Contingency Tables

Andy Grogan-Kaylor

4 May 2020 14:53:07

Key Concepts and Commands

Flipping Two Coins

Setup

```
. clear all

. set seed 3846
```

Good value labels are **key** here.

```
. label define nickel ///
> 1 "heads for nickel" ///
> 0 "tails for nickel" // define value label

. label define quarter ///
> 1 "heads for quarter" ///
> 0 "tails for quarter" // define value label

. set obs 1000 // 1000 observations
number of observations (_N) was 0, now 1,000

. * curiously it takes around 1000 obs for the proportions
. * below to "take hold"

. generate nickel = rbinomial(1, .75) // unfair nickel

. generate quarter = rbinomial(1, .5) // fair quarter

. label values nickel nickel // assign value label

. label values quarter quarter // assign value label
```

Crosstabulation

```
. tabulate nickel quarter, row col
```

Key
frequency *row percentage* *column percentage*

nickel	quarter tails for	heads for	Total
tails for nickel	104	140	244
	42.62	57.38	100.00
	21.62	26.97	24.40
heads for nickel	377	379	756
	49.87	50.13	100.00
	78.38	73.03	75.60
Total	481	519	1,000
	48.10	51.90	100.00
	100.00	100.00	100.00

Graphing (Mosaic Plot)

```
. * ssc install spineplot // mosaicplots (spineplots)

. * ssc install scheme-burd, replace // BuRd graph scheme

. spineplot nickel quarter, scheme(burd)

. graph export nickel-quarter.png, width(500) replace
(file nickel-quarter.png written in PNG format)
```

Bar Chart

Does a bar chart work to visualize these relationships?

```
. graph bar, over(quarter) over(nickel) scheme(burd)

. graph export nickel-quarter-bar1.png, width(500) replace
(file nickel-quarter-bar1.png written in PNG format)
```

Option `asyvars` adds a crucial color element.

```
. graph bar, over(quarter) over(nickel) scheme(burd) asyvars

. graph export nickel-quarter-bar2.png, width(500) replace
(file nickel-quarter-bar2.png written in PNG format)
```

And `hbar` may improve legibility even more.

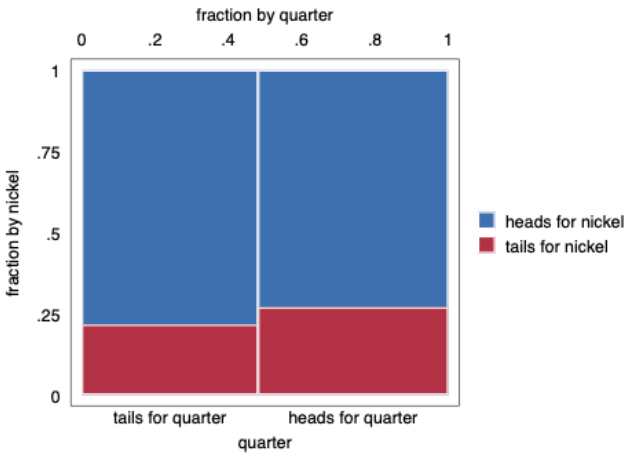

Figure 1: Mosaic Plot

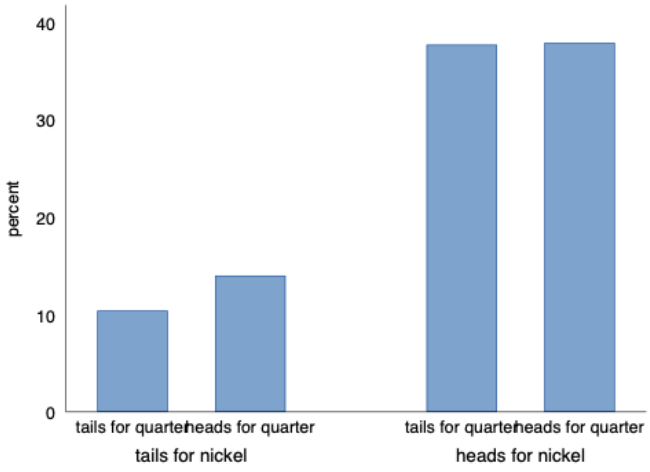

Figure 2: Bar Chart 1

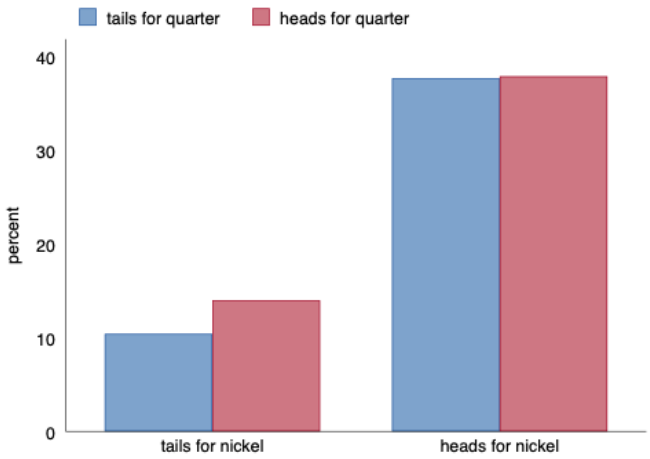

Figure 3: Bar Chart 2

```
. graph hbar, over(quarter) over(nickel) scheme(burd) asyvars

. graph export nickel-quarter-bar3.png, width(500) replace
(file nickel-quarter-bar3.png written in PNG format)
```

