This document contains links to Internet resources, so view it rather than print it.

Future conferences

CSCC XIII *28 – 30/8/2013*

The premier, international credit scoring conference, Credit Scoring and Credit Control, will be held from 28th to 30th August 2013, in Edinburgh, Scotland. Brochures and registration information are available from http://www.business-school.ed.ac.uk/crc/conferences.

If you wish to present a paper at the conference you must submit a 200 to 400 word abstract via the conference website by 15 April. Past papers and presentations are available at http://www.business-school.ed.ac.uk/crc/conferences/conference-archive.

In early 2011 Veda ran a competition for a trip to the CSCC conference, which was won by Jason Humphrey (Head of Retail Risk, ANZ New Zealand). Here are his impressions of the conference:

> The conference was made up of four streams, modelling, lending, methodology and general topics. As advice to future attendees, you need to ensure pre-reading of papers for the sessions you are attending to ensure you get the most out of the session. As an example, the modelling sessions contain a lot of calculus and algebraic equations so why attend these sessions if I'm not a model builder or maths nut?
>
> Well, from the northern hemisphere it is quite clear (to me at least) that modelling is still extremely valuable but little progressed in the southern hemisphere over the last 10 to 15 years. This, I believe, was similar in the northern hemisphere and has led to some of the problems they have encountered - with an increase in unexpected losses coming from segments that have been good performers in the past, based on benign development periods.
>
> The issue that regulators and financial institutions have uncovered is that models built over different economic cycles yield different results in other cycles. (Well, that's not news to anyone.) But what have we really done about this? The answer is, very little, because the cycles we have been through haven't yielded shifts significant enough to cause major re-thinks in our modelling. The typical answer has usually been to just rebuild a model with these new observed bad segments or re-calibrate models. The GFC (or for the northern hemisphere, FC) showed shifts in the PD curves and results that broke what we know and rely on in modelling. The conference was full of decision, capital and profitability modelling examples of what occurred and how to fix the issue - typically from the use of economic data allowing for automated calibration of risk dependent on the cycles.
>
> The Australian conference (ARCRC 2011) held at the start of this year was a good mirror of what occurs in Edinburgh. Perhaps that is why only 18 or 19 people out of the 400 who attended the Edinburgh conference came from the Australia and New Zealand region. The biggest difference between the conferences is the volume of discussion papers not available to us in Australia and NZ on actual impacts and issues currently being faced in Europe and the US. Therefore, it is important that we maintain the quality of papers and presenters that are made available to the greater risk and business

population to ensure solutions to crisis and problems can be addressed both from a practical and modelling perspective.

All in all for a once every two year conference, I highly recommend attendance to both model owners and model developers or just those of us who want to progress risk strategy beyond what we are using today.

Technical Notes

Achievable bad rates and approval rates

Alison Heron & Ross Gayler

A frequently asked question in credit scoring is: How predictive does my application scorecard need to be to get the approval rate and accepted bad rate expected by the business? There is an intrinsic relationship between Gini, approval rate, population bad rate, and approval bad rate, so that sometimes the model Gini might not be high enough to achieve the business's desired approval rate and approval bad rate on the given population.

There are three scenarios, corresponding to degrees of knowledge about the scorecard Gini, where exploring this relationship between Gini, population bad rate, approval rate and approval bad rate can be applied:

- The population bad rate is fixed and the scorecard Gini is known

  This is the typical scenario after scorecard development. We can explore the relationship between the approval rate and approval bad rate. The score distribution gives the trade-off between approval rate and approval bad rate.

- The population bad rate is fixed and 100% scorecard Gini is assumed

  This is just a special case of the preceding scenario, so we can map out the trade-off between approval rate and approval bad rate. Of course the Gini can't be 100%, but this provides an upper bound on the results that can be achieved in practice. If the desired approval rate and approval bad rate can't be achieved with 100% Gini there is absolutely no way it can be achieved with a realistic Gini. Assuming a 100% Gini makes the maths particularly easy, so we will give an example below.

- The population bad rate is fixed and the scorecard Gini is unknown

  For a given population bad rate there is a fixed relationship between Gini, approval rate and approval bad rate. Then, for the desired approval rate and approval bad rate we can find the minimum Gini that is needed. Obviously, we can query that relationship in different ways (e.g. plot the approval bad rate as a function of Gini, given a fixed approval rate). The maths for this scenario is more complicated than the preceding scenario, so we expand a little on it below but leave a detailed discussion for a future issue of CRAON.

We concentrate here on the second scenario – Gini assumed to be 100%. This is useful because it places an upper bound on the performance and the maths is very simple.

A 100% Gini means that the Goods and Bads are perfectly separated (i.e. all the Bads score below all the Goods). Only the ordering of the scores matters so let's assume that the scores are uniformly distributed. Figure 1 provides an illustrative example with different cut-off options on a perfectly separated population with a population bad rate of 20%.

