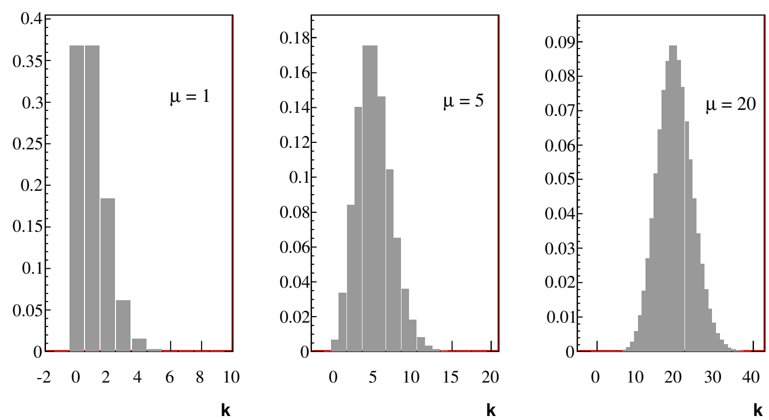


Rysunek 10: Przykłady rozkładu Poissona dla różnych wartości parametru μ

Definicja 13.4 (Proces Poissona)**.**
Proces Poissona o intensywności λ jest zbiorem $X_1, X_2, ...$ punktów losowych o własnościach:

- $$N_A = \text{Poi}\,(\mu = \lambda|A|)$$
  gdzie N_A jest ilością punktów losowych zawartych w zbiorze A, λ jest intensywnością procesu a $|A|$ jest miarą zbioru A.
- Dla rozłącznych zbiorów $A_1, A_2, ..., A_n$ zmienne losowe $N_{A_1}, N_{A_2}, ..., N_{A_n}$ są niezależne.

Proces Poissona modeluje najbardziej losowy sposób dystrybucji punktów losowych w czasie lub przestrzeni (długości, powierzchni, objętości,...).

Definicja 13.5 (Rozkład jednostajny)**.**
Jeśli X ma gęstość prawdopodobieństwa:

$$f(x) = \frac{1}{\beta - \alpha} \quad x \in [\alpha, \beta]$$

to mówimy, że ma rozkład jednostajny na przedziale $[\alpha, \beta]$

- oznaczmy go jako unif (α, β);
- X opisuje losowy wybór liczby z przedziału $[\alpha, \beta]$

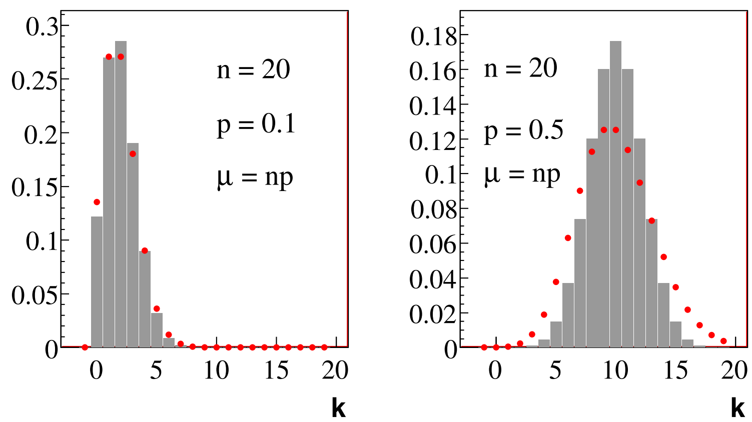
Twierdzenie 13.5 (Rozkład jednostajny)**.**
Niech $X \sim \mathit{unif}\,(\alpha, \beta)$

$$E[X] = \frac{\alpha + \beta}{2}$$
$$\mathit{Var}\,[X] = \frac{(\beta - \alpha)^2}{12}$$

Definicja 13.6 (Rozkład wykładniczy)**.**
Jeśli X ma gęstość prawdopodobieństwa:

