufności

Przedziałem ufności dla parametru $\theta$ na poziomie ufności $1-\alpha$ nazywamy przedział $[\widehat{\theta}_L, \widehat{\theta}_P]$ taki, że $\widehat{\theta}_L$ i $\widehat{\theta}_P$ są statystykami wyznaczonymi z próby, oraz:

$$P\left(\widehat{\theta}_L \leqslant \theta \leqslant \widehat{\theta}_P\right) = 1 - \alpha$$

- Z prawdopodobieństwem $1 - \alpha$ prawdziwa wartość parametru znajduje się w przedziale
- Uwaga: to przedział jest losowy, a nie wartość parametru!
- Czasem $\alpha$ (zamiast $1 - \alpha$) nazywa się poziomem ufności
- Czasem używam się przedziału otwartego $(\widehat{\theta}_L, \widehat{\theta}_P)$
- Przykład: W problemie szacowania poparcia kandydata $A$, estymacja przedziałowa dała przedział ufności na poziomie 95% równy $[0.58, 0.62]$, tzn. z prawdopodobieństwem 95% prawdziwy parametr $p$ znajduje się między 0.58 a 0.62.

# Metody wyznaczania przedziałów ufności

Zwykle rozpinamy przedział symetrycznie wokół pewnego estymatora punktowego $\widehat{\theta}$:

$$[\widehat{\theta} - \Delta,\ \widehat{\theta} + \Delta], \qquad \text{tzn. } \ \widehat{\theta}_L = \widehat{\theta} - \Delta, \quad \widehat{\theta}_P = \widehat{\theta} + \Delta$$

Rozpiętość przedziału $\Delta$ wyznaczana jest w oparciu o docelowy poziom ufności, tzn. rozwiązując równanie (ze względu na $\Delta$):

$$1 - \alpha \ = \ P\left(\widehat{\theta} - \Delta \ \leqslant \ \theta \ \leqslant \ \widehat{\theta} + \Delta\right)$$

# Metody wyznaczania przedziałów ufności

Zwykle rozpinamy przedział symetrycznie wokół pewnego estymatora punktowego $\widehat{\theta}$:

$$[\widehat{\theta} - \Delta,\ \widehat{\theta} + \Delta], \qquad \text{tzn. } \widehat{\theta}_L = \widehat{\theta} - \Delta, \quad \widehat{\theta}_P = \widehat{\theta} + \Delta$$

Rozpiętość przedziału $\Delta$ wyznaczana jest w oparciu o docelowy poziom ufności, tzn. rozwiązując równanie (ze względu na $\Delta$):

$$\begin{aligned} 1 - \alpha &= P\left(\widehat{\theta} - \Delta \ \leqslant\ \theta \ \leqslant\ \widehat{\theta} + \Delta\right) \\ &= P\left(-\Delta \ \leqslant\ \theta - \widehat{\theta} \ \leqslant\ \Delta\right) \end{aligned}$$

# Metody wyznaczania przedziałów ufności

Zwykle rozpinamy przedział symetrycznie wokół pewnego estymatora punktowego $\widehat{\theta}$:

$$[\widehat{\theta} - \Delta,\ \widehat{\theta} + \Delta], \qquad \text{tzn. } \ \widehat{\theta}_L = \widehat{\theta} - \Delta, \quad \widehat{\theta}_P = \widehat{\theta} + \Delta$$

Rozpiętość przedziału $\Delta$ wyznaczana jest w oparciu o docelowy poziom ufności, tzn. rozwiązując równanie (ze względu na $\Delta$):

$$\begin{aligned} 1 - \alpha &= P\left(\widehat{\theta} - \Delta \ \leqslant\ \theta \ \leqslant\ \widehat{\theta} + \Delta\right) \\ &= P\left(-\Delta \ \leqslant\ \theta - \widehat{\theta} \ \leqslant\ \Delta\right) \\ &= P\left(-\Delta \ \leqslant\ \widehat{\theta} - \theta \ \leqslant\ \Delta\right) \end{aligned}$$

# Przykład: rozkład normalny ze znaną wariancją

Niech $X \sim N(\mu, \sigma^2)$, gdzie $\sigma^2$ jest znana

# Przykład: rozkład normalny ze znaną wariancją

Niech $X \sim N(\mu, \sigma^2)$, gdzie $\sigma^2$ jest znana

Konstruujemy estymator wartości oczekiwanej $\mu$ na podstawie próby $X_1, \ldots, X_n$:

$$\widehat{\mu} \;=\; \overline{X}_n \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i$$

# Przykład: rozkład normalny ze znaną wariancją

Niech $X \sim N(\mu, \sigma^2)$, gdzie $\sigma^2$ jest znana

Konstruujemy estymator wartości oczekiwanej $\mu$ na podstawie próby $X_1, \ldots, X_n$:

$$\widehat{\mu} \ = \ \overline{X}_n \ = \ \frac{1}{n} \sum_{i=1}^{n} X_i$$

Ponieważ $\overline{X}_n$ jest kombinacją liniową niezależnych zmiennych o rozkładzie normalnym, $\overline{X}_n$ ma rozkład normalny z parametrami:

$$E(\overline{X}_n) \ = \ \mu, \qquad D^2(\overline{X}_n) \ = \ \frac{\sigma^2}{n}, \qquad \text{czyli} \ \ \overline{X}_n \ \sim \ N\left(\mu, \frac{\sigma^2}{n}\right)$$

# Przykład: rozkład normalny ze znaną wariancją

Niech $X \sim N(\mu, \sigma^2)$, gdzie $\sigma^2$ jest znana

Konstruujemy estymator wartości oczekiwanej $\mu$ na podstawie próby $X_1, \ldots, X_n$:

$$\widehat{\mu} \;=\; \overline{X}_n \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i$$

