Credit Scoring

2019 - Assignment 2

This assignment is to be submitted by 17:00, on Wednesday March 27.

It can be submitted either in the box outside the M.Sc. Hub in the JCMB or on Learn.

If submitting on Learn, please submit your assignment as a pdf file.

In all cases, please ensure that your name appears on it.

This is an individual assignment.

Students are expected to work on this assignment independently, and for the work submitted to be their own. Further, while this is not a test of the English language, students are expected to express ideas in their own words; and not simply re-use phrases from the lecture notes.

Answer the questions using full sentences and do not just write the calculations and the final result. The style in which you present your answers is part of the marking.

There are 10 questions. You should attempt all questions.

This assignment contributes 65% of the grade for the course.

**Please include appropriate print-out’s of your calculations, manipulations, data modelling, etc.**

**Please also ensure that your name appears at the top of your submission.**

|  |  |  |
| --- | --- | --- |
| **Question** | **Score** | **Max Score** |
| **1** |  | **4** |
| **2** |  | **5** |
| **3** |  | **5** |
| **4** |  | **12** |
| **5** |  | **8** |
| **6** |  | **4** |
| **7** |  | **5** |
| **8** |  | **2** |
| **9** |  | **7** |
| **10** |  | **13** |

**Question 1** (4 marks)

In a), scorecard development and b), scorecard monitoring and tracking, we use the Kolmogorov-Smirnoff (K-S) statistic to assess the differences between two distributions.

Is there a difference in how we calculate the K-S statistics for these two uses? If so, what is the difference?

Is there a difference in the way we interpret the K-S statistics for these two uses? If so, what is the difference?

**Question 2** (5 marks)

Some organisations have many scorecards in use while others have very few.

King’s Buildings Bank, a small local lender, uses the same behavioural scorecard for their personal loan and their credit card businesses.

Explain one major disadvantage of this?

It should be clear that, with only one scorecard, the bank’s analysis and reporting will take less time to compile and review.

Explain two other possible advantages.

**Question 3** (5 marks)

In monitoring a scorecard, we have the Population Drift graph above, plotted for the Irish data over the 12 months of applications in the development data, applying one of the models that were considered.

Write a **brief** note on your findings from this display: what changes can you identify, what might be causing these changes, and what these changes might mean for the business?

**Question 4**  (12 marks)

We have already applied a four-characteristic model to the Irish data set. Here, we are going to add a fifth characteristic. This is as follows:

Current Account - Months of Unauthorised Debit Interest L6M

|  |  |
| --- | --- |
| Attribute Grouping | Score |
| 0 | 4 |
| 1 | -9 |
| 2 | -14 |
| 3 | -19 |
| 4 | -26 |
| 5 | -32 |
| 6 | -37 |

One way to do this is as follows:

*In the Irish data set Applying the Model Tab, insert a new column between Column AM and AN. Give it an appropriate label. Write the Excel command to apply the new characteristic, perhaps using the syntax in columns AJ, AK, or AM as an example, or set up a new Lookup table, which is how we calculated the scores in Column AL. Change the commands in the old Column AN to include this new characteristic’s scores.*

*The old scores, i.e. the scores before adding the new characteristic, for the first five records are 142, 164, 122, 180, and 94. Do not change the order of the records.*

Derive the new total scores for all of the data.

Include the first few rows of your spreadsheet where you have applied the model.

Applying this new model to the data, calculate the Kolmogorov-Smirnov statistic and draw the K-S graph.

Also, include your K-S graph with appropriate labels on the axes.

Using the same data and the new model, calculate the Gini coefficient.

What are your conclusions about whether this new characteristic should be included in the model?

Your results are very likely to be biased. What do we need to do to eliminate that bias?

**Question 5** (8 marks)

The methodology and principles of credit scoring have been applied and used in many areas beyond lending money.

In the UK (and many other countries), people in prison may be released for a short period, often 48 hours. This is often where they ask for a temporary release for specific compassionate reasons, for example to attend a family funeral. However, it could also be that the request is to help the prisoner with the rehabilitation process.

