

Esportare l'output di tabstat in Microsoft Excel e in \LaTeX : i comandi tabstatxls e tabstattex

NICOLA TOMMASI

CENTRO INTERDIPARTIMENTALE DI DOCUMENTAZIONE ECONOMICA

UNIVERSITÀ DEGLI STUDI DI VERONA



Indice

1	INTRODUZIONE	1
2	INSTALLAZIONE	1
3	SINTASSI	2
3.1	Sintassi del comando <code>tabstat</code>	2
3.2	Sintassi del comando <code>tabstattex</code>	4
3.3	Sintassi del comando <code>tabstatxls</code>	6
4	ESEMPI	9
4.1	Versione semplificata senza l'opzione <code>by</code>	9
4.2	Versione con l'opzione <code>by</code>	12
5	ESEMPI CON ALTRE OPZIONI PER <code>tabstattex</code>	15
6	ESEMPI CON ALTRE OPZIONI PER <code>tabstatxls</code>	20
7	Appendice	24
	Riferimenti bibliografici	25

1 INTRODUZIONE

Lo scopo di questi due nuovi comandi (`tabstatxls` e `tabstattex`) è l'esportazione dell'output del comando `tabstat` in Excel e in \LaTeX . Si tenga presente che `tabstat` può considerarsi anche un sostituto del comando `summarize`. Per quanto possibile ho cercato riprodurre la sintassi del comando Stata tralasciando quelle opzioni che non sono riuscito a riprogrammare o che non hanno nessuna attinenza con l'output. Alcune parti sono ancora incomplete e ci possono essere degli errori per cui faccio affidamento sul feedback degli utenti.

Questi comandi si inseriscono in un progetto più ampio che, con lo stesso scopo, comprende altri comandi Stata. Al momento sono in fase più o meno avanzata i seguenti comandi:

`arraytex` e `arrayxls`: non esiste in Stata un apposito comando ma servono per esportare le statistiche descrittive delle domande di tipo array (chiamate anche batterie di item)

`fretex` e `frexls`: esportano l'output del comando `user written fre` e che può considerarsi anche come sostituto del comando `tabulate oneway`

`tabtex` e `tabxls`: esportano l'output del comando `tabulate twoway`

`mrtabtex` e `mrtabxls`: esportano l'output del comando `user written mrtab`

`tabletex` e `tablexls`: esportano l'output del comando `table`

Lo spunto per scrivere questi comandi è nato dalla “fatica” di dover produrre report statistici per varie ricerche e quindi al momento mi sono concentrato solo su quegli output che sono funzionali alla soluzione di questo problema. Per esempio, tutte le tabelle di [questo lavoro](#) sono state generate usando i comandi descritti sopra ¹. Il layout delle tabelle è stato ispirato dal lavoro di Simon Fear “Publication quality tables in \LaTeX ” ([\[3\]](#)). Per l'esportazione in Excel invece sono stati utili gli articoli pubblicati su The STATA Blog “Export tables to Excel” ([\[1\]](#)) e “Export tabulation results to Excel - Update” ([\[2\]](#)) di Kevin Crow.

Questo documento è prodotto usando il pacchetto `texdoc`².

Il presente lavoro è organizzato come segue: nel paragrafo [2](#) si descrive come configurare \LaTeX e installare i due comandi; nel paragrafo [3](#) si descrive la sintassi del comando `tabstat` e dei comandi derivati `tabstattex` e `tabstatxls`; nel paragrafo [4](#) si presentano una serie di esempi presi direttamente dall'help del comando `tabstat`; nel paragrafo [5](#) una serie di esempi specifici per il comando `tabstattex` e nel paragrafo [6](#) una serie di esempi specifici per il comando `tabstatxls`. In Appendice un piccolo script per produrre un file Excel con i colori e i background patterns previsti da Stata.

2 INSTALLAZIONE

Per riprodurre le tabelle in \LaTeX vi mostro l'elenco dei pacchetti che devono essere presenti nel preambolo:

```
\usepackage{booktabs}
\usepackage{caption}
\usepackage{hyperref}
\usepackage[utf8]{inputenc}
\usepackage{pdfpages}
\usepackage{rotating}
\usepackage{tabularx}
\usepackage{array}
\usepackage{gensymb}
\usepackage{multirow}
\usepackage{pdfscape}
\usepackage{etoolbox}
\usepackage{color}
\usepackage{colortbl}
```

¹Indagine sulla qualità della vita degli/le studenti dell'Ateneo di Verona (2017)

²<http://repec.sowi.unibe.ch/stata/texdoc/>

e ci devono essere anche queste definizioni:

```
\renewcommand{\arraystretch}{1.1}
\newcolumntype{Z}{>{\centering\arraybackslash}X}

%% required for bold rows
\newcommand\setrow[1]{\gdef\rowmac{#1}#1\ignorespaces}
\newcolumntype{$}{>{\global\let\currentrowstyle\relax}}
\newcolumntype{^}{>{\currentrowstyle}}
\newcommand{\rowstyle}[1]{\gdef\currentrowstyle{#1}#1\ignorespaces}
```

Per installare `tabstattex` digitate nella command bar di Stata il seguente comando:

```
. net from https://raw.githubusercontent.com/NicolaTommasi8/tabstattex/master/

https://raw.githubusercontent.com/NicolaTommasi8/tabstattex/master/
(no title)

PACKAGES you could -net describe-:
    tabstattex      Comando per esportare l'output del comando tabstat in LaTeX.
```

quindi cliccate su `tabstattex` e alla pagina successiva su ([click here to install](#)).

Per `tabstatxls` digitate il seguente comando:

```
. net from https://raw.githubusercontent.com/NicolaTommasi8/tabstatxls/master/

https://raw.githubusercontent.com/NicolaTommasi8/tabstatxls/master/
(no title)

PACKAGES you could -net describe-:
    tabstatxls      Comando per esportare l'output del comando tabstat in Microsoft Excel.
```

quindi cliccate su `tabstatxls` e alla pagina successiva su ([click here to install](#)).

Gli aggiornamenti dei due comandi potranno essere installati in maniera automatica con il comando `adoupdate`, `update`. Entrambi i comandi si appoggiano sul comando `fre`³; durante l'esecuzione viene verificata la presenza del comando e nel caso mancasse, viene fatta, in automatico, l'installazione.

3 SINTASSI

3.1 Sintassi del comando `tabstat`

Questa è la sintassi del comando Stata `tabstat`⁴ con le opzioni pertinenti per `tabstattex` e `tabstatxls`:

```
tabstat varlist [ if ] [ in ] [ weight ] [, options]
```

dove *varlist* è lista delle variabili e le *options* sono:

- `by(varname)` variabile categorica per cui condizionare le statistiche delle variabili indicate in *varlist*
- `statistics(statname)` statistiche da calcolare per le variabili indicate in *varlist*
- `columns(variables|statistics)` indica come verrà visualizzata la tabella. `columns(variables)` visualizza le variabili nelle colonne della tabella (è il default), `columns(statistics)` visualizza le statistiche nelle colonne della tabella.
- `format(%fmt)` formato numerico delle statistiche, il default è `%9.0g`
- `casewise` cancella le osservazioni missing (?)

³Ben Jann. FRE: Stata module to display one-way frequency table [4]

⁴`help tabstat` dalla command bar se volete vedere l'help completo del comando.

- `nototal` non riporta il totale delle statistiche quando si usa l'opzione `by()`
- `missing` riporta le statistiche per i valori missing della variabile usata nell'opzione `by()`

le possibili statistiche da inserire in `statistics(statname)` sono:

- `mean` media
- `n` osservazioni non missing
- `sum` sommatoria
- `max` massimo
- `min` minimo
- `sd` deviazione standard
- `variance` varianza
- `cv` coefficiente di variazione ($sd/mean$)
- `semean` errore standard della media (sd/\sqrt{n})
- `skewness` (indice di simmetria o indice di Pearson)
- `kurtosis` (indice di curtosi o di appiattimento)
- `p1` primo percentile
- `p5` quinto percentile
- `p10` decimo percentile
- `p25` venticinquesimo percentile
- `median` mediana
- `p50` cinquantesimo percentile
- `p75` settantacinquesimo percentile
- `p90` novantesimo percentile
- `p95` novantacinquesimo percentile
- `p99` novantanovesimo percentile
- `range` range ($max - min$)
- `iqr` = range interquartile ($p75 - p25$)
- `q` equivalente a `p25` `p50` `p75`

