Statistics for Computing

MA4413 Lecture 4A

Kevin O'Brien

Kevin.obrien@ul.ie

Dept. of Mathematics & Statistics,
University *of* Limerick

Autumn Semester 2011

Today's Class: Continuous Distributions

- The Uniform Distribution
- The Exponential Distribution
- The Normal Distribution

In the last class, we had a look at the continuous uniform distribution. It is very useful in constructing simulations. Briefly we will look at some relevant `R` function. The distribution has two parameters: i.e `min` and `max`. (Here chosen as 5 and 10 respectively)

```
># Generate Four Random Number
> runif(4, min=5,max=10)
[1] 9.709372 7.884805 5.571331 5.017549
>
># Compute Density
> dunif(4:11,min=5,max=10)
[1] 0.0 0.2 0.2 0.2 0.2 0.2 0.2 0.0
>
> #Compute distribution of
> punif(4:11,min=5,max=10)
[1] 0.0 0.0 0.2 0.4 0.6 0.8 1.0 1.0
```

The Binomial Probability Distribution

- The number of independent trials is denoted n.
- The probability of a ‘success’ is p
- The expected number of ‘successes’ from n trials is $E(X) = np$

Continuous Random Variables

- Probability Density Function
- Cumulative Density Function

If X is a continuous random variable then we can say that the probability of obtaining a **precise** value x is infinitely small, i.e. close to zero.

$$P(X = x) \approx 0$$

Consequently, for continuous random variables (only), $P(X \leq x)$ and $P(X < x)$ can be used interchangeably.

$$P(X \leq x) \approx P(X < x)$$

The Memoryless property

The most interesting property of the exponential distribution is the ***memoryless*** property. By this , we mean that if the lifetime of a component is exponentially distributed, then an item which has been in use for some time is a good as a brand new item with regards to the likelihood of failure.
The exponential distribution is the only distribution that has this property.

Continuous Random Variables

- Probability Density Function
- Cumulative Density Function

If X is a continuous random variable then we can say that the probability of obtaining a **precise** value x is infinitely small, i.e. close to zero.

$$P(X = x) \approx 0$$

Consequently, for continuous random variables (only), $P(X \leq x)$ and $P(X < x)$ can be used interchangeably.

$$P(X \leq x) \approx P(X < x)$$

Continuous Uniform Distribution

A random variable X is called a continuous uniform random variable over the interval (a,b) if it's probability density function is given by

$$f_X(x) = \frac{1}{b-a} \qquad \text{when } a \leq x \leq b$$

The corresponding cumulative density function is

$$F_x(x) = \frac{x-a}{b-a} \qquad \text{when } a \leq x \leq b$$

The mean of the continuous uniform distribution is

$$E(X) = \frac{a+b}{2}$$

$$V(X) = \frac{(b-a)^2}{12}$$

The Memoryless property

The most interesting property of the exponential distribution is the ***memoryless*** property. By this , we mean that if the lifetime of a component is exponentially distributed, then an item which has been in use for some time is a good as a brand new item with regards to the likelihood of failure.
The exponential distribution is the only distribution that has this property.

Random Variables

A pair of dice is thrown. Let X denote the minimum of the two numbers which occur. Find the distributions and expected value of X.

A fair coin is tossed four times. Let X denote the longest string of heads. Find the distribution and expectation of X.

Random Variables

A fair coin is tossed until a head or five tails occurs. Find the expected number E of tosses of the coin.

A coin is weighted so that $P(H) = 0.75$ and $P(T) = 0.25$

The coin is tossed three times. Let X denote the number of heads that appear.

- (a) Find the distribution f of X.
- (b) Find the expectation $E(X)$.

- Now consider an experiment with only two outcomes. Independent repeated trials of such an experiment are called Bernoulli trials, named after the Swiss mathematician Jacob Bernoulli (1654-1705).
- The term ***independent trials*** means that the outcome of any trial does not depend on the previous outcomes (such as tossing a coin).
- We will call one of the outcomes the "success" and the other outcome the "failure".

- Let p denote the probability of success in a Bernoulli trial, and so $q = 1 - p$ is the probability of failure. A binomial experiment consists of a fixed number of Bernoulli trials.
- A binomial experiment with n trials and probability p of success will be denoted by

$$B(n,p)$$

Probability Mass Function

- a probability mass function (pmf) is a function that gives the probability that a discrete random variable is exactly equal to some value.
- The probability mass function is often the primary means of defining a discrete probability distribution

Thirty-eight students took the test. The X-axis shows various intervals of scores (the interval labeled 35 includes any score from 32.5 to 37.5). The Y-axis shows the number of students scoring in the interval or below the interval.

cumulative frequency distributionA can show either the actual frequencies at or below each interval (as shown here) or the percentage of the scores at or below each interval. The plot can be a histogram as shown here or a polygon.

Notation for Poisson Distribution

A discrete random variable X is said to follow a Poisson distribution with parameter m, written $X \sim \mathrm{Po}(m)$, if it has probability distribution

$$P(X=k) = e^{-m}\frac{m^k}{k!}$$

where

- $k = 0, 1, 2, \ldots$
- $m > 0.$