Probability and Statistics

Justin Pounders

January 24, 2016

Contents

Consider a real and continuous *random variable*, ξ. This variable will take a value somewhere on the interval $[a, b]$, and let the probability of the random variable occurring between x_1 and x_2 be given by $P\left[x_1 \leq \xi \leq x_2\right]$. We can define a *probability density function*, $f(x)$, such that $f(x)dx$ is the probability that the continuous random variable, ξ, will have a value between x and $x + dx$ in the limit of $dx \rightarrow 0$. In other words,

$$f(x)dx = P\left[x \leq \xi \leq x + dx\right] \quad . \tag{1}$$

The probability of the random variable ξ having a value somewhere between x_1 and x_2 is then given by

$$\int_{x_1}^{x_2} f(x)dx = P\left[x_1 \leq \xi \leq x_2\right] \quad . \tag{2}$$

Because the random variable ξ must take a value *somewhere* on the interval $[a, b]$, the density function $f(x)$ must be normalized to unity:

$$\int_a^b f(x)dx = 1 \quad . \tag{3}$$

This normalization guarantees (with probability one) that the random variable will have a value somewhere in $[a, b]$. Note that a and b are allowed to go to plus or minus infinity, respectively. The interval $[a, b]$ is called the *support* of $f(x)$.

The *mean* value of a probability density function is defined by

$$\langle x \rangle = \int_a^b x f(x)dx \quad . \tag{4}$$

In addition to the probability density function, the *cumulative probability distribution function* can be defined as the probability that the random variable ξ will take a value less than or equal to x: %

$$F(x) = P\left[\xi \leq x\right] \quad . \tag{5}$$

% The cumulative distribution can be defined in terms of the density function by %

$$F(x) = \int_a^x f(x')dx' \quad . \tag{6}$$

%