Midterm Exam

- 1300hr on Tuesday (Week 7)
- Worth 15% of overall mark for module.
- The test is comprised of 15 short questions, or components of compound questions.
- Topics:
  - Basic Probability
  - The Binomial Distribution
  - The Poisson Distribution
  - The Normal Distribution
- Revision lecture on Monday of Week 7.

Today's Class

- Introduction to Continuous Random Variables
- Probability Density Functions
- Density Curves
- Cumulative Distribution Functions
- The Continuous Uniform Distribution
- The Exponential Distribution
- Introduction to the Normal Distribution
- (Tomorrow: More on the Normal Distribution)

Continuous Random Variable

A continuous random variable is one which takes an infinite number of possible values. Continuous random variables are usually measurements. Examples include height, weight, the amount of sugar in an orange, the time required to run a mile.

In this module, we will look at random variables from the following distributions

- Continuous uniform probability distribution
- Exponential probabilty distribution
- Normal probability distribution

Before this, we will look at some important concepts.

Functions and Definite integrals

Integration is not part of the syllabus, and it is assumed that students have are not familiar with how to compute definite integrals.

However, it is useful to know what the purpose of definite integrals are, because we will be using the results derived from definite integrals.
It is assumed that students are familiar with functions.

Functions

IMAGE : 5A Functions

Definite Integral

IMAGES : 5A Definite Integrals

Definite Integral

Definite integrals are used to compute the ”area under curves”.
The area under the curve between X=1 and X=2 is depicted in grey. Using definite integrals

Probability Density Function

The probability density function of a continuous random variable is a function which can be integrated to obtain the probability that the random variable takes a value in a given interval.

In probability theory, a probability density function (pdf), or density of a continuous random variable is a function that describes the relative likelihood for this random variable to occur at a given point.

The probability for the random variable to fall within a particular region is given by the integral of this variables density over the region. The probability density function is non-negative everywhere, and its integral over the entire space is equal to one.

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

Density Curves

Any curve that is always on or above the horizontal axis and has total are underneath equal to one is a density curve.

- Area under the curve in a range of values indicates the proportion of values in that range.
- Come in a variety of shapes, but the “normal family of familiar bell-shaped densities is commonly used.
- Remember the density is only an approximation, but it simplies analysis and is generally accurate enough for practical use.

The Cumulative Distribution Function

The cumulative distribution function (CDF), or just distribution function, describes the probability that a real-valued random variable X with a given probability distribution will be found at a value less than or equal to x. Intuitively, it is the “area so far” function of the probability distribution.

$$F_X(x) = P(X \leq x)$$

Exact Probability

Remarks: This is for continuous distributions only.

- The probability that a continuous random variable will take an exact value is infinitely small. We will usually treat it as if it was zero.
- When we write probabilities for continuous random variables in mathematical notation, we often retain the equality component (i.e. the "...or equal to..").
  For example, we would write expressions $P(X \leq 2)$ or $P(X \geq 5)$.
- Because the probability of an exact value is almost zero, these two expression are equivalent to $P(X < 2)$ or $P(X > 5)$.
- The complement of $P(X \geq k)$ can be written as $P(X \leq k)$.

The Continuous Uniform Distributions

IMAGE : 5A uniform

Parameters

The continuous uniform distribution is characterised by the following parameters

- The lower limit a
- The upper limit b

It is not possible to have an outcome that is lower than a or larger than b.

$$P(X < a) = P(X > b) = 0$$

- The only possible outcomes are between a and b. Suppose $a = 3$ and $b = 6$.
- The following values are possible outcomes: 3.14, 3.78, 4.66, 5.8, 5.9999.
- The probability of being exactly equal to 3 or 6 can be assumed to be zero.
- The following outcomes are not possible, either because they are too high or too low. 1.67, 2, 67, 7.14, 8.78.

Example

- Suppose there is a platform in a subway station in a large large city.
- Subway trains arrive **every three minutes** at this platform.
- What is the shortest possible time a passenger would have to wait for a train?
- What is the longest possible time a passenger will have to wait?

