**Statistics 747 – Project Part 2**

In this part of the research process, imagine we have presented a design to 200 beer drinkers and recorded their responses to 36 different scenarios. The data collected can be found on Cecil called ‘Heineken2010.csv’ (note: this data probably may have different design for what you presented in Part 1 of this project – however, for the sale of consistency we will use this design for all subsequent work).

**Your Tasks**

In order to present this project, you will need to:

1. Create an Excel **I**nteraction **D**ecision **T**ool (IDT) to present the results dynamically (i.e. as price changes for each brand).

An example of such an IDT, from a different study, can be investigated in (via Cecil) [‘Choice.xls](http://www.stat.auckland.ac.nz/~reilly/Choice.xls)’

2) Present, in person, a PowerPoint presentation along with the IDT in the week commencing 19th October 2010 with no more than 12 slides. The presentation should take no longer than 20 minutes.

 This presentation will have:

1. An introduction stating the research objectives and methodology employed
2. What you discovered using this IDT along with other analyses
3. Your overall conclusions and how they relate to the initial objectives outlined in Part 1 of this project for DB Breweries.

Again the emphasis will be on clarity and meeting the overall marketing objectives for DB breweries.

Hand in a copy of your slides, along with your program code and relevant output, by **Friday 22nd October 2010.** The program code should show what analyses were done and why; include enough comments to make this clear. Also email a copy of your IDT to Andrew Balemi. (a.balemi@auckland.ac.nz).

**Notes for Analysis**

The file Heineken2010.csv presents the 36 different scenarios with the price for each product offered and the frequency of purchase for each brand for all 200 people, across rows.

In order to produce this you will need to use the R function SAS.DCM.FORMAT and create the relevant data/design matrix data for SAS’s **PROC PHREG** via the R commands and code found in the file ‘Proj2.R’:

**HeinekenData2010<-SAS.DCM.FORMAT (Heineken2010)**

**Note the first 18 scenarios are for the Heineken XLN format – the remaining 18 for the current format.** You’ll need to make a dummy variable in SAS to reflect this (in the code statements below.

The resulting data should look something like the data in the file called

‘HeinekenData2010 Abridged.csv’ before it is imported into a SAS data set and analysed in SAS (note that I have only made Scenarios 1 to 6 available for your inspection).

Note: make sure that imported SAS data set has correct column labels.

Some useful SAS code for getting estimates:

\*\* create XLN dummy variable;

**data** temp;

 set hold.Heineken2010Data;

 xln=**0**;

 if set<=**18** then xln=hn;

 PR\_hn2 = PR\_hn\*\***2**;

 PR\_sa2 = PR\_sa\*\***2**;

 PR\_sn2 = PR\_sn\*\***2**;

 PR\_xg2 = PR\_xg\*\***2**;

**run**;

**proc** **phreg** data=temp;

model t\*t(**2**)=

 hn xln sa sn xg op ms

 PR\_hn PR\_sa PR\_sn PR\_xg

 PR\_hn2 PR\_sa2 PR\_sn2 PR\_xg2

 hn\_sa hn\_sn hn\_xg hn\_op hn\_ms

 sa\_hn sa\_sn sa\_xg sa\_op sa\_ms

 sn\_hn sn\_sa sn\_xg sn\_op sn\_ms

 xg\_hn xg\_sa xg\_sn xg\_op xg\_ms

 op\_hn op\_sa op\_sn op\_xg op\_ms

 ms\_hn ms\_sa ms\_sn ms\_xg ms\_op/selection=forward sle=**.3** include=**11** ties=breslow;

 strata set;

 freq freq;

**run**;

**proc** **phreg** data=temp;

model t\*t(**2**)=

 hn xln sa sn xg op ms

 PR\_hn PR\_sa PR\_sn PR\_xg

 PR\_hn2 PR\_sa2 PR\_sn2 PR\_xg2

 hn\_sa hn\_sn hn\_xg hn\_op hn\_ms

 sa\_hn sa\_sn sa\_xg sa\_op sa\_ms

 sn\_hn sn\_sa sn\_xg sn\_op sn\_ms

 xg\_hn xg\_sa xg\_sn xg\_op xg\_ms

 op\_hn op\_sa op\_sn op\_xg op\_ms

 ms\_hn ms\_sa ms\_sn ms\_xg ms\_op /selection=backward sls=**.3** include=**11** ties=breslow;

 strata set;

 freq freq;

 **run**;