3.2 Sintassi del comando tabstattex

Per adesso riprendo direttamente dall'help

```
. type `c(sysdir_personal)'tabstattex.sthlp
```

help for **tabstattex**

Esportare l'output di tabstat in LaTeX

tabstattex varlist [if] [in] [, tabstat_options] latex_options

Description

tabstattex permette di esportare in LaTeX l'output del comando **tabstat**. varlist è la lista delle variabili di cui si vogliono esportare le statistiche. Per funzionare correttamente nel preambolo del documento LaTeX devono essere specificate le seguenti direttive:

```
\usepackage{booktabs}
\usepackage{caption}
\usepackage{hyperref}
\usepackage[utf8]{inputenc}
\usepackage{pdfpages}
\usepackage{rotating}
\usepackage{tabularx}
\usepackage{array}
\usepackage{gensymb}
\usepackage{multirow}
\usepackage{pdflscape}
\usepackage{etoolbox}
\usepackage{color}
\usepackage{colortbl}
```

e ci devono essere anche queste definizioni:

```
\renewcommand{\arraystretch}{1.1}
\newcolumntype{Z}{>{\centering\arraybackslash}X}
%% required for bold rows
\newcommand\setrow[1]{\gdef\rowmac{#1}#1\ignorespaces}
\newcolumntype{$}{>{\global\let\currentrowstyle\relax}}
\newcolumntype{^}{>{\currentrowstyle}}
\newcommand{\rowstyle}[1]{\gdef\currentrowstyle{#1}#1\ignorespaces}
```

tabstat options

by(varname): specifica che le statistiche delle variabili specificate in varlist devono essere visualizzate condizionando per la variabile specificata in **by(varname)**

statistics(statname): specifica quali statistiche devono essere visualizzate. Se non si specifica nulla viene calcolata la sola media. Ogni statistica deve essere separata da uno spazio. Le possibili statistiche sono:

statname	Definizione
mean	media
count	numero di osservazioni non missing
n	uguale a count
sum	sommatoria
max	massimo
min	minimo
range	range = max - min
sd	standard deviation
variance	varianza
cv	coefficiente di variazione (sd/mean)
semean	standard error della media (sd/sqrt(n))
skewness	simmetria
kurtosis	curtosi
p1	1° percentile
p5	5° percentile
p10	10° percentile
p25	25° percentile
median	mediana (equivalente a p50)
p50	50° percentile (equivalente a median)
p75	75° percentile
p90	90° percentile
p95	95° percentile
p99	99° percentile
iqr	range interquartile = p75 - p25
q	equivale a specificare p25 p50 p75

columns(variables|statistics): specifica cosa deve essere visualizzato in colonna. **variables** visualizza le variabili di varlist (opzione di default), **statistics** visualizza le statistiche specificate nell'opzione **statistics(statname)**.

format(%fmt) specifica il formato generale di visualizzazione delle statistiche. Il formato di default è **%12.2gc**.

nototal non visualizza le statistiche generali; si usa sole se viene specificata l'opzione **by(varname)**.

missing visualizza le statistiche anche per i valori missing della variabile **by(varname)**.

latex options

- texfile(filename)**: specifica il file .tex (ed eventuale percorso) in cui salvare il codice della tabella. Questa opzione è obbligatoria.
- caption(string)**: specifica il testo da inserire nell'opzione `\caption{}` del pacchetto table di LaTeX. Di default è vuoto.
- label(string)**: specifica il testo da inserire nell'opzione `\label{}` del pacchetto table di LaTeX. Il comando prevede il prefisso `tbl:` per cui l'opzione **label(Tab1)** produce il codice LaTeX `\label{tbl:Tab1}`.
- position(string)**: specifica la posizione della tabella secondo le regole LaTeX. Il default è **position(!htp)**
- intc1(string)**: specifica il testo da inserire come descrizione della prima colonna della tabella. In assenza dell'opzione **by(varname)** nella prima colonna ci possono essere le variabili o le statistiche, dipende da cosa specificato nell'opzione **columns()** e in questo caso di default **intc1()** è vuoto. Se viene specificata l'opzione **by(varname)**, nella prima colonna ci sono i valori della variabile **varname** e di default in **intc1()** c'è la descrizione associata a **varname**.
- intc2(string)**: specifica il testo da inserire come descrizione della seconda colonna della tabella e si applica solo nel caso in cui sia specificata l'opzione **by(varname)**. Se l'opzione **columns()** è **variables** il default è **intc2(Statistiche)**, se l'opzione **columns()** è **statistics** il default è **intc2(Variabili)**.
- wintr1(number)**: specifica la larghezza della prima colonna della tabella. DA TESTARE
- note(string)**: specifica il testo da inserire come nota a piè di tabella. Di default è vuoto.
- widthtable(string)**: specifica la larghezza della tabella da esprimersi con un valore numerico ed una unità di misura (ad esempio **widthtable(5cm)**). Di default vale `\textwidth` e significa che la larghezza della tabella equivale ai margini della pagina. Le unità di misura possibili sono pt, mm, cm, in, ex ed em.
- landscape**: specifica che la tabella deve essere orientata in orizzontale rispetto alla pagina.
- fontsize(string)**: specifica la dimensione del font da usare nella tabella. I valori ammessi sono quelli di LaTeX, cioè Huge, huge, LARGE, Large, large, normalsize (default), small, footnotesize, scriptsize e tiny. Si veda la documentazione di LaTeX per maggiori informazioni.
- vardisp(varlabel|varname)**: specifica come visualizzare le variabili specificate in varlist. **vardisp(varlabel)** visualizza la descrizione associata a ciascuna variabile, **vardisp(varname)** visualizza solo il nome della variabile. **vardisp(varlabel)** è il default.
- s1(string)...s10(string)**: specifica la descrizione delle statistiche indicate nell'opzione **statistics(statname)**. L'ordine deve essere quello di statname, ovvero **s1()** indica la descrizione della prima statistica di **statistics(statname)**, **s2()** indica la descrizione della seconda statistica di **statistics(statname)** e così via. Queste sono le descrizioni di default:

statname	Descrizione
mean	Media
count	Numero di osservazioni
n	Numero di osservazioni
sum	Sommatoria
max	Massimo
min	Minimo
range	Massimo - Minimo
sd	Deviazione standard
variance	Varianza
cv	Coefficiente di variazione
semean	Errore standard della media
skewness	Simmetria
kurtosis	Curtosi
p1	1° percentile
p5	5° percentile
p10	10° percentile
p25	25° percentile
median	Mediana
p50	50° percentile
p75	75° percentile
p90	90° percentile
p95	95° percentile
p99	99° percentile
iqr	Range interquartile

- dfs1(string)...dfs10(string)**: specifica il formato numerico delle statistiche indicate nell'opzione **statistics(statname)**. L'ordine deve essere quello di statname, ovvero **dfs1()** indica il formato della prima statistica di **statistics(statname)**, **dfs2()** indica il formato della seconda statistica di **statistics(statname)** e così via. La sintassi del formato è la medesima di Stata per i formati numerici. A tal prosito si veda l'help di format.

Examples

- ```
. tabstat price weight mpg rep78, stat(n mean cv q) col(stat) texfile(ex1.tex) dfs1(%9.0f) dfs2(%9.2f)
 dfs3(%9.2f) dfs4(%9.0f) dfs5(%9.0f) dfs6(%9.0f)

. tabstat price weight mpg rep78, stat(n mean cv q) col(stat) col(v) texfile(ex1.tex) dfs1(%9.0f) dfs2(%9.2f)
 dfs3(%9.2f) dfs4(%9.0f) dfs5(%9.0f) dfs6(%9.0f)

. tabstat price weight mpg rep78, stat(n mean cv q) col(v) texfile(ex2bis.tex) ///
 s1(Nonmissing observations) s2(Mean) s3(Coefficient of variation) s4(25th percentile) s5(50th percentile)
 s6(75th percentile) ///
 dfs1(%9.0f) dfs2(%9.2f) dfs3(%9.2f) dfs4(%9.0f) dfs5(%9.0f) dfs6(%9.0f)

. tabstat price weight mpg rep78, by(foreign) stat(mean sd min max) texfile(ex3.tex) dfs1(%9.2f) dfs2(%9.2f)
```

```

 dfs3(%9.0f) dfs4(%9.0f)
. tabstat tex price weight mpg rep78, by(foreign) stat(mean sd min max) c(s) texfile(ex3bis.tex) dfs1(%9.2f)
 dfs2(%9.2f) dfs3(%9.0f) dfs4(%9.0f)

```

#### Author

Nicola Tommasi  
 nicola.tommasi@univr.it  
 nicola.tommasi@gmail.com

#### Acknowledgments

#### References

#### Also see

Online: `tabstat`, `tabstatxls` (if installed)

## 3.3 Sintassi del comando `tabstatxls`

Per adesso riprendo direttamente dall'help

```
. type `c(sysdir_personal)'tabstatxls.sthlp
```

---

help for **tabstatxls**

---

#### Esportare l'output di `tabstat` in Microsoft Excel

```
tabstatxls varlist [if] [in] [, tabstat_options] excel_options
```

#### Description

**tabstatxls** permette di esportare in LaTeX l'output del comando **tabstat**. `varlist` è la lista delle variabili di cui si vogliono esportare le statistiche. Il comando usa la classe mata `xl()` per esportare in Excel 1997/2003 i files di estensione `.xls` e in Excel 2007/2013 i files di estensione `.xlsx`. Per funzionare è richiesta la presenza del comando **fre**.

