IB120/201 - Lab 5

Probability Distributions

Due Date: February 28, 2020

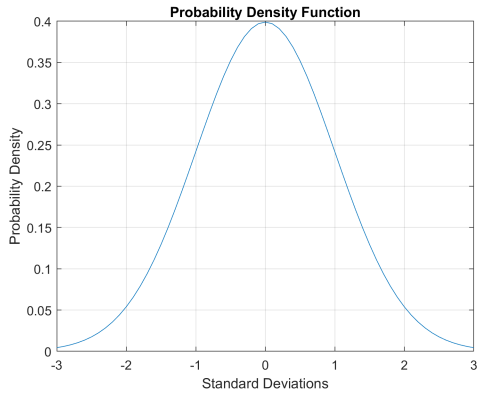
University of California, Berkeley *GSI: Naveed Ziari*

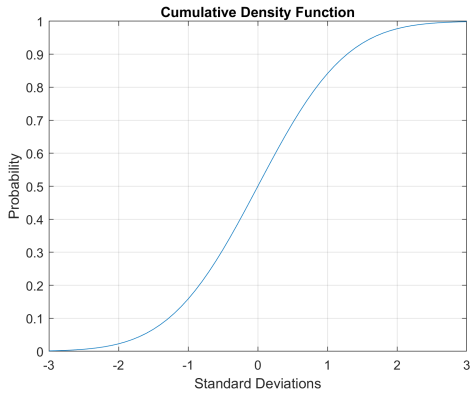
In this lab, we will familiarize ourselves with the different probability distributions:

Background

A probability distribution is a function which gives the probability of occurrence of an event confined within one experiment. The most famous and widely used probability distribution is the normal (or Gaussian) distribution. The central limit theorem states that the sum of independent random variables will resemble a normal distribution as the number of samples increases. For example, the distribution of students' heights at UC Berkeley resembles a bell curve, which is the normal distribution.

When the random variable is continuous (not discretized), we call it a *probability density function*. A *cumulative distribution function* shows the probability that a random variable will take a value less than or equal to x, which in the graphs below have units of standard deviations.





Questions

1. From the example in class, what does the y-axis of the histogram respresent (probability of ___) and how does it relate to the binomial distribution?
2. From the example in class, please make a plot of the probabilities (exactly how you define that is up to you) as a function of θ. Please set n (the number of draws) to be sufficiently large (e.g. 1000) and plot the results. What does this plot represent?
3. If $f(x)$ is a PDF of a certain distribution and $F(x)$ is its corresponding CDF, please define $F(x)$ in terms of $f(x)$.
4. Please list the parameters of the normal distribution and what each of them determine.
5. Come up with an example of when it is apt to use the quantile function.

6. Instead of using R's built-in function for a cumulative distribution function, please code up the CDF as a function of any probability density function of your choosing and its associated parameters and plot both the PDF and CDF.

7. Please explain in your own words what a survival function is.

8. Using the `rnorm` function, please make 100 random draws from the normal distribution and plot the results. Compare it to your plots from questions 1 and 2.