**Qus 1: What is a Cumulative Distribution Function, and how does it work?**

CDF is like PDF but CDF will sum up the previous values. In other way we can say that it’s running total.

**Qus 2: When should we use t-test vs z-test?**

There are two conditions:

1. When population std is given
2. When sample size is n>=30

If these 2 condition meets then we will use Z-test otherwise we will use t-test

**Qus 3: How do we examine two category characteristics?**

If there is categorical data then we should Chi-Square test.

**Qus 4: Explain the concept of Chebyshev's Inequality.**

When a random variable **Y is not belong to Gaussian Dist**

Then this come under the category of Chebyshev inequality.

Pr( population\_mean – k\*population\_std <= Y <= population\_mean+ k\*population\_std )

Where k>1

**Qus 5: Explain the concept of Pareto Distribution.**

Data that follows the power law(80-20 rule) is known as Pareto distribution.

To transform pareto into Gaussian we need to apply boxcox transformation.