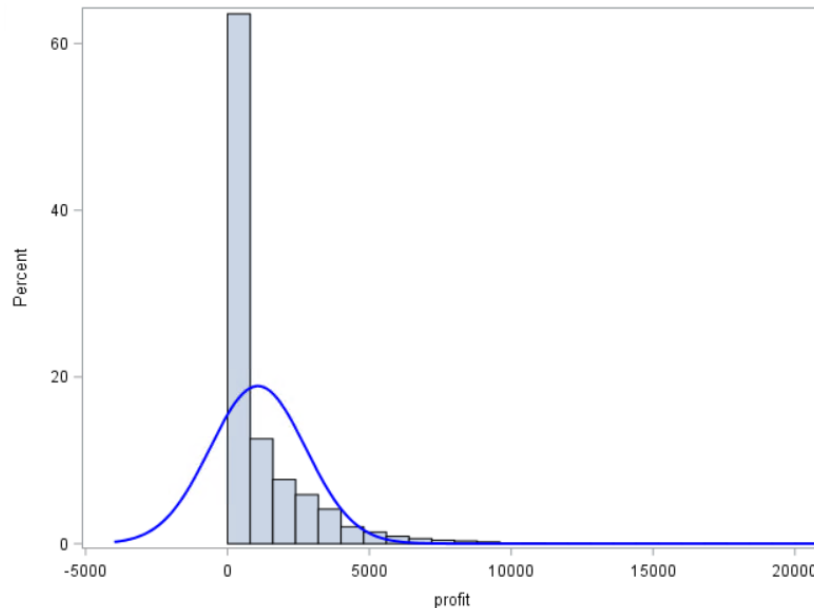


### Assignment-6 (Group 11)

1. As per our profit tobit model, we initially set the lower bound as 0 and the upper bound as 5000 based on the below histogram.



A very low value for the log likelihood and a very high value for the AIC as well as BIC was observed when the upper bound censored value for profit was set at 5000. This shows that there is further scope for improvement in the model and, the model is not doing a very good job in explaining the profit to be earned.

Looking into each variable:

- Age has a negative effect on the profit, and it is very significant. For every value increase in the age, the profit decreases by \$19 approximately.
- Rewards has a negative effect on the profit and is significant, wherein, when a customer has a reward, the profit decreases by \$300 approximately against a person who does not have a reward and this does not make much sense as a customer who has a reward is more likely to spend more and thus a higher profit can be earned.
- The number of cards a customer has, causes the profit to increase by a value of \$28 approximately. But this variable is not significant.
- For the different modes of acquisition, When the customer is acquired through direct selling, the profit decreases by \$838 as compared to that of direct mail. The profit decreases by a value of \$817 when customer is acquired through telephone selling as compared to direct mail and by \$154 when acquired through internet as compared to direct mail, but the net variable is insignificant and hence the overall effect cannot be gauged.
- Based on the type of card, all card types are significant and have an overall negative effect on the profit.
- Based on the type of affinity cards, except for SectorC, SectorE and SectorF, every other variable is insignificant and all of them have negative effects on the profit.

2. The selection model focusses on the case where the model is checked for the case of the active customers and then if the customers are active, how much profit can be estimated for each case.

The log of the profit has been taken and that is the dependent variable in this case. In comparison to the first model, the AIC and BIC is much lower, and the log-likelihood is much higher.

