Understanding Statistical Models: Predicting the Future from Data

Clara Maj[*]

Department of Mathematics, Jomo Kenyatta University, Kenya

clara_maj@gmail.com

Description

In hypothesis testing, a null hypothesis (usually representing the status quo) is compared with an alternative hypothesis, and statistical tests determine whether there is enough evidence to reject the null hypothesis in Favor of the alternative. Confidence intervals are another tool used in inferential statistics to estimate the range within which a population parameter is likely to lie. For example, a researcher might use a confidence interval to estimate the true average income of a population based on a sample of incomes. Probability theory is the foundation of inferential statistics. It allows statisticians to quantify uncertainty and draw conclusions based on the likelihood of certain events occurring. In everyday life, we are often faced with uncertainty, such as the probability of rain tomorrow or the chances of winning a lottery. In statistics, probability is used to model and analyse random events, which have a wide range of applications. For instance, in healthcare, probability is used to assess the likelihood of a patient developing a certain disease based on risk factors like age, gender, and lifestyle. In business, companies use probability models to predict customer behaviour, market trends, and sales figures. In sports, probability is used to assess the chances of a team winning a game based on past performance and other factors. Statistical sampling is a technique used to select a subset of individuals or observations from a larger population for the purpose of making inferences about the population as a whole. The key to effective sampling is ensuring that the sample is representative of the population. This means that the sample should mirror the diversity and characteristics of the population from which it is drawn. There are various methods of sampling, including random sampling, stratified sampling, and cluster sampling. In random sampling, every member of the population has an equal chance of being selected. Stratified sampling involves dividing the population into subgroups based on certain characteristics, and then randomly selecting samples from each subgroup. Cluster sampling is used when the population is geographically dispersed, and involves selecting random groups or clusters of individuals rather than individual members. Proper sampling techniques are crucial for obtaining reliable and valid results. If the sample is biased or unrepresentative, the conclusions drawn from the analysis may not accurately reflect the true characteristics of the population. Statistics is a versatile tool that is used in virtually every field, ranging from the social sciences to natural sciences, business, healthcare, government, and beyond. In medicine, statistics is used to analyse clinical trials, assess treatment effectiveness, and identify health risks. For example, the effectiveness of a new drug is often measured through randomized controlled trials, which rely heavily on statistical analysis to determine whether the drug works better than a placebo. In business and economics, statistical methods are employed to analyse market trends, forecast demand, and evaluate the success of marketing campaigns. Financial analysts use statistics to assess the risk of investments, while economists use data to study unemployment rates, inflation, and economic growth. In social sciences, researchers use statistical methods to analyse survey data, conduct experiments, and test theories.

Acknowledgement

None.

Conflict of Interest

The authors are grateful to the journal editor and the anonymous reviewers for their helpful comments and suggestions.