Unit 7 Notes for Parametric Tests →

These tests are widely used in research to analyse data and infer conclusions about the population. The foundation of parametric tests lies in their reliance on certain assumptions regarding the data's distribution, scale, and other properties.

Assumptions of Parametric Tests:

1. **Normal Distribution**: The data should be approximately normally distributed. This assumption is crucial for the validity of the test results. However, with large sample sizes, the Central Limit Theorem often justifies the use of parametric tests even if the data are not perfectly normal.
2. **Interval or Ratio Scale:** The data should be measured on an interval or ratio scale, providing meaningful figures and allowing for the calculation of differences and ratios.
3. **Homogeneity of Variances:** For tests comparing two or more groups, the variances within these groups should be approximately equal. This condition is known as homoscedasticity.
4. **Independence:** Observations must be independent of each other, meaning the data collected from one participant or observation does not influence or is influenced by data from another.

Common Types of Parametric Tests

t-tests: Used to compare the means of two groups (independent samples t-test) or the mean difference between paired observations (paired samples t-test).

ANOVA (Analysis of Variance): Used to compare the means among three or more groups. There are several types of ANOVA, including one-way ANOVA (for one independent variable) and two-way ANOVA (for two independent variables).

Linear Regression: Assesses the relationship between two continuous variables, estimating how one variable changes in relation to the other.

Pearson Correlation Coefficient: Measures the strength and direction of the linear relationship between two continuous variables.

Conducting parametric test in R-Studio, using box plots and regression lines

r Copy code

```r
library(ggplot2)

# For a box plot
ggplot(df, aes(x=group, y=value)) + geom_boxplot()

# For a scatter plot with a regression line
ggplot(df, aes(x=variable1, y=variable2)) + geom_point() + geom_smooth(method=lm)
```