empirical distribution function

Canonical name	EmpiricalDistributionFunction
Date of creation	2013-03-22 14:33:27
Last modified on	2013-03-22 14:33:27
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	7
Author	CWoo (3771)
Entry type	Definition
Classification	msc 62G30

Let $X_1, \dots, X_n$ be random variables with realizations $x_i = X_i(\omega) \in \mathbb{R}$, $i = 1, \dots, n$. The *empirical distribution function* $F_n(x, \omega)$ based on $x_1, \dots, x_n$ is

$$F_n(x, \omega) = \frac{1}{n} \sum_{i=1}^{n} \chi_{A_i}(x, \omega),$$

where χ_{A_i} is the indicator function (or characteristic function) and $A_i = \{(x, \omega) \mid x_i \leq x\}$. Note that each indicator function is itself a random variable.

The empirical function can be alternatively and equivalently defined by using the order statistics $X_{(i)}$ of X_i as:

$$F_n(x, \omega) = \begin{cases} 0 & \text{if } x < x_{(1)}; \\ \frac{1}{n} & \text{if } x_{(1)} \leq x < x_{(2)},\ 1 \leq k < 2; \\ \frac{2}{n} & \text{if } x_{(2)} \leq x < x_{(3)},\ 2 \leq k < 3; \\ \vdots & \\ \frac{i}{n} & \text{if } x_{(i)} \leq x < x_{(i+1)},\ i \leq k < i+1; \\ \vdots & \\ 1 & \text{if } x \geq x_{(n)}; \end{cases}$$

where $x_{(i)}$ is the realization of the random variable $X_{(i)}$ with outcome ω.