

Statistical Analysis of Non Banking Mobile App Users

B.K.S. Ishanka
Computer Science Department
University Of Moratuwa
Katubedda, Sri Lanka
shehan.20@cse.mrt.ac.lk

Abstract—Today, banks are providing the facility to customers to fulfill their banking needs through mobile banking applications without coming to the banks physically. But some of the customers are not still using those due to different reasons. Here I attempt to identify various reasons through a survey conducted on non banking mobile application users and provide statistical analysis on it. Several statistical hypothesis tests are carried out to improve the statistics for the population.

Keywords—Hypothesis testing, P-value, Confidence Interval, Normal Distribution, T- Distribution

I. INTRODUCTION

With the rapid development of technology, most of the people tend to use novel technological thing to ease their daily routine works. Today most of the things that we did earlier manually has been replaced with mobile application based system. This paves the path to the users to engage in their essential works without traveling. One such most useful application is mobile application facility for bank users. Bank customers does not need to go to the bank anymore to do their transaction activities such as take money and do payments, transfer money across different accounts, loan applications etc. This mobile applications has made people's life much easier than earlier. Today almost every bank in Sri Lanka provides this service without any charges and they even persuade customers to use more the mobile applications since both parties can save time and additional cost. Some examples of mobile applications that provides digital banking facility are People's Wave in People's Bank, B App in Bank of Ceylon, Sampath Wishwa in Sampath Bank, ComBankDigital in Commercial Bank. It is hard to find a bank which does not provide such facility. Further more banks have been cooperately intergrated with other business models to pay bills through the banking apps and it makes these applications special and rather useful. Even the mobile applications provide the functionality transfer money across different accounts in different banks.

Eventhough this type of assistance exists, some customer still tend to physically come to the bank and do their transactions. Typically, this kind of behaviour prevails in elderly people. We can assume that the elderly community is not very much familiar with new technologies with compared to young community. Sometimes, most of the elderly people go to the banks to collect their pension and also they make it a place to meet their old companions. These are some main factors which elderly people could not be trained to use mobile technology. Eventhough elderly community accounts for most of proportion of the non mobile banking application users, there is a considerable no of proportion of young community who are reluctant to use them can be identified. Foremost reason for this is low computer literacy. Other than that some customers feel more secured when transferring money at a bank physically rather

than virtually through mobile applications. These facts raise that it is needed to followup on this matter and do a proper analysis.

In statistical inference, it attempts to provide insights regarding a population from a sample. Practically it is hard to acquire data regarding a whole population. Since, the data is collected for a small proportion of the population which is a sample, and then from the sample, the statistics of the population is estimated. Statistical hypothesis is assumed regarding population parameter from sample parameter. In hypothesis testing, it is attempted to testify whether the assumption is valid or not. For this procedure, it includes null hypothesis and alternative hypothesis. The null hypothesis, denoted by H_0 , is usually the hypothesis that sample observations result purely from chance. The alternative hypothesis, denoted by H_1 or H_a , is the hypothesis that sample observations are influenced by some non-random cause. Hypothesis test can have two probable outcomes. They are accepting the null hypothesis or rejecting the null hypothesis. Failure to reject implies that the data are not sufficiently persuasive for us to prefer the alternative hypothesis over the null hypothesis.

In this procedure, there are four steps which are state the hypotheses, formulate an analysis plan, analyze sample data and interpret results. In formulating analysis plan phase, significance level is selected and proper testing method is selected. Typically, the test method involves a test statistic and a sampling distribution. Computed from sample data, the test statistic might be a mean score, proportion, difference between means, difference between proportions, z-score, t statistic, chi-square, etc. Given a test statistic and its sampling distribution, a researcher can assess probabilities associated with the test statistic. If the test statistic probability is less than the significance level, the null hypothesis is rejected. Then using sample data, perform computations called for in the analysis plan. Then if the sample findings are unlikely, given the null hypothesis, the researcher rejects the null hypothesis. Typically, this involves comparing the P-value to the significance level, and rejecting the null hypothesis when the P-value is less than the significance level.