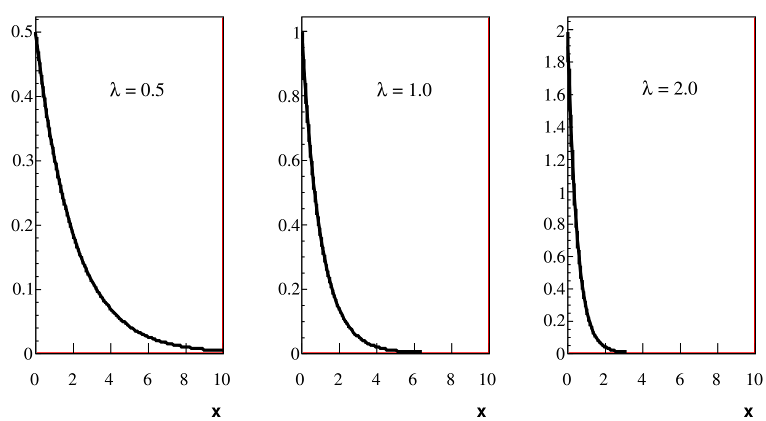
$$f(x) = \lambda e^{-\lambda x} \quad x > 0$$

to mówimy, że ma rozkład wykładniczy o parametrze λ ($\lambda > 0$)



Rysunek 11: Porównanie rozkładu dwumianowego (histogram) z rozkładem Poissona (punkty) dla różnych wartości parametru p

- oznaczmy go jako $\exp(\lambda)$;
- X opisuje czas życia obiektów które się nie starzeją
- λ jest odwrotnością średniego czasu życia obiektu.



Rysunek 12: Przykłady rozkładu wykładniczego dla różnych wartości parametru λ

Twierdzenie 13.6 (Rozkład wykładniczy)**.**
Niech $X \sim exp\ (\lambda)$

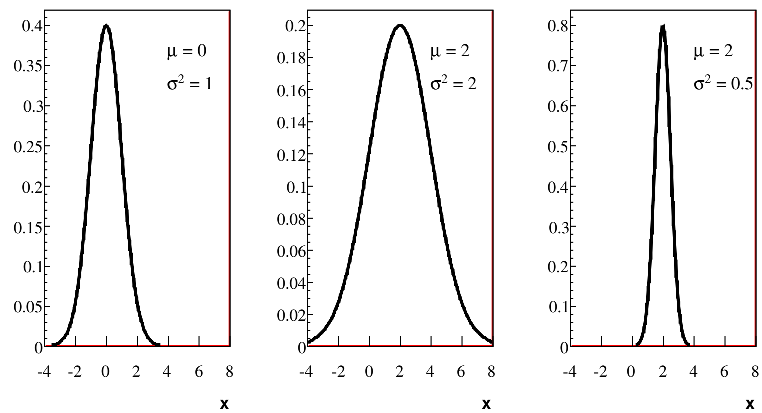
$$E[X] = \frac{1}{\lambda}; \qquad Var\,[X] = \frac{1}{\lambda^2}$$

Definicja 13.7 (Rozkład normalny(Gaussa))**.**
Jeśli X ma gęstość prawdopodobieństwa:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$$

to mówimy, że ma rozkład normalny o parametrach μ, σ^2 ($|\mu| < \infty; \sigma \neq 0$)

- oznaczmy go jako N (μ, σ^2);
- X opisuje wiele sytuacji w których wielkość o wartości μ ulega "rozmyciu" z powodu np. błędów pomiarowych, fluktuacji statystycznych, szumów, ...



Rysunek 13: Przykłady rozkładu normalnego dla różnych wartości parametrów μ i σ

Twierdzenie 13.7 (Rozkład normalny)**.**
Niech $X \sim N(\mu, \sigma^2)$

$$E[X] = \mu; \qquad Var[X] = \sigma^2$$

Twierdzenie 13.8.
Niech $X \sim N(\mu, \sigma^2)$ *a* $Z = \frac{X - \mu}{\sigma}$*. Wtedy*

$$Z \sim N(0, 1)$$

Rozkład $N(0, 1)$ *nazywamy* *standardowym* *rozkładem normalnym*

Dowód.

$$\begin{aligned} F_Z(z) = P(Z < z) &= P\left(\frac{X-\mu}{\sigma} < z\right) \\ &= P(X < \mu + z\sigma) \\ &= \int_{-\infty}^{\mu+z\sigma} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \\ & \quad y = \frac{t-\mu}{\sigma} \\ &= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y)^2} dy = \Phi(z) \end{aligned}$$

