The Stata Journal (2009)
9, Number 1, pp. 169–170

Stata tip 74: firstonly, a new option for tab2

Roberto G. Gutierrez
StataCorp
College Station, TX
rgutierrez@stata.com

Peter A. Lachenbruch[1]
Department of Public Health
Oregon State University
Corvallis, OR
peter.lachenbruch@oregonstate.edu

In many research contexts, we have several categorical variables and correspondingly want to look at several contingency tables. The `tab2` command allows us to do this. Using the `auto` dataset (after categorizing `mpg` to `mpgcat` and `price` to `cost`), we have

```
. sysuse auto, clear
(1978 Automobile Data)
. quietly egen mpgcat = cut(mpg), at(0, 15, 25, 35, 45) label
. quietly egen cost = cut(price), at(0, 4000, 6000, 10000, 16000) label
. tab2 foreign mpgcat cost rep78, chi2
-> tabulation of foreign by mpgcat
```

Car type	mpgcat 0-	15-	25-	35-	Total
Domestic	7	37	8	0	52
Foreign	1	10	8	3	22
Total	8	47	16	3	74

```
          Pearson chi2(3) =  12.9821   Pr = 0.005
```

(*output omitted*)

```
-> tabulation of cost by rep78
```

cost	Repair Record 1978 1	2	3	4	5	Total
0-	0	1	4	2	3	10
4000-	2	5	17	8	6	38
6000-	0	1	2	8	1	12
10000-	0	1	7	0	1	9
Total	2	8	30	18	11	69

```
         Pearson chi2(12) =  18.5482   Pr = 0.100
```

With 4 variables, we have 6 tables ($4 \times 3/2 = 6$). With more variables, the number of tables explodes quadratically. Even with 10 variables, we end up with 45 tables, which is likely to be more than really interests us. We may well want finer control.

Often there is one response variable of special interest. Our first focus may then be to relate that response to possible predictors. Suppose we wish to study if domestic or foreign cars differ on some variables. Thus we are interested in the three tables of `foreign` versus `mpgcat`, `cost`, and `rep78`. The `firstonly` option added to `tab2` in the update of 15 October 2008 allows us to get just the contingency table of the first-named variable versus the others.

1. Peter Lachenbruch's work was partially supported by a grant from the Cure JM Foundation.


```
. tab2 foreign mpgcat cost rep78, firstonly chi2
-> tabulation of foreign by mpgcat
```

Car type	mpgcat 0-	15-	25-	35-	Total
Domestic	7	37	8	0	52
Foreign	1	10	8	3	22
Total	8	47	16	3	74

Pearson chi2(3) = 12.9821 Pr = 0.005

```
-> tabulation of foreign by cost
```

Car type	cost 0-	4000-	6000-	10000-	Total
Domestic	7	31	6	8	52
Foreign	4	9	7	2	22
Total	11	40	13	10	74

Pearson chi2(3) = 5.3048 Pr = 0.151

```
-> tabulation of foreign by rep78
```

Car type	Repair Record 1978 1	2	3	4	5	Total
Domestic	2	8	27	9	2	48
Foreign	0	0	3	9	9	21
Total	2	8	30	18	11	69

Pearson chi2(4) = 27.2640 Pr = 0.000

Here the number of tables is reduced from 6 to 3, a small change. However, for 10 variables (say, one response and nine predictors), the change is from 45 to 9.

This could have been programmed fairly easily with a `foreach` loop (see [P] **foreach**), but the new `firstonly` option makes life even a little easier.