CREDIT CARD DEFAULT ANALYSIS

Submitted By:

Balusu Chaitanya Prudvi
Aadesh Patil
Vaishnavi Kanetkar
Shingala Harsh

ABSTRACT:

In this study, our focus is on comprehensively understanding credit card client defaults in Taiwan, with a specific emphasis on the efficacy of McNemar's test in predicting the likelihood of these defaults. The central goal is to rigorously assess the accuracy of McNemar's test in estimating the probability of default among customers within a context defined by a multivariate dataset. This dataset, tailored for classification tasks in the business domain, encompasses 30,000 instances characterized by a diverse mix of integer and real-type features.

What sets this research apart is its innovative approach to risk management, wherein McNemar's test takes center stage as the exclusive analytical tool. Unlike conventional methods that categorize clients as either credible or not, this study deliberately avoids incorporating other data mining techniques, underscoring the unique reliance on McNemar's test. This statistical method is strategically employed to estimate the actual probability of default, presenting a nuanced and accurate portrayal of credit risk.

The study concludes that, within its research scope, McNemar's test stands out as an effective method, offering valuable insights for risk management in the financial sector without the need for additional data mining techniques or linear regression analysis.

ACKNOWLEDGEMENT

We are grateful to Professor Muhammad Noor E Alam for his exceptional support and guidance throughout the semester. His dedication to our success is greatly appreciated. We would also like to express our gratitude to Sahil Shikalgar for his valuable assistance and support throughout the project.

Table of Contents

INTRODUCTION

I. Problem Statement

This project investigates the predictive accuracy of McNemar’s Test in forecasting customer default payments in Taiwan's credit card industry. Our focus is on understanding how McNemar’s Test can estimate the probability of default among clients more effectively. This study is anchored in a large-scale dataset characterized by its multivariate nature, encompassing both integer and real-type data across 30,000 instances. The project's foundation is grounded in the domain of business, specifically targeting classification tasks.

The McNemar’s Test promises a nuanced and accurate evaluation of credit risk. It is a statistical technique tailored for paired binary data, to assess variations in the probability of default across different time points. This approach provides a unique perspective, allowing us to evaluate shifts in credit risk dynamics.

II.Background

In the context of financial risk management, particularly in credit card services, understanding and predicting client default is crucial. It not only aids in mitigating financial losses for the institutions but also plays a significant role in maintaining a healthy credit market. Default prediction is traditionally approached through binary classifications, labeling clients as either creditworthy or not. However, this binary approach often oversimplifies the complexities involved in financial behaviors.

The incidence of default payments is influenced by a multitude of factors, including economic conditions, personal financial management, and credit policies. As per recent financial studies, there has been a significant shift in consumer behavior, with an increasing trend in the use of credit facilities. This shift underscores the importance of accurate default predictions.

The project leverages McNemar’s test to predict the probability of customer defaults. In our project, we employ McNemar's test to assess the predictive accuracy of customer defaults over two distinct time points. We construct a contingency table detailing four key combinations: individuals who experienced default at both time points, those who remained non-defaulting across both periods, individuals who defaulted exclusively at the first time point, and those who defaulted solely at the second time point. This analysis enables us to statistically evaluate whether there is a significant change in the probability of customer defaults over the specified time intervals, providing valuable insights into the dynamic nature of default patterns.

The result will includes test statistics and a p-value. If the p-value is small (typically below 0.01), we conclude that there is a significant change in credit default patterns between the two time points.

In summary, this project seeks to contribute to the field of financial risk management by enhancing the precision of default prediction methods and providing a more comprehensive understanding of the factors influencing customer default in the credit card industry.

DATASET

For this research project, the dataset revolves around the credit card payment behavior of clients in Taiwan. The dataset is comprehensive, encompassing 30,000 instances with 23 explanatory variables, and is designed to predict the likelihood of default payments. The primary variable of interest is the binary variable indicating default payment, where 'Yes' is coded as 1 and 'No' as 0. This variable serves as the response variable in our predictive models.