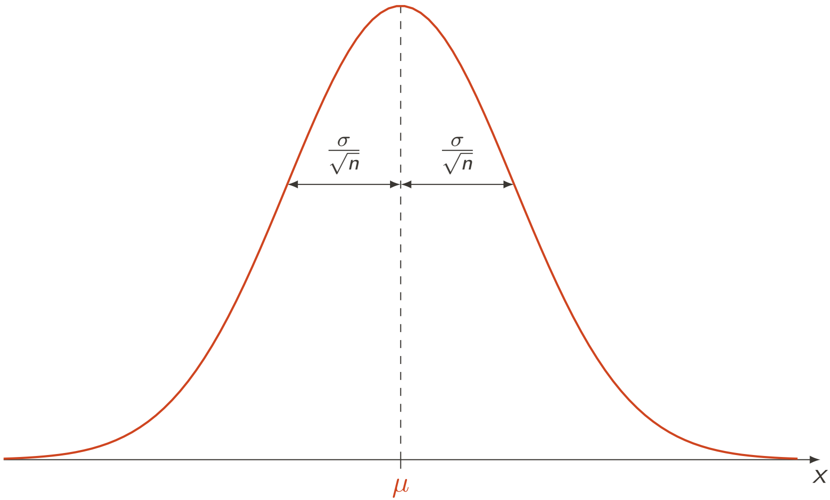
Ponieważ $\overline{X}_n$ jest kombinacją liniową niezależnych zmiennych o rozkładzie normalnym, $\overline{X}_n$ ma rozkład normalny z parametrami:

$$E(\overline{X}_n) \;=\; \mu, \qquad D^2(\overline{X}_n) \;=\; \frac{\sigma^2}{n}, \qquad \text{czyli} \;\; \overline{X}_n \;\sim\; N\left(\mu, \frac{\sigma^2}{n}\right)$$

Konstruujemy następnie przedział ufności na poziomie ufności $1-\alpha$:

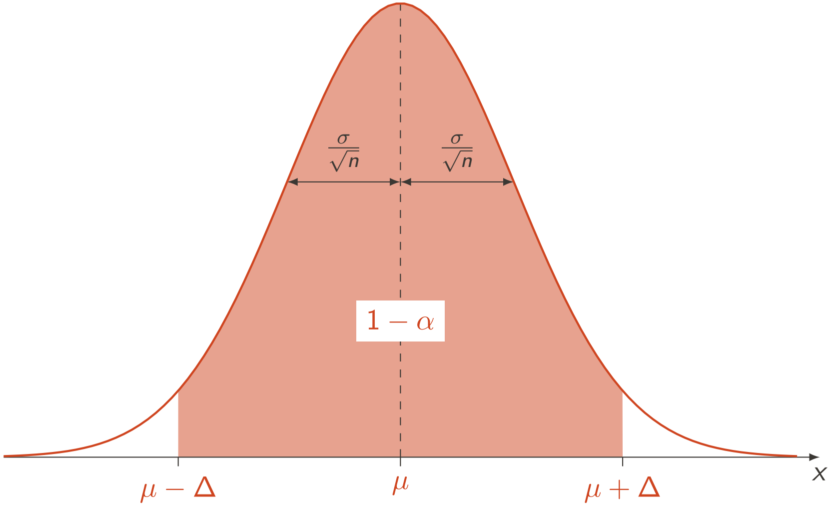
$$1-\alpha \;=\; P\left(-\Delta \;\leqslant\; \overline{X}_n - \mu \;\leqslant\; \Delta\right)$$

# Rozkład estymatora $\overline{X}_n$



# Rozkład estymatora $\overline{X}_n$

$$P\left(\mu - \Delta \leqslant \overline{X}_n \leqslant \mu + \Delta\right) \;=\; 1 - \alpha$$



# Przykład: rozkład normalny ze znaną wariancją

Skoro $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, to $Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

# Przykład: rozkład normalny ze znaną wariancją

Skoro $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, to $Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$1 - \alpha = P\left(-\Delta \leqslant \overline{X}_n - \mu \leqslant \Delta\right)$$

# Przykład: rozkład normalny ze znaną wariancją

Skoro $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, to $Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$\begin{aligned} 1-\alpha &= P\left(-\Delta \leqslant \overline{X}_n - \mu \leqslant \Delta\right) \\ &= P\left(-\frac{\Delta}{\sigma}\sqrt{n} \leqslant \frac{\overline{X}_n - \mu}{\sigma}\sqrt{n} \leqslant \frac{\Delta}{\sigma}\sqrt{n}\right) \end{aligned}$$

# Przykład: rozkład normalny ze znaną wariancją

Skoro $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, to $Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$
\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \overline{X}_n - \mu \leqslant \Delta\right) \\
&= P\left(-\frac{\Delta}{\sigma}\sqrt{n} \leqslant \frac{\overline{X}_n - \mu}{\sigma}\sqrt{n} \leqslant \frac{\Delta}{\sigma}\sqrt{n}\right) \\
&= P\left(-\frac{\Delta}{\sigma}\sqrt{n} \leqslant Z \leqslant \frac{\Delta}{\sigma}\sqrt{n}\right)
\end{aligned}
$$

# Przykład: rozkład normalny ze znaną wariancją

Skoro $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, to $Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \overline{X}_n - \mu \leqslant \Delta\right) \\
&= P\left(-\frac{\Delta}{\sigma}\sqrt{n} \leqslant \frac{\overline{X}_n - \mu}{\sigma}\sqrt{n} \leqslant \frac{\Delta}{\sigma}\sqrt{n}\right) \\
&= P\left(-\frac{\Delta}{\sigma}\sqrt{n} \leqslant Z \leqslant \frac{\Delta}{\sigma}\sqrt{n}\right) \\
&= \Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) - \Phi\left(-\frac{\Delta}{\sigma}\sqrt{n}\right)
\end{aligned}$$

# Przykład: rozkład normalny ze znaną wariancją

$$\text{Skoro } \overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right), \qquad \text{to } Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \overline{X}_n - \mu \leqslant \Delta\right) \\
&= P\left(-\frac{\Delta}{\sigma}\sqrt{n} \leqslant \frac{\overline{X}_n - \mu}{\sigma}\sqrt{n} \leqslant \frac{\Delta}{\sigma}\sqrt{n}\right) \\
&= P\left(-\frac{\Delta}{\sigma}\sqrt{n} \leqslant Z \leqslant \frac{\Delta}{\sigma}\sqrt{n}\right) \\
&= \Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) - \Phi\left(-\frac{\Delta}{\sigma}\sqrt{n}\right) \\
&= 2\Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) - 1
\end{aligned}$$

