Paper PP20

Overview of Receiver Operating Characteristic (ROC) Curves in SAS PROC LOGISTIC

Paul Constantin, Quanticate, Hitchin, UK
Sue McKendrick, Quanticate, Edinburgh, UK

ABSTRACT

The repeated dosing of some drugs can induce injury to the human liver. There is interest in developing new biomarkers that are more specific than the standard tests [e.g. Alanine Transaminase (ALT)] commonly used.

A receiver operating characteristic (ROC) analysis is a powerful statistical analysis tool that is used to assess the ability of a test to correctly classify diseased (sensitivity) and non-diseased (specificity) subjects. The sensitivity and specificity rates are used to construct the ROC curve which is used to visually inspect the ability of the test to discriminate between patients' true status of disease. The most widely used summary statistic is the area under the ROC curve (AUROC).

We present recent enhancements (SAS® v9.2) to PROC LOGISTIC for constructing ROC curves and comparing AUROCs between biomarkers with standard errors and 95% confidence intervals (CIs). The code and output are shown using simulated data on two candidate biomarkers.

ADDITIONAL NOTES

DERIVATION OF THE ROC CURVE

Sensitivity is defined as the proportion of truly positive subjects that are correctly identified as positive. Specificity is defined as the proportion of truly negative subjects that are correctly identified as negative. The example table shows the results of classifying subjects as positive or negative for DILI based on whether their biomarker data are above or below a threshold. Here, sensitivity is 74/100 = 0.74 and the false positive rate is 12/100 = 0.12 is plotted as a single point of the ROC curve. By considering each possible threshold value, we generate a series of paired values of sensitivity and specificity from which we can construct the ROC curve. A measure of the performance of the binary test is the area under the ROC curve (AUROC) and in this example is 0.8857 for BM1. A high AUROC value (close to one) is associated with a good discriminator and a ROC curve that tends towards the top left corner of the graph.

NEW FEATURES IN PROC LOGISTIC

- PLOT = ROC option which produces a ROC curve for each of the biomarkers that are included in the model (**Figure 1**);
- ROC statement provides an AUROC estimate with standard error and CI for the biomarkers of interest (**Output 1**);
- ROCCONTRAST statement is used to calculate AUROC differences from a standard reference (in this case ALT) with standard errors, CIs and a p-value (**Output 2**).

CONCLUSIONS

ROC analysis is now much easier with new features of SAS PROC LOGISTIC introduced in SAS V9.2. There is scope for further enhancements to estimate the best threshold and partial AUROCs, and add validation tools. A word of caution: we may not know the true status because definitive clinical diagnosis requires biopsy. If clinical diagnosis relies on established biomarkers such as ALT, AUROC of ALT will be overestimated.

CONTACT INFORMATION

Paul Constantin
Quanticate
9-11 Bevan House
Bancroft Court
Hitchin,SG4 0DA
Email: paul.constantin@quanticate.com

Sue Mckendrick
Quanticate
9-10 St Andrew Square
Edinburgh, EH2 2AF
Email: sue.mckendrick@quanticate.com