WIKIPEDIA

Law of the unconscious statistician

In probability theory and statistics, the **law of the unconscious statistician (LOTUS)** is a theorem used to calculate the expected value of a function $g(X)$ of a random variable X when one knows the probability distribution of X but one does not know the distribution of $g(X)$. The form of the law can depend on the form in which one states the probability distribution of the random variable X. If it is a discrete distribution and one knows its probability mass function f_X (but not $f_{g(X)}$), then the expected value of $g(X)$ is

$$\operatorname{E}[g(X)] = \sum_x g(x) f_X(x),$$

where the sum is over all possible values x of X. If it is a continuous distribution and one knows its probability density function f_X (but not $f_{g(X)}$), then the expected value of $g(X)$ is

$$\operatorname{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) \, \mathrm{d}x$$

If one knows the cumulative probability distribution function F_X (but not $F_{g(X)}$), then the expected value of $g(X)$ is given by a Riemann–Stieltjes integral

$$\operatorname{E}[g(X)] = \int_{-\infty}^{\infty} g(x) \, \mathrm{d}F_X(x)$$

(again assuming X is real-valued).[1][2][3][4]

Contents

Etymology

This proposition is known as the law of the unconscious statistician because of a purported tendency to use the identity without realizing that it must be treated as the result of a rigorously proved theorem, not merely a definition.[4]