$\Phi(-x) = 1 - \Phi(x)$

# Przykład: rozkład normalny ze znaną wariancją

Skoro $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, to $Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$1 - \alpha = 2\Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) - 1$$

# Przykład: rozkład normalny ze znaną wariancją

$$\boxed{\text{Skoro } \overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right), \qquad \text{to } Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)}$$

$$\begin{aligned} 1-\alpha &= 2\Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) - 1 \\ \Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) &= 1 - \frac{\alpha}{2} \end{aligned}$$

# Przykład: rozkład normalny ze znaną wariancją

$$\text{Skoro } \overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right), \qquad \text{to } Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

$$\begin{aligned} 1 - \alpha &= 2\Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) - 1 \\ \Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) &= 1 - \frac{\alpha}{2} \\ \frac{\Delta}{\sigma}\sqrt{n} &= \underbrace{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}_{z_{1-\alpha/2}} \end{aligned}$$

# Przykład: rozkład normalny ze znaną wariancją

Skoro $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, to $Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$\begin{aligned}
1-\alpha &= 2\Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) - 1 \\
\Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) &= 1 - \frac{\alpha}{2} \\
\frac{\Delta}{\sigma}\sqrt{n} &= \underbrace{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}_{z_{1-\alpha/2}} \\
\Delta &= z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}
\end{aligned}$$

# Przykład: rozkład normalny ze znaną wariancją

Skoro $\overline{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, to $Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

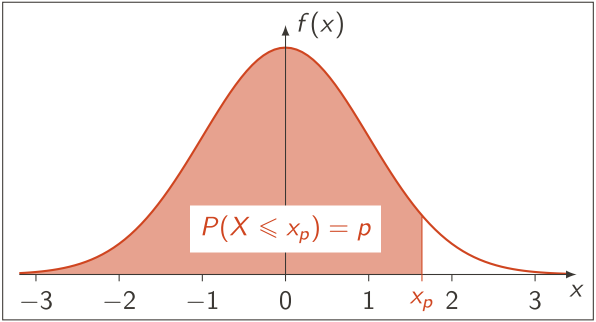
$$\begin{aligned}
1-\alpha &= 2\Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) - 1 \\
\Phi\left(\frac{\Delta}{\sigma}\sqrt{n}\right) &= 1 - \frac{\alpha}{2} \\
\frac{\Delta}{\sigma}\sqrt{n} &= \underbrace{\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}_{z_{1-\alpha/2}} \\
\Delta &= z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}
\end{aligned}$$

Uwaga: $z_{1-\alpha/2}$ jest kwantylem rzędu $1 - \frac{\alpha}{2}$ rozkładu normalnego standardowego

# Kwantyl (uproszczona definicja!)

Rozważmy zmienną losową $X$ ze ściśle rosnącą dystrybuantą.
Kwantylem rzędu $p$ zmiennej $X$ nazywamy liczbę $x_p \in \mathbb{R}$ taką, że:

$$F_X(x_p) \; = \; P(X \leqslant x_p) \; = \; p, \qquad \text{czyli } x_p = F_X^{-1}(p)$$




Kwantyl $x_p$ otrzymujemy poprzez wyznaczenia wartości dystrybuanty odwrotnej (tzw. funkcji kwantylowej) w $p$

# Kwantyl $z_p$ dla $Z \sim N(0,1)$ a tablica wartości $\Phi(z)$

| ***z*** | ***0.00*** | ***0.01*** | ***0.02*** | ***0.03*** | ***0.04*** | ***0.05*** | ***0.06*** | ***0.07*** | ***0.08*** | ***0.09*** |
|---|---|---|---|---|---|---|---|---|---|---|
| ***0.00*** | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| ***0.10*** | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5674 | 0.5714 | 0.5753 |
| ***0.20*** | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| ***0.30*** | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| ***0.40*** | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| ***0.50*** | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| ***0.60*** | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| ***0.70*** | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| ***0.80*** | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| ***0.90*** | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.6315 | 0.8340 | 0.8365 | 0.8389 |
| ***1.00*** | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| ***1.10*** | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| ***1.20*** | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| ***1.30*** | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| ***1.40*** | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| ***1.50*** | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| ***1.60*** | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| ***1.70*** | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| ***1.80*** | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| ***1.90*** | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| ***2.00*** | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| ***2.10*** | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| ***2.20*** | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| ***2.30*** | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| ***2.40*** | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| ***2.50*** | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| ***2.60*** | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| ***2.70*** | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| ***2.80*** | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| ***2.90*** | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| ***3.00*** | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| ***3.10*** | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9992 | 0.9992 | 0.9992 | 0.9992 | 0.9993 | 0.9993 |
| ***3.20*** | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9995 | 0.9995 | 0.9995 |
| ***3.30*** | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9997 |
| ***3.40*** | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |

# Kwantyl $z_p$ dla $Z \sim N(0,1)$ a tablica wartości $\Phi(z)$

| *z* | *0.00* | *0.01* | *0.02* | *0.03* | *0.04* | *0.05* | *0.06* | *0.07* | *0.08* | *0.09* |
|---|---|---|---|---|---|---|---|---|---|---|
| ***0.00*** | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| ***0.10*** | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | | | | | |
| ***0.20*** | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | | | | | |
| ***0.30*** | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | | | | | |
| ***0.40*** | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | | | | | |
| ***0.50*** | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | | | | | |
| ***0.60*** | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | | | | | |
| ***0.70*** | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | | | | | |
| ***0.80*** | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | | | | | |
| ***0.90*** | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | | | | | |
| ***1.00*** | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | | | | | |
| ***1.10*** | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | | | | | |
| ***1.20*** | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | | | | | |
| ***1.30*** | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | | | | | |
| ***1.40*** | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | | | | | |
| ***1.50*** | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | | | | | |
| ***1.60*** | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| ***1.70*** | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| ***1.80*** | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| ***1.90*** | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| ***2.00*** | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| ***2.10*** | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| ***2.20*** | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| ***2.30*** | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| ***2.40*** | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| ***2.50*** | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| ***2.60*** | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| ***2.70*** | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| ***2.80*** | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| ***2.90*** | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| ***3.00*** | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| ***3.10*** | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9992 | 0.9992 | 0.9992 | 0.9992 | 0.9993 | 0.9993 |
| ***3.20*** | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9995 | 0.9995 | 0.9995 |
| ***3.30*** | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9997 |
| ***3.40*** | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |

