

table - 要約統計情報のテーブル化 【評価版】

table コマンドはカテゴリ変数の組合せごとに要約統計情報を算出し、結果をテーブル形式で出力します。要約統計情報の種類としては数多くのもが用意されています。Superrows, supercolumns を用いたテーブル構成にも対応しています。

1. table コマンド	
2. 一元表	用例 1
3. 二元表	用例 2
4. 三元表	用例 3

1. table コマンド

table コマンドを使用する場合にキーとなるのはカテゴリ変数です。その数が n 個の場合には n 元の表 (n -way table) が作成されます (ただし $n \leq 7$)。一方、カテゴリ変数の値の組合せに対して個々のテーブルセルが対応することになるわけですが、そこに表示できる要約統計情報としては次のようなものが用意されています。

- 度数 (frequency)
- 平均値 (mean)
- 標準偏差 (standard deviation)
- 最大値 (maximum)
- 最小値 (minimum)
- 中央値 (median)
- 四分位範囲 (interquantile range)
- パーセンタイル (percentile)

それぞれのセルに複数の統計情報を出力させることもできます。

2. 一元表

▷ 用例 1

最初に一元の場合についてその用例を示します。使用する Example データセットは auto.dta です。

```
. sysuse auto.dta *1
(1978 Automobile Data)
```

このデータセット中には 1978 年に米国で販売された 74 車種のデータが収納されています。変数 displacement は排気量を表す連続変数ですが単位が立方インチでわかりにくいので、最初にそれをリッターに変換します。

```
. generate disp1 = 16.4*displacement/1000 *2
```

変数 disp1 の最大値と最小値をチェックしておく

```
. summarize disp1 *3
```

Variable	Obs	Mean	Std. Dev.	Min	Max
disp1	74	3.235676	1.50613	1.2956	6.97

より、排気量の値は 1.3l から 7.0l までの範囲に及びことがわかります。この変数 disp1 から次のようなカテゴリ変数 class を生成してみます。

class	排気量区分
0	$\text{disp1} \leq 2.0$
1	$2.0 < \text{disp1} \leq 3.0$
2	$3.0 < \text{disp1} \leq 4.0$
3	$4.0 < \text{disp1}$

```
. generate class = irecode(disp1, 2.0, 3.0, 4.0) *4
```

参考までに先頭から 5 台の車種について disp1 と class の値をリスト出力しておきます。

*1 メニュー操作 : File ▷ Example datasets ▷ Example datasets installed with Stata と操作しロードする。

*2 メニュー操作 : Data ▷ Create or change data ▷ Create new variable

*3 メニュー操作 : Statistics ▷ Summaries, tables and tests ▷ Summary and descriptive statistics ▷ Summary statistics

*4 関数 irecode() については [FN] *Programming functions (mwp-076)* を参照

```
. list make disp1 class in 1/5
```

	make	disp1	class
1.	AMC Concord	1.9844	0
2.	AMC Pacer	4.2312	3
3.	AMC Spirit	1.9844	0
4.	Buick Century	3.2144	2
5.	Buick Electra	5.74	3

このカテゴリ変数 `class` を用いて一元の表を作成してみましょう。その際、変数 `mpg` (miles per gallon) に関する要約統計情報をいくつか出力させることにします。

- Statistics > Summaries, tables, and tests > Other tables > Flexible table of summary statistics と操作
- Main タブ: Row variable: `class`
 - Statistics 1: Count nonmissing `mpg`
 - Statistics 2: Mean `mpg`
 - Statistics 3: Standard deviation `mpg`

