Stata users meeting, London, May 2001

Triangular plots

Nicholas J. Cox

Department of Geography, University of Durham,
Durham City, DH1 3LE, UK
n.j.cox@durham.ac.uk

Suppose we have two variables which have constant sum, such as y = proportion female and x = proportion male, so that $y + x = 1$. A plot of data for those variables shows all points lying on the line defined by that constraint. In practice, we would not draw that plot, unless by accident: we would recognise a bivariate situation as essentially univariate and examine the distribution of (e.g.) y.

More interesting is the case of three variables with constant sum, say three proportions with sum $z + y + x = 1$. This constraint defines a plane in (x, y, z) space. If also each variable lies in $[0, \textit{constant}]$, as is usually true, then points are confined to a triangular subset of the plane. Putting these constraints together, data points can be plotted on a triangular graph without loss of information: the trivariate situation is essentially bivariate.

`triplot` produces a triangular plot of three variables, which are plotted on the left, right and bottom sides of an equilateral triangle. Each should have values between 0 and some maximum value (default 1) and the sum of the three variables should be equal to that maximum (within rounding error). Most commonly, three fractions or proportions add to 1, or three percents add to 100. `triplot` may be installed now from SSC-IDEAS.

Triangular plots appear under various names in the literature, including trilinear, triaxial, three-element maps, ternary, reference triangles, percentage triangles, mixture, barycentric. Common geological applications are to sedimentary facies or particle form: hence more specific terms such as facies and form triangles.

Some options

`max()` indicates upper limit of each variable and sum of all three variables

`label()` specifies a list of numeric labels

`vertices()` specifies that only the vertices should be shown and not the complete bounding triangle. The argument is the fraction of each side that is shown

`y` specifies that the Y is to be drawn that divides the triangle into regions in which each variable is greater than the other two. Some political scientists call the spokes of the Y 'win lines'

`anti` specifies that axis scales should increase anti-clockwise

`ptsize()` controls point symbol size

`fontr()` and `fontc()` control the font used for symbols

`connect()` specifies that successive data points are to be connected. `connect(l)` specifies use of a straight line and `connect(a)` specifies use of an arrow

`ltext()`, `rtext()` and `btext()` control text on the sides of the plot; `bltext()`, `brtext()`, `ttext()` control text by the vertices of the triangle

`key()` indicates that values are to be subdivided into classes. A key will be shown if `key()` determines at least 2 and no more than 12 classes