Przykładowe wartości:

| $p$ | $z_p = \Phi^{-1}(p)$ |
|---|---|
| 0.95 | 1.64 |
| 0.975 | 1.96 |
| 0.99 | 2.33 |
| 0.995 | 2.58 |

# Kwantyl $z_p$ dla $Z \sim N(0,1)$ a tablica wartości $\Phi(z)$

| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| **0.00** | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| **0.10** | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | | | | | |
| **0.20** | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | | | | | |
| **0.30** | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | | | | | |
| **0.40** | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | | | | | |
| **0.50** | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | | | | | |
| **0.60** | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | | | | | |
| **0.70** | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | | | | | |
| **0.80** | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | | | | | |
| **0.90** | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | | | | | |
| **1.00** | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | | | | | |
| **1.10** | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | | | | | |
| **1.20** | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | | | | | |
| **1.30** | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | | | | | |
| **1.40** | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | | | | | |
| **1.50** | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | | | | | |
| **1.60** | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| **1.70** | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| **1.80** | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9 | | | | | |
| **1.90** | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9 | | | | | |
| **2.00** | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9 | | | | | |
| **2.10** | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9 | | | | | |
| **2.20** | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9 | | | | | |
| **2.30** | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9 | | | | | |
| **2.40** | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9 | | | | | |
| **2.50** | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9 | | | | | |
| **2.60** | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9 | | | | | |
| **2.70** | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9 | | | | | |
| **2.80** | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9 | | | | | |
| **2.90** | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9 | | | | | |
| **3.00** | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9 | | | | | |
| **3.10** | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9 | | | | | |
| **3.20** | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9 | | | | | |
| **3.30** | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9997 |
| **3.40** | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |

Przykładowe wartości:

| $p$ | $z_p = \Phi^{-1}(p)$ |
|---|---|
| 0.95 | 1.64 |
| 0.975 | 1.96 |
| 0.99 | 2.33 |
| 0.995 | 2.58 |

Przykładowe wartości:

| $\alpha$ | $z_{1-\alpha/2} = \Phi^{-1}(1-\alpha/2)$ |
|---|---|
| 0.1 | 1.64 |
| 0.05 | 1.96 |
| 0.02 | 2.33 |
| 0.01 | 2.58 |

# Przykład: rozkład normalny ze znaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym ze znaną wariancją $\sigma^2 = 16$ dała wartość statystyki $\overline{x}_n = 2$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

# Przykład: rozkład normalny ze znaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym ze znaną wariancją $\sigma^2 = 16$ dała wartość statystyki $\overline{x}_n = 2$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

Rozwiązanie: Przedział ufności ma postać:

$$[\overline{x}_n - \Delta, \ \overline{x}_n + \Delta], \qquad \text{gdzie} \ \ \Delta = z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

# Przykład: rozkład normalny ze znaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym ze znaną wariancją $\sigma^2 = 16$ dała wartość statystyki $\overline{x}_n = 2$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

Rozwiązanie: Przedział ufności ma postać:

$$[\overline{x}_n - \Delta, \ \overline{x}_n + \Delta], \qquad \text{gdzie} \ \ \Delta = z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$z_{1-\alpha/2} \ = \ z_{0.975} \ = \ \Phi^{-1}(0.975) \ = \ 1.96$$

# Przykład: rozkład normalny ze znaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym ze znaną wariancją $\sigma^2 = 16$ dała wartość statystyki $\overline{x}_n = 2$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

Rozwiązanie: Przedział ufności ma postać:

$$[\overline{x}_n - \Delta, \ \overline{x}_n + \Delta], \qquad \text{gdzie} \ \ \Delta = z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\begin{aligned} z_{1-\alpha/2} &= z_{0.975} = \Phi^{-1}(0.975) = 1.96 \\ \Delta &= 1.96\frac{4}{\sqrt{16}} = 1.96 \end{aligned}$$

# Przykład: rozkład normalny ze znaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym ze znaną wariancją $\sigma^2 = 16$ dała wartość statystyki $\overline{x}_n = 2$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

Rozwiązanie: Przedział ufności ma postać:

$$[\overline{x}_n - \Delta, \ \overline{x}_n + \Delta], \qquad \text{gdzie} \ \ \Delta = z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$\begin{aligned}
z_{1-\alpha/2} &= z_{0.975} = \Phi^{-1}(0.975) = 1.96 \\
\Delta &= 1.96 \frac{4}{\sqrt{16}} = 1.96 \\
[\overline{x}_n - \Delta, \ \overline{x}_n + \Delta] &= [2 - 1.96, \ 2 + 1.96] = [0.04, 3.96]
\end{aligned}$$

# Przykład: rozkład dwupunktowy

Niech $X \sim B(p)$. Znajdziemy przedział ufności dla parametru $p$.

# Przykład: rozkład dwupunktowy

Niech $X \sim B(p)$. Znajdziemy przedział ufności dla parametru $p$.

Estymator parametru $p$ na podstawie próby $X_1, \ldots, X_n$:

$$\widehat{p} \; = \; \overline{X}_n \; = \; \frac{1}{n} \sum_{i=1}^{n} X_i \; = \; \frac{\#\text{ sukcesów}}{n}$$

# Przykład: rozkład dwupunktowy

Niech $X \sim B(p)$. Znajdziemy przedział ufności dla parametru $p$.

Estymator parametru $p$ na podstawie próby $X_1, \ldots, X_n$:

$$\widehat{p} \;=\; \overline{X}_n \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i \;=\; \frac{\#\ \text{sukcesów}}{n}$$

$$E(\widehat{p}) \;=\; p, \qquad D^2(\widehat{p}) \;=\; \frac{D^2(X)}{n} \;=\; \frac{p(1-p)}{n},$$

# Przykład: rozkład dwupunktowy

Niech $X \sim B(p)$. Znajdziemy przedział ufności dla parametru $p$.

