RA2211047010020

Gokul C J

AI-A

Alzheimer's Disease Analysis: A Case Study in Inferential Statistics

Abstract

This report presents an inferential statistical analysis of an Alzheimer's disease dataset, aiming to explore key relationships between patient health indicators. The study's objective was to apply and interpret common statistical tests, including the one-sample t-test, two-sample t-test, and one-way ANOVA, to draw conclusions about the dataset's population. We found that the mean systolic blood pressure of the dataset's population is significantly higher than the standard reference value of 120 mmHg. However, no significant statistical difference was observed in mean BMI levels between smokers and non-smokers, nor in mean MMSE scores across different ethnic groups. These findings provide insights into the interdependencies of various health metrics within an Alzheimer's patient population and demonstrate the utility of inferential statistics in medical data analysis.

Introduction & Methods

Introduction

This report analyzes an Alzheimer's disease dataset to explore relationships between health indicators using inferential statistics. The goal is to apply and interpret one-sample t-tests, two-sample t-tests, and one-way ANOVA to identify significant differences and patterns within the data.

Methods

Dataset Description

The analysis used an Alzheimer's disease dataset with continuous variables (SystolicBP, BMI, MMSE) and categorical variables (Smoking, Ethnicity). Missing values were handled appropriately before analysis.

Hypotheses & Analysis

Three statistical tests were performed:

• One-Sample T-Test: Compared mean SystolicBP to a reference value of 120 mmHg. • Two-Sample T-Test: Compared mean BMI between smokers and non-smokers. • One-Way ANOVA: Compared mean MMSE scores across different ethnic groups, followed by a Tukey HSD post-hoc test if significant.

Results

One-Sample T-Test for Systolic Blood Pressure

The one-sample t-test compared the mean SystolicBP of the dataset to a reference value of 120 mmHg.

• Mean Systolic BP: 134.26

• T-statistic: 25.48

• P-value: <0.001

Since the p-value is far less than $\alpha=0.05$, we reject the null hypothesis. The results provide strong evidence that the mean systolic blood pressure in this dataset is significantly higher than the reference value of 120 mmHg.

Two-Sample T-Test for BMI (Smokers vs Non-Smokers)

An independent two-sample t-test was used to compare mean BMI levels between smokers and non-smokers.

• Mean BMI (Non-smokers): 27.56

• Mean BMI (Smokers): 27.89

• T-statistic: -0.947

• P-value: 0.344

The p-value is greater than $\alpha=0.05$. Therefore, we fail to reject the null hypothesis. The analysis provides evidence that mean BMI is not significantly different between smokers and non-smokers.

One-Way ANOVA for MMSE by Ethnicity

A one-way ANOVA was conducted to test for differences in mean MMSE scores across ethnic groups.

- F-statistic: 0.208

- P-value: 0.891

With a p-value much greater than $\alpha=0.05$, the results indicate no significant difference in mean MMSE scores across the ethnicity groups. Thus, we fail to reject the null hypothesis.

Discussion

The statistical analysis provided meaningful insights into the Alzheimer's dataset. The significant t-test results for systolic blood pressure align with known clinical risk factors, confirming that this patient cohort exhibits elevated blood pressure compared to the healthy reference. The lack of significant differences in BMI between smokers and non-smokers, and in MMSE across ethnic groups, may reflect dataset characteristics rather than true absence of effect. These findings highlight the importance of using inferential statistics to validate hypotheses while acknowledging dataset limitations.

Limitations

This study has several limitations. The dataset's single origin may limit the generalizability of these findings. Handling of missing data through imputation can introduce bias. As an observational study, causal relationships cannot be established. Additionally, some subgroups may have small sample sizes, reducing statistical power.

Conclusion

In conclusion, this case study successfully applied inferential statistics to an Alzheimer's disease dataset. The analysis confirmed that systolic blood pressure was significantly higher than the standard reference value, while no significant differences were observed for BMI between smokers and non-smokers or for MMSE across ethnic groups. These results demonstrate the utility of statistical methods in confirming general trends and assessing specific relationships within clinical data, offering insights that could inform further research.