Exercise 2

Exploratory Data Analysis - Titanic Data Set

In this exercise, we want to know if there is any association between the Survival Rate as our variable of interest and the following variables: Passenger Class, Gender, and Age.

The analysis obtained is described by performing the statistics and plots as follows:

1) Determine if the Survival Rate is associated to the Class of Passenger

Interpretation of the graph:

When we examine the side-by-side graph, it becomes evident that there is no apparent independence between Survival Rate and Passenger Class. This suggests a likely association between the two variables. To confirm and formally test our hypothesis, let's proceed with a Chi-Square analysis.

Hypotheses

H0: Survival Rate and Passenger Class are independent among all subjects in the population

Ha: Survival Rate and Passenger Class are not independent among all subjects in the population

Conclusion:

Since the Chi-square value of 102.889 is greater than the decision point DP = 5.99 with 2 Degrees of Freedom (Based on the Chi-Square Decision Point Table), we have enough evidence to reject the null hypothesis and conclude that the Survival Rate and Passenger Class are not significantly independent of each other among all subjects in the population.

In conclusion, there is an association between Survival Rate and Passenger Class

2) Determine if the Survival Rate is associated with the Gender

Interpretation of the graph:

When we examine the side-by-side graph, it becomes evident that there is no apparent independence between Survival Rate and Gender. This suggests a likely association between the two variables. To confirm and formally test our hypothesis, let's proceed with a Chi-Square analysis.

Hypotheses

H0: Survival Rate and Gender are independent among all subjects in the population

Ha: Survival Rate and Gender are not independent among all subjects in the population

---

Conclusion:

Since the Chi-square value of 260.717 is greater than the critical value Decision Point of 3.84 for 1 Degree of Freedom (as per the Chi-Square Decision Point Table), we have sufficient evidence to reject the null hypothesis. Therefore, we can conclude that the Survival Rate and Gender are not significantly independent among all subjects in the population.

In conclusion, there is an association between Survival Rate and Gender

3) Determine if the Survival Rate is associated to the Age

Interpretation of the graph:

When we examine the side-by-side graph, it becomes evident that there is no apparent independence between Survival Rate and Age. This suggests a likely association between the two variables. To confirm and formally test our hypothesis, let's proceed with a Chi-Square analysis.

Hypotheses

H0: Survival Rate and Age are independent among all subjects in the population

Ha: Survival Rate and Age are not independent among all subjects in the population

Conclusion:

Since the Chi-square value of 6.891 is greater than the critical value Decision Point of 5.99 for 2 Degrees of Freedom (as per the Chi-Square Decision Point Table), we have sufficient evidence to reject the null hypothesis. Therefore, we can conclude that the Survival Rate and Gender are not significantly independent of each other among all subjects in the population.

In conclusion, there is an association between Survival Rate and Age