Estymator parametru $p$ na podstawie próby $X_1, \ldots, X_n$:

$$\widehat{p} \;=\; \overline{X}_n \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i \;=\; \frac{\#\ \text{sukcesów}}{n}$$

$$E(\widehat{p}) \;=\; p, \qquad D^2(\widehat{p}) \;=\; \frac{D^2(X)}{n} \;=\; \frac{p(1-p)}{n},$$

Z Centralnego Twierdzenia Granicznego mamy:

$$\frac{\widehat{p} - E(\widehat{p})}{D(\widehat{p})} \;=\; \frac{\widehat{p} - p}{\sqrt{p(1-p)}}\sqrt{n} \;\xrightarrow{D}\; Z \;\sim\; N(0,1)$$

# Przykład: rozkład dwupunktowy

Niech $X \sim B(p)$. Znajdziemy przedział ufności dla parametru $p$.

Estymator parametru $p$ na podstawie próby $X_1, \ldots, X_n$:

$$\widehat{p} = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{\#\text{ sukcesów}}{n}$$

$$E(\widehat{p}) = p, \qquad D^2(\widehat{p}) = \frac{D^2(X)}{n} = \frac{p(1-p)}{n},$$

Z Centralnego Twierdzenia Granicznego mamy:

$$\frac{\widehat{p} - E(\widehat{p})}{D(\widehat{p})} = \frac{\widehat{p} - p}{\sqrt{p(1-p)}}\sqrt{n} \xrightarrow{D} Z \sim N(0,1)$$

Z kolei z Prawa Wielkich Liczb mamy $\widehat{p} \xrightarrow{P} p$, a tym samym*:

$$\frac{\widehat{p} - p}{\sqrt{\widehat{p}(1-\widehat{p})}}\sqrt{n} \xrightarrow{D} Z \sim N(0,1)$$

Założymy, że $n$ jest wystarczająco duże aby skorzystać z tego przybliżenia.

---

*nietrywialne: jeśli $X_n \xrightarrow{D} X$ i $Y_n \xrightarrow{P} c$ to $X_n f(Y_n) \xrightarrow{D} X f(c)$ dla ciągłej $f$.

# Przykład: rozkład dwupunktowy

Konstruujemy przedział ufności na poziomie ufności $1 - \alpha$:

$$1 - \alpha \quad = \quad P\left(-\Delta \ \leqslant \ \widehat{p} - p \ \leqslant \ \Delta\right)$$

# Przykład: rozkład dwupunktowy

Konstruujemy przedział ufności na poziomie ufności $1-\alpha$:

$$
\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \widehat{p}-p \leqslant \Delta\right) \\
&= P\left(-\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}} \leqslant \underbrace{\frac{\widehat{p}-p}{\sqrt{\widehat{p}(1-\widehat{p})}}\sqrt{n}}_{\simeq Z} \leqslant \frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right)
\end{aligned}
$$

# Przykład: rozkład dwupunktowy

Konstruujemy przedział ufności na poziomie ufności $1-\alpha$:

$$\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \widehat{p}-p \leqslant \Delta\right) \\
&= P\Bigg(-\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}} \leqslant \underbrace{\frac{\widehat{p}-p}{\sqrt{\widehat{p}(1-\widehat{p})}}\sqrt{n}}_{\simeq Z} \leqslant \frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\Bigg) \\
&\simeq \Phi\left(\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) - \Phi\left(-\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) = 2\Phi\left(\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) - 1
\end{aligned}$$

# Przykład: rozkład dwupunktowy

Konstruujemy przedział ufności na poziomie ufności $1-\alpha$:

$$
\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \widehat{p}-p \leqslant \Delta\right) \\
&= P\Bigg(-\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}} \leqslant \underbrace{\frac{\widehat{p}-p}{\sqrt{\widehat{p}(1-\widehat{p})}}\sqrt{n}}_{\simeq Z} \leqslant \frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\Bigg) \\
&\simeq \Phi\left(\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) - \Phi\left(-\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) = 2\Phi\left(\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) - 1
\end{aligned}
$$

Rozwiązując równanie otrzymujemy:

$$
\Delta = z_{1-\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}
$$

# Przykład: rozkład dwupunktowy

Konstruujemy przedział ufności na poziomie ufności $1-\alpha$:

$$
\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \widehat{p}-p \leqslant \Delta\right) \\
&= P\left(-\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}} \leqslant \underbrace{\frac{\widehat{p}-p}{\sqrt{\widehat{p}(1-\widehat{p})}}\sqrt{n}}_{\simeq Z} \leqslant \frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) \\
&\simeq \Phi\left(\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) - \Phi\left(-\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) = 2\Phi\left(\frac{\Delta\sqrt{n}}{\sqrt{\widehat{p}(1-\widehat{p})}}\right) - 1
\end{aligned}
$$

Rozwiązując równanie otrzymujemy:

$$
\Delta = z_{1-\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}}
$$

**Zadanie 2**

Rozwiąż powyższe zagadnienie krok po kroku

# Przykład: rozkład dwupunktowy

W celu oszacowania poparcia kandydata $A$ w drugiej turze wyborów prezydenckich pobrano losowo próbę o liczności $n = 1000$ osób, odnotowując czy dana osoba głosuje na kandydata $A$ ($X = 1$), czy na jego konkurenta $B$ ($X = 0$). Uzyskano 550 głosów za $A$ i 450 głosów za $B$. Wyznacz przedział ufności dla poparcia kandydata $A$ dla $\alpha = 0.05$.

# Przykład: rozkład dwupunktowy

W celu oszacowania poparcia kandydata $A$ w drugiej turze wyborów prezydenckich pobrano losowo próbę o liczności $n = 1000$ osób, odnotowując czy dana osoba głosuje na kandydata $A$ ($X = 1$), czy na jego konkurenta $B$ ($X = 0$). Uzyskano 550 głosów za $A$ i 450 głosów za $B$. Wyznacz przedział ufności dla poparcia kandydata $A$ dla $\alpha = 0.05$.