Consider the possibility of credit scoring being used in the prison system to decide who is to be allowed out for this 48-hour period.

We need to consider how this might work.

Specifically, consider this as a scorecard development and propose suitable definitions for:

The applicant population

Good’s, Bad’s, and Rejects

Also, consider the situation and suggest eight plausible (and quite different) characteristics for the scorecard development phase on which we would try to capture data.

**Question 6** (4 marks)

We have a simple analysis using a transition matrix which applies to the changes that can happen to a card account at the end of each month in a credit card business.

This simple analysis uses only six states. Accounts are either Closed Good, Inactive, Active and 0 payments delinquent, Active and 1-2 payments delinquent, Active and 3-4 payments delinquent, or Closed Bad.

The two states of Closed Good and Closed Bad are absorbing states.

We would like to use this matrix to represent the transitions at the end of each month. However, it is recognised that it is unlikely that the process of the movements of these accounts is a stationary process. Briefly, provide (at least) four reasons why the Stationarity property is unlikely to hold.

**Question 7** (5 marks)

The above graph presents some alignment analysis. For two attributes, we have plotted the performance against score.

Can you tell if this characteristic is in the scorecard?

Why should these two curves be aligned?

The scorecard cut-off has been 260. One option that might help to resolve this situation would be to add 10 points to the score for an application coming from an applicant who has had a Previous Loan. This would align the two curves. State one consequence of this that we should consider.

**Question 8** (2 marks)

In developing a scorecard for a Payday loan product, the developers are not usually concerned about the length of the performance window for the development sample. Why is this?

**Question 9** (7 marks)

We are analysing a scorecard being used on the credit product from which we got the Irish data.

This scorecard has been developed using the usual PDO scale of 20.

We are looking at the scorecard characteristic Loan Term Requested.

This has been classed into four attributes: 3-12 months, 13-48 months, 49-60 months, 61-180 months.

The current scorecard weights for this characteristic are in the table below.

Analysis has been carried out on a sample of 8000 accounts, opened since the scorecard was developed and implemented.

We have the actual number of Goods and Bads in this sample.

We also have the expected number of Goods and Bads (using the data from the scorecard development).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Loan term (Months) | 3-12 | 13-48 | 49-60 | 61-180 |
| Score | 51 | 27 | 18 | 0 |
| Actual Goods | 936 | 4093 | 1926 | 396 |
| Actual Bads | 40 | 323 | 234 | 52 |
| Expected Goods | 942.6 | 4088.6 | 1952.6 | 380.1 |
| Expected Bads | 33.4 | 327.4 | 207.4 | 67.9 |

Calculate the changes we could make to the scores for the attributes of this characteristic so that it is aligned better with the actual performance, and comment on the proposed changes.

**Question 10** (13 marks)

This question has 13 multiple choice questions. For each part, you are given five possible answers, labelled A, B, C, D, and E, only one of which is correct.

There is no penalty for selecting an incorrect answer. A correct answer gets one mark.