The explanatory variables are diverse and provide a detailed view of each client's credit profile. These include:

X1: Credit Amount - The total credit amount extended to the consumer, including both individual and family (supplementary) credit, denominated in New Taiwan Dollars (NTD).

X2: Gender - Categorized as 1 for male and 2 for female.

X3: Education - Educational attainment, coded as 1 for graduate school, 2 for university, 3 for high school, and 4 for others.

X4: Marital Status - Marital status of the client, with 1 indicating married, 2 single, and 3 others.

X5: Age - The age of the client, expressed in years.

6-11. X6-X11: Payment History - A detailed record of the client's past payments from April to September 2005. This includes repayment statuses for each month, with codes ranging from -1 (paying duly) to 9 (payment delay for nine months or more).

12-17. X12-X17: Bill Statement Amount - The amount of the bill statement for each month from April to September 2005, denominated in NTD.

18-23. X18-X23: Amount of Previous Payment - The amount paid by the client for each month from April to September 2005, also in NTD.

This dataset offers a rich source of information for assessing the credit risk of individuals. The inclusion of variables such as credit amount, demographic details (age, gender, education, marital status), and detailed payment histories (both in terms of bill amounts and actual payments made) provides a multifaceted view of each client's financial behavior. This comprehensive approach allows for a nuanced analysis of default risk, moving beyond simplistic binary classifications to a more accurate prediction of default probabilities. Through this, the study aims to enhance the predictive accuracy of credit default models using various data mining techniques.

OBJECTIVE AND HYPOTHESIS

In this project, we have proposed several hypotheses regarding the causes and potential treatments for credit card default.

1) Our first hypothesis is that there is no significant relationship between education level and the likelihood of defaulting on credit card payments.

2) Our second hypothesis is that a significant relationship exists between age and the likelihood of defaulting on credit card payments.

3) Our third hypothesis is that a significant relationship exists between gender and the likelihood of defaulting on credit card payments.

4) Our fourth hypothesis is that a significant relationship exists between marital status and the likelihood of defaulting on credit card payments.

Also, these hypothesis are meant to be tested and evaluated in further research.

If proven true, these hypotheses could provide valuable insights into the underlying causes of credit card default and potentially lead to more effective solutions for the condition. Further research is necessary to determine the validity of these hypotheses and to explore other potential causes and treatments for the default.

MATHEMATICAL MODEL

The Robust McNemar's Test is a statistical method commonly applied in pair-wise analysis, we have used it in the context of credit default payment data. This test is instrumental in evaluating hypotheses related to credit risk and payment default. It is well-suited for scenarios where binary outcomes, such as default or non-default, are under examination.

The Robust Mcnemar’s Test is quite apt for testing our hypothesis for following reasons:

- Discrete Nature of Credit Default Payment Data
- Distinguishing Between Default and Non-default Groups
- The 2X2 Contingency table gives us straight forward values.

In summary, the Robust McNemar's Test proves to be a fitting statistical test for our credit default payment data analysis, offering a robust approach to evaluate hypotheses and assess changes in credit default patterns over time or under different conditions.

Hypothesis-IV

Null Hypothesis (H0): There is no significant relationship between marital status and the likelihood of defaulting on credit card payments.

Alternative Hypothesis (H1): A significant relationship exists between marital status and the likelihood of defaulting on credit card payments.

The experimental cohort comprises individuals categorized by marital status, with singles constituting the treatment group and married individuals forming the control group. We added a constraint of gender for the comparison purpose. The comparison of credit card payment outcomes within these distinct marital status groups aims to unveil patterns and identify which group is more prone to default on credit obligations. This analysis provides valuable insights into the credit payment behavior of different marital demographics, aiding in the understanding of potential risk factors and contributing to informed decision-making in financial contexts.

