Discrete Random Variables

Model Description

A discrete random variable has a discrete sample space. Consider a random variable denoted by uppercase X, with outcomes, i.e., realizations, denoted by lowercase x. The probability of occurrence of each outcome of the discrete random variable is given by the probability mass function (PMF):

$$p_X(x) \equiv P(X = x) \tag{1}$$

The PMF has the following property:

$$\sum_{i=1}^{N} p_X(x_i) = 1 \tag{2}$$

where N is the number of possible outcomes. An alternative presentation of the probability distribution is the cumulative distribution function (CDF):

$$F_X(x) \equiv P(X \le x) = P(X \le x_i) = \sum_{j=1}^{i} p_X(x_j) \tag{3}$$

The CDF has the two properties F(-∞)=0 and F(∞)=1. Yet another representation of the probability distribution is the complementary CDF (CCDF):

$$G_X(x) = 1 - F_X(x) \tag{4}$$

which has the properties G(-∞)=1 and F(∞)=0.

Model Parameters

A random variable is completely defined by its probability distribution. However, "partial descriptors" are useful in lieu of having the complete distribution. The partial descriptors are equal to or related to the parameters of the probability distributions that are listed later in this document. The partial descriptors are also related to the statistical moments of the probability distribution. The first moment of the distribution is the mean of the random variable:

$$\mu_X = E[X] = \sum_{i=1}^{N} x_i \cdot p_X(x_i) \tag{5}$$

The second moment is called the mean square of the random variable:

$$E[X^2] = \sum_{i=1}^{N} x_i^2 \cdot p_X(x_i) \tag{6}$$

Conversely, central moments are taken about the mean of the random variable. As a result, the first central moment is zero. The second central moment is the variance of the random variable, which is the square of the standard deviation:

$$\sigma_X^2 = \text{Var}[X] = E[(x-\mu_X)^2] = \sum_{i=1}^{N} (x_i - \mu_X)^2 \cdot p_X(x_i) \tag{7}$$

Several concepts for discrete random variables, such as coefficient of variation and coefficient of skewness are the same for continuous random variables. Therefore, further details are provided in the document on continuous random variables.

Theorem of Total Probability Applied to a Random Variable

Recall that the rule of total probability is applicable only when the conditioning is on mutually exclusive and collectively exhaustive events. The rule of total probability to obtain the probability of an event, having probability values conditioned upon the outcomes of a discrete random variable:

$$\text{P}(A) = \sum_{i=1}^{N} \text{P}(A \mid x_i) \cdot p(x_i) \tag{8}$$

The rule of total probability to obtain a probability distribution, having the distribution conditioned on the outcomes of another random variable or some other discrete events:

$$p_X(x) = \sum_{i=1}^{N} p_X(x \mid y_i) \cdot p_Y(y_i) \tag{9}$$

Common Distribution Models

Bernoulli

Consider a discrete random variable, X, with two possible outcomes: failure and success, i.e., 0 and 1, respectively. The probability of success is denoted p. Consequently, the probability of failure is 1-p and the Bernoulli PMF is thus defined:

$$p(x) = \begin{cases} 1-p & \text{for} \quad x=0 \\ p & \text{for} \quad x=1 \end{cases} \tag{10}$$

Using earlier formulas, the mean is p and the variance is $p(1-p)$. To specify that a random variable, X, has the Bernoulli distribution, one writes: X~Bernoulli(p).

Binomial

Consider a sequence of mutually independent Bernoulli trials with constant success probability, p. Let the random variable X denote the number of successes in n trials. The PMF for this random variable is the binomial distribution

$$p(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} \tag{11}$$

The mean of X is np and its variance is $np(1-p)$. To specify that a random variable, X, has the binomial distribution, one writes: X~Binomial(p,n).

Geometric

The number of trials, S, until success, and the number of trials between successes, in a Bernoulli sequence is given by the geometric distribution:

$$p(s) = p \cdot (1-p)^{s-1} \tag{12}$$

The mean recurrence time, sometimes called return period, is $1/p$. The variance is $(1-p)/p^2$. To specify that a random variable, S, has the geometric distribution, one writes: S~Geometric(p).

Negative Binomial

The number of Bernoulli trials, W, until k occurrences of success is

$$W = S_1 + S_2 + \cdots + S_k \tag{13}$$

where S_i is the number of trials between success number i-1 and success number i. The distribution of S is geometric. Combined with the fact that Eq. (13) is a sum of random variables, the mean and variance of W is

$$\mu_W = k \cdot \mu_S = k \cdot \frac{1}{p} \tag{14}$$

$$\sigma_W^2 = k \cdot \sigma_S^2 = k \cdot \frac{1-p}{p^2} \tag{15}$$

The distribution type for W is the negative binomial distribution:

$$p(w) = \binom{w-1}{k-1} \cdot p^k \cdot (1-p)^{w-k} \tag{16}$$

To specify that a random variable, W, has the Bernoulli distribution, one writes: W~NegativeBinomial(p,k).

Poisson

In situations where the number of Bernoulli trials is infinite, such as when every time instant is considered a trial, the Poisson distribution gives the number of successes, x:

$$p(x) = \frac{(\lambda \cdot T)^x}{x!} e^{-\lambda \cdot T} \tag{17}$$

where λ is the rate of occurrence of success per unit time and T is the time period under consideration. The mean number of occurrences is λT, which is also equal to the variance. To specify that a random variable, X, has the Poisson distribution, one writes: X~Poisson(λ,T).