In this research I attempt to provide a statistical analysis with several hypothesis tests regarding the users who does not use mobile applications for banking transactions.

The dataset is prepared by doing a survey, asking questions from different non banking mobile application users. The statistical analysis¹ is done using R language. Section II describes about how the dataset is prepared and Section III describes about the statistical analysis. Under section III preliminary analysis about dataset and statistical hypothesis testings are provided.

¹Code is available is at
<https://github.com/ShehanIshanka/non-banking-mobile-app-user-analysis>

II. DATASET

The dataset was prepared from survey which was done for about two weeks. Data is collected through Google Forms by sending the questionnaire to non mobile banking application users through social media. The dataset includes 73 records.

The questionnaire was prepared from both English(Table I) and Sinhala Language(Table II) targeting mainly sinhala people. The questions can be divide into two subsections; demographics questions and non banking mobile application user questions. The answers for the questions were provided so that it is needed only to select the answer and it deprives the necessity of preprocessing answers. All the questions are mandatory except the last question which is optional.

Demographic questions are focused on user's details. First 5 questions are the demographic questions. It generally focuses on non banking mobile application users' age, gender, education qualification, home town and occupation. Possible answers are provided for each of the question. For age, age groups of less than 16, 16 – 25, 25 – 45, 45 – 60, More than 60 are provided to capture young, mid age and elderly community. For gender, 3 types of answers are provided. They are Male, Female and Prefer not to say. For education qualifications, Grade 5 Scholarship, Ordinary Level, Advanced Level, Diploma, Degree and Prefer not to say are provided as the answers. For home town, the 26 districts of Sri Lanka are provided as the answers.

Non banking mobile application user questions are more focused on computer literacy and people's knowledge about the new technology to measure one aspect of not using mobile applications for banking processes. Final 7 questions are related to this type of questions and these questions have answers of Yes, No and No idea. The last question is not a multiple choice question and it is a simply long answer question to capture any other missed facts about non banking mobile application usage.

Only 16 responses are captured for the final question. The most of the answers for the final optional question is regarding security issues. Most of the bank users are afraid of both security glitches of mobile applications and trustworthiness of bank processes. They would not still be compelled to trust a mobile applications to ease their works. Some users feel that these mobile applications are less

TABLE I. SINHALA QUESTIONS

| | Sinhala Questions |
|----|--|
| 1 | ඔබගේ ස්ත්‍රී පුරුෂ භාවය කුමක්ද? |
| 2 | ඔබ අයත් වයස් කාණ්ඩය තෝරන්න. |
| 3 | ඔබගේ වර්තමාන ජීවත්වන නගරය තෝරන්න. |
| 4 | ඔබගේ ඉහළම අධ්‍යාපන සුදුසුකම කුමක්ද? |
| 5 | ඔබේ රැකියාව කුමක්ද? |
| 6 | ඔබ මීට පෙර බැංකු ජංගම දුරකථන යෙදුම් ගැන අසා තිබේද? |
| 7 | ඔබ මීට පෙර බැංකු ජංගම යෙදුම් භාවිතා කර තිබේද? |
| 8 | ඉහත ප්‍රශ්නයට ඔව් නම්, ඔබේ බැංකුව / බැංකු එම පහසුකම් සපයයිද? |
| 9 | ඔබේ බැංකු ක්‍රියාදාමයන් පහසු කිරීමට බැංකු ජංගම යෙදුම් ඔබට උපකාරීවනු ඇතැයි ඔබ සිතනවාද? |
| 10 | ඔබට ස්මාර්ට් ජංගම දුරකතනයක් තිබේද? |
| 11 | ඔබගේ දුරකථනයේ ජංගම බැංකු යෙදුමක් ස්ථාපනය කර එය ඔබේ බැංකු ක්‍රියාටයින් සඳහා භාවිතා කළ හැකි යැයි ඔබ සිතනවාද? |
| 12 | ඉහත කරුණු හැරුණු විට, ජංගම බැංකු යෙදුම් භාවිතා නොකිරීමට හේතුව / හේතු මොනවාද? |