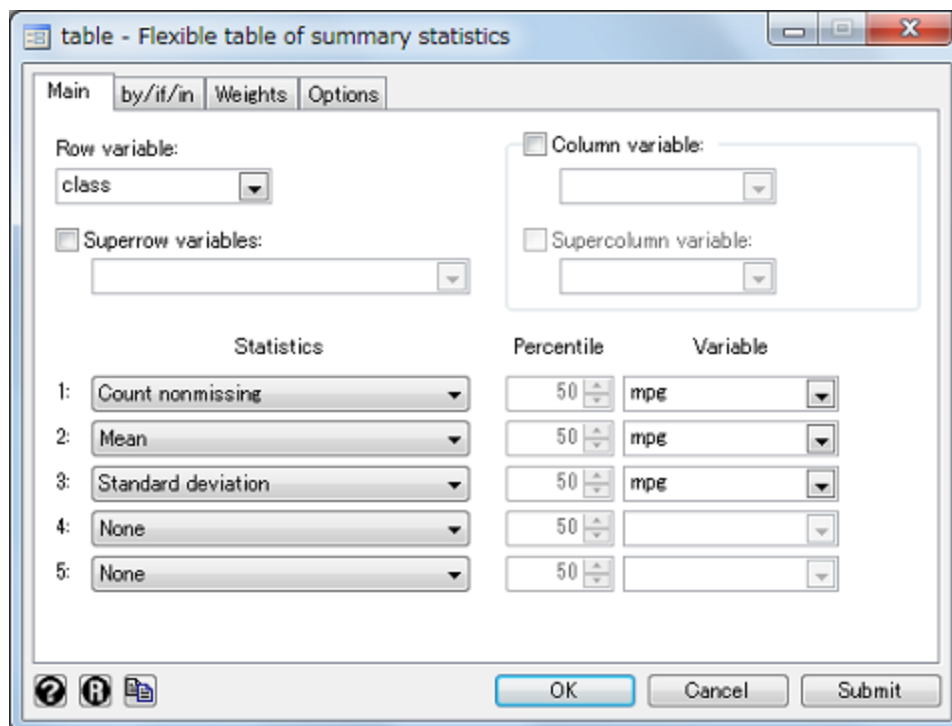


図1 table ダイアログ- Main タブ

```
. table class, contents(count mpg mean mpg sd mpg )
```

class	N(mpg)	mean(mpg)	sd(mpg)
0	23	27.2609	5.100957
1	13	20.6154	3.819652
2	19	19.4211	2.734873
3	19	16.4211	3.48514

定義された排気量クラスごとに変数 mpg の度数、平均値、標準偏差が出力されていることがわかります。なお、変数 class に対し値ラベルを設定してやればよりわかりやすい出力とすることができます。

```
. label define class 0 "<2000cc" 1 "2000-3000cc" 2 "3000-4000cc" 3 ">4000cc"
. label values class class
. table class, contents(count mpg mean mpg sd mpg )
```

```
. table class, contents(count mpg mean mpg sd mpg )
```

class	N(mpg)	mean(mpg)	sd(mpg)
<2000cc	23	27.2609	5.100957
2000-3000cc	13	20.6154	3.819652
3000-4000cc	19	19.4211	2.734873
>4000cc	19	16.4211	3.48514



値ラベルの設定については
 tb024: 値ラベルの設定
 をご参照ください。

さらに table ダイアログの Options タブ上で表示書式を明示してやれば数値の表示様式も調整できます。

- Options タブ: Override display format for numbers in cells: %9.2f

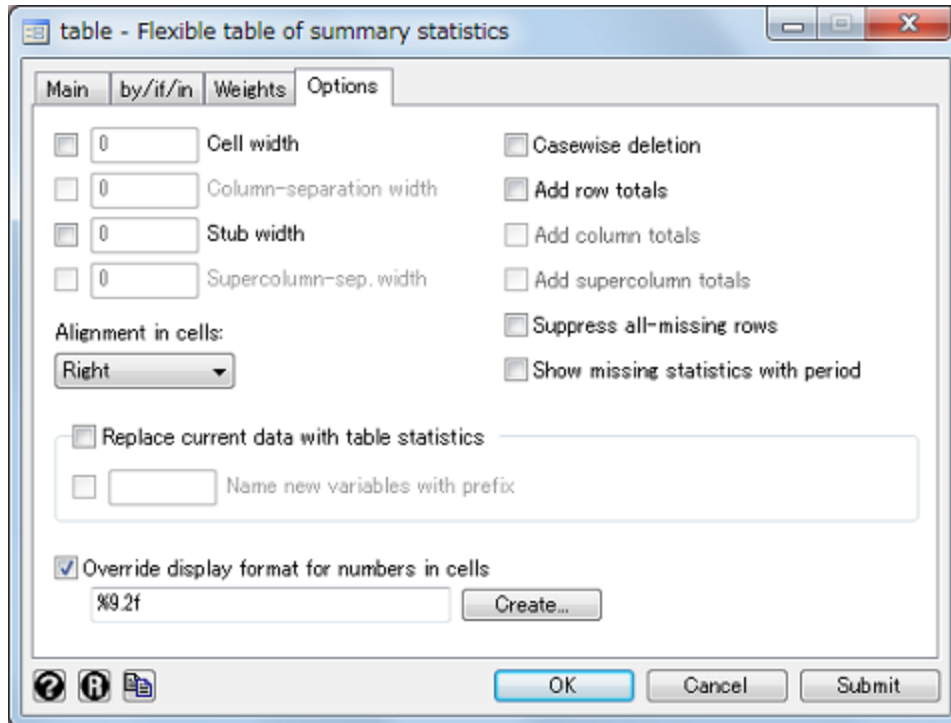


図 2 table ダイアログ- Options タブ

```
. table class, contents(count mpg mean mpg sd mpg ) format(%9.2f)
```

class	N(mpg)	mean(mpg)	sd(mpg)
<2000cc	23	27.26	5.10
2000-3000cc	13	20.62	3.82
3000-4000cc	19	19.42	2.73
>4000cc	19	16.42	3.49

◁

3. 二元表

▷ 用例 2

評価版では割愛しています。

4. 三元表

▷ 用例 3

評価版では割愛しています。