Rozwiązanie:

$$\widehat{p} = \frac{550}{1000} = 0.55$$

# Przykład: rozkład dwupunktowy

W celu oszacowania poparcia kandydata $A$ w drugiej turze wyborów prezydenckich pobrano losowo próbę o liczności $n = 1000$ osób, odnotowując czy dana osoba głosuje na kandydata $A$ ($X = 1$), czy na jego konkurenta $B$ ($X = 0$). Uzyskano 550 głosów za $A$ i 450 głosów za $B$. Wyznacz przedział ufności dla poparcia kandydata $A$ dla $\alpha = 0.05$.

Rozwiązanie:

$$
\begin{aligned}
\widehat{p} &= \frac{550}{1000} = 0.55 \\
z_{1-\alpha/2} &= z_{0.975} = \Phi^{-1}(0.975) = 1.96
\end{aligned}
$$

# Przykład: rozkład dwupunktowy

W celu oszacowania poparcia kandydata $A$ w drugiej turze wyborów prezydenckich pobrano losowo próbę o liczności $n = 1000$ osób, odnotowując czy dana osoba głosuje na kandydata $A$ ($X = 1$), czy na jego konkurenta $B$ ($X = 0$). Uzyskano 550 głosów za $A$ i 450 głosów za $B$. Wyznacz przedział ufności dla poparcia kandydata $A$ dla $\alpha = 0.05$.

Rozwiązanie:

$$\begin{aligned}
\widehat{p} &= \frac{550}{1000} = 0.55 \\
z_{1-\alpha/2} &= z_{0.975} = \Phi^{-1}(0.975) = 1.96 \\
\Delta &= z_{1-\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = 1.96\sqrt{\frac{0.55 \cdot 0.45}{1000}} \simeq 0.031
\end{aligned}$$

# Przykład: rozkład dwupunktowy

W celu oszacowania poparcia kandydata $A$ w drugiej turze wyborów prezydenckich pobrano losowo próbę o liczności $n = 1000$ osób, odnotowując czy dana osoba głosuje na kandydata $A$ ($X = 1$), czy na jego konkurenta $B$ ($X = 0$). Uzyskano 550 głosów za $A$ i 450 głosów za $B$. Wyznacz przedział ufności dla poparcia kandydata $A$ dla $\alpha = 0.05$.

Rozwiązanie:

$$\begin{aligned}
\widehat{p} &= \frac{550}{1000} = 0.55 \\
z_{1-\alpha/2} &= z_{0.975} = \Phi^{-1}(0.975) = 1.96 \\
\Delta &= z_{1-\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}} = 1.96\sqrt{\frac{0.55\cdot 0.45}{1000}} \simeq 0.031 \\
[\widehat{p}-\Delta,\ \widehat{p}+\Delta] &= [0.55-0.031,\ 0.55+0.031] = [0.519,\ 0.581]
\end{aligned}$$

# Przykład: rozkład normalny z nieznaną wariancją

Niech $X \sim N(\mu, \sigma^2)$, ale teraz nie znamy $\sigma^2$

# Przykład: rozkład normalny z nieznaną wariancją

Niech $X \sim N(\mu, \sigma^2)$, ale teraz nie znamy $\sigma^2$

Jak poprzednio, jako estymator wartości oczekiwanej $\mu$ wybieramy $\widehat{\mu} = \overline{X}_n$

# Przykład: rozkład normalny z nieznaną wariancją

Niech $X \sim N(\mu, \sigma^2)$, ale teraz nie znamy $\sigma^2$

Jak poprzednio, jako estymator wartości oczekiwanej $\mu$ wybieramy $\widehat{\mu} = \overline{X}_n$

Ponieważ nie znamy wariancji $\sigma^2$, estymujemy ją z próby:

$$\widehat{\sigma}^2_* = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$$

# Przykład: rozkład normalny z nieznaną wariancją

Niech $X \sim N(\mu, \sigma^2)$, ale teraz nie znamy $\sigma^2$

Jak poprzednio, jako estymator wartości oczekiwanej $\mu$ wybieramy $\widehat{\mu} = \overline{X}_n$

Ponieważ nie znamy wariancji $\sigma^2$, estymujemy ją z próby:

$$\widehat{\sigma}_*^2 \; = \; \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$$

Wprowadzamy nową zmienną:

$$T \; = \; \frac{\overline{X}_n - \mu}{\widehat{\sigma}_*}\sqrt{n}$$

# Przykład: rozkład normalny z nieznaną wariancją

Niech $X \sim N(\mu, \sigma^2)$, ale teraz nie znamy $\sigma^2$

Jak poprzednio, jako estymator wartości oczekiwanej $\mu$ wybieramy $\widehat{\mu} = \overline{X}_n$

Ponieważ nie znamy wariancji $\sigma^2$, estymujemy ją z próby:

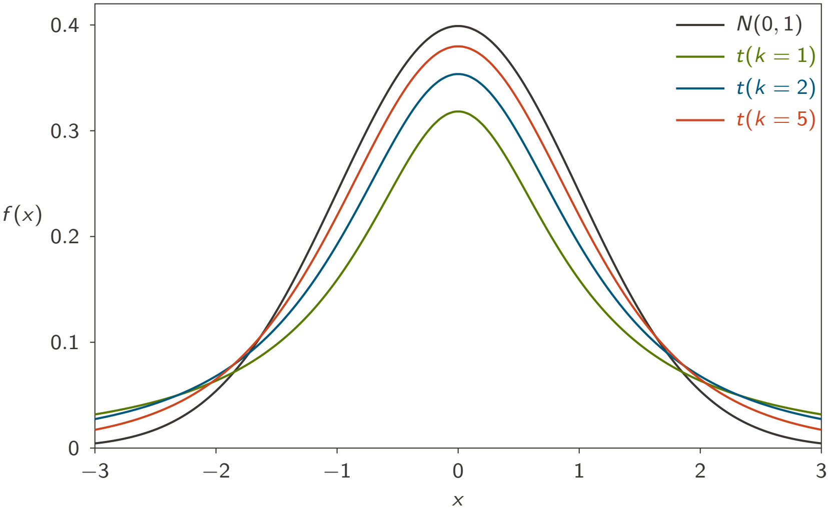
$$\widehat{\sigma}^2_* = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}_n)^2$$

