ANALYSIS OF VARIANCE

ANOVA

ANOVA

- Introduction
- Different types
- Assumptions of ANOVA
- Python implementation

ASSUMPTIONS

- Distribution of sample means is normally distributed
- Independent errors
- Absence of outliers
- Homogeneity of Variance

ANOVA

- Analysis of variance (ANOVA) is a statistical method used to compare the means of two or more groups
- Must have one or more independent categorial variable
  - Factors (variables)
  - Levels

0mg	50mg	100mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

  - Factor - Dosage
  - Levels - 0mg, 50mg, 100mg

TYPES OF ANOVA

One- way ANOVA

Two- way ANOVA

One- way ANOVA

- One factor with at least two levels, levels are independent

0mg	50mg	100mg
9	7	4
8	6	3
7	6	2
8	7	3
8	8	4
9	7	3
8	6	2

- Null hypothesis: there is no difference between the groups and means
- Alternative hypothesis: there is a difference between the means of the groups

One- way ANOVA assumptions

- Normality – That each sample is taken from a normally distributed population
- Sample independence – that each sample has been drawn independently of the other samples
- Variance Equality – That the variance of data in the different groups should be the same
- Your dependent variable should be continuous

Two- way ANOVA

- The two-way ANOVA therefore examines the effect of two factors (month and gender) on a dependent variable

- Hypothesis:
  - H0: The means of all month groups are equal
  - H1: The mean of at least one month group is different

  - H0: The means of the gender groups are equal
  - H1: The means of the gender groups are different

  - H0: There is no interaction between the month and gender
  - H1: There is interaction between the month and gender

Two - way ANOVA assumptions

- Our dependent variable should be continuous
- Two independent variables should be in categorical, independent groups
- Sample independence – that each sample has been drawn independently of the other samples
- Variance Equality – That the variance of data in the different groups should be the same
- Normality – That each sample is taken from a normally distributed population

Thank You!