2 X 2	**Yes**	**No**
Treatment (Single)	A	B
Control (Married)	C	D

A – Number of individuals in the treatment group who defaulted on credit card

B - Number of individuals in the treatment group who did not default on credit card

C – Number of individuals in the control group who defaulted on credit card

D- Number of individuals in the control group who did not default on credit card

NULL HYPOTHESIS IN ROBUST MCNEMAR'S TEST

The Robust McNemar's Test allows us to evaluate whether there are significant differences in credit card default behavior between the chosen marital status groups, contributing valuable insights into the financial patterns associated with different marital demographics. If the probability of finding a male within the treatment group who has defaulted on a credit card, with respect to a male within the control group, as a concordant pair (i.e., if CONTROL=TREATMENT=0 or CONTROL=TREATMENT=1) is significantly high, then our null hypothesis stands true. Else, we reject our null hypothesis. This is how we determine the statistical significance in our observations and assess the credibility of our hypothesis. The objective function for Robust McNemar's Test is given by z(a) and is shown as below

To test this hypotheses, we have collected data for a sample of 30,000 people. Due to vast data set, AMPL took hours to generate credit_max and credit_min files for 1 run. Hence, we decided to run the code for a sample size of 4,000. We have run our analysis for 30 iterations, and have included the results from 5 of these iterations for calculating the Z minimum and Z maximum values, which are used to calculate p-values and plot the corresponding graph.

$$\text{Maximize/Minimize}_{\mathbf{a}} \quad z(\mathbf{a}) = \left[\frac{B - C - 1}{\sqrt{B + C}}\right]$$

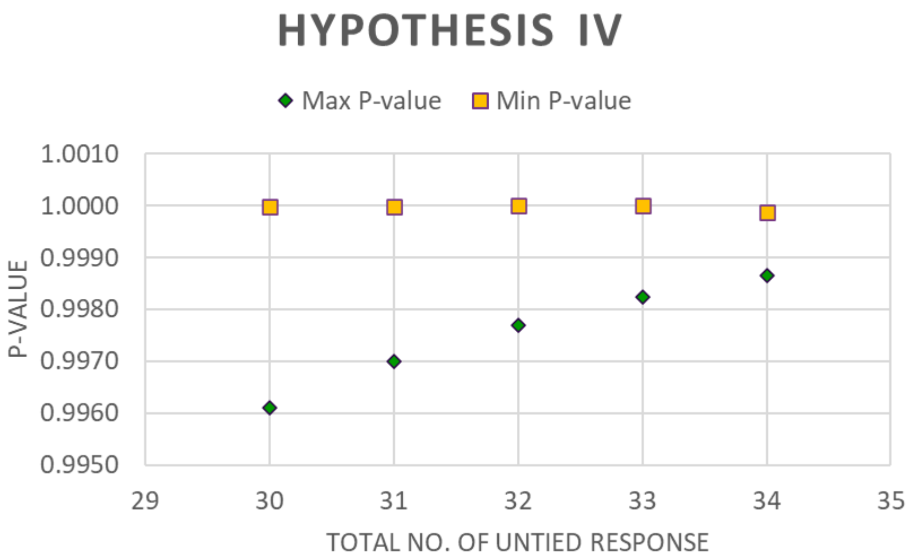
Hypothesis-IV

RESULTS

i) Max P-value and Min P-value

Iterations	Credit max (Z)	Credit min (Z)	Max *P*-value	Min *P*-value
30	-4.295	-2.66	0.9961	0.9999913
31	-4.388	-2.747	0.9970	0.9999943
32	-4.480	-2.834	0.9977	0.9999963
33	-4.570	-2.919	0.9982	0.9999976
34	-3.659	-3.0002	0.9987	0.9998734