Wprowadzamy nową zmienną:

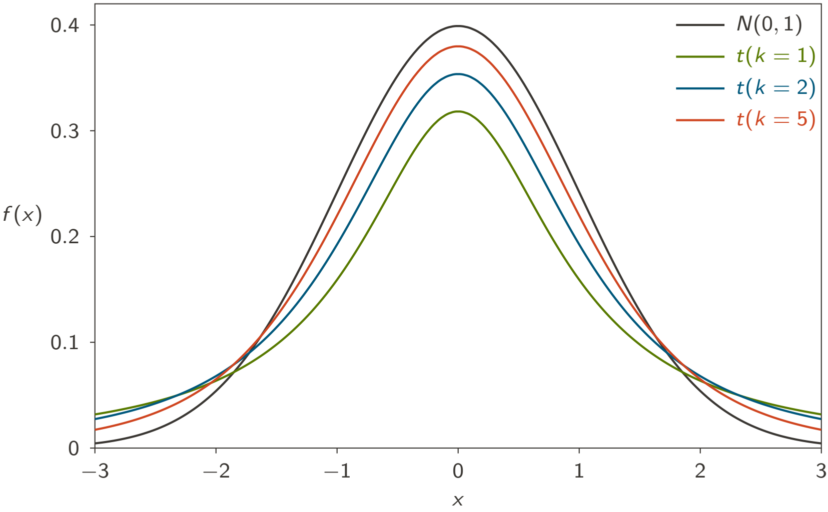
$$T = \frac{\overline{X}_n - \mu}{\widehat{\sigma}_*} \sqrt{n}$$

Można pokazać, że zmienna $T$ ma rozkład $t$-Studenta o $n-1$ stopniach swobody

# Rozkład $t$-Studenta o $k$ stopniach swobody



# Rozkład $t$-Studenta o $k$ stopniach swobody




Dla $k \to \infty$ rozkład $t$ zbiega do rozkładu $N(0,1)$

# Przykład: rozkład normalny z nieznaną wariancją

$$1 - \alpha \quad = \quad P\left(-\Delta \ \leqslant \ \overline{X}_n - \mu \ \leqslant \ \Delta\right)$$

# Przykład: rozkład normalny z nieznaną wariancją

$$
\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \overline{X}_n - \mu \leqslant \Delta\right) \\
&= P\left(-\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n} \leqslant \frac{\overline{X}_n - \mu}{\widehat{\sigma}_*}\sqrt{n} \leqslant \frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right)
\end{aligned}
$$

# Przykład: rozkład normalny z nieznaną wariancją

$$
\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \overline{X}_n - \mu \leqslant \Delta\right) \\
&= P\left(-\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n} \leqslant \frac{\overline{X}_n - \mu}{\widehat{\sigma}_*}\sqrt{n} \leqslant \frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) \\
&= P\left(-\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n} \leqslant T \leqslant \frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right)
\end{aligned}
$$

# Przykład: rozkład normalny z nieznaną wariancją

$$\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \overline{X}_n - \mu \leqslant \Delta\right) \\
&= P\left(-\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n} \leqslant \frac{\overline{X}_n - \mu}{\widehat{\sigma}_*}\sqrt{n} \leqslant \frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) \\
&= P\left(-\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n} \leqslant T \leqslant \frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) \\
&= F_T\left(\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) - F_T\left(-\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) = 2F_T\left(\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) - 1,
\end{aligned}$$

gdzie użyliśmy symetrii rozkładu $t$-Studenta: $F_T(-x) = 1 - F_T(x)$.

# Przykład: rozkład normalny z nieznaną wariancją

$$
\begin{aligned}
1-\alpha &= P\left(-\Delta \leqslant \overline{X}_n - \mu \leqslant \Delta\right) \\
&= P\left(-\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n} \leqslant \frac{\overline{X}_n - \mu}{\widehat{\sigma}_*}\sqrt{n} \leqslant \frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) \\
&= P\left(-\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n} \leqslant T \leqslant \frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) \\
&= F_T\left(\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) - F_T\left(-\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) = 2F_T\left(\frac{\Delta}{\widehat{\sigma}_*}\sqrt{n}\right) - 1,
\end{aligned}
$$

gdzie użyliśmy symetrii rozkładu $t$-Studenta: $F_T(-x) = 1 - F_T(x)$.

Postępując analogicznie do przypadku znanej wariancji dostajemy:

$$\Delta = t_{1-\alpha/2;n-1}\frac{\widehat{\sigma}_*}{\sqrt{n}}, \qquad \text{gdzie} \quad t_{1-\alpha/2;n-1} = F_T^{-1}\left(1-\frac{\alpha}{2}\right)$$

jest kwantylem rzędu $1-\frac{\alpha}{2}$ rozkładu $t$-Studenta o $n-1$ stopniach swobody

# Przykładowa tablica kwantyli rozkładu $t$-Studenta

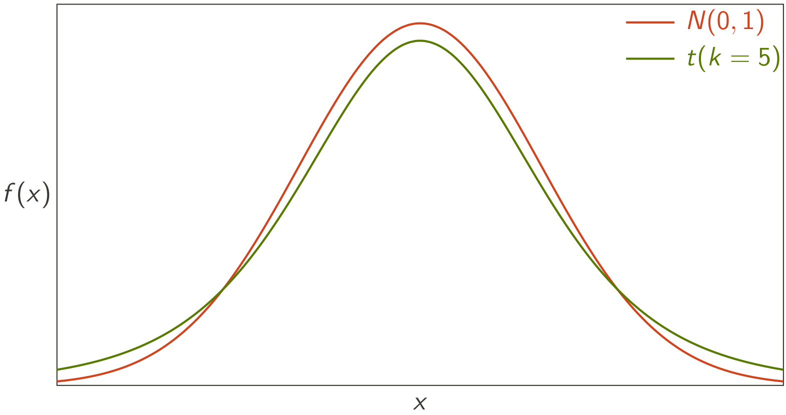
| $df$ | $t_{.90}$ | $t_{.95}$ | $t_{.975}$ | $t_{.99}$ | $t_{.995}$ |
|---|---|---|---|---|---|
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 |
| 32 | 1.309 | 1.694 | 2.037 | 2.449 | 2.738 |
| 34 | 1.307 | 1.691 | 2.032 | 2.441 | 2.728 |
| 36 | 1.306 | 1.688 | 2.028 | 2.434 | 2.719 |
| 38 | 1.304 | 1.686 | 2.024 | 2.429 | 2.712 |
| $\infty$ | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 |

