Sure, here are the brief explanations for each of your questions:

1. What is a Cumulative Distribution Function, and how does it work?

A Cumulative Distribution Function (CDF) is a function that describes the probability that a random variable will take a value less than or equal to a specific value. For a given random variable \( X \), the CDF, denoted as \( F(x) \), is defined as \( F(x) = P(X \leq x) \). It provides a way to understand the distribution of the variable by showing the cumulative probability up to each point. The CDF is non-decreasing and ranges from 0 to 1 as \( x \) moves from the smallest to the largest possible values.

2. When should we use a t-test vs a z-test?

A t-test is used when the sample size is small (typically less than 30) and/or when the population standard deviation is unknown. It is most appropriate for comparing the means of two groups or a sample mean to a known value. On the other hand, a z-test is used when the sample size is large (typically 30 or more) and the population standard deviation is known. It is suitable for testing hypotheses about the population mean or comparing sample means when the variances are known.

3. How do we examine two category characteristics?

To examine the relationship between two categorical variables, we typically use a contingency table (or cross-tabulation) which displays the frequency distribution of variables. Statistical tests such as the Chi-square test of independence can then be used to determine if there is a significant association between the two variables. This test compares the observed frequencies in each category to the frequencies we would expect if there was no association between the variables.

4. Explain the concept of Chebyshev's Inequality.

Chebyshev's Inequality is a statistical theorem that states for any real-valued random variable with a finite mean and variance, the probability that the variable deviates from its mean by more than \( k \) standard deviations is at most \( \frac{1}{k^2} \). Formally, for a random variable \( X \) with mean \( \mu \) and standard deviation \( \sigma \), \( P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2} \). This inequality holds for any distribution shape, providing a conservative estimate of the spread of the distribution.

5. Explain the concept of Pareto Distribution.

The Pareto Distribution is a power-law probability distribution used to describe a phenomenon where a small number of occurrences account for a large proportion of the effect. It is characterized by the "80/20 rule," which states that roughly 80% of the effects come from 20% of the causes. The distribution is defined by a shape parameter \( \alpha \) (also called the Pareto index) and a scale parameter \( x\_m \). It is often used in economics to describe the distribution of wealth, where a small percentage of the population controls a large portion of the wealth.