#### tabstat options

**by(varname)**: specifica che le statistiche delle variabili specificate in `varlist` devono essere visualizzate condizionando per la variabile specificata in **by(varname)**

**statistics(statname)**: specifica quali statistiche devono essere visualizzate. Se non si specifica nulla viene calcolata la sola media. Ogni statistica deve essere separata da uno spazio. Le possibili statistiche sono:

| statname        | Definizione                                      |
|-----------------|--------------------------------------------------|
| <b>mean</b>     | media                                            |
| <b>count</b>    | numero di osservazioni non missing               |
| <b>n</b>        | uguale a <b>count</b>                            |
| <b>sum</b>      | sommatoria                                       |
| <b>max</b>      | massimo                                          |
| <b>min</b>      | minimo                                           |
| <b>range</b>    | range = <b>max</b> - <b>min</b>                  |
| <b>sd</b>       | standard deviation                               |
| <b>variance</b> | varianza                                         |
| <b>cv</b>       | coefficiente di variazione ( <b>sd/mean</b> )    |
| <b>semean</b>   | standard error della media ( <b>sd/sqrt(n)</b> ) |
| <b>skewness</b> | simmetria                                        |
| <b>kurtosis</b> | curtosi                                          |
| <b>p1</b>       | 1° percentile                                    |
| <b>p5</b>       | 5° percentile                                    |
| <b>p10</b>      | 10° percentile                                   |
| <b>p25</b>      | 25° percentile                                   |
| <b>median</b>   | mediana (equivalente a <b>p50</b> )              |
| <b>p50</b>      | 50° percentile (equivalente a <b>median</b> )    |
| <b>p75</b>      | 75° percentile                                   |
| <b>p90</b>      | 90° percentile                                   |
| <b>p95</b>      | 95° percentile                                   |
| <b>p99</b>      | 99° percentile                                   |
| <b>iqr</b>      | range interquartile = <b>p75</b> - <b>p25</b>    |
| <b>q</b>        | equivale a specificare <b>p25 p50 p75</b>        |

---

**columns(variables|statistics)**: specifica cosa deve essere visualizzato in colonna. **variables** visualizza le variabili di `varlist` (opzione di default), **statistics** visualizza le statistiche specificate nell'opzione **statistics(statname)**.

**format(%fmt)** specifica il formato generale di visualizzazione delle statistiche. Il formato di default è **%12.2gc**.

**nototal** non visualizza le statistiche generali; si usa solo se viene specificata l'opzione **by(varname)**.

**missing** visualizza le statistiche anche per i valori missing della variabile **by(varname)**.

#### excel options

**xlsfile(filename.ext)**: specifica il file `.xls` o `.xlsx` (ed eventuale percorso) in cui salvare il codice della tabella. Questa opzione e l'estensione del file sono obbligatori.



**sheet(sheetname)**: specifica il nome del foglio in cui scrivere l'output. Di default si usa "Foglio 1".

**replace**: specifica di sovrascrivere il file indicato in **texfile(filename.ext)**.

**sheetreplace**: specifica di sovrascrivere il foglio indicato in **sheet(sheetname)**.

**sheetmodify**: specifica di modificare il foglio indicato in **sheet(sheetname)**.

**cell**: specifica la cella da cui iniziare l'output. Di default si usa A1. Usare solo la notazione lettera e numero.

**caption(string)**: specifica il testo da inserire come titolo della tabella. Di default è vuoto.

**note(string)**: specifica il testo da inserire come nota a piè di tabella. Di default è vuoto.

**intc1(string)**: specifica il testo da inserire come descrizione della prima colonna della tabella. In assenza dell'opzione **by(varname)** nella prima colonna ci possono essere le variabili o le statistiche, dipende da cosa specificato nell'opzione **columns()** e in questo caso di default **intc1()** è vuoto. Se viene specificata l'opzione **by(varname)**, nella prima colonna ci sono i valori della variabile **varname** e di default in **intc1()** c'è la descrizione associata a **varname**.

**intc2(string)**: specifica il testo da inserire come descrizione della seconda colonna della tabella e si applica solo nel caso in cui sia specificata l'opzione **by(varname)**. Se l'opzione **columns()** è **variables** il default è **intc2(Statistiche)**, se l'opzione **columns()** è **statistics** il default è **intc2(Variabili)**.

**wintr1(number)**: specifica la larghezza della prima colonna della tabella. In assenza dell'opzione **by(varname)** nella prima colonna ci possono essere le variabili o le statistiche, se l'opzione **by(varname)** è specificata nella prima colonna ci sono i valori della variabile **varname**. Di default il valore è pari a 40.

**wintr2(number)**: specifica la larghezza della seconda colonna della tabella e si applica solo nel caso in cui sia specificata l'opzione **by(varname)**. Se l'opzione **columns()** è **variables** nella seconda colonna ci sono le statistiche, se l'opzione **columns()** è **statistics** ci sono le variabili. Di default il valore è pari a 30.

**intc\_size(number)**: specifica l'altezza della prima riga della tabella. La prima riga contiene la descrizione delle variabili o delle statistiche a seconda di cosa è specificato in **columns()**. Di default il valore è pari a 30.

**resc\_size(number)**: specifica la larghezza delle colonne del corpo della tabella cioè delle colonne con i risultati delle statistiche specificate in **statistics(statname)**. Di default il valore è 16.

**nodisplast**: sopprime la dicitura della statistica nelle intestazioni di riga. Questa opzione è abilitata solo quando vengono specificate le opzioni **columns(variables)** e **by(varname)**.

**rows\_size(number)**: specifica l'altezza delle righe del corpo della tabella. Di default il valore è 15.

**fontname(string)**: specifica il font da usare nella tabella. Il default è **fontname(Calibri)**.

**fontsize(number)**: specifica la dimensione del font usato nella tabella. Il default è 11.

**bcolor\_intc(string)**: specifica il colore di sfondo della prima riga della tabella. I colori possono essere indicati nel formato RGB all'interno di virgolette (**pattern\_intc("255 255 255")**) o usando uno dei colori predefiniti da Stata per l'esportazione in Excel, vedi [M-5] **xl()** alla sezione Format colors. Di default non è previsto nessun colore.

**pattern\_intc(string)**: specifica il pattern di riempimento della prima riga della tabella. Vedi [M-5] **xl()** alla sezione Codes for fill pattern styles. Di default non è solid.

**vardisp(varlabel|varname)**: specifica come visualizzare le variabili specificate in varlist. **vardisp(varlabel)** visualizza la descrizione associata a ciascuna variabile, **vardisp(varname)** visualizza solo il nome della variabile. **vardisp(varlabel)** è il default. DA VERIFICARE !!

**bold**: specifica di formattare in bold la prima riga della tabella (intestazioni delle colonne).

**s1(string)...s10(string)**: specifica la descrizione delle statistiche indicate nell'opzione **statistics(statname)**. L'ordine deve essere quello di statname, ovvero **s1()** indica la descrizione della prima statistica di **statistics(statname)**, **s2()** indica la descrizione della seconda statistica di **statistics(statname)** e così via. Queste sono le descrizioni di default:

| statname        | Descrizione                 |
|-----------------|-----------------------------|
| <b>mean</b>     | Media                       |
| <b>count</b>    | Numero di osservazioni      |
| <b>n</b>        | Numero di osservazioni      |
| <b>sum</b>      | Sommatoria                  |
| <b>max</b>      | Massimo                     |
| <b>min</b>      | Minimo                      |
| <b>range</b>    | Massimo - Minimo            |
| <b>sd</b>       | Deviazione standard         |
| <b>variance</b> | Varianza                    |
| <b>cv</b>       | Coefficiente di variazione  |
| <b>semean</b>   | Errore standard della media |
| <b>skewness</b> | Simmetria                   |
| <b>kurtosis</b> | Curtosi                     |
| <b>p1</b>       | 1° percentile               |
| <b>p5</b>       | 5° percentile               |
| <b>p10</b>      | 10° percentile              |
| <b>p25</b>      | 25° percentile              |
| <b>median</b>   | Mediana                     |
| <b>p50</b>      | 50° percentile              |
| <b>p75</b>      | 75° percentile              |
| <b>p90</b>      | 90° percentile              |
| <b>p95</b>      | 95° percentile              |
| <b>p99</b>      | 99° percentile              |
| <b>iqr</b>      | Range interquartile         |

**dfs1(string)...dfs10(string)**: specifica il formato numerico delle statistiche indicate nell'opzione **statistics(statname)**. L'ordine deve essere quello di statname, ovvero **dfs1()** indica il formato della prima

statistica di **statistics(statname)**, **dfs2()** indica il formato della seconda statistica di **statistics(statname)** e così via. La sintassi del formato è la medesima di Mata per i formati numerici nell'esportazione in Excel (vedi [M-5] **xl()** alla sezione Codes for numeric formats).

| Formato | Esempio |
|---------|---------|
|---------|---------|

|                                |                |
|--------------------------------|----------------|
| <b>number</b>                  | 1000           |
| <b>number_d2</b>               | 1000.00        |
| <b>number_sep</b>              | 100,000        |
| <b>number_sep_d2</b>           | 100,000.00     |
| <b>number_sep_negbra</b>       | (1,000)        |
| <b>number_sep_negbrared</b>    | (1,000)        |
| <b>number_d2_sep_negbra</b>    | (1,000.00)     |
| <b>number_d2_sep_negbrared</b> | (1,000.00)     |
| <b>currency_negbra</b>         | (\$4000)       |
| <b>currency_negbrared</b>      | (\$4000)       |
| <b>currency_d2_negbra</b>      | (\$4000.00)    |
| <b>currency_d2_negbrared</b>   | (\$4000.00)    |
| <b>account</b>                 | 5,000          |
| <b>accountcur</b>              | \$ 5,000       |
| <b>account_d2</b>              | 5,000.00       |
| <b>account_d2_cur</b>          | \$ 5,000.00    |
| <b>percent</b>                 | 75%            |
| <b>percent_d2</b>              | 75.00%         |
| <b>scientific_d2</b>           | 10.00E+1       |
| <b>fraction_onedig</b>         | 10 1/2         |
| <b>fraction_twodig</b>         | 10 23/95       |
| <b>date</b>                    | 3/18/2007      |
| <b>date_d_mon_yy</b>           | 18-Mar-07      |
| <b>date_d_mon</b>              | 18-Mar         |
| <b>date_mon_yy</b>             | Mar-07         |
| <b>time_hmm_AM</b>             | 8:30 AM        |
| <b>time_HMMSS_AM</b>           | 8:30:00 AM     |
| <b>time_HMM</b>                | 8:30           |
| <b>time_HMMSS</b>              | 8:30:00        |
| <b>time_MMSS</b>               | 30:55          |
| <b>time_H0MMSS</b>             | 20:30:55       |
| <b>time_MMSS0</b>              | 30:55.0        |
| <b>date_time</b>               | 3/18/2007 8:30 |
| <b>text</b>                    | this is text   |

È possibile usare anche i formati numerici personalizzati, vedi [M-5] **xl()** alla sezione Custom formatting.

