Title

stata.com

xttab — Tabulate xt data

Description

`xttab`, a generalization of `tabulate` (see [R] **tabulate oneway**), performs one-way tabulations and decomposes counts into between and within components in panel data.

`xttrans`, another generalization of `tabulate` (see [R] **tabulate oneway**), reports transition probabilities (the change in one categorical variable over time).

Quick start

Overall, between, and within one-way tabulation of `v1` using `xtset` data

```
xttab v1
```

Report transition probabilities for `v2`

```
xttrans v2
```

Add frequency of transitions

```
xttrans v2, freq
```

As above, but for each level of `catvar`

```
bysort catvar: xttrans v2, freq
```

Menu

xttab

Statistics > Longitudinal/panel data > Setup and utilities > Tabulate xt data

xttrans

Statistics > Longitudinal/panel data > Setup and utilities > Report transition probabilities

Syntax

```
xttab varname [if]

xttrans varname [if] [, freq]
```

A panel variable must be specified; use `xtset`; see [XT] **xtset**.

`by` is allowed with `xttab` and `xttrans`; see [D] **by**.

Option

Main

`freq`, allowed with `xttrans` only, specifies that frequencies as well as transition probabilities be displayed.

Remarks and examples

stata.com

If you have not read [XT] **xt**, please do so.

▷ Example 1: xttab

Using the `nlswork` dataset described in [XT] **xt**, variable `msp` is 1 if a woman is married and her spouse resides with her, and 0 otherwise:

```
. use http://www.stata-press.com/data/r15/nlswork
(National Longitudinal Survey.  Young Women 14-26 years of age in 1968)
. xttab msp
```

	Overall		Between		Within
msp	Freq.	Percent	Freq.	Percent	Percent
0	11324	39.71	3113	66.08	62.69
1	17194	60.29	3643	77.33	75.75
Total	28518	100.00	6756	143.41	69.73
			(n = 4711)		

The overall part of the table summarizes results in terms of person-years. We have 11,324 person-years of data in which `msp` is 0 and 17,194 in which it is 1—in 60.3% of our data, the woman is married with her spouse present. Between repeats the breakdown, but this time in terms of women rather than person-years; 3,113 of our women ever had `msp` 0 and 3,643 ever had `msp` 1, for a grand total of 6,756 ever having either. We have in our data, however, only 4,711 women. This means that there are women who sometimes have `msp` 0 and at other times have `msp` 1.

The within percent tells us the fraction of the time a woman has the specified value of `msp`. If we take the first line, conditional on a woman ever having `msp` 0, 62.7% of her observations have `msp` 0. Similarly, conditional on a woman ever having `msp` 1, 75.8% of her observations have `msp` 1. These two numbers are a measure of the stability of the `msp` values, and, in fact, `msp` 1 is more stable among these younger women than `msp` 0, meaning that they tend to marry more than they divorce. The total within of 69.73% is the normalized between weighted average of the within percents, that is, $(3113 \times 62.69 + 3643 \times 75.75)/6756$. It is a measure of the overall stability of the `msp` variable.

A time-invariant variable will have a tabulation with within percents of 100:

```
. xttab race
```

	Overall		Between		Within
race	Freq.	Percent	Freq.	Percent	Percent
white	20180	70.72	3329	70.66	100.00
black	8051	28.22	1325	28.13	100.00
other	303	1.06	57	1.21	100.00
Total	28534	100.00	4711	100.00	100.00
			(n = 4711)		

◁

▷ Example 2: xttrans

`xttrans` shows the transition probabilities. In cross-sectional time-series data, we can estimate the probability that $x_{i,t+1} = v_2$ given that $x_{it} = v_1$ by counting transitions. For instance

```
. xttrans msp
```

1 if married, spouse present	1 if married, spouse present 0	1	Total
0	80.49	19.51	100.00
1	7.96	92.04	100.00
Total	37.11	62.89	100.00

The rows reflect the initial values, and the columns reflect the final values. Each year, some 80% of the `msp` 0 persons in the data remained `msp` 0 in the next year; the remaining 20% became `msp` 1. Although `msp` 0 had a 20% chance of becoming `msp` 1 in each year, the `msp` 1 had only an 8% chance of becoming (or returning to) `msp` 0. The `freq` option displays the frequencies that go into the calculation:

```
. xttrans msp, freq
```

1 if married, spouse present	1 if married, spouse present 0	1	Total
0	7,697	1,866	9,563
	80.49	19.51	100.00
1	1,133	13,100	14,233
	7.96	92.04	100.00
Total	8,830	14,966	23,796
	37.11	62.89	100.00

◁

❑ Technical note

The transition probabilities reported by `xttrans` are not necessarily the transition probabilities in a Markov sense. `xttrans` counts transitions from each observation to the next once the observations have been put in t order within i. It does not normalize for missing periods. `xttrans` does pay attention to missing values of the variable being tabulated, however, and does not count transitions from nonmissing to missing or from missing to nonmissing. Thus if the data are fully rectangularized, `xttrans` produces (inefficient) estimates of the Markov transition matrix. `fillin` will rectangularize datasets; see [D] **fillin**. Thus the Markov transition matrix could be estimated by typing

```
. fillin idcode year
. xttrans msp
```
(*output omitted*)

❑

Stored results

`xttab` stores the following in `r()`:

Scalars
- `r(n)` number of panels

Matrices
- `r(results)` results matrix

Also see

[XT] **xtdescribe** — Describe pattern of xt data

[XT] **xtsum** — Summarize xt data