Example

- What is the shortest possible time a passenger would have to wait for a train?
- If the passenger arrives just before the doors close, then the waiting time is zero.

$$a = 0 \text{ minutes}$$

Example

- What is the longest possible time a passenger will have to wait?
- If the passenger arrives just after the doors close, and missing the train, then he or she will have to wait the full three minutes for the next one.

$$b = 3 \text{ minutes } = 180 \text{ seconds}$$

The Expected Value

We are told that, for waiting times, the lower limit a is 0, and the upper limit b is 3 minutes.

The expected waiting time $\mathrm{E}[X]$ is computed as follows

$$\mathrm{E}[X] = \frac{b+a}{2} = \frac{3+0}{2} = 1.5 \text{ minutes}$$

Interval Probability

- We wish to compute the probability of an outcome being within a range of values.
- We shall call this lower bound of this range L and the upper bound U.
- Necessarily L and U must be possible outcomes.
- The probability of X being between L and U is denoted $P(L \leq X \leq U)$.

$$P(L \leq X \leq U) = \frac{U - L}{b - a}$$

The probability density function is given as

$$f(x) = \frac{1}{b-a} \text{ for } a \leq x \leq b$$

For any value “c” between the minimum value a and the maximum value b

$$P(X \geq c) = \frac{b-c}{b-a}$$

here b is the upper bound while c is the lower bound

$$P(X \leq c) = \frac{c-a}{b-a}$$

here c is the upper bound while a is the lower bound

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.

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

- L :lower bound of an interval
- U: upper bound of an interval

Probability of an outcome being between lower bound L and upper bound U

$$P(L \leq X \leq U) = \frac{U - L}{b - a}$$

Reminder "$\leq$" is less than or equal to.
"$\geq$" is greater than or equal to.
$L \leq X \leq U$ xan be verbalized as X between L and U. simply states that X is between L and U inclusively. ("inclusively" mean that X could be exactly L or U also, although the probability of this is extremely low)

Continuous Uniform Distribution

- The Uniform distributions model (some) continuous random variables and (some) discrete random variables.
- The values of a uniform random variable are uniformly distributed over an interval.
- For example, if buses arrive at a given bus stop every 15 minutes, and you arrive at the bus stop at a random time, the time you wait for the next bus to arrive could be described by a uniform distribution over the interval from 0 to 15.

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

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

Continuous Uniform Distribution

The mean of the continuous uniform distribution is

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

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

Uniform Distribution: Variance

The variance of the continuous uniform distribution, denoted $\text{Var}[X]$, is computed using the following formula

$$\text{Var}[X] = \frac{(b-a)^2}{12}$$

For our previous example this is

$$\text{Var}[X] = \frac{(3-0)^2}{12} = \frac{3^2}{12} = \frac{9}{12} = 0.75$$

The Exponential Distribution

A continuous random variable having p.d.f. f(x), where: $f(x) = \lambda x e^{-\lambda x}$ is said to have an exponential distribution, with parameter λ. The cumulative distribution is given by: $F(x) = 1e^{\lambda x}$
Expectation and Variance $E(X) = 1/\lambda$ $V(X) = 1/\lambda^2$

Example

Suppose that the service time for a customer at a fast-food outlet has an exponential distribution with mean 3 minutes. What is the probability that a customer waits more than 4 minutes?

$$P(X \leq 4) = 1 - e^{-4/3}$$

$$P(X \leq 4) = e^{-4/3} = 0.2636$$

Exponential Distribution Lifetimes

The average lifespan of a laptop is 5 years. You may assume that the lifespan of computers follows an exponential probability distribution.

- (3 marks) What is the probability that the lifespan of the laptop will be at least 6 years?
- (3 marks) What is the probability that the lifespan of the laptop will not exceed 4 years?
- (3 marks) What is the probability of the lifespan being between 5 years and 6 years?

Suppose the lifetime of a PC is exponentially distributed with mean $\mu = 5$ We should be told the average lifetime μ.

$$P(X \geq x_o) = e^{\frac{-x_o}{\mu}}$$

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.