#### Examples

- . tabstatxls price weight mpg rep78, stat(n mean cv q) col(stat) xlsfile(tabstat.xlsx) wintr1(20) replace sheet(sheet1) cell(A1) dfs1(number) dfs2(number\_d2) dfs3(number\_d2) dfs4(number) dfs5(number) dfs6(number) resc\_size(14)
- . tabstatxls price weight mpg rep78, stat(n mean cv q) col(stat) xlsfile(tabstat.xlsx) wintr1(20) replace sheet(sheet1) cell(A1) dfs1(number) dfs2(number\_d2) dfs3(number\_d2) dfs4(number) dfs5(number) dfs6(number) resc\_size(14)
- . tabstatxls price weight mpg rep78, stat(n mean cv q) col(v) xlsfile(tabstat.xlsx) sheetmodify sheet(sheet1) cell(I1) dfs1(number) dfs2(number\_d2) dfs3(number\_d2) dfs4(number) dfs5(number) dfs6(number)
- . tabstatxls price weight mpg rep78, stat(n mean cv q) col(v) by(foreign) xlsfile(tabstat.xlsx) sheetmodify cell(A80) sheet(esempi excel) wintr1(12) wintr2(23) s1(Nonmissing observations) s2(Mean) s3(Coefficient of variation) s4(25th percentile) s5(50th percentile) s6(75th percentile)
- . tabstatxls price weight mpg rep78, stat(mean sd min max) c(s) by(foreign) xlsfile(tabstat.xlsx) sheetmodify cell(A40) sheet(esempi excel) wintr1(12) wintr2(20) fontname(Times New Roman) fontsize(9) pattern\_intc(silver)

#### Author

Nicola Tommasi  
nicola.tommasi@univr.it  
nicola.tommasi@gmail.com

#### Acknowledgments

#### References

#### Also see

Online: tabstat, tabstattex (if installed)

## 4 ESEMPI

### 4.1 Versione semplificata senza l'opzione by

In questo paragrafo vi mostrerò sempre l'esempio in Stata seguito dalla versione L<sup>A</sup>T<sub>E</sub>X e poi dalla versione Excel. `tabstat` nella sua forma più semplice, cioè senza l'utilizzo dell'opzione `by`, può essere usato come un `summarize` potenziato poichè possono essere visualizzate un numero superiore di statistiche di una serie di variabili. Questa è la versione standard di Stata con `col(stat)`

```
. tabstat price weight mpg rep78, stat(n mean cv q) col(stat)
```

| variable | N  | mean     | cv       | p25  | p50    | p75  |
|----------|----|----------|----------|------|--------|------|
| price    | 74 | 6165.257 | .478406  | 4195 | 5006.5 | 6342 |
| weight   | 74 | 3019.459 | .2573949 | 2240 | 3190   | 3600 |
| mpg      | 74 | 21.2973  | .2716543 | 18   | 20     | 25   |
| rep78    | 69 | 3.405797 | .290661  | 3    | 3      | 4    |

e vi mostro anche il comando `summarize` per facilitare il confronto

```
. summarize price weight mpg rep78
```

| Variable | Obs | Mean     | Std. Dev. | Min  | Max   |
|----------|-----|----------|-----------|------|-------|
| price    | 74  | 6165.257 | 2949.496  | 3291 | 15906 |
| weight   | 74  | 3019.459 | 777.1936  | 1760 | 4840  |
| mpg      | 74  | 21.2973  | 5.785503  | 12   | 41    |
| rep78    | 69  | 3.405797 | .9899323  | 1    | 5     |

Versione L<sup>A</sup>T<sub>E</sub>X (Tabella 1)

```
. tabstattex price weight mpg rep78, stat(n mean cv q) col(stat) texfile(ex1.tex) caption(" ") ///
> dfs1(%9.0f) dfs2(%9.2f) dfs3(%9.2f) dfs4(%9.0f) dfs5(%9.0f) dfs6(%9.0f) label(tab1)
```

Tabella 1:

|                    | Numero di osservazioni | Media   | Coefficiente di variazione | 25° percentile | 50° percentile | 75° percentile |
|--------------------|------------------------|---------|----------------------------|----------------|----------------|----------------|
| Price              | 74                     | 6165.26 | 0.48                       | 4195           | 5007           | 6342           |
| Weight (lbs.)      | 74                     | 3019.46 | 0.26                       | 2240           | 3190           | 3600           |
| Mileage (mpg)      | 74                     | 21.30   | 0.27                       | 18             | 20             | 25             |
| Repair Record 1978 | 69                     | 3.41    | 0.29                       | 3              | 3              | 4              |

Versione Excel (Figura 1)

```
. tabstatxls price weight mpg rep78, stat(n mean cv q) col(stat) xlsfile(tabstat.xlsx) wintr1(20) replace ///
> sheet(sheet1) cell(A1) dfs1(number) dfs2(number_d2) dfs3(number_d2) dfs4(number) dfs5(number) dfs6(number) ///
> resc_size(14)
```

Apri il file excel: **tabstat.xlsx**

|   | A                  | B                      | C       | D                          | E              | F              | G              |
|---|--------------------|------------------------|---------|----------------------------|----------------|----------------|----------------|
| 1 |                    |                        |         |                            |                |                |                |
| 2 |                    | Numero di osservazioni | Media   | Coefficiente di variazione | 25° percentile | 50° percentile | 75° percentile |
| 3 | Price              | 74                     | 6165,26 | 0,48                       | 4195           | 5007           | 6342           |
| 4 | Weight (lbs.)      | 74                     | 3019,46 | 0,26                       | 2240           | 3190           | 3600           |
| 5 | Mileage (mpg)      | 74                     | 21,30   | 0,27                       | 18             | 20             | 25             |
| 6 | Repair Record 1978 | 69                     | 3,41    | 0,29                       | 3              | 3              | 4              |

Figura 1:

Questa è la versione standard di Stata con `col(variables)`. Notate che rispetto all'esempio precedente le variabili sono passate in colonna e le statistiche sulle righe

```
. tabstat price weight mpg rep78, stat(n mean cv q) col(v)
```

| stats | price    | weight   | mpg      | rep78    |
|-------|----------|----------|----------|----------|
| N     | 74       | 74       | 74       | 69       |
| mean  | 6165.257 | 3019.459 | 21.2973  | 3.405797 |
| cv    | .478406  | .2573949 | .2716543 | .290661  |
| p25   | 4195     | 2240     | 18       | 3        |
| p50   | 5006.5   | 3190     | 20       | 3        |
| p75   | 6342     | 3600     | 25       | 4        |

Versione L<sup>A</sup>T<sub>E</sub>X (Tabella 2)

```
. tabstattex price weight mpg rep78, stat(n mean cv q) col(v) texfile(ex2.tex) caption(" ") ///
> dfs1(%9.0f) dfs2(%9.2f) dfs3(%9.2f) dfs4(%9.0f) dfs5(%9.0f) dfs6(%9.0f) label(tab2)
```

Tabella 2:

|                            | Price   | Weight (lbs.) | Mileage (mpg) | Repair Record<br>1978 |
|----------------------------|---------|---------------|---------------|-----------------------|
| Numero di osservazioni     | 74      | 74            | 74            | 69                    |
| Media                      | 6165.26 | 3019.46       | 21.30         | 3.41                  |
| Coefficiente di variazione | 0.48    | 0.26          | 0.27          | 0.29                  |
| 25° percentile             | 4195    | 2240          | 18            | 3                     |
| 50° percentile             | 5007    | 3190          | 20            | 3                     |
| 75° percentile             | 6342    | 3600          | 25            | 4                     |

Versione Excel (Figura 2)

```
. tabstatxls price weight mpg rep78, stat(n mean cv q) col(v) xlsfile(tabstat.xlsx) ///
> sheetmodify sheet(sheet1) cell(I1) dfs1(number) dfs2(number_d2) dfs3(number_d2) ///
> dfs4(number) dfs5(number) dfs6(number)
10
```

Apri il file excel: **tabstat.xlsx**

| I                          | J       | K             | L             | M                     |
|----------------------------|---------|---------------|---------------|-----------------------|
|                            | Price   | Weight (lbs.) | Mileage (mpg) | Repair Record<br>1978 |
| Numero di osservazioni     | 74      | 74            | 74            | 69                    |
| Media                      | 6165,26 | 3019,46       | 21,30         | 3,41                  |
| Coefficiente di variazione | 0,48    | 0,26          | 0,27          | 0,29                  |
| 25° percentile             | 4195    | 2240          | 18            | 3                     |
| 50° percentile             | 5007    | 3190          | 20            | 3                     |
| 75° percentile             | 6342    | 3600          | 25            | 4                     |