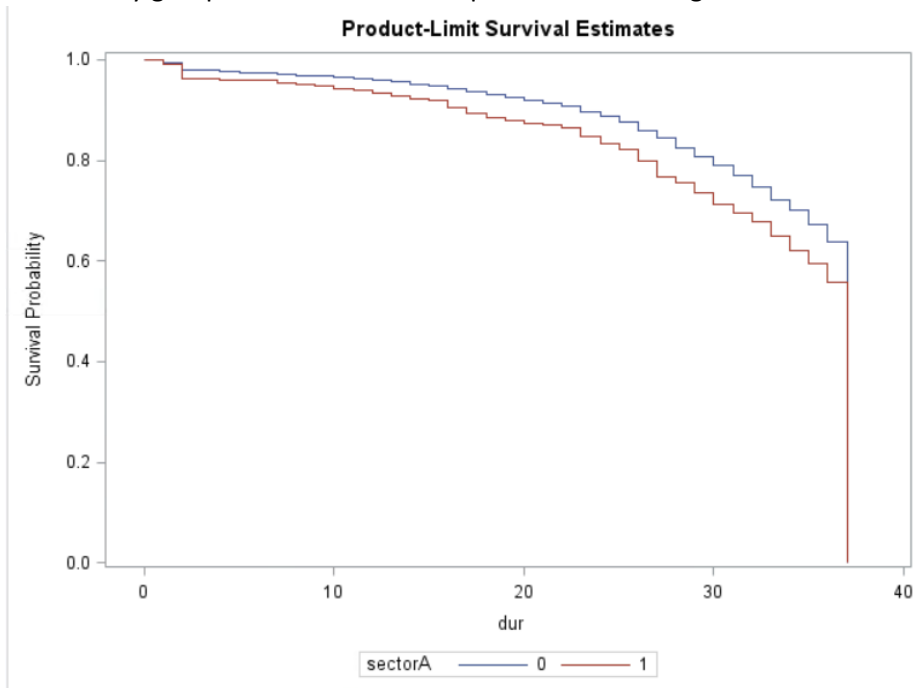
Considering the effect of each variable:

- The age variable is highly significant and for each value increase in the age, the overall profit decreases by 1.4%.
- The profit decreases by 21% if the person has a reward as compared to person who does not have one.
- The number of cards a person is not significant in determining the profit.
- Each mode of acquisition is insignificant in determining the profit.
- The different types of card a customer holds is significant in estimating the profit and each of the variables have a highly negative effect in the profit as a whole.
- With the exception of SectorB, each of the types of affinity cards are significant in determining the profit. Each case, the values are highly negative.

The model as a whole, is much better than the previous tobit model as the overall AIC and BIC has decreased and the log likelihood has increased. Hence the selection model, where only the effect of the active customers is considered. The Rho value in this case is insignificant and hence a logit/linear regression model should provide a better insight.

3. After initially deleting all the customers who are inactive, the proportional hazards model was run with the duration being censored to 37 months. We find the AIC as well as the log likelihood values to be large. Looking into each of the independent variables, the total transactions, the limit, the standard card as well as the SectorA are significant in determining the duration for the hazard model. The total transaction has a hazard ratio of 1, which means there is no effect. A similar case is observed for the limit variable as well. For the case of the standard card type, the hazard ratio is 0.8, which means that the overall hazard for a customer with a standard credit card decreases by 18% as compared to a person without one. SectorA is a person who does not have an affinity card and the hazard rate in this case is 1.2 which implies that for a person without an affinity card, the overall hazard increases by 20% as compared to a person with one.
4. Plotting the parametric hazard model for the duration, we obtain an AIC value of 961 and a BIC of 1078, which shows us that the model is much better than the previous tobit as well as selection models. The log-likelihood has also increased. Analyzing the important effects, we find most variables to be highly insignificant with the exception of the standard card type and the SectorA(customer does have an affinity card). For a customer with a standard card type, the expected ratio is  $\exp(0.0329) = 1.03$ , which means that the survival of a customer is 1.03 times higher than a customer who does not have a standard card type. The survival ratio for SectorA is  $\exp(-0.0259) = 0.97$  which means that a person who has an affinity card, the chance of survival is 0.97 times that of a person that does not have one. So overall, it is better to have some kind of an affinity or have a standard card type to have a better survival for the company as a customer.

5. As per the results, the survivor function for affinity groups are significantly different from that of non-affinity groups based on the Chi-Square test of the log rank.



A customer who does not have an affinity card has a lower chance of survival against a customer with an affinity card based on the above graph.

It can be concluded that the survivor function for affinity groups are significantly different from that of non-affinity groups and the since the customers that do have an affinity card have a lower chance of survival, that is, they would stop using the card, having an affinity card therefore gives the company a better chance a retaining a customer than not having one.