$\Phi(z)$ jest dystrybuantą $\mathrm{N}(0,1)$ $\square$

Twierdzenie 13.9 (Rozkład normalny standardowy)**.**
Niech $X \sim N(\mu, \sigma^2)$ *dla dowolnych* $a, b \in \mathbb{R}$ *zachodzi*

$$F_X(x) = P(X \leqslant x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$P(a \leqslant X \leqslant b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Dowód.

$$\begin{aligned} F_X(x) = P(X \leqslant x) &= P(X - \mu \leqslant x - \mu) \\ &= P\left(\frac{X-\mu}{\sigma} \leqslant \frac{x-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right) \end{aligned}$$

$\square$

Twierdzenie 13.10 (Rozkład normalny standardowy)**.**
Dla $x \in \mathbb{R}$

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

Dowód.

$$\begin{aligned} \Phi(-x) &= \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t)^2} dt \\ &= \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t)^2} dt \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t)^2} dt - \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(t)^2} dt \\ &= 1 - \Phi(x) \end{aligned}$$

$\square$

Twierdzenie 13.11 (Kombinacja liniowa rozkładu normalnego)**.**
Niech $X \sim N(\mu, \sigma^2)$ *a* $a, b \in \mathbb{R}$ *oraz* $Y = aX + b$*. Wtedy*

$$Y \sim N(a\mu + b, a^2\sigma^2)$$

Przykład 13.1 (Rozkład normalny)**.**
Przyjmuje się, że iloraz inteligencji,$IQ \sim \mathrm{N}(100, 15^2)$. Jakie jest pr., że losowo wybrany osobnik ma $IQ > 130$? Dla jakiej wartości x 99% populacji ma IQ przynajmniej x?

Rozwiązanie.

$$\begin{aligned} P(IQ > 130) &= 1 - P(IQ \leqslant 130) \\ &= 1 - \Phi\left(\frac{130 - 100}{15}\right) \\ &= 1 - \Phi(2) \approx 1 - 0,9772 \approx 0,023 \end{aligned}$$

Rozwiązanie.

$$\begin{aligned} 0,99 = P(IQ > x) &= 1 - P(IQ \leqslant x) \\ &= 1 - \Phi\left(\frac{x - 100}{15}\right) \end{aligned}$$

$$\Phi\left(\frac{x - 100}{15}\right) = 0,01 \Rightarrow \frac{x - 100}{15} = \Phi^{-1}(0,01)$$

$$x = 100 + 15\Phi^{-1}(0,01) \approx 100 + 15(-2,31) \approx 65,4$$

14 Twierdzenia graniczne

- Twierdzenia graniczne dotyczą własności funkcji zmiennych losowych

$$f(X_1, X_2, ..., X_n) \quad \text{dla} \quad n \to \infty$$

- W praktyce korzystamy z granicy do przybliżenia dokładnej wartości którą trudno jest policzyć

$$f(X_1, X_2, ..., X_k) \approx \lim_{n\to\infty} f(X_1, X_2, ..., X_n) \quad \text{dla} \quad 1 \ll k < \infty$$

- Częstą praktyką jest powtarzanie pomiaru wielokrotnie w tych samych warunkach i prezentowanie wyniku jako wartość średnią z uzyskanych pomiarów.

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

- Niech X_k będą niezależnymi zmiennymi losowi o średniej μ i wariancji σ^2

$$E[\bar{X}] = E[\frac{1}{n}\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \frac{E[X_i]}{n} = \mu$$

$$Var[\bar{X}] = Var[\frac{1}{n}\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \frac{Var[X_i]}{n^2} = \frac{\sigma^2}{n}$$

- $$Var[\bar{X}] = \frac{\sigma^2}{n} \to 0 \quad \text{dla} \quad n \to \infty$$

$$E[\bar{X}] = \mu \Rightarrow \bar{X} \to \mu \quad \text{dla} \quad n \to \infty$$

- dla liczb

$$a_n \to a \Rightarrow \epsilon > 0, \quad |a_n - a| < \epsilon \quad \text{dla } n > n_0$$