# Przykładowa tablica kwantyli rozkładu $t$-Studenta

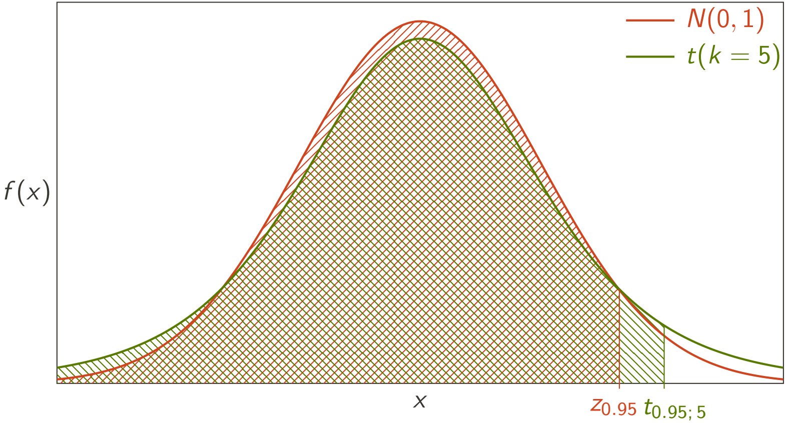
| $df$ | $t_{.90}$ | $t_{.95}$ | $t_{.975}$ | $t_{.99}$ | $t_{.995}$ |
|---|---|---|---|---|---|
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 |
| 2 | [illegible] | [illegible] | [illegible] | 6.965 | 9.925 |
| 3 | [illegible] | [illegible] | [illegible] | 4.541 | 5.841 |
| 4 | [illegible] | [illegible] | [illegible] | 3.747 | 4.604 |
| 5 | [illegible] | [illegible] | [illegible] | 3.365 | 4.032 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 |
| 32 | 1.309 | 1.694 | 2.037 | 2.449 | 2.738 |
| 34 | 1.307 | 1.691 | 2.032 | 2.441 | 2.728 |
| 36 | 1.306 | 1.688 | 2.028 | 2.434 | 2.719 |
| 38 | 1.304 | 1.686 | 2.024 | 2.429 | 2.712 |
| $\infty$ | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 |

Liczba stopni swobody $k$
(df – *degrees of freedom*)

# Rozkład $t$-Studenta



# Rozkład $t$-Studenta




Kwantyle rozkładu $t$-Studenta są nieznacznie większe od odpowiadających im kwantyli rozkładu normalnego (rozkład $t$-Studenta ma grubsze ogony).

Jest to spowodowane estymowaniem wariancji z danych (zwiększa niepewność przy konstrukcji przedziału ufności)

# Przykład: rozkład normalny z nieznaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym statystyki $\overline{x}_n = 1$ oraz $\widehat{s}_*^2 = 9$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

# Przykład: rozkład normalny z nieznaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym statystyki $\overline{x}_n = 1$ oraz $\widehat{s}_*^2 = 9$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

Rozwiązanie: Przedział ufności ma postać:

$$\left[\overline{x}_n - \Delta,\ \ \overline{x}_n + \Delta\right], \qquad \text{gdzie}\ \ \Delta = t_{1-\alpha/2;n-1}\frac{\widehat{s}_*}{\sqrt{n}}$$

# Przykład: rozkład normalny z nieznaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym statystyki $\overline{x}_n = 1$ oraz $\widehat{s}_*^2 = 9$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

Rozwiązanie: Przedział ufności ma postać:

$$[\overline{x}_n - \Delta,\ \ \overline{x}_n + \Delta], \qquad \text{gdzie} \ \ \Delta = t_{1-\alpha/2;n-1} \frac{\widehat{s}_*}{\sqrt{n}}$$

$$t_{1-\alpha/2;n-1} \ = \ t_{0.975;15} \ = \ 2.131$$

# Przykład: rozkład normalny z nieznaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym statystyki $\overline{x}_n = 1$ oraz $\widehat{s}_*^2 = 9$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

Rozwiązanie: Przedział ufności ma postać:

$$\left[\overline{x}_n - \Delta,\ \ \overline{x}_n + \Delta\right], \qquad \text{gdzie} \ \ \Delta = t_{1-\alpha/2;n-1} \frac{\widehat{s}_*}{\sqrt{n}}$$

$$\begin{aligned} t_{1-\alpha/2;n-1} &= t_{0.975;15} = 2.131 \\ \Delta &= 2.145 \frac{3}{\sqrt{16}} = 1.609 \end{aligned}$$

# Przykład: rozkład normalny z nieznaną wariancją

Próba o liczności $n = 16$ pomiarów pewnej cechy o rozkładzie normalnym statystyki $\overline{x}_n = 1$ oraz $\widehat{s}_*^2 = 9$. Wyznacz przedział ufności dla wartości oczekiwanej na poziomie ufności $1 - \alpha$ dla $\alpha = 0.05$.

Rozwiązanie: Przedział ufności ma postać:

$$[\overline{x}_n - \Delta, \ \ \overline{x}_n + \Delta], \qquad \text{gdzie} \ \ \Delta = t_{1-\alpha/2;n-1} \frac{\widehat{s}_*}{\sqrt{n}}$$

$$\begin{aligned} t_{1-\alpha/2;n-1} &= t_{0.975;15} = 2.131 \\ \Delta &= 2.145 \frac{3}{\sqrt{16}} = 1.609 \\ [\overline{x}_n - \Delta, \ \ \overline{x}_n + \Delta] &= [1 - 1.609, \ 1 + 1.609] = [-0.609, 2.609] \end{aligned}$$