TABLE II. ENGLISH QUESTIONS

| | English Questions |
|----|--|
| 1 | What is your gender? |
| 2 | Select the age group that you belong to. |
| 3 | Select your current living town. |
| 4 | What is your highest educational qualification? |
| 5 | What is your occupation? |
| 6 | Have you heard of banking mobile apps before? |
| 7 | Have you earlier used banking mobile apps before? |
| 8 | If yes to above question, do/does your bank/banks provide the facility? |
| 9 | Do you think, Banking Mobile Apps would help you to make your bank processes comfortable? |
| 10 | Do you own a smart phone? |
| 11 | Do you think you could install a mobile banking app and use it for your banking processes? |
| 12 | Other than the above points, what is/are the reason/reasons for not using mobile banking apps? |

efficient and not convenient access with respect to banking applications in other countries. Further there were some reasons like poor internet connection, hidden charges, annual fee for mobile/internet banking facilities, filling multiple forms without activating the service etc.

After extracting data, dataset was preprocessed to remove sinhala wordings to adhere it to use in statistical analysis. Any user identification data has not been also extracted.

III. ANALYSIS

Analysis is carried out in two ways; preliminary analysis and statistical hypothesis test analysis.

A. Preliminary Analysis

Under this, an analysis is carried out to understand about the dataset. The demographic questions are selected for this analysis and each question related feature is separately analysed.

1) Age

Although there are 5 age groups, the responses are from only 3 groups. It can be assumed that most of the people above age 45 are not familiar with electronic media because the survey is sent as a google form (But this may be not the case also). Further the age group of 25 – 45 has the most no of responses as expected.

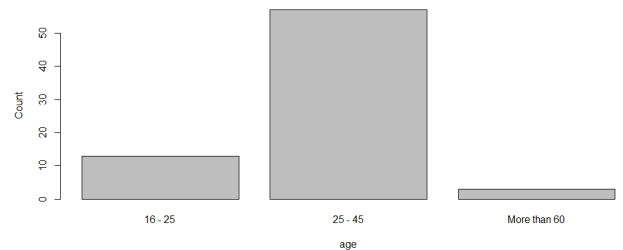


Fig. 1. Age Distribution

2) Education Qualification

Although there are 6 answers for this question, the responses are only related to 3 educational qualifications. The people with a degree have responded more than others with other qualifications.

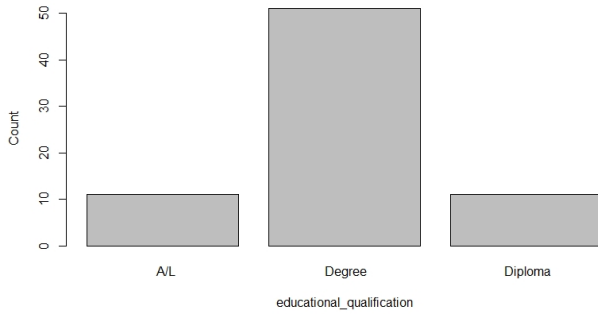


Fig. 2. Education Qualification Distribution

3) Gender

All the responders have selected either male or female and no one selected the answer which is “prefer not to say”. Most of responders are male responders.

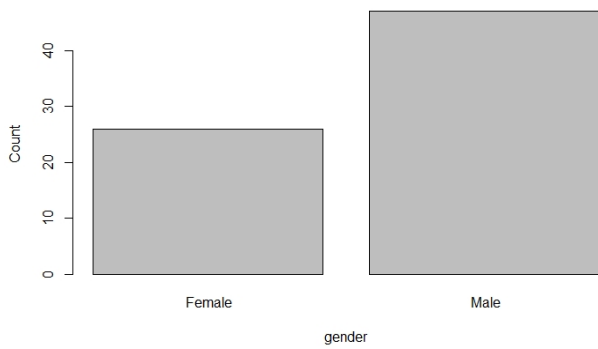


Fig. 3. Gender Distribution

4) Hometown

For the hometown question, 26 districts of Sri Lanka have provided. But there are responses only from 10 districts. The survey was mainly targeted on Matara district. Since Most of the records are from Matara District.

