Stata tip 67: J() now has greater replicating powers

Nicholas J. Cox
Department of Geography
Durham University
Durham City, UK
n.j.cox@durham.ac.uk

The Mata standard function J() was generalized in the Stata update of 25 February 2008. This tip flags its greater replicating powers. Note that J() in Stata's original matrix language remains as it was.

Users of Mata will have become accustomed to the role of J() in creating matrices of constants. For example, once within Mata,

```
: J(5,5,1)
[symmetric]
       1   2   3   4   5

  1    1
  2    1   1
  3    1   1   1
  4    1   1   1   1
  5    1   1   1   1   1
```

You may be used to thinking of the way J() works like this: I want a 5×5 matrix, all of whose elements are the scalar 1. Another way of thinking about it is this: Give me 5 replicates or copies rowwise and 5 copies columnwise of the scalar 1. The results are identical when scalars are being replicated.

The second way of thinking about it helps in understanding the generalization now in place. What is to be replicated can now be a matrix, naturally including not only scalars but also vectors as special cases.

The help file gives full technical details and a variety of examples, but here is another. My Speaking Stata column in this issue (Cox 2008) mentions the bias on Fisher's z scale when estimating correlation r from sample size n of $2r/(n-1)$. The question is thus how big this is for a variety of values of r and n. We can quickly get a table from Mata:

```
: r = J(5, 1, (.1, .3, .5, .7, .9))
: r
       1    2    3    4    5

  1    .1   .3   .5   .7   .9
  2    .1   .3   .5   .7   .9
  3    .1   .3   .5   .7   .9
  4    .1   .3   .5   .7   .9
  5    .1   .3   .5   .7   .9
```

```
: n = J(1, 5, (10, 20, 50, 100, 200)´)
: n
           1     2     3     4     5

    1     10    10    10    10    10
    2     20    20    20    20    20
    3     50    50    50    50    50
    4    100   100   100   100   100
    5    200   200   200   200   200

: 2 * r :/ (n :- 1)
                1              2              3              4              5

    1    .0222222222    .0666666667    .1111111111    .1555555556             .2
    2    .0105263158    .0315789474    .0526315789    .0736842105    .0947368421
    3    .0040816327     .012244898    .0204081633    .0285714286    .0367346939
    4     .002020202    .0060606061    .0101010101    .0141414141    .0181818182
    5    .0010050251    .0030150754    .0050251256    .0070351759    .0090452261
```

Notice again how the first two arguments of J() are the numbers of replicates or copies, rowwise and columnwise, and not necessarily the numbers of rows and columns in the resulting matrix.

Reference

Cox, N. J. 2008. Speaking Stata: Correlation with confidence, or Fisher's z revisited. *Stata Journal* 8: 413–439.