The University of Manchester

MANCHESTER 1824

WHAT ARE RISK SCORES?

Presentation to methods@manchester Seminar, 25 March 2010.

IAN PLEWIS
SOCIAL STATISTICS

What are Receiver Operating Characteristic (ROC) curves?

What is longitudinal data analysis?

Using risk scores (a.k.a. statistical prediction rules) is an aid to decision making in conditions of uncertainty.

Should an intervention designed for young children and families in order to reduce subsequent educational disadvantage be targeted and, if so, at whom?

Might the problem of attrition in longitudinal studies be lessened by redirecting field work resources to those respondents more likely to drop out?

Child protection issues – how should social workers allocate their scarce resources in order to prevent neglect, abuse and murder?

Many of the ideas about risk scores originate in signal detectability/signal detection theory – voltage varying with time: Peterson et al. (1954).

The ideas have been widely applied in epidemiology, much less so in social science.

A well known example in medicine comes from the famous Framingham Heart Study; predicting coronary heart disease using risk factor categories generated from age, diabetes, smoking, blood pressure and cholesterol levels.

Examples in social science often come from criminology: predicting recidivism (Copas et al., 1996); allocating scarce probation resources (Berk et al., 2009).

How likely is that a person with a set of socio-medical characteristics will have a heart attack and what are the implications for treatment?

How likely is it that this person will commit a crime again after leaving prison?

How likely is it that someone on probation will commit murder?

Let’s assume we have a binary outcome (disease/not disease; did or did not reoffend etc.)

We also have a set of variables (‘risk’ variables) that predict the outcome.

We can estimate a function that relates the probability of a ‘good’ or ‘poor’ outcome to these risk variables, for example by using logistic or probit regression.

How accurate is this function, both in terms of predicting the ‘disease’ but also in predicting the ‘not disease’?

Where should the cut-off point be located?

This is a rather different way of thinking about the statistical analysis of longitudinal data.

Suppose we have three predictors of a binary outcome: x_1 and x_2 can have three values (1, 2, 3) and x_3 has two (0, 1).

P is the probability of a poor outcome and this probability increases as each of x_1, x_2 and x_3 increases.

Suppose the function that links them is:

$$\text{logit}(P) = -3 + 0.5x_1 + 2x_2 + 0.1x_3$$

Then the predicted probabilities of a poor outcome vary from 0.38 (x_1 =1; x_2 = 1 and x_3 = 0) to 0.99 (x_1 =3; x_2 = 3 and x_3 = 1)

Suppose we take our cut-off as a predicted probability of 0.9 – most of our cases defined this way will be have the poor outcome but there is a risk that many with the poor outcome will be missed. If we lower the cut-off to 0.7 say then we will miss fewer cases with poor outcomes but will include (i.e. ‘treat’) more cases with good outcomes.

We can construct a table ('truth table') for any cut-off point:

OUTCOME	TEST		
	Below Threshold	Above Threshold	Total
Poor	a	b	a + b
Good	c	d	c + d
Total	a + c	b + d	n

True Positive Fraction = b/(a + b). This is often called the sensitivity of the test.

False Positive Fraction = d/(c + d): 1 – FPF is often called the specificity of the test.

