http://www.iohk.com/UserPages/erictang/probability.html   Go

36 captures
21 Sep 03 - 22 Jul 15

MAY **JUL** AUG
◀ **22** ▶
2014 2015 2016

Close

Help

PROBABILITY DISTRIBUTIONS

[HOME]

[Quantum Computation]

Cryptography

Mathematical Study of Severe Acute Respiratory Syndrome (SARS) in HK NEW

Some Useful Probability Distributions

Consider the probability of success in a single trial is p, for 0 < p < 1, and the unsuccess is (1-p).

The density function, the mean, and the variance are represented by $f(x), \mu, \sigma^2$ correspondingly. And the random-number generator generates U=uniform(0,1) for the samples.

Discrete Uniform Distribution

The outcome x of the discrete uniform distribution is the finding of the single unsuccess event. Given a range of possible outcomes [a, b]. All outcomes are equally likely.

Continuous Uniform Distribution

The continuous equivalent of the discrete uniform distribution. It is used when no information about the distribution is known, other than its bounds and the fact it is continuous.

Bernoulli Distribution

The outcomes x of a Bernoulli distribution are 0 and 1. It is the chance of success in a sample size of 1.

Binomial Distribution

The outcome x corresponds to the number of successes out of n trials. The binomial distribution is the sum of n independent Bernoulli distributions.

Geometric Distribution

The outcome x of geometric distribution is the number of trials that are successful up to and including the first failure. x = 1,2, 3, ... The simplest method of generating geometric smaples uses the inverse-transform method.

Poisson Distribution

The Poisson distribution is the limiting form of the binomial distribution as n becomes very large with np fixed. It is used to model the number of events that occur within a given interval. It is closely related to the exponential distribution in that, if the time between events follows an exponential distribution, the number of events that occur within a given interval follows a Poisson

distribution.

Exponential Distribution

The exponential distribution is the continuous approximation of the geometric distribution. It is used to model the time between successive events, or the time required to service an event. The exponential distribution has the property of have no memory, i.e., knowing the time that the last event occured is in no way helpful in predicting when the next event may occur. It is very common to assume that the interarrival times of events in a computer system follow an exponential distribution. If the times between events are independent with an exponential distribution with mean $\mu = \beta$, then the number of events within time T is Poisson distributed with mean equal to

$$\mu = \lambda = T / \beta.$$

Gaussian (Normal) Distribution

The Gaussian distribution is also called normal distribution. It is the bell-shaped curve symmetric around the mean with the width determined by the variance. The standard normal distribution is a Gaussian distribution with mean = 0 and variance = 1.

Erlang Distribution

The Erlang distribution can be thought of as a generalization of the exponential distribution. There are two parameters to describe the density function, the scale parameter, a >0 and m, is a positive integer as the shape parameter.

Pareto Distribution

The Pareto distribution is the generalized power function. It has a single shape parameter, a

The Pareto distribution is the generalized power function. It has a single shape parameterm a.
Given a set of observations {x}, it can shown that

Summary

	Density Function	Mean & Variance
Discrete Uniform Distribution	$f(x) = \dfrac{1}{b-a+1}$ where x=a, a+1, ... , b. Random-number generation - Use the inverse-transformation method, x =a + [(b-a+1)U].	$\mu = \dfrac{(b+a)}{2}$ $\sigma^2 = \dfrac{(b-a+1)^2}{12} - 1$
Continuous Uniform Distribution	$f(x) = \dfrac{1}{b-a}, a \leq x \leq b$ Random-number generation - Use the inverse-transformation method, x = a + (b-a)U	$\mu = \dfrac{(b+a)}{2}$ $\sigma^2 = \dfrac{(b-a)^2}{12}$
Bernoulli Distribution	$f(x) = p^x(1-p)^{1-x}, x = 0,1.$ Random-number generation - x = 1, if U =< p; otherwise x = 0.	$\mu = p, \sigma^2 = p(1-p)$
Binomial Distribution	$f(x) = \binom{n}{x}p^x(1-p)^{n-x}, x = 0,1,....,n; 0 < p < 1.$ Random-number generation - Sum of n independent Bernoulli distributions. Count the number of U samples out of n that are less than or equal to p.	$\mu = np, \sigma^2 = np(1-p)$
	$f(x) = p^{x-1}(1-p), x = 1,2,3$	$\dfrac{1}{\quad}$ p

Geometric Distribution	$f(x) = p \quad (1-p), x = 1,2,3\ldots$ Random-number generation - Use the inverse-transformation method, x = [lnU / ln(1-p)]	$\mu = \dfrac{}{(1-p)}, \sigma^2 = \dfrac{}{(1-p)^2}$

Poisson Distribution	$f(x) = \dfrac{\lambda^x e^{-\lambda}}{x!}, x = 0,1,2,\ldots$
	Random-number generation - One sample of a Poisson distribution is the number of samples from an exponential distribution with mean equal to 1 that must be added together to exceed the value mean of the Poisson distribution. Algorithm:
	sum = 0, i = -1
	while (sum <= lambda)
	y = U, z = -ln (y)
	sum = sum + z
	i = i + 1
	endwhile
	return (i)
		$\mu = \lambda, \sigma^2 = \lambda$

Exponential Distribution	$f(x) = \dfrac{1}{\beta} e^{-x/\beta}, x \geq 0.$ Random-number generation - Use the inverse-transformation method, x = - beta lnU.	$\mu = \beta, \sigma^2 = \beta^2$
Gaussian (Normal) Distribution	$f(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}, -\infty \leq x \leq \infty.$ Random-number generation - a generalization of the inverse - transformation method to two dimensions, known as the Box-Muller method, generates two samples at a time from a Gauaaian distribution. For U1 and U2, we have $x_1 = \mu + \sigma\cos(2\pi U_1)\sqrt{-2\ln U_2}$	$\mu, \sigma^2 > 0$

	$x_2 = \mu + \sigma \sin(2\pi U_1)\sqrt{-2\ln U_2}$	
Erlang Distribution	$f(x) = \dfrac{x^{m-1}e^{-x/a}}{(m-1)!\,a^m}, 0 \le x \le \infty$ where m is a positive integer. Random-number generation - Use the convolution technique, generating m values, U1, U2, ,,, Um. x = -a ln(U1 x U2 x ... x Um).	$\mu = am, \sigma^2 = a^2 m$
Pareto Distribution	$f(x) = ax^{-(a+1)}; a = \dfrac{n}{\sum_{i=1}^{n}\ln x_i}$ Random-number generation - Use the inverse-transformation method, x = 1 / U^(1/a)	$\mu = \dfrac{a}{(a-1)}, a > 1; \sigma^2 = \dfrac{a}{(a-1)^2(a-2)}, a > 2$

Reference:

1. David J. Lilja, Measuring Computer Performance - A practitioner's guide (2000), Cambridge University Press.