# Statistics

# Assignment 6

1. **What is a Cumulative Distribution Function, and how does it work?**
   **Ans:** The cumulative distribution function (CDF) calculates the cumulative probability for a given x-value. It is used to describe the probability distribution of random variables. It can be used to describe the probability for a discrete, continuous or mixed variable. It is obtained by summing up the probability density function and getting the cumulative probability for a random variable.
2. **When should we use a t-test vs a z-test?**
   **Ans:** Both z-test and t-test are used to compare a sample mean with the population mean. When we have population standard deviation available and the sample size is larger than 30, we can use the z-test. Otherwise, we can use the t-test.
3. **How do we examine two category characteristics??**
   **Ans:** If I understood this correctly, the question is about two types of data, i.e. quantitative and qualitative. To test quantitative data, we can use z-test, t-test, ANOVA test etc. To test qualitative data, we use the Chi-square test.
4. **Explain the concept of Chebyshev's Inequality?**
   **Ans:** Chebyshev's inequality, also known as Bienayme-Chebyshev inequality states that the probability that an observation is k standard deviation from the mean will be greater than or equal to $1-(1/k^2)$
   For example, the probability of a value being 2 standard deviations away (k = 2) from the mean will be greater than or equal to 75%. The calculation is as follows: $1-1/k^2 = 1-1/2^2 = 3/4 = 0.75$
5. **Explain the concept of Pareto Distribution.**
   **Ans:** The Pareto distribution was first stated by an Italian economist Vilfredo Pareto. He observed that 80% of the wealth is in the hands of 20% of the population. From this principle, the Pareto principle (also called the 80-20 rule) originated. Pareto distribution is a skewed distribution with heavy, or "slowly decaying" tails. Much of the data is in the tails. In another word, we can say 80% of the distribution in the Y axis is in the 20% of the X axis.