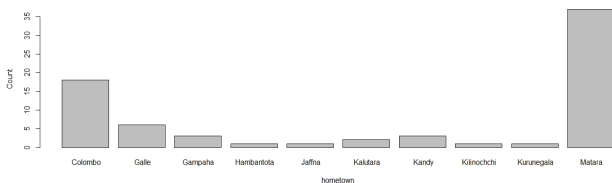


Fig. 4. Home Town Distribution

5) Occupation

For occupation question, 6 answers are provided but responses are only for 4 answers. Out of them, responses with private sector is the largest.

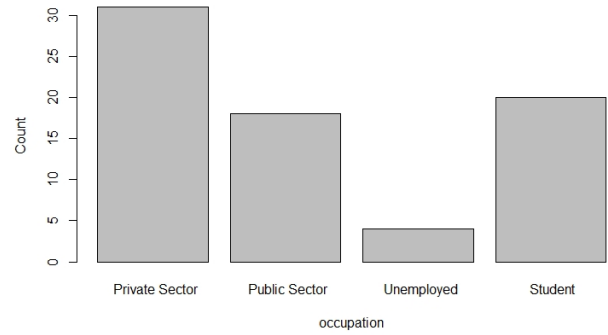


Fig. 5. Education Qualification Distribution

B. Statistical Analysis

Under this several hypothesis tests are carried out to verify different insights.

All the hypothesis tests are proportion tests and first five tests (1-5) are single proportion tests and the next two tests (6,7) are related to distributions of differences in proportions. The final two tests (8,9) are association tests.

1) Most of the people have heard of mobile banking applications

For this question, the answers are “Yes” and “No”.

TABLE III. DISTRIBUTION OF THE ANSWERS

| Yes | No |
|-------|-------|
| 72 | 1 |
| 0.986 | 0.014 |

p = Proportion of people who have heard of mobile banking applications.

$$H_0: p = 72/73 = 0.986$$

$$H_a: p \neq 1/73 = 0.014$$

Since bootstrap distribution is skewed and not a normal distribution, it is erroneous to perform hypothesis tests. But we can argue that it is almost everyone has heard about mobile banking applications since the people whom that the survey is filled should have a mobile application because form is a google form and probably they have heard of mobile banking applications.

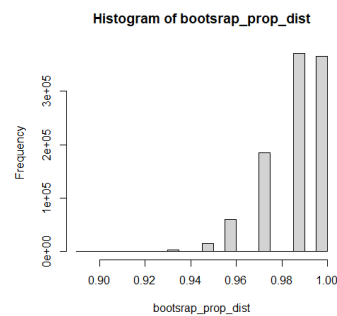


Fig. 6. Bootstrap Distribution

- 2) *Most of the people have earlier used banking mobile apps.*

For this question, the answers are “Yes” and “No”.

TABLE IV. DISTRIBUTION OF THE ANSWERS

| Yes | No |
|------|------|
| 39 | 34 |
| 0.53 | 0.47 |

p = Proportion of people have earlier used banking mobile apps.

$$H_0: p = 0.5$$

$$H_a: p < 0.5$$

Conditions of Test for a Single Proportion : $n \cdot p = 39 > 10$,
 $n \cdot (1-p) = 34 > 10$

p -value = 0.28

Since the P-value (0.28) is greater than the significance level (0.05), we cannot reject the null hypothesis. Since that it can be assumed that most of the people has been not satisfied with the functionalities of mobile banking applications.

- 3) *Most of the people think Mobile Apps would help you to make your bank processes comfortable*

For this question, the answers are “Yes”, “No” and “No idea”.

