Lessons Not Learned at University #4

Pushing statistics too far?

The Consultant was in that fraught country Burma when it changed its name to Myanmar, but that was the least of his worries.

First, consider the task. It was to design and implement a system to process the National Forest Inventory (NFI) for the country, and to do it within four months. Then, consider the bureaucratic entanglements. It was a Tripartite project, jointly managed by the United Nations Food and Agriculture Organization (FAO), the United Nations Development Programme (UNDP), and the Forest Department of the Government of Peoples Republic of Burma.

The forest inventory was a classical systematic sample with **L**-shaped 1.05 ha plots consisting of five sub-plots, with trees felled and measured for volume on some plots. The plots were 3300 yards (approximately 3000 m) apart on a regular grid. The Chief Technical Adviser (*The CTA*) was an Austrian nearing retirement age but he knew how to run an inventory and he managed the project with Teutonic thoroughness. The Burmese Assistant Director running the field crews had an MSc from the ANU Forestry Department and he ran very efficient teams. The plot measurements were as good as one could hope to achieve in a tropical forest inventory.

By the time *The Consultant* arrived they had measured some 11,000 plots for their NFI. Burma is one of the few countries in the world still using imperial measures. *The CTA* was proudly Austrian and he, and FAO, had insisted on measuring everything in metric.

The Forest Department of Burma recognised some 500+ tree species but not all these were found on the 11,000 inventory plots. Certainly the number of different tree species felled for tree volume table development was far less. On the Bago Yoma, that central spine of forest on the hills in between the Irrawaddy and Salween rivers, some 3727 trees had been measured for volume with a Relaskop and of these some 529 trees had been felled and remeasured by sectional measurements. There were 110 different species.

For the NFI to be processed they needed a individual tree volume prediction model for each of the 500+ species. The problem was to find the best way to develop the models.

A multiple linear regression model was developed based on the 529 felled trees and this was used to revise the volume estimate for the 3727 trees based on their Relaskop measurements. The correction was quite small and there were no differences between species in terms of the correction.

The data were summarised and it was found that only 27 species had more than 30 samples measured, this being accepted as the minimum number that would give good tree volume equations. Anyone interested in why the number 30 was chosen should read Chapter 4 of Kim Iles book for an erudite discussion of the appropriate statistics.[1]

For each of these 27 species a polynomial regression in dbhob was developed and pairwise regression analysis was used to see which species were significantly different. This enabled an initial grouping of species to be determined. Of the 2805 trees of the 27 species 485 were Teak and 837 were of Pyinkado. These are the two most important tree species and so it was decided that tree volume equations would be developed for these species separately, regardless of any statistical analysis indicating possible groupings. The *Burmese Counterpart* and *The CTA* were happy with this approach.

The Consultant returned home and consulted a friendly *Practical Statistician.* He determined, just as *The Consultant* had expected, that there was a better way of grouping the species and used Principal Coordinate Analysis to analyse the 'distance' between species in a regression sense, the Behrens-Fisher analogue of Hotelling's T^2 (whatever that really is). This provided a grouping of species that made sense, was statistically robust and could be implemented as a consistent protocol.

On the second mission equations were developed for the species with 5-29 observations and analyses determined which group they best fitted. Each species was not significantly different from at least one group. Species with less than 5 observations were then analysed to determine which group

[1] Iles, Kim (2009) The Compassman, The Nun, and the Steakhouse Statistician. Kim Iles and Associates Ltd., 359pp.

they fitted best and the group regressions were then reanalysed. The final step was to look at the 400 or so species that were not represented in the data. An expert panel of senior Burmese forestry staff assigned each species to a group. Let's say the discussion was robust.

The group models were used to predict the volume for all the trees in the inventory. The technique might be considered somewhat arbitrary but it did enable the NFI to be completed.

The Visiting Biometrician from FAO then arrived. He was either in transit between countries A and B and had just happened to call in, or he wanted to visit Burma as a tourist and arranged his visit accordingly. To justify his visit he asked to review the whole statistical procedure. Shock, horror, it did not meet his statistical criteria. Teak and Pyinkado had been arbitrarily kept separate when they could have been combined with other more minor species. Regression models had been developed for species with as few as 5 trees and these had been assigned to a group. Now that was obviously not sound statistically. Even worse, species without any samples at all had been assigned to a group. No amount of discussion would shift him from his theoretical approach, his idealistic statistical base. Or do I mean bias?

Even *The CTA,* normally so pragmatic, was hard pressed trying to work out how to pacify him. Then *The Consultant*'s Burmese counterpart asked the innocuous question "then how would you do it". "Fell and measure more trees" was the immediate answer. "But it might take years to comb the forest looking for those trees, there might be only a handful in all of Burma, and they might not even be on the Bago Yoma", was *The Burmese Counterpart*'s response, "so that is simply not a viable option". This did not faze *The Visiting Biometrician*, he continued to hold the view that the approach was invalid and so the processing of the NFI was going to be invalid, and that FAO couldn't use the information. The discussion ended in a stalemate. Fortunately other staff in FAO later agreed that what had been done was completely appropriate.

The Consultant was so happy with the process that he decided to try to publish a paper, jointly authored with the *Practical Statistician* and *The Counterpart.* The three reviewers of the paper all asked for the details of the Principal Coordinate Analysis, stating that the analyses in the paper did not go far enough.

In his halcyon undergraduate days *The Consultant* had lived in a University College where Professor Sir Ronald Fisher (yes that doyen of statistics and well known for his Fisher's F-test) was residing, having been retrenched from Cambridge at the ripe young age of 75. Fisher held mulled wine suppers inviting at various stages all members of the college. Living in the room opposite this august person *The Consultant* had often been called in to make up the numbers and he had sat at the great man's feet bewailing the fact that his statistical knowledge was too inadequate to take real advantage of the great opportunity that he was experiencing. However some comments had been written down, much as lecture notes can be taken without understanding what they mean. One pseudo-quotation was about Principal Component Analysis and Principal Coordinate Analysis (PCA). Fisher had said in his delightful English accent something like "PCA, yes, well it is perfectly correct mathematically, after all I developed it myself, but what does it mean?"

Fortunately the *Practical Statistician* had worked with Fisher, and he actually understood statistics, so *The Consultant* replied to the Editor with the pseudo-quotation. It prompted the expected reply "did you actually meet Fisher?" When the answer was a terse "yes" the paper was re-reviewed and if the memory is correct the changes required to meet publication were a couple of commas and the breaking of a paragraph into two.

In retrospect, there were a number of ways that the inventory could have been improved, but there could be no argument that the data were in any way biased. They had been well measured, and it was simply a matter of analysing the data in the best way possible in the time-frame available. The approach used a judicious mixture of pragmatism, understanding of the forest management context and statistical analysis, but did not push the statistical analysis too far to the exclusion of the other aspects.

Lesson: Statistical purity is for the purists. Statistics can be pushed to the point where they lose their point. Keep in mind the real objective of the exercise.

Jerry Leech

jleech@ozemail.com.au