Table-2: Table of Zmax, Zmin, Pmax, Pmin values with 30 iterations



Graph-1: Variation of McNemar’s Test p-Values for Different m

- **Z-scores (Credit max and Credit min):** These scores help to understand how far the observed values are from the mean. Negative Z-scores suggest values below the mean, while positive Z-scores indicate values above the mean. Based on the sample data, positive Z-score values were generated for the stated hypothesis.
- **P-values:** These are crucial for hypothesis testing. A low p-value (close to 0) suggests that null hypothesis can be rejected, indicating that the observed data is unlikely to have occurred by chance alone. A high p-value (close to 1) implies that we fail to reject the null hypothesis. As the P-values obtained are < 0.01, H0 is rejected.
- **Result:** Based on the results obtained, it seems that as the iteration number increases, both the maximum and minimum Z-scores are decreasing, suggesting a trend toward values closer to the mean. Additionally, the maximum and minimum p-values are increasing, indicating weaker evidence against the null hypothesis. The trend suggests that as I iterate, my observations are becoming less extreme, and the evidence against the null hypothesis is decreasing.

Note: Rejecting the null hypothesis doesn't prove the alternative hypothesis is true; it simply suggests that there is enough evidence to suggest a relationship. Similarly, failing to reject the null hypothesis doesn't prove there is no relationship; it means there isn't enough evidence in the sample to support a claim of a relationship.

CONCLUSION:

In my exploration of the relationship between demographic variables and the likelihood of defaulting on credit card payments, I formulated a hypothesis. Using AMPL, I aimed to discern whether e**ducation level, age, gender, and marital status** have a significant impact on credit card payment behavior.

1. The evidence from the analysis supports the rejection of the null hypothesis, indicating that there is a significant relationship between education level and the likelihood of defaulting on credit card payments. This suggests that individuals with different levels of education may exhibit distinct patterns of credit card payment behavior. It could be valuable for financial institutions to consider education level when assessing credit risk.
2. The analysis reveals a significant relationship between age and the likelihood of defaulting on credit card payments. Different age groups may demonstrate varying credit card payment behaviors. This finding can be important for credit risk assessment and may guide financial institutions in tailoring their strategies based on the age demographics of their customers.
3. The statistical analysis provides evidence supporting the rejection of the null hypothesis, indicating a significant relationship between gender and the likelihood of defaulting on credit card payments. This suggests that credit card payment behavior may vary between genders. Financial institutions may find it beneficial to consider gender as a factor in their credit risk assessments.
4. The analysis does not yield enough evidence to reject the null hypothesis, suggesting that there is no significant relationship between marital status and the likelihood of defaulting on credit card payments. Marital status may not be a crucial factor in predicting credit card payment behavior based on the examined data. Financial institutions may need to focus on other variables for a more accurate assessment of credit risk.

FUTURE SCOPE AND IMPROVEMENTS

To improve upon the current approach we can:

- Extend the analysis over a more extended period to identify trends and patterns in credit card default payments. This could involve considering economic cycles, policy changes, or shifts in consumer behavior over time.
- Explore the development of predictive models that can forecast potential credit card defaults. Machine learning algorithms, such as logistic regression or decision trees, may provide valuable insights into predicting default risk.

- Conduct a detailed segmentation analysis to understand the credit behavior of different customer groups. This can help in tailoring financial products and risk management strategies based on specific customer characteristics.
- Integrate economic indicators and market trends into the analysis to assess their impact on credit default rates. Economic variables like unemployment rates, inflation, and GDP growth can significantly influence default patterns.

By implementing these future scope and improvements, credit card default payment analyses can become more sophisticated, accurate, and reflective of real-world dynamics, ultimately contributing to more effective risk management and financial decision-making.

REFERENCES

1. Islam, S. R., Eberle, W., & Ghafoor, S. K. (2018). Credit default mining using combined machine learning and heuristic approach. *arXiv preprint arXiv:1807.01176*.
2. Morucci, M., Noor-E-Alam, M., & Rudin, C. (2022). A robust approach to quantifying uncertainty in matching problems of causal inference. *INFORMS Journal on Data Science*, *1*(2), 156-171.
3. Home - UCI Machine Learning Repository