Figura 2:

...variazione sul tema, con cambio della descrizione delle statistiche (Tabella 3)

```
. tabstattex price weight mpg rep78, stat(n mean cv q) col(v) texfile(ex2bis.tex) caption(" ") ///
> s1(Nonmissing observations) s2(Mean) s3(Coefficient of variation) s4(25th percentile) ///
> s5(50th percentile) s6(75th percentile) dfs1(%9.0f) dfs2(%9.2f) dfs3(%9.2f) ///
> dfs4(%9.0f) dfs5(%9.0f) dfs6(%9.0f) label(tab3)
```

Tabella 3:

|                          | Price   | Weight (lbs.) | Mileage (mpg) | Repair Record<br>1978 |
|--------------------------|---------|---------------|---------------|-----------------------|
| Nonmissing observations  | 74      | 74            | 74            | 69                    |
| Mean                     | 6165.26 | 3019.46       | 21.30         | 3.41                  |
| Coefficient of variation | 0.48    | 0.26          | 0.27          | 0.29                  |
| 25th percentile          | 4195    | 2240          | 18            | 3                     |
| 50th percentile          | 5007    | 3190          | 20            | 3                     |
| 75th percentile          | 6342    | 3600          | 25            | 4                     |

## 4.2 Versione con l'opzione by

Partiamo da questo esempio, ricordando che non specificando l'opzione `columns()` in colonna vengono messe le variabili:

```
. tabstat price weight mpg rep78, by(foreign) stat(mean sd min max)
Summary statistics: mean, sd, min, max
by categories of: foreign (Car type)
```

| foreign  | price    | weight   | mpg      | rep78    |
|----------|----------|----------|----------|----------|
| Domestic | 6072.423 | 3317.115 | 19.82692 | 3.020833 |
|          | 3097.104 | 695.3637 | 4.743297 | .837666  |
|          | 3291     | 1800     | 12       | 1        |
|          | 15906    | 4840     | 34       | 5        |
| Foreign  | 6384.682 | 2315.909 | 24.77273 | 4.285714 |
|          | 2621.915 | 433.0035 | 6.611187 | .7171372 |
|          | 3748     | 1760     | 14       | 3        |
|          | 12990    | 3420     | 41       | 5        |
| Total    | 6165.257 | 3019.459 | 21.2973  | 3.405797 |
|          | 2949.496 | 777.1936 | 5.785503 | .9899323 |
|          | 3291     | 1760     | 12       | 1        |
|          | 15906    | 4840     | 41       | 5        |

Versione L<sup>A</sup>T<sub>E</sub>X (Tabella 4):

```
. tabstat price weight mpg rep78, by(foreign) stat(mean sd min max) texfile(ex3.tex) caption(" ") ///
> dfs1(%9.2f) dfs2(%9.2f) dfs3(%9.0f) dfs4(%9.0f) label(tab4)
```

Tabella 4:

| Car type | Statistiche         | Price    | Weight (lbs.) | Mileage (mpg) | Repair Record<br>1978 |
|----------|---------------------|----------|---------------|---------------|-----------------------|
| Domestic | Media               | 6072.42  | 3317.12       | 19.83         | 3.02                  |
|          | Deviazione standard | 3097.10  | 695.36        | 4.74          | 0.84                  |
|          | Minimo              | 3291     | 1800          | 12            | 1                     |
|          | Massimo             | 15906    | 4840          | 34            | 5                     |
| Foreign  | Media               | 6384.68  | 2315.91       | 24.77         | 4.29                  |
|          | Deviazione standard | 2621.92  | 433.00        | 6.61          | 0.72                  |
|          | Minimo              | 3748     | 1760          | 14            | 3                     |
|          | Massimo             | 12990    | 3420          | 41            | 5                     |
| Totale   | Media               | 6165.26  | 3019.46       | 21            | 3                     |
|          | Deviazione standard | 2949.50  | 777.19        | 6             | 1                     |
|          | Minimo              | 3291.00  | 1760.00       | 12            | 1                     |
|          | Massimo             | 15906.00 | 4840.00       | 41            | 5                     |

Versione Excel (Figura 3):

```
. tabstatxls price weight mpg rep78, by(foreign) stat(mean sd min max) xlsfile(tabstat.xlsx) ///
> sheetreplace sheet(sheet2) cell(A1) wintr1(10) wintr2(19)
Apri il file excel: tabstat.xlsx
```

|    | A        | B                   | C         | D             | E             | F                     |
|----|----------|---------------------|-----------|---------------|---------------|-----------------------|
| 1  |          |                     |           |               |               |                       |
| 2  |          |                     | Price     | Weight (lbs.) | Mileage (mpg) | Repair Record<br>1978 |
| 3  | Domestic | Media               | 6.072,42  | 3.317,12      | 19,83         | 3,02                  |
| 4  |          | Deviazione standard | 3.097,10  | 695,36        | 4,74          | 0,84                  |
| 5  |          | Minimo              | 3.291,00  | 1.800,00      | 12,00         | 1,00                  |
| 6  |          | Massimo             | 15.906,00 | 4.840,00      | 34,00         | 5,00                  |
| 7  | Foreign  | Media               | 6.384,68  | 2.315,91      | 24,77         | 4,29                  |
| 8  |          | Deviazione standard | 2.621,92  | 433,00        | 6,61          | 0,72                  |
| 9  |          | Minimo              | 3.748,00  | 1.760,00      | 14,00         | 3,00                  |
| 10 |          | Massimo             | 12.990,00 | 3.420,00      | 41,00         | 5,00                  |
| 11 | Totale   | Media               | 6.165,26  | 3.019,46      | 21,30         | 3,41                  |
| 12 |          | Deviazione standard | 2.949,50  | 777,19        | 5,79          | 0,99                  |
| 13 |          | Minimo              | 3.291,00  | 1.760,00      | 12,00         | 1,00                  |
| 14 |          | Massimo             | 15.906,00 | 4.840,00      | 41,00         | 5,00                  |
| 15 |          |                     |           |               |               |                       |

Figura 3:

Come fatto in precedenza faccio vedere anche la versione con inversione delle variabili e delle statistiche

```
. tabstat price weight mpg rep78, by(foreign) stat(mean sd min max) c(s)
```

Summary for variables: price weight mpg rep78  
by categories of: foreign (Car type)

| foreign  | mean     | sd       | min  | max   |
|----------|----------|----------|------|-------|
| Domestic | 6072.423 | 3097.104 | 3291 | 15906 |
|          | 3317.115 | 695.3637 | 1800 | 4840  |
|          | 19.82692 | 4.743297 | 12   | 34    |
|          | 3.020833 | .837666  | 1    | 5     |
| Foreign  | 6384.682 | 2621.915 | 3748 | 12990 |
|          | 2315.909 | 433.0035 | 1760 | 3420  |
|          | 24.77273 | 6.611187 | 14   | 41    |
|          | 4.285714 | .7171372 | 3    | 5     |
| Total    | 6165.257 | 2949.496 | 3291 | 15906 |
|          | 3019.459 | 777.1936 | 1760 | 4840  |
|          | 21.2973  | 5.785503 | 12   | 41    |
|          | 3.405797 | .9899323 | 1    | 5     |

Versione  $\text{\LaTeX}$  (Tabella 5):

Tabella 5:

| Car type | Variabili          | Media   | Deviazione<br>standard | Minimo | Massimo |
|----------|--------------------|---------|------------------------|--------|---------|
| Domestic | Price              | 6072.42 | 3097.10                | 3291   | 15906   |
|          | Weight (lbs.)      | 3317.12 | 695.36                 | 1800   | 4840    |
|          | Mileage (mpg)      | 19.83   | 4.74                   | 12     | 34      |
|          | Repair Record 1978 | 3.02    | 0.84                   | 1      | 5       |
| Foreign  | Price              | 6384.68 | 2621.92                | 3748   | 12990   |
|          | Weight (lbs.)      | 2315.91 | 433.00                 | 1760   | 3420    |
|          | Mileage (mpg)      | 24.77   | 6.61                   | 14     | 41      |
|          | Repair Record 1978 | 4.29    | 0.72                   | 3      | 5       |
| Totale   | Price              | 6165.26 | 2949.50                | 3291   | 15906   |
|          | Weight (lbs.)      | 3019.46 | 777.19                 | 1760   | 4840    |
|          | Mileage (mpg)      | 21.30   | 5.79                   | 12     | 41      |
|          | Repair Record 1978 | 3.41    | 0.99                   | 1      | 5       |

```
. tabstattex price weight mpg rep78, by(foreign) stat(mean sd min max) c(s) texfile(ex3bis.tex) caption(" ") ///
> dfs1(%9.2f) dfs2(%9.2f) dfs3(%9.0f) dfs4(%9.0f) label(tab5)
```

Versione Excel (Figura 4):

```
. tabstatxls price weight mpg rep78, by(foreign) stat(mean sd min max) c(s) xlsfile(tabstat.xlsx) sheetmodify ///
> sheet(sheet2) cell(I1) wintr1(10) wintr2(19)
```