TABLE V. DISTRIBUTION OF THE ANSWERS

| Yes | Rest |
|--------|--------|
| 67 | 6 |
| 0.9178 | 0.0822 |

p = Proportion of people who think Mobile Apps would help you to make your bank processes comfortable

$$H_0: p = 67/73 = 0.9178$$

$$H_a: p \neq 6/73 = 0.0822$$

It can be assumed that bootstrap distribution is approximately normal distribution. From that it can be stated that we are 95% confident that proportion of people who think Mobile Apps would help you to make your bank processes comfortable is between 0.9171 and 0.9181. (Null hypothesis can not be rejected)

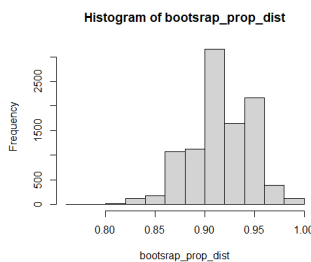


Fig. 7. Bootstrap Distribution

- 4) *Most of the people use smart phone*

For this question, the answers are “Yes” and “No”.

TABLE VI. DISTRIBUTION OF THE ANSWERS

| Yes | No |
|------|------|
| 70 | 3 |
| 0.96 | 0.04 |

p = Proportion of people who use smart phone

$$H_0: p = 70/73 = 0.96$$

$$H_a: p \neq 3/73 = 0.04$$

Since bootstrap distribution is skewed and not a normal distribution, it is erroneous to perform hypothesis tests. But we can argue that it is almost everyone has a smart phone since the people whom that the survey is filled should have a mobile application because form is a google form and probably they have heard of mobile banking applications.

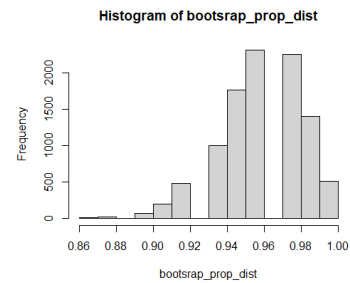


Fig. 8. Bootstrap Distribution

- 5) *Most of the people could install a mobile banking app and use it for your banking processes*

For this question, the answers are “Yes”, “No” and “No idea”.

TABLE VII. DISTRIBUTION OF THE ANSWERS

| Yes | Rest |
|--------|--------|
| 67 | 6 |
| 0.9178 | 0.0822 |

p = Proportion of people who could install a mobile banking app and use it for your banking processes

$$H_0: p = 67/73 = 0.9178$$

$$H_a: p \neq 6/73 = 0.0822$$

It can be assumed that bootstrap distribution is approximately normal distribution. From that it can be stated that we are 95% confident that proportion of people who could install a mobile banking app and use it for your banking processes is between 0.9173 and 0.9184. (Null hypothesis can not be rejected)

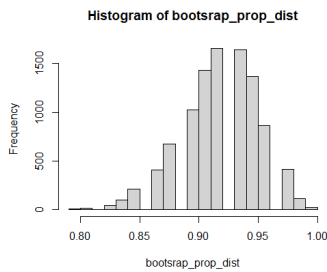


Fig. 9. Bootstrap Distribution

- 6) *Most of the males have heard of mobile banking applications.*

TABLE VIII. PROPORTION VALUES

| | Male | Female | Total |
|-------|------|--------|-------|
| Yes | 31 | 8 | 39 |
| No | 16 | 18 | 34 |
| Total | 47 | 26 | 73 |

p_1 = Proportion of males who have heard of mobile banking applications

p_2 = Proportion of males who have heard of mobile banking applications

$$H_0 : p_1 = p_2$$

$$H_a : p_1 > p_2$$

Conditions of Test for a Difference Proportion : $n_1 * p_1 = 31 > 10$, $n_1 * (1 - p_1) = 8 < 10$, $n_2 * p_2 = 16 > 10$, $n_2 * (1 - p_2) = 18 > 10$

Since one condition is not full filled, this hypothesis test cannot be continued.