For each question, make your choice and record it in the appropriate box at the end of the question.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Qu.** |  | **A** | **B** | **C** | **D** | **E** |
| 11A | Which of these might be a suitable scenario to use in stress testing for Basel capital adequacy? | A change in leader or government in a major country such as the USA, Russia, or China. | An increase in Eurozone inflation from 1.6% to 8%. | Heavy rain causing one quarter of Europe to be under water. | New York Stock Exchange to lose 80% of its value. | The arrival of creatures from another planet. |
| 11B | PIH is the? | People’s Interdependent Harmony | Permanent Income Hypothesis | Personal Independence at Home | Personal Insurance for Health | Platform for Independence under Hosmer-Lemeshow |
| 11C | Which one of these is not a commonly-used feature of Big Data? | Value | Velocity | Veracity | Volatility | Volume |
| 11D | The three parties involved in P2P lending are usually known as the … | Depositor – Host -Borrower | Depositor – Lender – Borrower | Investor – Host – Borrower | Investor – Platform - Debtor | Saver – Lender – Borrower |
| 11E | Which of these are possible reasons for a bank to be removed from a stress testing review by the central bank? | The bank has been bought by a bank from another country | The bank has moved the location of its Head Office | The bank has stopped granting new credit | The bank is no longer a systemic bank | The bank reported a loss for the most recent financial year |
| 11F | Discussion of Quality of Earnings relates to | The consistency of a bank’s profits | The penalty fees paid by a bank’s customers for making late payments | The proportion of a bank’s accounts that are in the higher score bands | The proportion of bank notes that are new | The quality of jobs that a bank’s customers have |
| 11G | Refer to Table 10.1 below  For a proposed scorecard cut-off of 181, what is the expected reject rate? | 25.00% | 25.15% | 25.30% | 29.95% | 30.43% |
| **Qu.** |  | **A** | **B** | **C** | **D** | **E** |
| 11H | In a portfolio of £14,000 personal loans, the PD is 2.4%, the LGD is 57%, and we expect the balance to have reduced by 35% by the time the case goes into default. What is the Expected Loss for an account in this segment? | £50.57 | £67.03 | £93.91 | £124.49 | £191.52 |
| 11I | Reject Inference is a process | where we ask our applicants who rejected their applications before they applied to us | where we estimate how rejected applications would have performed if we had accepted them | where we estimate how the current applicants will perform if we reject them and they are then accepted by another lender | where we estimate which customers will reject our offer if we accept them but only at a higher price | where we work out which rejected applications would also have been rejected by our competitors |
| 11J | In an application process where risk-based pricing is used, there is an extra step to | Calculate the expected return | Confirm the bureau information | Consider the value of the security | Discover if the applicant will borrow at the price offered | Discover what other prices the applicant has been offered |
| 11K | Again referring to Table 10.1 below, the organisation is also considering a cut-off of 179. If the average profit from a Good account is £625 and the average loss from a Bad account is £3100, calculate the additional number of Goods and Bads if they adopt 179 rather than 181. What is the additional profit? | £16,475 | £27,025 | £27,050 | £28,350 | £41,525 |
| 11L | Referring to Table 10.2, calculate the effect on the average application score from the individual attributes and the score overall. | The effect from Holiday is negative and the effect from Education is positive | The effect from Holiday is negative and the effect from the characteristic overall is negative | The effect from Holiday is negative and the effect from the characteristic overall is negligible | The effect from Holiday is positive and the effect from Car is positive | The effect from Holiday is positive and the overall effect from the characteristic is negligible |
| 11M | Referring to Table 10.3, this product has a scorecard cut-off of 220. The overall accept rate and HSO rate are respectively | 60.8% and 11.7% | 60.8% and 14.2% | 72.4% and 6.0% | 72.4% and 11.7% | 85.8% and 14.2% |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Question | 11A | 11B | 11C | 11D | 11E | 11F | 11G | 11H | 11I | 11J | 11K | 11L | 11M |
| Answer |  |  |  |  |  |  |  |  |  |  |  |  |  |



**Table 10.1**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Expected | Current Sample | Score Weight |
| Home Improvement | 38.60% | 42.90% | 19 |
| Car | 33.40% | 29.40% | 7 |
| **Holiday** | **21.90%** | **19.20%** | **-2** |
| Education | 6.10% | 8.50% | 5 |

**Table 10.2**

|  |  |  |  |
| --- | --- | --- | --- |
| **Score** | **Applications** | **Accepts** | **Rejects** |
| **Up to 209** | **901** | **40** | **861** |
| **210-219** | **762** | **59** | **703** |
| **220-229** | **833** | **603** | **230** |
| **230-239** | **611** | **488** | **123** |
| **240-249** | **474** | **430** | **44** |
| **250+** | **1739** | **1615** | **124** |

**Table 10.3**