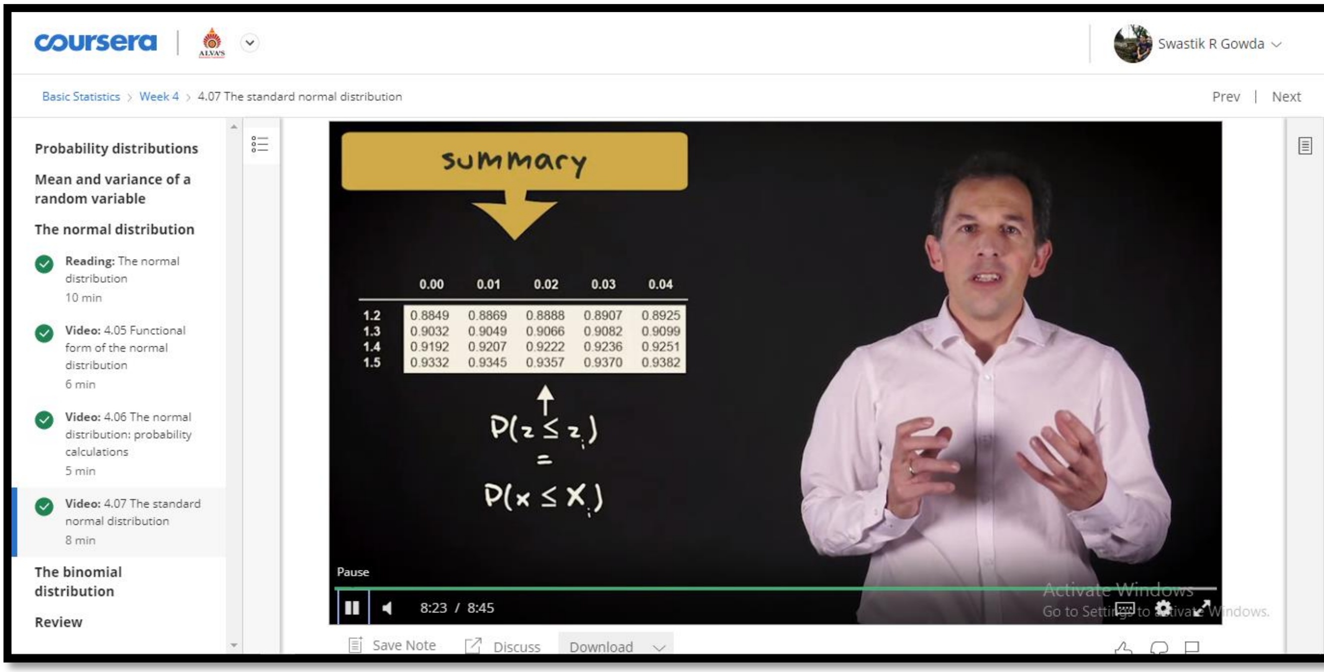
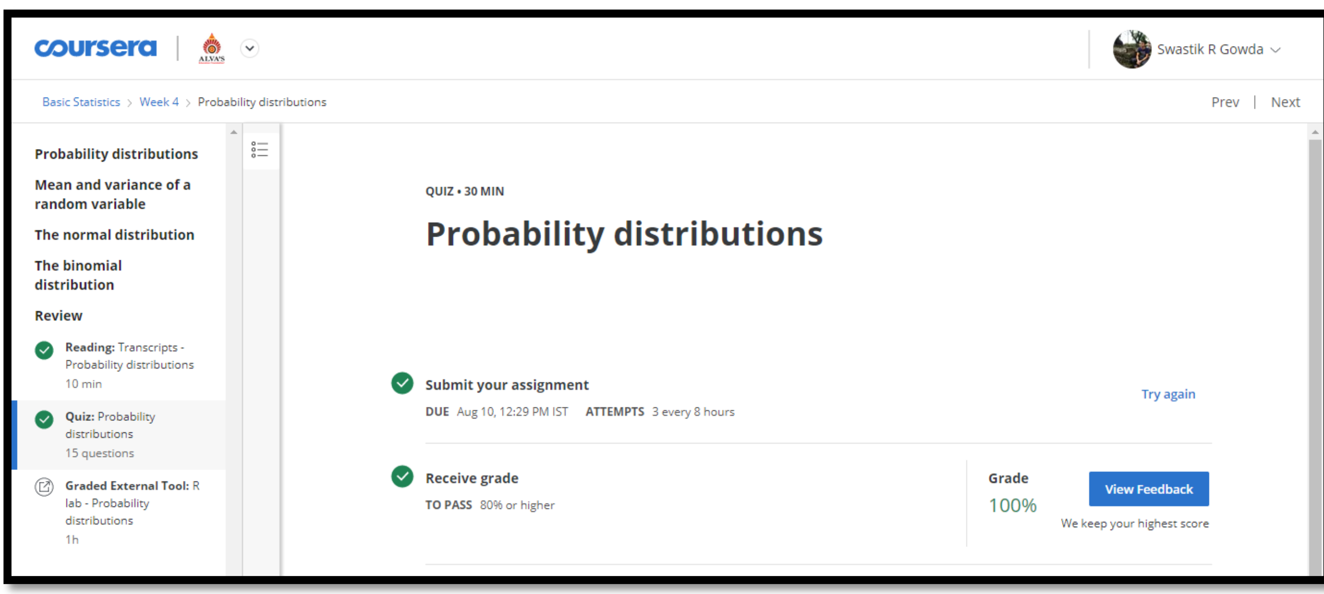
DAILY ASSESSMENT

Date:	**27-July-2020**	**Name:**	**Swastik R Gowda**
Course:	**Coursera - Basic Statistics**	**USN:**	**4AL17EC091**
Topic:	**Week 4**	**Semester & Section:**	**6^{th} Sem 'B' Sec**
Github Repository:	**swastik-gowda**		

FORENOON SESSION DETAILS

Image of session





Report – Report can be typed or hand written for up to two pages.