Apri il file excel: **tabstat.xlsx**

|          |                    | I | J | K        | L                   | M        | N         |
|----------|--------------------|---|---|----------|---------------------|----------|-----------|
|          |                    |   |   | Media    | Deviazione standard | Minimo   | Massimo   |
| Domestic | Price              |   |   | 6.072,42 | 3.097,10            | 3.291,00 | 15.906,00 |
|          | Weight (lbs.)      |   |   | 3.317,12 | 695,36              | 1.800,00 | 4.840,00  |
|          | Mileage (mpg)      |   |   | 19,83    | 4,74                | 12,00    | 34,00     |
|          | Repair Record 1978 |   |   | 3,02     | 0,84                | 1,00     | 5,00      |
| Foreign  | Price              |   |   | 6.384,68 | 2.621,92            | 3.748,00 | 12.990,00 |
|          | Weight (lbs.)      |   |   | 2.315,91 | 433,00              | 1.760,00 | 3.420,00  |
|          | Mileage (mpg)      |   |   | 24,77    | 6,61                | 14,00    | 41,00     |
|          | Repair Record 1978 |   |   | 4,29     | 0,72                | 3,00     | 5,00      |
| Totale   | Price              |   |   | 6.165,26 | 2.949,50            | 3.291,00 | 15.906,00 |
|          | Weight (lbs.)      |   |   | 3.019,46 | 777,19              | 1.760,00 | 4.840,00  |
|          | Mileage (mpg)      |   |   | 21,30    | 5,79                | 12,00    | 41,00     |
|          | Repair Record 1978 |   |   | 3,41     | 0,99                | 1,00     | 5,00      |

Figura 4:



## 5 ESEMPI CON ALTRE OPZIONI PER tabstattex

Esempio con uso delle opzioni caption(), label() e note()

```
. tabstat price weight mpg rep78, by(foreign)
Summary statistics: mean
by categories of: foreign (Car type)
foreign | price weight mpg rep78
Domestic | 6072.423 3317.115 19.82692 3.020833
Foreign | 6384.682 2315.909 24.77273 4.285714
Total | 6165.257 3019.459 21.2973 3.405797

.
. tabstattex price weight mpg rep78, by(foreign) texfile(ex1p2732.tex) caption(Esempio di caption) ///
> label(tab_ex1p2732) note(Fonte: i dati provengono dal file auto.dta)
```

Tabella 6: Esempio di caption

| Car type | Statistiche | Price | Weight (lbs.) | Mileage (mpg) | Repair Record<br>1978 |
|----------|-------------|-------|---------------|---------------|-----------------------|
| Domestic | Media       | 6,072 | 3,317         | 20            | 3                     |
| Foreign  | Media       | 6,385 | 2,316         | 25            | 4.3                   |
| Totale   | Media       | 6,165 | 3,019         | 21            | 3.4                   |

Fonte: i dati provengono dal file auto.dta

In questo modo posso richiamare la tabella tramite l'opzione `\ref{}` di  $\text{\LaTeX}$ . Esempio: Si veda la tabella 6 di pagina 15.

Ora modifico l'intestazione delle prime 2 colonne (Tabella 7):

```
. tabstattex price weight mpg rep78, by(foreign) texfile(ex1p2732bis.tex) intc1(Descrizione Colonna 1) ///
> intc2(Descrizione Colonna 2) caption(" ") label(tab6)
```

Tabella 7:

| Descrizione Colonna 1 | Descrizione Colonna 2 | Price | Weight<br>(lbs.) | Mileage<br>(mpg) | Repair<br>Record<br>1978 |
|-----------------------|-----------------------|-------|------------------|------------------|--------------------------|
| Domestic              | Media                 | 6,072 | 3,317            | 20               | 3                        |
| Foreign               | Media                 | 6,385 | 2,316            | 25               | 4.3                      |
| Totale                | Media                 | 6,165 | 3,019            | 21               | 3.4                      |

Modifica della larghezza della tabella e modifica della dimensione del font (Tabella 8):

```
. tabstat price weight mpg rep78, by(foreign) stat(mean sd min max)
Summary statistics: mean, sd, min, max
by categories of: foreign (Car type)
```

| foreign  | price    | weight   | mpg      | rep78    |
|----------|----------|----------|----------|----------|
| Domestic | 6072.423 | 3317.115 | 19.82692 | 3.020833 |
|          | 3097.104 | 695.3637 | 4.743297 | .837666  |
|          | 3291     | 1800     | 12       | 1        |
|          | 15906    | 4840     | 34       | 5        |
| Foreign  | 6384.682 | 2315.909 | 24.77273 | 4.285714 |
|          | 2621.915 | 433.0035 | 6.611187 | .7171372 |
|          | 3748     | 1760     | 14       | 3        |
|          | 12990    | 3420     | 41       | 5        |
| Total    | 6165.257 | 3019.459 | 21.2973  | 3.405797 |
|          | 2949.496 | 777.1936 | 5.785503 | .9899323 |
|          | 3291     | 1760     | 12       | 1        |
|          | 15906    | 4840     | 41       | 5        |

```
. tabstat price weight mpg rep78, by(foreign) stat(mean sd min max) texfile(ex2p2732.tex) ///
> widthtable(15cm) fontsize(small) caption(" ") label(tab7)
```

Tabella 8:

| Car type | Statistiche         | Price  | Weight<br>(lbs.) | Mileage<br>(mpg) | Repair<br>Record 1978 |
|----------|---------------------|--------|------------------|------------------|-----------------------|
| Domestic | Media               | 6,072  | 3,317            | 20               | 3                     |
|          | Deviazione standard | 3,097  | 695              | 4.7              | .84                   |
|          | Minimo              | 3,291  | 1,800            | 12               | 1                     |
|          | Massimo             | 15,906 | 4,840            | 34               | 5                     |
| Foreign  | Media               | 6,385  | 2,316            | 25               | 4.3                   |
|          | Deviazione standard | 2,622  | 433              | 6.6              | .72                   |
|          | Minimo              | 3,748  | 1,760            | 14               | 3                     |
|          | Massimo             | 12,990 | 3,420            | 41               | 5                     |
| Totale   | Media               | 6,165  | 3,019            | 21               | 3.4                   |
|          | Deviazione standard | 2,949  | 777              | 5.8              | .99                   |
|          | Minimo              | 3,291  | 1,760            | 12               | 1                     |
|          | Massimo             | 15,906 | 4,840            | 41               | 5                     |

La stessa della precedente ma con posizionamento in orizzontale della tabella (vedi pagina successiva (Tabella 9)):

```
. tabstat price weight mpg rep78, by(foreign) stat(mean sd min max) texfile(ex2p2732bis.tex) ///
> landscape widthtable(18cm) caption(" ") label(tab8)
```

Tabella 9:

| Car type | Statistiche         | Price  | Weight (lbs.) | Mileage (mpg) | Repair Record<br>1978 |
|----------|---------------------|--------|---------------|---------------|-----------------------|
| Domestic | Media               | 6,072  | 3,317         | 20            | 3                     |
|          | Deviazione standard | 3,097  | 695           | 4.7           | .84                   |
|          | Minimo              | 3,291  | 1,800         | 12            | 1                     |
|          | Massimo             | 15,906 | 4,840         | 34            | 5                     |
| Foreign  | Media               | 6,385  | 2,316         | 25            | 4.3                   |
|          | Deviazione standard | 2,622  | 433           | 6.6           | .72                   |
|          | Minimo              | 3,748  | 1,760         | 14            | 3                     |
|          | Massimo             | 12,990 | 3,420         | 41            | 5                     |
| Totale   | Media               | 6,165  | 3,019         | 21            | 3.4                   |
|          | Deviazione standard | 2,949  | 777           | 5.8           | .99                   |
|          | Minimo              | 3,291  | 1,760         | 12            | 1                     |
|          | Massimo             | 15,906 | 4,840         | 41            | 5                     |

Modifico l'opzione `column()` imponendo che le statistiche siano sulle colonne (Tabella 10)

```
. tabstat price weight mpg rep78, by(foreign) stat(min mean max) col(stat) long
foreign variable	min mean max
Domestic price | 3291 6072.423 15906
 weight | 1800 3317.115 4840
 mpg | 12 19.82692 34
 rep78 | 1 3.020833 5
-----|-----
Foreign price | 3748 6384.682 12990
 weight | 1760 2315.909 3420
 mpg | 14 24.77273 41
 rep78 | 3 4.285714 5
-----|-----
Total price | 3291 6165.257 15906
 weight | 1760 3019.459 4840
 mpg | 12 21.2973 41
 rep78 | 1 3.405797 5
-----|-----

. tabstattex price weight mpg rep78, by(foreign) stat(min mean max) col(stat) texfile(ex2p2733.tex) ///
> caption(" ") label(tab9)
```

Tabella 10:

| Car type | Variabili          | Minimo | Media | Massimo |
|----------|--------------------|--------|-------|---------|
| Domestic | Price              | 3,291  | 6,072 | 15,906  |
|          | Weight (lbs.)      | 1,800  | 3,317 | 4,840   |
|          | Mileage (mpg)      | 12     | 20    | 34      |
|          | Repair Record 1978 | 1      | 3     | 5       |
| Foreign  | Price              | 3,748  | 6,385 | 12,990  |
|          | Weight (lbs.)      | 1,760  | 2,316 | 3,420   |
|          | Mileage (mpg)      | 14     | 25    | 41      |
|          | Repair Record 1978 | 3      | 4.3   | 5       |
| Totale   | Price              | 3,291  | 6,165 | 15,906  |
|          | Weight (lbs.)      | 1,760  | 3,019 | 4,840   |
|          | Mileage (mpg)      | 12     | 21    | 41      |
|          | Repair Record 1978 | 1      | 3.4   | 5       |

