Preliminaria

Próba: niezależne zmienne losowe $X_1, X_2, \ldots, X_n$ o jednakowym rozkładzie P_θ z dystrybuantą F_θ i gęstością p_θ

Dystrybuanta empiryczna:

$$F_n(t) = \frac{\#\{1 \le j \le n : X_j \le t\}}{n}$$

Ważne fakty

Dla każdego $t \in \mathbb{R}$ zachodzi

$$E_\theta F_n(t) = F_\theta(t)$$

Dla każdego $t \in \mathbb{R}$ zachodzi

$$P_\theta\{\lim_{n\to\infty} F_n(t) = F_\theta(t)\} = 1$$

Dla każdego $t \in \mathbb{R}$ rozkład zmiennej losowej

$$\sqrt{n}\frac{F_n(t) - F_\theta(t)}{\sqrt{F_\theta(t)(1 - F_\theta(t))}}$$

dąży do rozkładu $N(0, 1)$ przy $n \to \infty$.

Twierdzenie 3.1 *(Podstawowe twierdzenie statystyki matematycznej; twierdzenie Gliwienki–Cantelliego)*
Jeżeli próba $X_1, X_2, \ldots, X_n$ pochodzi z rozkładu F_θ, to zmienna losowa

$$D_n = \sup_{-\infty<t<\infty} |F_n(t) - F_\theta(t)|$$

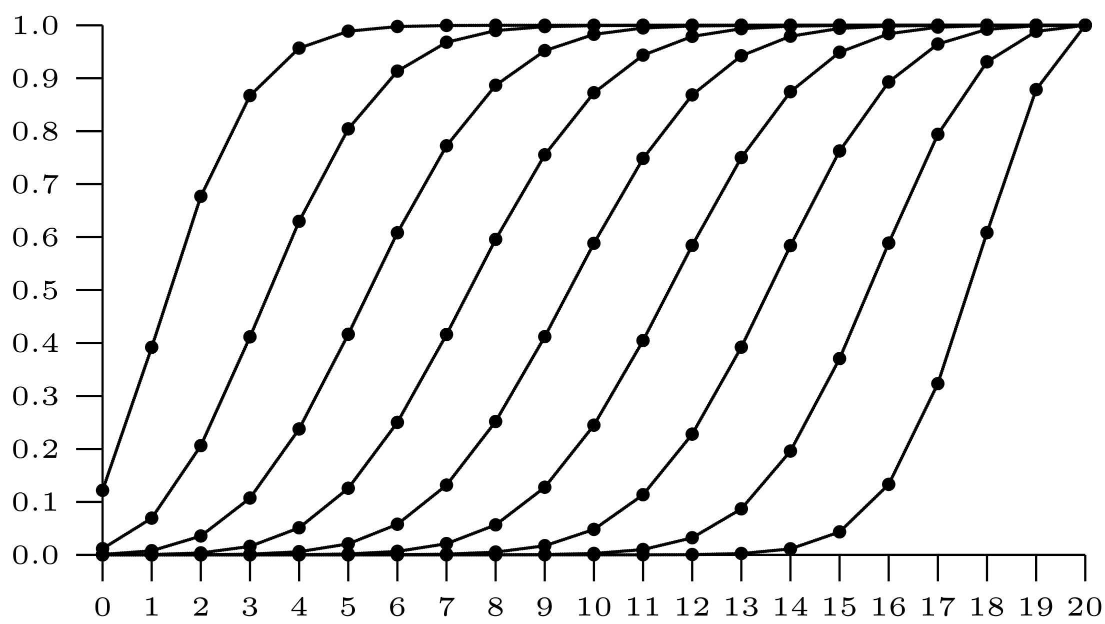
dąży do zera z prawdopodobieństwem 1. ■

Przykład 1. Model statystyczny

$$(\{0, 1, \ldots, 20\}, \{B\,(20, \theta)\,,\ \theta \in [0, 1]\})$$

„Dystrybuanty” dla

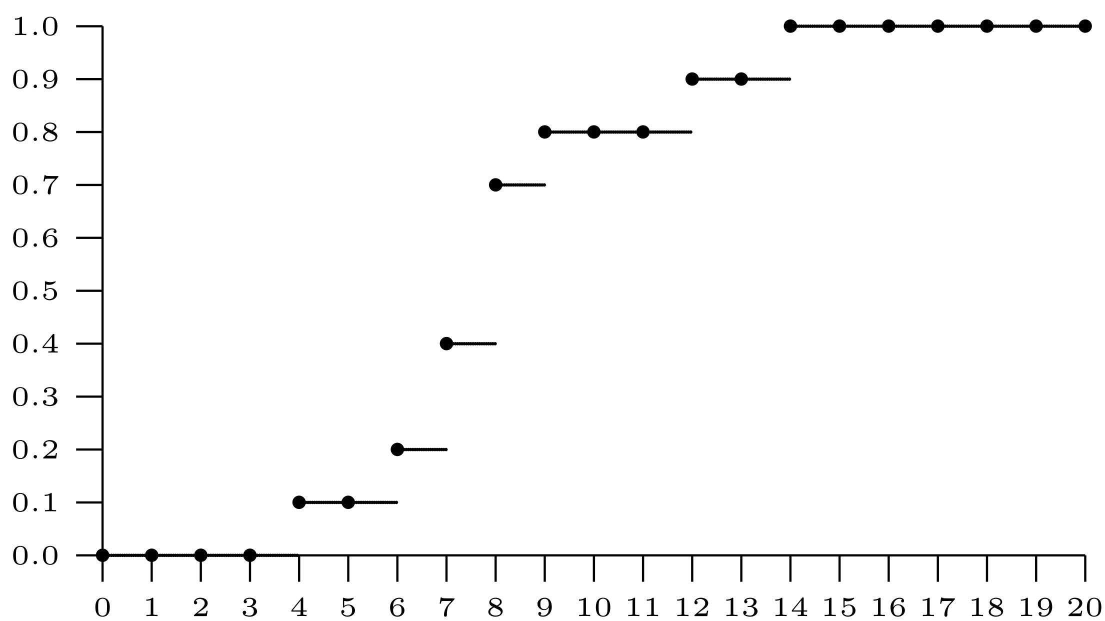
$$\theta = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$$



