**[Que-34] - What is the statistical hypothesis? Explain the errors in hypothesis testing.**

**(b) Explain the Sample. What are Large Samples & Small Samples?**

### **Statistical Hypothesis**

A statistical hypothesis is a statement or assumption about a population parameter, such as the population mean or population proportion. It is used as a basis for making inferences about the population from a sample. There are two types of statistical hypotheses:

1. **Null Hypothesis (H0 )**: This hypothesis states that there is no effect or no difference, and it serves as the default or baseline assumption. For example, H0 might state that the population mean is equal to a specific value.
2. **Alternative Hypothesis (H1 or Ha )**: This hypothesis states that there is an effect or a difference. It is what the researcher aims to support. For example, H1 might state that the population mean is different from a specific value.

### **Errors in Hypothesis Testing**

When performing hypothesis testing, two types of errors can occur:

1. **Type I Error (α)**: This occurs when the null hypothesis is rejected when it is actually true. It is also known as a false positive or an error of the first kind. The probability of making a Type I error is denoted by the significance level α, which is usually set at 0.05 or 0.01.
2. **Type II Error (β)**: This occurs when the null hypothesis is not rejected when it is actually false. It is also known as a false negative or an error of the second kind. The probability of making a Type II error is denoted by β. The power of a test, which is 1−β, measures the test's ability to correctly reject a false null hypothesis.

### **Sample**

A sample is a subset of individuals or observations selected from a population. The purpose of sampling is to draw conclusions about the population without examining every member. Samples should ideally be representative of the population to ensure accurate and generalizable results.

### **Large Samples vs. Small Samples**

1. **Large Samples**:
   * Typically, a sample is considered large if it has 30 or more observations (n≥30).
   * Large samples are often preferred because the sampling distribution of the sample mean approximates a normal distribution due to the Central Limit Theorem, regardless of the shape of the population distribution.
   * With large samples, estimates of population parameters tend to be more accurate, and hypothesis tests tend to be more reliable.
2. **Small Samples**:
   * A sample is considered small if it has fewer than 30 observations (n<30).
   * When dealing with small samples, the t-distribution is often used instead of the normal distribution, especially if the population standard deviation is unknown.
   * Small samples require more careful consideration of the sample selection process and the underlying distribution of the population. The results from small samples are more susceptible to variability and may be less reliable.