STA 3180 Statistical Modelling: Robust Statistics

Robust Statistics

Robust statistics is a form of statistical analysis that is used to analyze data sets that contain outliers or contain data points that are not normally distributed. It is used to reduce the effects of outliers on the results of the analysis. Robust statistics is also used to provide more accurate estimates of population parameters when the data is not normally distributed.

Key Concepts

• Robust statistics is a form of statistical analysis that is used to analyze data sets that contain outliers or contain data points that are not normally distributed.

• Robust statistics is used to reduce the effects of outliers on the results of the analysis.

• Robust statistics is also used to provide more accurate estimates of population parameters when the data is not normally distributed.

• Robust statistics can be used to identify outliers in a data set and to adjust the results of the analysis to account for the outliers.

• Robust statistics can also be used to identify patterns in data that may not be apparent when using traditional statistical methods.

Definitions

• Outlier: An outlier is an observation that is significantly different from the other observations in the data set.

• Non-normally distributed data: Data that does not follow a normal distribution.

• Population parameter: A population parameter is a measure of the characteristics of a population, such as the mean or standard deviation.

Rules

• Robust statistics should be used when the data is not normally distributed or when there are outliers present in the data.

• Robust statistics should be used to identify outliers in the data and to adjust the results of the analysis to account for the outliers.

• Robust statistics should be used to identify patterns in the data that may not be apparent when using traditional statistical methods.

Examples

• A researcher is studying the heights of people in a population. The data is not normally distributed and there are several outliers present. The researcher can use robust statistics to identify the outliers and to adjust the results of the analysis to account for the outliers.

• A researcher is studying the salaries of people in a population. The data is not normally distributed and there are several outliers present. The researcher can use robust statistics to identify the outliers and to adjust the results of the analysis to account for the outliers. The researcher can also use robust statistics to identify patterns in the data that may not be apparent when using traditional statistical methods.