## Wykład 2: Rozkład normalny (gaussowski)
## Rozkłady pojawiające się w testach statystycznych

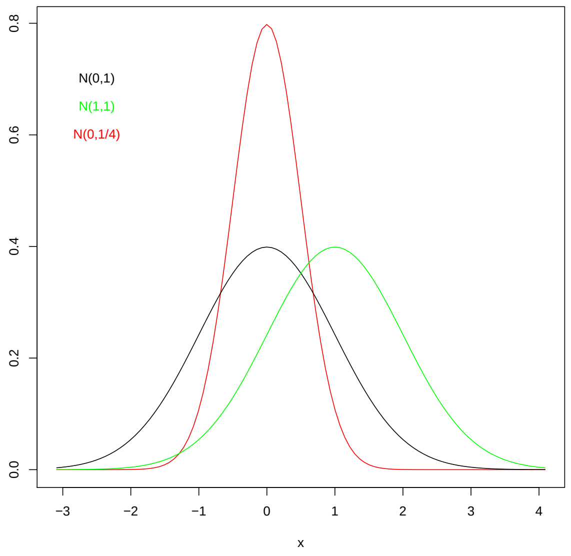
### ROZKŁAD NORMALNY

**Definicja.** *Rozkład normalny* (zwany także *rozkładem gaussowskim*) o parametrach $\mu$ i $\sigma^2$, oznaczany symbolem $\mathcal{N}(\mu, \sigma^2)$, to rozkład ciągły o gęstości

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right), \quad x \in \mathbb{R}, \text{ gdzie } \mu \in \mathbb{R}, \ \sigma > 0.$$

Własności rozkładu normalnego:

- Wykres gęstości rozkładu normalnego to tzw. *krzywa gaussa*:




- Jeśli $X \sim \mathcal{N}(\mu, \sigma^2)$, to $EX = \mu$ i $Var(X) = \sigma^2$.
- Rozkład normalny $\mathcal{N}(\mu, \sigma^2)$ jest rozkładem symetrycznym względem $\mu$.
- Jeśli $X \sim \mathcal{N}(\mu, \sigma^2)$, to $U = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$. Rozkład $\mathcal{N}(0, 1)$ nazywamy *standardowym rozkładem normalnym.*
- Ogólniej jeśli $X \sim \mathcal{N}(\mu, \sigma^2)$ i $a \neq 0, b$ są liczbami rzeczywistymi, to $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$.

**Twierdzenie** (Centralne twierdzenie graniczne Lindeberga-Lévy'ego)**.**
Załóżmy, że $X_1, X_2, \ldots$ jest ciągiem niezależnych zmiennych losowych o tym samym rozkładzie ze skończoną wariancją. Oznaczmy $\mu = E(X_1)$, $\sigma^2 = Var(X_1)$ i $S_n = \sum_{i=1}^n X_i$. Wówczas

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} U \sim \mathcal{N}(0, 1),$$

co oznacza, że

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leqslant x\right) = F_{\mathcal{N}(0,1)}(x) \text{ dla każdego } x \in \mathbb{R},$$

gdzie $F_{\mathcal{N}(0,1)}(x)$ to dystrybuanta standardowego rozkładu normalnego.

## ROZKŁADY POJAWIAJĄCE SIĘ W TESTACH STATYSTYCZNYCH

**Definicja.** *Rozkład chi-kwadrat o $n$ stopniach swobody*, który będziemy oznaczać symbolem $\chi^2_{[n]}$, to rozkład sumy kwadratów $n$ niezależnych zmiennych losowych o standardowym rozkładzie normalnym:

$$\text{jeśli } U_1, U_2, \ldots, U_n \text{ są niezależnymi zmiennymi losowymi o rozkładzie } \mathcal{N}(0,1)$$
$$\text{i } X = U_1^2 + U_2^2 + \ldots + U_n^2, \text{ to } X \sim \chi^2_{[n]}.$$

Własności rozkładu chi-kwadrat:

- Rozkład $\chi^2_{[n]}$ to rozkład ciągły o gęstości danej wzorem
$$f(x) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} x^{n/2-1} \exp(-x/2) & \text{dla } x > 0 \\ 0 & \text{dla } x \leqslant 0 \end{cases}.$$
- Rozkład $\chi^2_{[n]}$ to to samo co rozkład $Gamma(a = n/2, s = 2)$.

**Definicja.** *Rozkład F-Snedecora o stopniach swobody $m$ i $n$*, który będziemy oznaczać symbolem $F_{[m,n]}$, to rozkład ilorazu niezależnych zmiennych losowych o rozkładach chi-kwadrat $\chi^2_{[m]}$ i $\chi^2_{[n]}$, podzielonych przez ich stopnie swobody:

$$\text{jeśli } W \text{ i } Z \text{ są niezależnymi zmiennymi losowymi, takimi że } W \sim \chi^2_{[m]}, Z \sim \chi^2_{[n]}$$
$$\text{i } X = \frac{W/m}{Z/n}, \text{ to } X \sim F_{[m,n]}.$$

Własności rozkładu F-Snedecora:

- Rozkład $F_{[m,n]}$ to rozkład ciągły o gęstości danej wzorem
$$f(x) = \begin{cases} \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{n}\right)^{m/2} \frac{x^{(m-2)/2}}{\left(1+\frac{m}{n}x\right)^{(m+n)/2}} & \text{dla } x > 0 \\ 0 & \text{dla } x \leqslant 0 \end{cases}.$$
- Jeśli $X \sim F_{[m,n]}$, to
  1. $EX = \frac{n}{n-2}$ dla $n > 2$,
  2. $Var(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ dla $n > 4$.

**Definicja.** *Rozkład t-Studenta o $n$ stopniach swobody*, który będziemy oznaczać symbolem $t_{[n]}$, to rozkład ilorazu zmiennej losowej o standardowym rozkładzie normalnym $\mathcal{N}(0,1)$ i pierwiastka ze zmiennej losowej o rozkładzie chi-kwadrat$\chi^2_{[n]}$, podzielonej przez jej stopnie swobody, przy założeniu, że te zmienne losowe są niezależne:

$$\text{jeśli } U \text{ i } Z \text{ są niezależnymi zmiennymi losowymi, takimi że } U \sim \mathcal{N}(0,1), Z \sim \chi^2_{[n]}$$
$$\text{i } X = \frac{U}{\sqrt{Z/n}}, \text{ to } X \sim t_{[n]}.$$

Własności rozkładu t-Studenta:

- Rozkład $t_{[n]}$ to rozkład ciągły o gęstości danej wzorem
$$f(x) = \frac{\Gamma((n+1)/2)}{\Gamma(n/2)} \frac{1}{\sqrt{n\pi}} \frac{1}{(1+x^2/n)^{(n+1)/2}} \text{ dla } x \in \mathbb{R}.$$
- Jeśli $X \sim t_{[n]}$, to
  1. $EX = 0$ dla $n > 1$,
  2. $Var(X) = \frac{n}{n-2}$ dla $n > 2$.