1. How are you going to figure out the average heights of all the trees in Karnataka?

To figure out the average heights of all the trees in Karnataka, you would start by taking a representative sample of trees from different regions within Karnataka. This involves selecting random tree samples from various forests, parks, and urban areas. You would then measure the heights of these trees accurately and calculate the mean height of the sample. To ensure the sample accurately reflects the population, statistical methods like stratified sampling can be used to account for different tree species and environmental conditions across the state. The mean height from the sample can then be used as an estimate for the average height of all trees in Karnataka.

2. What is hypothesis testing and how does it work?

Hypothesis testing is a statistical method used to make decisions about the properties of a population based on a sample. It involves formulating two hypotheses: the null hypothesis (H0), which represents a statement of no effect or status quo, and the alternative hypothesis (H1), which represents a statement of effect or difference. Data is collected and a test statistic is computed, which is then compared to a critical value derived from a probability distribution. Based on this comparison, a decision is made to either reject the null hypothesis or fail to reject it, helping to infer whether the observed data provides enough evidence to support the alternative hypothesis.

3. Explain the differences between Alpha and Beta errors. Which inaccuracy is the most hazardous?

Alpha (α) error, also known as a Type I error, occurs when the null hypothesis is incorrectly rejected when it is true. Beta (β) error, or Type II error, occurs when the null hypothesis is incorrectly accepted when the alternative hypothesis is true. The most hazardous inaccuracy depends on the context. In medical testing, a Type I error (false positive) might lead to unnecessary treatment, while a Type II error (false negative) might result in a missed diagnosis. Generally, Type I errors are often considered more hazardous because they can lead to the acceptance of a false effect, which can have significant negative consequences.

4. What is the significance of the p-value?

The p-value measures the probability of obtaining test results at least as extreme as the observed results, assuming that the null hypothesis is true. It helps determine the strength of the evidence against the null hypothesis. A low p-value (typically ≤ 0.05) indicates strong evidence against the null hypothesis, suggesting that the observed effect is statistically significant and unlikely to have occurred by chance. Conversely, a high p-value suggests that the observed data is consistent with the null hypothesis, and there is not enough evidence to support the alternative hypothesis.

5. What is the Probability Distribution Function and how does it work?

The Probability Distribution Function (PDF) describes the likelihood of different outcomes in a random variable. For a continuous random variable, the PDF is a function that defines the probability density of the variable at each point within its range. The area under the curve of the PDF over an interval represents the probability that the random variable falls within that interval. For discrete random variables, the probability mass function (PMF) serves a similar purpose, assigning probabilities to specific outcomes. The PDF or PMF provides a comprehensive description of the variable's distribution, aiding in understanding the probabilities of different outcomes and their likelihoods.