- dla zmiennych losowych nie można tak napisać

$$\bar{X} \to \mu \;\not\Rightarrow\; \epsilon > 0, \; |\bar{X} - \mu| < \epsilon \quad \text{dla } n > n_0$$

$$\bar{X} \to \mu \;\Rightarrow\; \epsilon > 0, \; P(|\bar{X} - \mu| > \epsilon) \to 0 \quad \text{dla } n \to \infty$$

Twierdzenie 14.1 (Prawo wielkich liczb)**.**
Niech $X_1, X_2, ...$ będą niezależnymi zmiennymi losowymi o wartości oczekiwanej μ a $\bar{X}$ będzie ich średnią. Wtedy dla dowolnego $\epsilon > 0$ zachodzi:

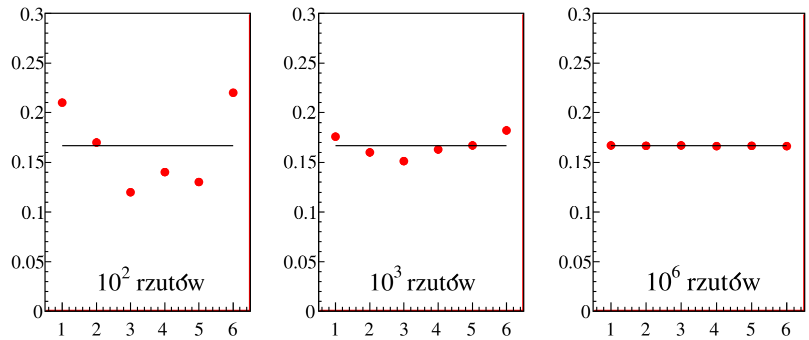
$$P(|\bar{X} - \mu| > \epsilon) \to 0 \quad \textit{dla } n \to \infty$$

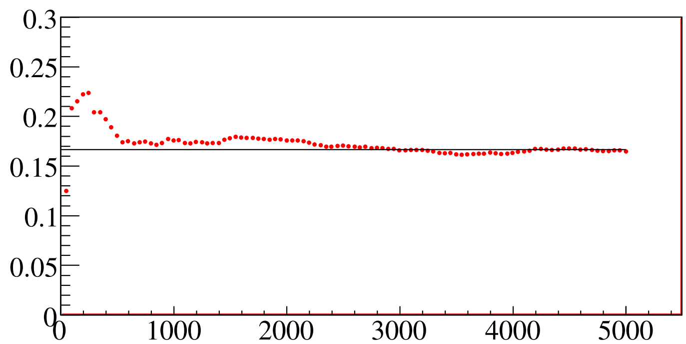
$$\bar{X} \xrightarrow{P} \mu \quad \textit{dla } n \to \infty$$

Zbieżność w sensie prawdopodobieństwa

W szczególności w eksperymecie gdzie zdarzenie A zachodzi z prawdopodobieństwem $P(A)$ powtarzanym niezależnie n razy zdarzenie A zaszło k razy to

$$\frac{k}{n} \to P(A) \quad \text{dla } n \to \infty$$





Symulacja rzutów kostką do gry

Prawo wielkich liczb nie określa jak dokładne jest przybliżenie

$$\bar{X} \approx \mu \quad \text{dla} \quad n \gg 1$$

Kolejne twierdzenia określa dokładność tego przybliżenia

Twierdzenie 14.2 (Centralne twierdzenie graniczne)**.**
Niech $X_1, X_2, \ldots$ będą niezależnymi zmiennymi losowymi o wartości oczekiwanej μ i wariancji $\sigma^2 < \infty$. Wtedy dla dowolnego z zachodzi:

$$P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z\right) \to \Phi(z) \;\; \textit{dla}\; n \to \infty$$

gdzie Φ jest dystrybuantą standardowego rozkładu normalnego.

$$P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z\right) = P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} < z\right)$$

Wnioski z centralnego twierdzenia granicznego:

$$P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z\right) \approx \Phi(z)$$

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$$

Zbieżność w sensie rozkładu.

$$\bar{X} \xrightarrow{D} N(\mu, \sigma^2/n)$$

$$\sum_{i=1}^{n} X_i \xrightarrow{D} N(n\mu, n\sigma^2)$$

W szczególności dla X o rozkładzie dwumianowym $bin(n, p)$

$$X \xrightarrow{D} N(\mu, \sigma^2) = N(np, np(1-p))$$