Sempre statistiche in colonna ma senza l'opzione `by()` (Tabella 11)

```
. tabstat price weight mpg rep78, stat(n mean cv q) col(stat)
variable | N mean cv p25 p50 p75
-----+-----
price | 74 6165.257 .478406 4195 5006.5 6342
weight | 74 3019.459 .2573949 2240 3190 3600
mpg | 74 21.2973 .2716543 18 20 25
rep78 | 69 3.405797 .290661 3 3 4

.
. tabstattex price weight mpg rep78, stat(n mean cv q) col(stat) texfile(ex3p2733.tex) ///
> caption(" ") label(tab10)
```

Tabella 11:

|                    | Numero di<br>osservazio-<br>ni | Media | Coefficiente<br>di<br>variazione | 25°<br>percentile | 50°<br>percentile | 75°<br>percentile |
|--------------------|--------------------------------|-------|----------------------------------|-------------------|-------------------|-------------------|
| Price              | 74                             | 6,165 | .48                              | 4,195             | 5,007             | 6,342             |
| Weight (lbs.)      | 74                             | 3,019 | .26                              | 2,240             | 3,190             | 3,600             |
| Mileage (mpg)      | 74                             | 21    | .27                              | 18                | 20                | 25                |
| Repair Record 1978 | 69                             | 3.4   | .29                              | 3                 | 3                 | 4                 |

## 6 ESEMPI CON ALTRE OPZIONI PER tabstatxls

Esempio con uso delle opzioni caption() e note() (Figura 5).

```
. tabstat price weight mpg rep78, stat(mean sd min max) c(s) by(foreign)
Summary for variables: price weight mpg rep78
by categories of: foreign (Car type)
```

| foreign  | mean     | sd       | min  | max   |
|----------|----------|----------|------|-------|
| Domestic | 6072.423 | 3097.104 | 3291 | 15906 |
|          | 3317.115 | 695.3637 | 1800 | 4840  |
|          | 19.82692 | 4.743297 | 12   | 34    |
|          | 3.020833 | .837666  | 1    | 5     |
| Foreign  | 6384.682 | 2621.915 | 3748 | 12990 |
|          | 2315.909 | 433.0035 | 1760 | 3420  |
|          | 24.77273 | 6.611187 | 14   | 41    |
|          | 4.285714 | .7171372 | 3    | 5     |
| Total    | 6165.257 | 2949.496 | 3291 | 15906 |
|          | 3019.459 | 777.1936 | 1760 | 4840  |
|          | 21.2973  | 5.785503 | 12   | 41    |
|          | 3.405797 | .9899323 | 1    | 5     |

```
. tabstatxls price weight mpg rep78, stat(mean sd min max) c(s) by(foreign) xlsfile(tabstat.xlsx) ///
> sheetreplace cell(A1) sheet(esempi excel) caption(Esempio di caption) ///
> wintr1(10) wintr2(19) note(Fonte: i dati provengono dal file auto.dta)
```

Apri il file excel: [tabstat.xlsx](#)

|    | A                                          | B                  | C        | D                   | E        | F         |
|----|--------------------------------------------|--------------------|----------|---------------------|----------|-----------|
| 1  | <b>Esempio di caption</b>                  |                    |          |                     |          |           |
| 2  |                                            |                    | Media    | Deviazione standard | Minimo   | Massimo   |
| 3  | Domestic                                   | Price              | 6.072,42 | 3.097,10            | 3.291,00 | 15.906,00 |
| 4  |                                            | Weight (lbs.)      | 3.317,12 | 695,36              | 1.800,00 | 4.840,00  |
| 5  |                                            | Mileage (mpg)      | 19,83    | 4,74                | 12,00    | 34,00     |
| 6  |                                            | Repair Record 1978 | 3,02     | 0,84                | 1,00     | 5,00      |
| 7  | Foreign                                    | Price              | 6.384,68 | 2.621,92            | 3.748,00 | 12.990,00 |
| 8  |                                            | Weight (lbs.)      | 2.315,91 | 433,00              | 1.760,00 | 3.420,00  |
| 9  |                                            | Mileage (mpg)      | 24,77    | 6,61                | 14,00    | 41,00     |
| 10 |                                            | Repair Record 1978 | 4,29     | 0,72                | 3,00     | 5,00      |
| 11 | Totale                                     | Price              | 6.165,26 | 2.949,50            | 3.291,00 | 15.906,00 |
| 12 |                                            | Weight (lbs.)      | 3.019,46 | 777,19              | 1.760,00 | 4.840,00  |
| 13 |                                            | Mileage (mpg)      | 21,30    | 5,79                | 12,00    | 41,00     |
| 14 |                                            | Repair Record 1978 | 3,41     | 0,99                | 1,00     | 5,00      |
| 15 | Fonte: i dati provengono dal file auto.dta |                    |          |                     |          |           |

Figura 5:

Ora modifico l'intestazione delle prime 2 colonne (Figura 6):

```
. tabstatxls price weight mpg rep78, stat(mean sd min max) c(s) by(foreign) xlsfile(tabstat.xlsx) ///
> sheetmodify cell(A20) sheet(esempi excel) wintr1(12) wintr2(20) intc1(Descrizione Colonna 1) ///
> intc2(Descrizione Colonna 2)
```

Apri il file excel: **tabstat.xlsx**

| 20 | Descrizione Colonna 1 | Descrizione Colonna 2 | Media    | Deviazione standard | Minimo   | Massimo   |
|----|-----------------------|-----------------------|----------|---------------------|----------|-----------|
| 21 |                       |                       |          |                     |          |           |
| 22 | Domestic              | Price                 | 6.072,42 | 3.097,10            | 3.291,00 | 15.906,00 |
| 23 |                       | Weight (lbs.)         | 3.317,12 | 695,36              | 1.800,00 | 4.840,00  |
| 24 |                       | Mileage (mpg)         | 19,83    | 4,74                | 12,00    | 34,00     |
| 25 |                       | Repair Record 1978    | 3,02     | 0,84                | 1,00     | 5,00      |
| 26 | Foreign               | Price                 | 6.384,68 | 2.621,92            | 3.748,00 | 12.990,00 |
| 27 |                       | Weight (lbs.)         | 2.315,91 | 433,00              | 1.760,00 | 3.420,00  |
| 28 |                       | Mileage (mpg)         | 24,77    | 6,61                | 14,00    | 41,00     |
| 29 |                       | Repair Record 1978    | 4,29     | 0,72                | 3,00     | 5,00      |
| 30 | Totale                | Price                 | 6.165,26 | 2.949,50            | 3.291,00 | 15.906,00 |
| 31 |                       | Weight (lbs.)         | 3.019,46 | 777,19              | 1.760,00 | 4.840,00  |
| 32 |                       | Mileage (mpg)         | 21,30    | 5,79                | 12,00    | 41,00     |
| 33 |                       | Repair Record 1978    | 3,41     | 0,99                | 1,00     | 5,00      |

Figura 6:

Modifica del font e dello sfondo della prima riga (Figura 7):

```
. tabstatxls price weight mpg rep78, stat(mean sd min max) c(s) by(foreign) xlsfile(tabstat.xlsx) ///
> sheetmodify cell(A40) sheet(esempi excel) wintr1(12) wintr2(20) fontname(Times New Roman) ///
> fontsize(9) bcolor_intc(silver)
```

Apri il file excel: **tabstat.xlsx**

| 40 |          |                    | Media    | Deviazione standard | Minimo   | Massimo   |
|----|----------|--------------------|----------|---------------------|----------|-----------|
| 41 |          |                    |          |                     |          |           |
| 42 | Domestic | Price              | 6.072,42 | 3.097,10            | 3.291,00 | 15.906,00 |
| 43 |          | Weight (lbs.)      | 3.317,12 | 695,36              | 1.800,00 | 4.840,00  |
| 44 |          | Mileage (mpg)      | 19,83    | 4,74                | 12,00    | 34,00     |
| 45 |          | Repair Record 1978 | 3,02     | 0,84                | 1,00     | 5,00      |
| 46 | Foreign  | Price              | 6.384,68 | 2.621,92            | 3.748,00 | 12.990,00 |
| 47 |          | Weight (lbs.)      | 2.315,91 | 433,00              | 1.760,00 | 3.420,00  |
| 48 |          | Mileage (mpg)      | 24,77    | 6,61                | 14,00    | 41,00     |
| 49 |          | Repair Record 1978 | 4,29     | 0,72                | 3,00     | 5,00      |
| 50 | Totale   | Price              | 6.165,26 | 2.949,50            | 3.291,00 | 15.906,00 |
| 51 |          | Weight (lbs.)      | 3.019,46 | 777,19              | 1.760,00 | 4.840,00  |
| 52 |          | Mileage (mpg)      | 21,30    | 5,79                | 12,00    | 41,00     |
| 53 |          | Repair Record 1978 | 3,41     | 0,99                | 1,00     | 5,00      |

Figura 7:

Font in bold per i titoli prima riga (Figura 8):

```
. tabstatxls price weight mpg rep78, stat(mean sd min max) c(s) by(foreign) xlsfile(tabstat.xlsx) ///
> sheetmodify cell(A60) sheet(esempi excel) wintr1(12) wintr2(20) bold
```