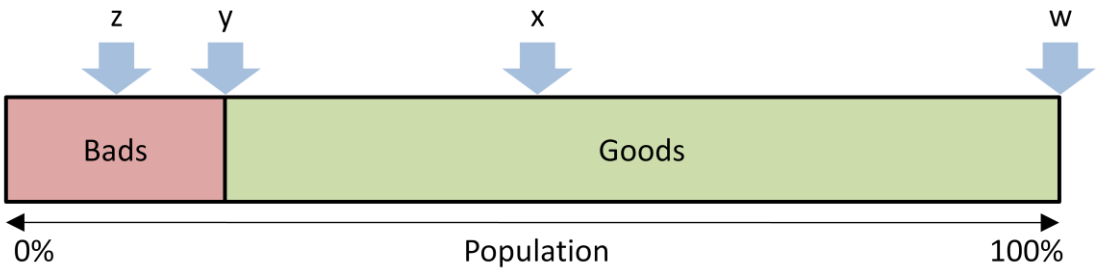

Figure 1. Example cut-offs on perfectly separated population

The figure assumes that the Goods and Bads are uniformly distributed along the horizontal axis. By considering the different cut-offs in Figure 1 we can observe the impact on approval rate and approval bad rate:

- At cut-off ***w***: 0% approval rate.

  There is no possibility of approved Bads because nobody is approved.

- At cut-off ***x***: 50% approval rate and 0% of approves are bad.

  The bad rate in the approvals must be 0% for all cut-offs with approval rates between 0% and 80% because the cut-off has not reached the highest scoring Bads.

- At cut-off ***y***: 80% approval rate and 0% of approves are bad.

  We have finally worked our way through all the Goods and all further approvals will be Bads.

- At cut-off ***z***: 90% approval rate and 11% of approves are bad.

  This is half way through the Bads. Assuming a total population of 100 we have 10 Bads (half the Bads) out of 90 approvals (10 Bads plus 80 Goods).

This can be displayed as a graph tracking the relationship between approval rate and percentage of approves that are Bads (Figure 2). For a realistic Gini the approval Bad rate must be higher at every approval rate and the approval rate must be lower at every approval bad rate. That is, with a Gini less than 100% the resulting curve can only be above the curve in Figure 2 on that given population. For example, we can see in Figure 2 a 90% approval rate with an approved bad rate of 5% is clearly not possible. In fact, at a 90% approval rate the approved bad rate must be at least 11% (and no more than 20%, corresponding to a 0% Gini). This method provides a simple way to identify the set of combinations of approval rate and approval bad rate that are absolutely not achievable on that population.

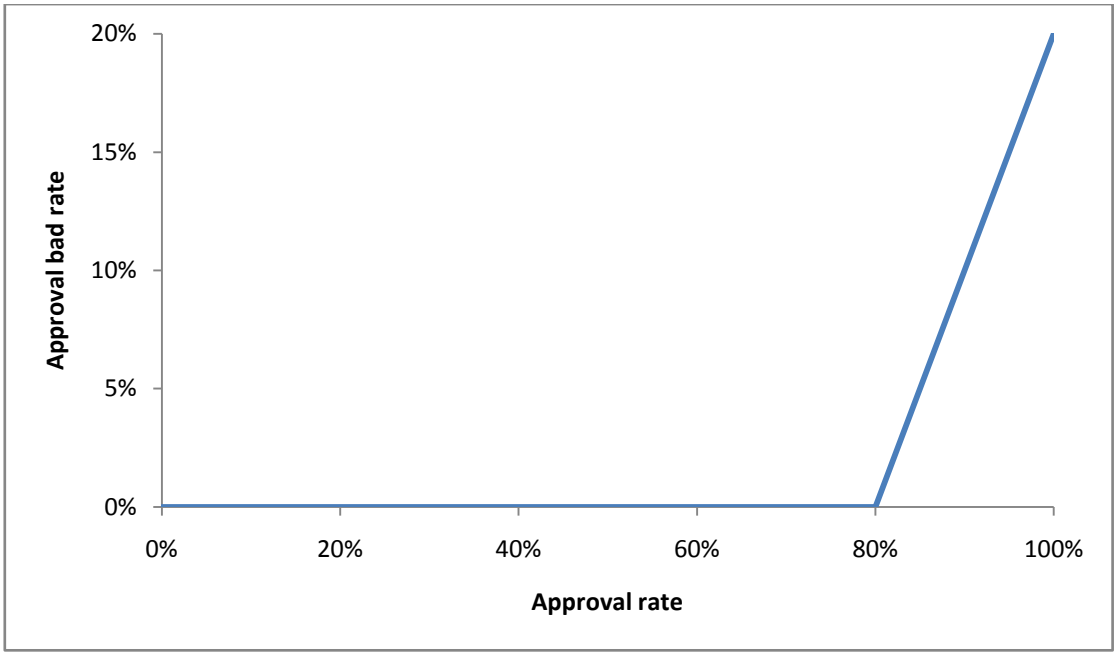

Figure 2: Example of approved bad rate as a function of approval rate

A more common question would be: If the model were to be redeveloped and an increase in Gini by ***X*** points is expected, how does that Gini increase translate into improved approval rate or reduced approval bad rate? Conversely, the question could be: Given the population with its known bad rate, what is the Gini required to obtain a particular approval rate and approval bad rate,?