After single proportion tests are carried out for each question, differences in proportions tests are continued. The differences tests cannot be carried out for the questions except the question, "Have you earlier used banking mobile apps before?", since all the other questions' answers are skewed to one particular answer. Further, only gender and occupation demographics have only significant distribution. So that for only those demographics, the differences in proportions tests are continued.

- 7) *Majority of private sector employees have earlier used banking mobile apps.*

TABLE IX. PROPORTION VALUES

| | Private sector | Rest | Total |
|-------|----------------|------|-------|
| Yes | 22 | 17 | 39 |
| No | 25 | 9 | 34 |
| Total | 47 | 26 | 73 |

p_1 = Proportion of private sector employees who have earlier used banking mobile apps

p_2 = Proportion of rest of the employees who have earlier used banking mobile apps

$$H_0 : p_1 = p_2$$

$$H_a : p_1 > p_2$$

Conditions of Test for a Difference Proportion : $n_1 * p_1 = 22 > 10$, $n_1 * (1 - p_1) = 17 > 10$, $n_2 * p_2 = 25 > 10$, $n_2 * (1 - p_2) = 9 < 10$

Since one condition is not full filled, this hypothesis test cannot be continued.

Same as earlier, for the question, "Have you earlier used banking mobile apps before?", the association tests are carried out with gender and occupation demographics.

- 8) *Association between earlier usage of banking mobile apps and gender.*

TABLE X. PROPORTION VALUES

| | Male | Female | Total |
|-------|-----------|-----------|-------|
| Yes | 31(25.11) | 8(13.89) | 39 |
| No | 16(21.89) | 18(12.11) | 34 |
| Total | 47 | 26 | 73 |

H_0 : Gender is not associated with earlier usage of banking mobile apps

H_a : Gender is associated with earlier usage of banking mobile apps

$$\chi^2 = 8.33$$

$$p\text{-value} = 0.004 \Rightarrow \text{Reject } H_0$$

It is evident that gender is associated with earlier usage of banking mobile apps.

- 9) *Association between earlier usage of banking mobile apps and education level.*

TABLE XI. PROPORTION VALUES

| | Degree | Diploma | A/L | Total |
|-------|-----------|---------|---------|-------|
| Yes | 34(27.25) | 3(5.88) | 2(5.88) | 39 |
| No | 17(23.75) | 8(5.12) | 9(5.12) | 34 |
| Total | 51 | 11 | 11 | 73 |

H_0 : Education level is not associated with earlier usage of banking mobile apps

H_a : Education level is associated with earlier usage of banking mobile apps

$$\chi^2 = 12.10828$$

$$p\text{-value} = 0.0023 \Rightarrow \text{Reject } H_0$$

It is evident that education level is associated with earlier usage of banking mobile apps.

IV. Conclusion

A statistical analysis has been carried out to understand why people are not using mobile banking applications despite of its huge advantages. The dataset is collected from responders and it seems that almost all the questions have received skewed answers. Since that many hypothesis analysis could not be carried out. This shows the importance of collecting data precisely with correct sampling to evaluate parameters for population

Several statistical hypothesis tests are carried out to derive final conclusions. They are

- Most of the people have heard of mobile banking applications
- Most of the people think mobile applications would help you to make your bank processes comfortable
- Most of the people could install a mobile banking app and use it for your banking processes

- Association between earlier usage of banking mobile apps and gender exists
- Association between earlier usage of banking mobile apps and education level exists

These conclusions can be reasoned for identification of the facts that cause to discourage the people to use mobile banking applications.

ACKNOWLEDGMENT

This research is completed on behalf of the semester 3 module, CS5651 - Statistical Inference, of MSc in Computer Science specialized in Data Science Course of batch 2020. I would like to thank Dr. R.T. Uthayasankar for the guidance to complete the research and the lectures on Statistical Inference.

REFERENCES

- [1] <https://stattrek.com/hypothesis-test/hypothesis-testing.aspx>