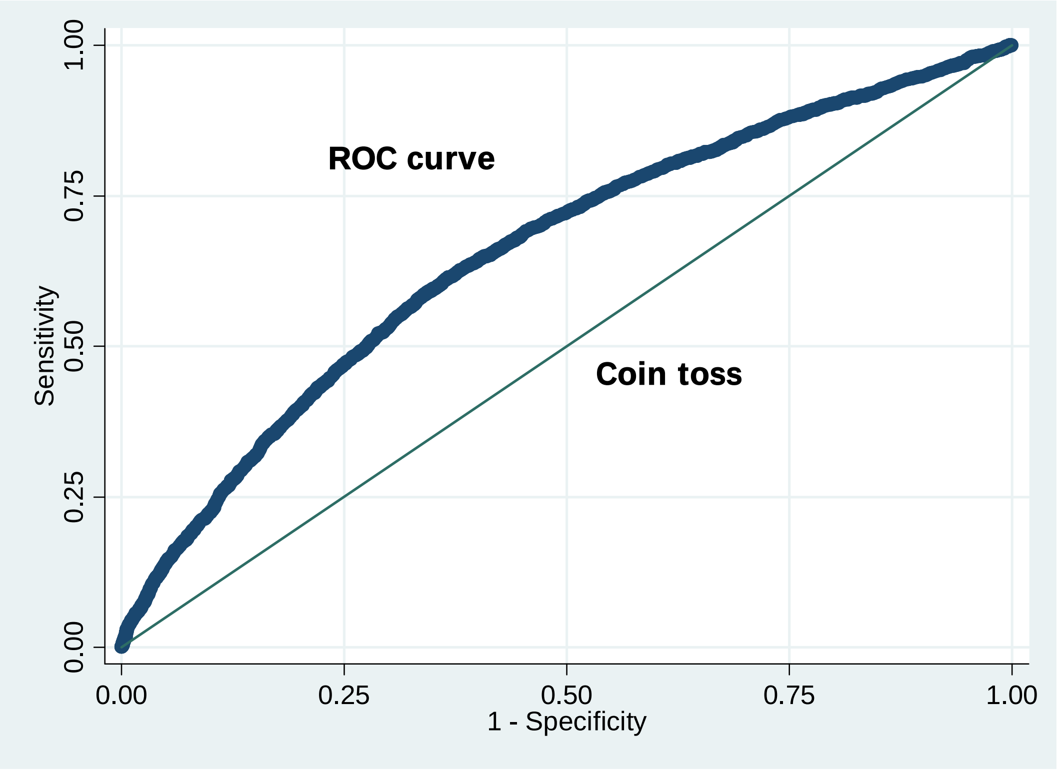
OUTCOME	TEST		
	Below Threshold 1	Above Threshold 1	Total
Poor	5	15	20
Good	45	35	80
Total	50	50	100

TPF = 15/20 = 0.75
FPF = 35/80 = 0.44

OUTCOME	TEST		
	Below Threshold 2	Above Threshold 2	Total
Poor	10	10	20
Good	70	10	80
Total	80	20	100

TPF = 10/20 = 0.5
FPF = 10/80 = 0.13

Summarising accuracy



Summarising accuracy

Summarising the ROC curve

Area under the curve (AUC): varies from 0.5 to 1

Gini coefficient: 2*AUC – 1: varies from 0 to 1

How should risk scores be used in decision making?

Where should the cut-point be for the risk score?

We need to assess the costs and benefits of decisions, for example about intervening to prevent poor outcomes.

How should risk scores be used in decision making?

The optimal cut point is determined from the slope of the ROC curve such that:

$$s^{o} = k * R$$

where:

k is the odds of a good outcome (and therefore *k* is usually greater than one),

R is the ratio of the costs of intervening when a good outcome would have occurred without intervention to the net costs of intervening to prevent a poor outcome

How should risk scores be used in decision making?

The value of *R* will depend on the efficacy of the intervention; it will be substantially less than one for effective interventions where the benefits exceed the costs but substantially greater than one for ineffective ones. Consequently, s^o will be smaller for effective interventions than it will be for ineffective ones.

A relatively ineffective intervention is targeted only at those few cases with a high probability of a poor outcome - where the slope of the ROC is high, sensitivity is low but specificity is high – whereas an effective intervention is made available to many more cases with consequent higher sensitivity but lower specificity.

References

Berk, R., Sherman, L., Barnes, G. et al. (2009). Forecasting murder within a population of probationers and parolees: A high stakes application of statistical learning. *Journal of the Royal Statistical Society, A*, 172, 191-211.

Copas, J. B., Marshall, P. and Tarling, R. (1996). *Predicting reoffending for discretionary conditional release*. Home Office Research Study 150. London: HMSO.

Krzanowski, W. J. and Hand, D. J. (2009). *ROC curves for continuous data*. Boca Raton, Fl.: Chapman and Hall/CRC.

Pepe, M. S. (2003). *The statistical evaluation of medical tests for classification and prediction*. Oxford: OUP.

Plewis, I. (2010) *Constructing and Applying Risk Scores in the Social Sciences*. CCSR Working Paper 2010-01. http://www.ccsr.ac.uk/publications/working/

Swets, J. A., Dawes, R. M. and Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, 1, 1-26.

Wilson, P. W., D'Agostino, R.B., Levy, D. et al. (1998) Prediction of coronary heart disease using risk factor categories. *Circulation,* 97, 1837–1847.