Apri il file excel: **tabstat.xlsx**

|          |                    | <b>Media</b> | <b>Deviazione standard</b> | <b>Minimo</b> | <b>Massimo</b> |
|----------|--------------------|--------------|----------------------------|---------------|----------------|
| Domestic | Price              | 6.072,42     | 3.097,10                   | 3.291,00      | 15.906,00      |
|          | Weight (lbs.)      | 3.317,12     | 695,36                     | 1.800,00      | 4.840,00       |
|          | Mileage (mpg)      | 19,83        | 4,74                       | 12,00         | 34,00          |
|          | Repair Record 1978 | 3,02         | 0,84                       | 1,00          | 5,00           |
| Foreign  | Price              | 6.384,68     | 2.621,92                   | 3.748,00      | 12.990,00      |
|          | Weight (lbs.)      | 2.315,91     | 433,00                     | 1.760,00      | 3.420,00       |
|          | Mileage (mpg)      | 24,77        | 6,61                       | 14,00         | 41,00          |
|          | Repair Record 1978 | 4,29         | 0,72                       | 3,00          | 5,00           |
| Totale   | Price              | 6.165,26     | 2.949,50                   | 3.291,00      | 15.906,00      |
|          | Weight (lbs.)      | 3.019,46     | 777,19                     | 1.760,00      | 4.840,00       |
|          | Mileage (mpg)      | 21,30        | 5,79                       | 12,00         | 41,00          |
|          | Repair Record 1978 | 3,41         | 0,99                       | 1,00          | 5,00           |

Figura 8:

Con cambio della descrizione delle statistiche (Figura 9)

```
. tabstatxls price weight mpg rep78, stat(n mean cv q) col(v) by(foreign) xlsfile(tabstat.xlsx) sheetmodify ///
> cell(A80) sheet(esempi excel) wintr1(12) wintr2(23) s1(Nonmissing observations) s2(Mean) ///
> s3(Coefficient of variation) s4(25th percentile) s5(50th percentile) s6(75th percentile)
```

Apri il file excel: **tabstat.xlsx**

|          |                          | <b>Price</b> | <b>Weight (lbs.)</b> | <b>Mileage (mpg)</b> | <b>Repair Record 1978</b> |
|----------|--------------------------|--------------|----------------------|----------------------|---------------------------|
| Domestic | Nonmissing observations  | 52,00        | 52,00                | 52,00                | 48,00                     |
|          | Mean                     | 6.072,42     | 3.317,12             | 19,83                | 3,02                      |
|          | Coefficient of variation | 0,51         | 0,21                 | 0,24                 | 0,28                      |
|          | 25th percentile          | 4.184,00     | 2.790,00             | 16,50                | 3,00                      |
|          | 50th percentile          | 4.782,50     | 3.360,00             | 19,00                | 3,00                      |
|          | 75th percentile          | 6.234,00     | 3.730,00             | 22,00                | 3,00                      |
| Foreign  | Nonmissing observations  | 22,00        | 22,00                | 22,00                | 21,00                     |
|          | Mean                     | 6.384,68     | 2.315,91             | 24,77                | 4,29                      |
|          | Coefficient of variation | 0,41         | 0,19                 | 0,27                 | 0,17                      |
|          | 25th percentile          | 4.499,00     | 2.020,00             | 21,00                | 4,00                      |
|          | 50th percentile          | 5.759,00     | 2.180,00             | 24,50                | 4,00                      |
|          | 75th percentile          | 7.140,00     | 2.650,00             | 28,00                | 5,00                      |
| Totale   | Nonmissing observations  | 74,00        | 74,00                | 74,00                | 69,00                     |
|          | Mean                     | 6.165,26     | 3.019,46             | 21,30                | 3,41                      |
|          | Coefficient of variation | 0,48         | 0,26                 | 0,27                 | 0,29                      |
|          | 25th percentile          | 4.195,00     | 2.240,00             | 18,00                | 3,00                      |
|          | 50th percentile          | 5.006,50     | 3.190,00             | 20,00                | 3,00                      |
|          | 75th percentile          | 6.342,00     | 3.600,00             | 25,00                | 4,00                      |

Figura 9:



Senza descrizione delle statistiche (Figura 10)

```
. tabstatxls price weight mpg length gear_ratio , stat(mean) col(v) by(rep78) xlsfile(tabstat.xlsx) sheetmodify ///
> cell(I1) sheet(esempi excel) wintr1(15) nodispstat nototal
```

Apri il file excel: **tabstat.xlsx**

|   | I                     | J        | K             | L             | M            | N          |
|---|-----------------------|----------|---------------|---------------|--------------|------------|
|   | Repair Record<br>1978 | Price    | Weight (lbs.) | Mileage (mpg) | Length (in.) | Gear Ratio |
| 1 |                       | 4.564,50 | 3.100,00      | 21,00         | 189,00       | 2,90       |
| 2 |                       | 5.967,63 | 3.353,75      | 19,13         | 199,38       | 2,69       |
| 3 |                       | 6.429,23 | 3.299,00      | 19,43         | 194,00       | 2,87       |
| 4 |                       | 6.071,50 | 2.870,00      | 21,67         | 184,83       | 3,18       |
| 5 |                       | 5.913,00 | 2.322,73      | 27,36         | 170,18       | 3,30       |

Figura 10:

## 7 Appendice

Questo script genera un file Excel con i colori e i background patterns previsti da Stata.

```
. capture erase mata_colors.xlsx
. mata
----- mata (type end to exit) -----
: b = xl()
: b.create_book("mata_colors", "colors", "xlsx")
: b.load_book("mata_colors")
: b.set_sheet("colors")
: b.set_mode("open")
:
: //b.set_sheet_gridlines("colors", "off")
: mata: b.put_string(1,1,"color name")
: cols = (1,19)
: b.set_font_bold(1,cols,"on")
: b.set_row_height(1,141,20)
: b.set_column_width(1,19,19)
:
: rows = (1,141)
: b.set_vertical_align(rows,cols,"center")
:
: end

.
. local col = 2
. foreach p in solid gray50 gray75 gray25 horstripe verstripe revdiagstripe diagstripe diagcrosshatch ///
> thickdiagcrosshatch thinhorstripe thinverstripe thinrevdiagstripe thindia stripe thinhorcrosshatch ///
> thindia crosshatch gray12p5 gray6p25 {
2. mata: b.put_string(1,'col',"p")
3. local col `++col'
4. }

.
. local row = 2
.
. foreach c in aliceblue antiquewhite aqua aquamarine azure beige bisque black blanchetalmond ///
> blue blueviolet brown burlywood cadetblue chartreuse chocolate coral cornflowerblue cornsilk ///
> crimson cyan darkblue darkcyan darkgoldenrod darkgray darkgreen darkkhaki darkmagenta ///
> darkolivegreen darkorange darkorchid darkred darksalmon darkseagreen darkslateblue ///
> darkslategray darkturquoise darkviolet deeppink deepskyblue dimgray dodgerblue firebrick ///
> floralwhite forestgreen fuchsia gainsboro ghostwhite gold goldenrod gray green greenyellow ///
> honeydew hotpink indianred indigo ivory khaki lavender lavenderblush lawngreen lemongrass ///
> lightblue lightcoral lightcyan lightgoldenrodyellow lightgray lightgreen lightpink lightsalmon ///
> lightseagreen lightskyblue lightslategray lightsteelblue lightyellow lime limegreen linen ///
> magenta maroon mediumaquamarine mediumblue mediumorchid mediumpurple mediumseagreen mediumslateblue ///
> mediumspringgreen mediumturquoise mediumvioletred midnightblue mintcream mistyrose moccasin ///
> navajowhite navy oldlace olive olivedrab orange orangered orchid palegoldenrod palegreen ///
> paleturquoise palevioletred papayawhip peachpuff peru pink plum powderblue purple red ///
> rosybrown royalblue saddlebrown salmon sandybrown seagreen seashell sienna silver skyblue ///
> slateblue slategray snow springgreen steelblue tan teal thistle tomato turquoise violet ///
> wheat white whitesmoke yellow yellowgreen {
2. local col = 2
3.
. foreach p in solid gray50 gray75 gray25 horstripe verstripe revdiagstripe diagstripe diagcrosshatch ///
> thickdiagcrosshatch thinhorstripe thinverstripe thinrevdiagstripe thindia stripe thinhorcrosshatch ///
> thindia crosshatch gray12p5 gray6p25 {
4. mata: b.set_fill_pattern(`row','col',"p","c")
5. local col `++col'
6. }
7. mata: b.put_string(`row',1,"c")
8. local row `++row'
9. }

.
. mata: b.close_book()
. **stlog close
```

## Riferimenti bibliografici

- [1] Kevin Crow. Export tables to excel, Sep 2013. <https://blog.stata.com/2013/09/25/export-tables-to-excel/>.
- [2] Kevin Crow. Export tabulation results to excel - update, Jun 2018. <https://blog.stata.com/2018/06/07/export-tabulation-results-to-excel-update/>.
- [3] Simon Fear. *Publication Quality Tables in L<sup>A</sup>T<sub>E</sub>X*, April 2016. <http://ctan.tug.org/tex-archive/macros/latex/contrib/booktabs>.
- [4] Ben Jann. FRE: Stata module to display one-way frequency table. Statistical Software Components, Boston College Department of Economics, April 2007.