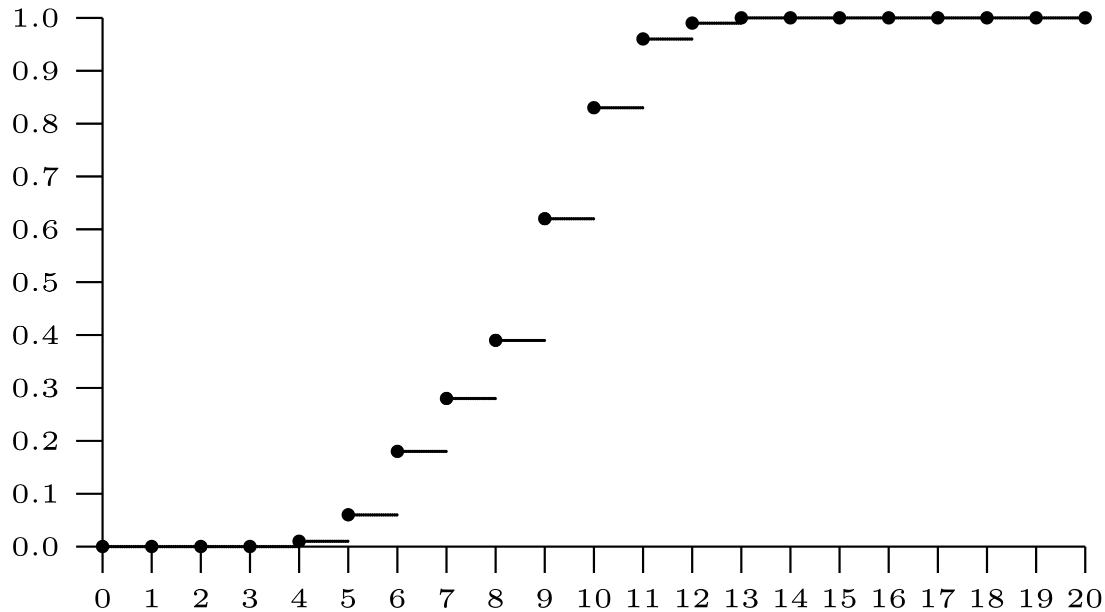
Próba: $14, 8, 12, 7, 8, 6, 8, 7, 4, 9$

$$F_{10}(t) = \frac{\#\{1 \le j \le n : X_j \le t\}}{10} = \begin{cases} 0, & t < 4, \\ 1/10, & t < 6, \\ 2/10, & t < 7, \\ 4/10, & t < 8, \\ 7/10, & t < 9, \\ 8/10, & t < 12, \\ 9/10, & t < 14, \\ 1, & t \ge 14, \end{cases}$$

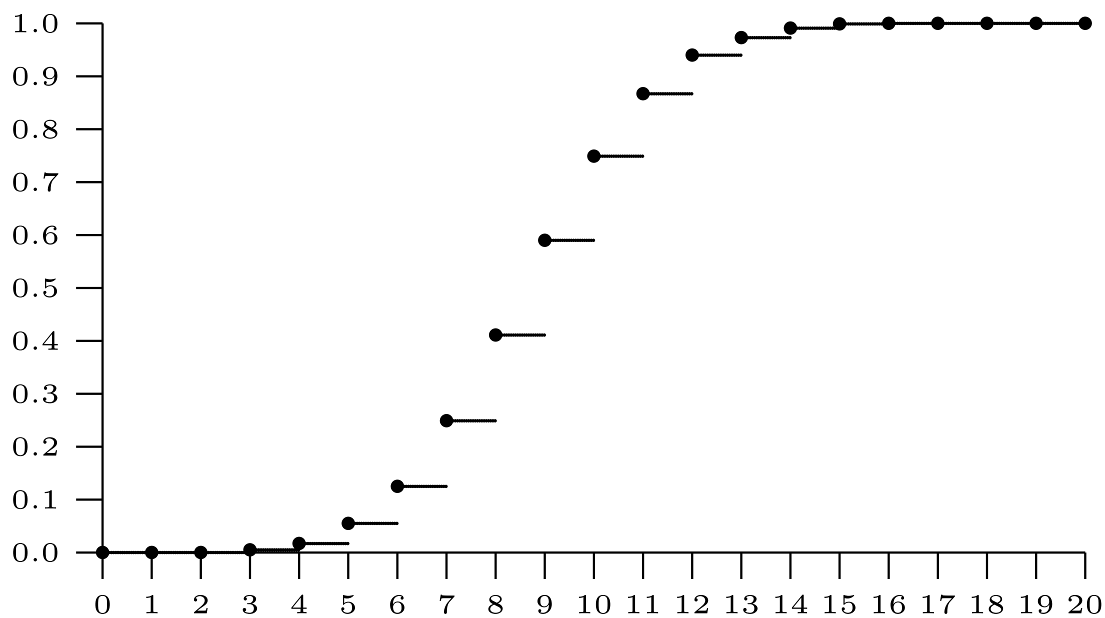
Dystrybuanta empiryczna dla próby o liczności 10



Dystrybuanta empiryczna dla próby o liczności 100



Dystrybuanta empiryczna dla próby o liczności 1000



. . . .