Probability distributions:

- How a probability distribution specifies the probabilities for each of the values that a random variable may take.
- It also illustrates how a probability distribution can take the form of a table, graph or equation.
- Finally, it explains how for a discrete random variable the probability distribution is called a probability mass function, giving probabilities, while for a continuous random variable the probability distribution is called a probability density function, giving probabilities per unit of the random variable. To obtain probabilities for a continuous random variable the sum (or integral) of all probabilities over an interval have to be considered.
- How a cumulative probability distribution can be obtained from a probability distribution by summing the probabilities in the latter from the smallest up to the largest value of the random variable.
- Also cumulative probability distributions can exist in the form of a table, graph or an equation. Interestingly, the difference between discrete and continuous variables disappears for cumulative probability distributions: for both variables the cumulative distribution gives cumulative probabilities: the probability of an event lower than or equal to the specified value of the random variable. The second video ends by illustrating how the cumulative probability distribution is useful to find a cumulative probability relating to a specified value of the random variable, but also the reverse: to find the threshold value of the random variable at a given probability level.

Random variables and probability distributions:

- When making observations on individuals or objects, you can observe several attributes per individual. These are called variables. Now imagine you've collected a dataset and you decide to repeat your study.
- You could perhaps find the same individuals or objects and measure the variables again. Or otherwise you could take a sample and measure similar individuals again. In any case you will find values for your variables that are different.
- Even if you were to measure, for example, a single person's length several times, the result would most likely deviate a few millimeters up to a centimeter, depending on time of day, the accuracy of your measuring tape, etc.

Cumulative probability distributions:

- Consider, for example, this probability density function. The corresponding cumulative distribution is given here.
- An interesting aspect of this step is that the y variable changes from a probability density to a probability, because in a cumulative distribution, it is the area from the smallest value of x, up to the value of interest in the probability, density function that is put on the y-axis.
- As you see, cumulative probability functions have continuously increasing values starting at zero and incrementing to a maximum of one.

- The sum of the probabilities for all the values that the random variable can take is one.
- The cumulative distribution, especially its graphical form, is very convenient because it can answer two questions. You can select a certain value of the random variable at the x-axis.

Mean and variance of a random variable:

- Once a probability distribution is defined it is possible to calculate summary statistics like a mean and standard deviation for a random variable, even if you did not have actual observations for that variable. These two videos show how these statistics can be calculated for discrete random variables.
- Next, it is shown what happens to the mean, variance and standard deviation when two random variables are added or subtracted and when a random variable is manipulated by adding or multiplying with a constant.
- Such operations occur frequently when dealing with real data. Examples of adding random variables are cases where the output of one model (e.g. a weather or an econometric model) is input for another (e.g. for hydrologic or macro-economic projections) or where different observations are combined to calculate a variable of interest (e.g. weight and height in a body-mass index).
- An examples of addition or multiplication with a constant could be applying a unit conversion to an observed variable (e.g. when converting a temperature in Celsius to Fahrenheit).

The Normal Distribution:

- The functional form of the normal distribution and the role of the two parameters in determining location (the mean) and the spread (the standard deviation) of the distribution. In addition it shows how the normal distribution not only exists as a statistical equation but is also a function that describes the outcome of many processes where some form of diffusion is important.
- The probability that a normally distributed random variable falls within a given range can be expressed in units of standard deviations (sigma) around the mean: the probability values of 0.68, 0.95 and almost 1 correspond with intervals of 1, 2 and 3 sigma around the mean respectively.
- By applying a z-transformation to a normally distributed variable. Probability statements can then be made for any value of the random variable (not just 1, 2 or 3 sigma around the mean) on the basis of the resulting z-scores, by using a table that lists cumulative probabilities with the corresponding z-values.
- It is shown how a cumulative probability can be found and interpreted for a given value of the random variable, and (reversely) how a threshold value of the random variable is found and interpreted for a given cumulative probability.
- The normal distribution and are able to use it effectively for probability calculations on normally distributed random variables.

The Binomial Distribution:

- The binomial probability distribution is introduced. It starts by explaining the type of elementary random variable to which the distribution relates: a variable with only two mutually exclusive outcomes and a fixed probability p to obtain one of the two outcomes (a Bernoulli trial) how the distribution gives the probability of observing x successes in n Bernoulli trials.
- The assumptions of independence among each outcome and a constant probability of success are specified, and the equation that describes this distribution is given with its two parameters p and n.
- The use of this equation is demonstrated with an example. Finally the equations for the mean and standard deviation of a binomial probability distribution are given and it is shown how the standard deviation of this distribution varies for different values of the parameters.