Both of these questions can be estimated by understanding the relationship between approval rate, population bad rate, Gini and approval bad rate. This is the third scenario listed above, which requires a more sophisticated approach than the back-of-the-envelope method we have just illustrated. To attack this problem we need to import some assumptions from Signal Detection Theory (SDT)

The graph in Figure 2 contains information about the distributions of scores of the Goods and Bads (the shapes and locations of the distributions) as well as information about the population bad rate (the relative sizes of the populations of Goods and Bads). The Receiver Operating Characteristic (ROC) curve is a related graph from SDT. It shows what proportion of the Good and Bad populations lie below any cut-off. That is, the ROC curve captures the information about the shapes and locations of the distributions of scores for the Goods and Bads while ignoring the relative sizes of those populations (which is captured separately as the population bad rate). By making some assumptions imported from SDT it is possible to draw the ROC curve for any chosen Gini without having actually developed the scorecard. Then, given the ROC curve, the approval bad rate versus the approval rate can be calculated for any given population bad rate. We will cover this in detail in a future issue of CRAON. In the mean time, if you want more background information see this presentation on SDT in credit scoring.

It is clear that the ability to estimate the relationship between approval rate, population bad rate, Gini and approval bad rate can be a valuable tool to use prior to embarking on model

redevelopment. This can avoid unnecessary business costs if the redevelopment cost is likely to exceed the benefit. It can also be used to understand if the desired business outcomes are achievable given the population.

Credit Scoring Groups

Melbourne Risk Analytics Group

The Melbourne Risk Analytics Group was founded in 2010 to provide an opportunity for retail credit risk analysts in Melbourne to get together for ongoing education and networking. The group is intended for retail credit modelling professionals from the various financial institutions, telcos, utilities, and consultancies based in Melbourne and the talks cover areas of interest such as credit scoring, Basel II modelling, and fraud modelling.

The lunchtime meetings are held quarterly with venues rotated between the various financial institutions located in central Melbourne. Typical attendance is around 50 people (out of approximately 200 members). Membership and meetings are free and you can join at http://www.meetup.com/Melbourne-Risk-Analytics-Group/. Meeting announcements are emailed automatically from the website, so you need to be a member to get the announcements.

Software

RStudio

If you use R (a free software environment for statistical computing and graphics) you should consider adding RStudio to your toolkit. RStudio is an Integrated Development Environment (IDE) for R. It provides integrated multi-window access to all the different tools you normally need when developing and using analytical code (syntax aware editor, command interpreter, plot browser, data object browser, file browser, documentation browser, etc.). Recent developments of RStudio have concentrated on support of literate statistics. This entails making it easy to mix text and R code in one document so that the results of the analyses are inserted directly into the document and typeset. The ideal is to create documents so that analyses can be accurately reproduced or modified on demand (known as reproducible research). Here is an example. We will include some credit scoring examples in a later issue.

About this newsletter

Intended audience, content & frequency

The Credit Risk Analytics Occasional Newsletter is intended for predictive modelling analysts working in retail credit risk management, for example, credit scoring, Basel II modelling, or fraud modelling. It aims to distribute topical, publicly available information that might otherwise be difficult to find because of being dispersed over a wide range of sources. The contents include information such as conference announcements and pointers to resources like research papers, software, and discussion groups. The timing of newsletter issues is expected to be irregular, with time-critical announcements being made as needed and other issues appearing approximately quarterly.

Subscribing & unsubscribing

This free newsletter is distributed by email. To subscribe, unsubscribe, or update your email address, send an email to Ross Gayler. Please provide a secondary email or other means of contact. If the newsletter is undeliverable to your primary email address you will be removed from the mailing list if we have no other way to contact you.

Contacting us

Please email us with comments, requests, and suggestions for content.

Disclaimer

This newsletter is an informal communications channel for the Veda Credit Analytics staff to share information with their industry colleagues. This information reflects the personal opinions of individual staff and does not represent any position of Veda or constitute professional advice. Whilst we have endeavoured to ensure that the content of the newsletter is free from error, we do not warrant its accuracy, adequacy or completeness or that it is suitable for your intended use. Your use of the newsletter, especially any internet sites or software mentioned, is at your own risk. By using this information, you agree that we are not responsible for any damage or harm arising out of that use. You also agree that all warranties, conditions and representations in respect of the information (express or implied) are excluded to the fullest extent permitted by law.

Copyright & re-use

This work is licensed under the Creative Commons Attribution-Noncommercial-Share Alike 3.0 Australia License. For author attribution purposes, the author of this newsletter is “Veda Credit Analytics Staff” except where specifically stated otherwise in the text.