1961 French Skiiers

```
. clear all
```

Define Matrix

```
. matrix input FrenchSkiiers = (31, 109 \ 17, 122)

. matrix rownames FrenchSkiiers = Placebo AscorbicAcid

. matrix colnames FrenchSkiiers = Cold NoCold

. matrix list FrenchSkiiers
FrenchSkiiers[2,2]
               Cold  NoCold
    Placebo      31     109
AscorbicAcid     17     122
```

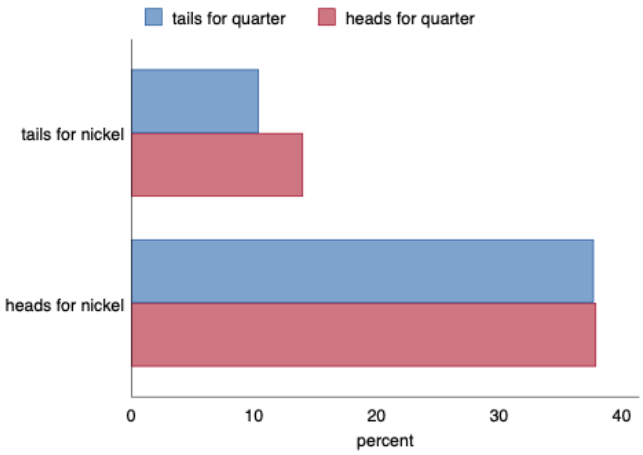

Figure 4: Bar Chart 3

Theme Music

Try Making a Data Set From Matrix

```
. svmat FrenchSkiiers, name(count)
number of observations will be reset to 2
Press any key to continue, or Break to abort
number of observations (_N) was 0, now 2

. list
```

	count1	count2
1.	31	109
2.	17	122

Enter Data By Hand

There are many alternative commands to do this, but the easiest way is using `edit`.

I have already done this. Note the structure of the data is different from above.

```
. use "FrenchSkiiers.dta", clear

. list // list the data
```

	Tx	Outcome	Count

1.	Ascorbic Acid	Cold	17
2.	Ascorbic Acid	No Cold	122
3.	Placebo	Cold	31
4.	Placebo	No Cold	109

Mosaic Plot

```
. spineplot Tx Outcome, scheme(burd)

. graph export FrenchSkiiers1.png, width(500) replace
(file FrenchSkiiers1.png written in PNG format)
```

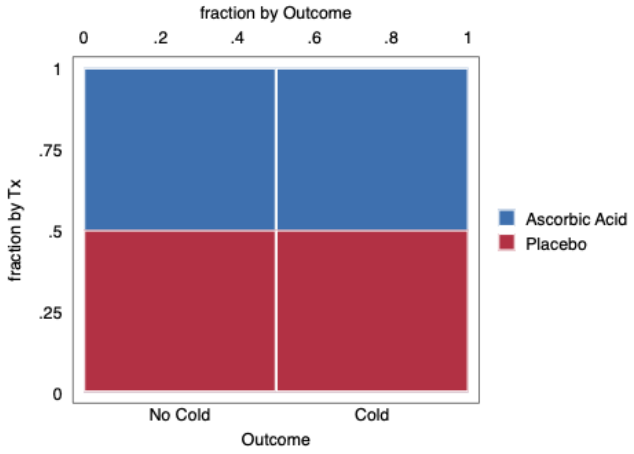

Figure 5: Mosaic Plot Attempt 1

spineplot Outcome Tx [fweight=Count], scheme(burd) // order matters to interpretability

graph export FrenchSkiiers2.png, width(500) replace

Definitions and Notation

Counts

c_{ij} c_{ij} $c_{i\bullet}$
c_{ij} c_{ij} $c_{i\bullet}$
$c_{\bullet j}$ $c_{\bullet j}$ $c_{\bullet\bullet}$

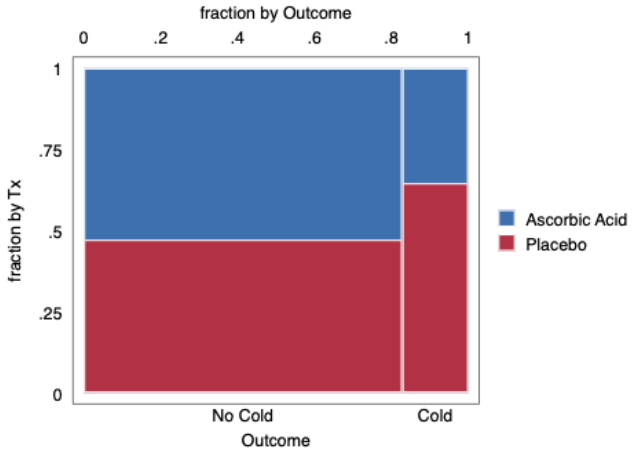

Figure 6: Mosaic Plot Attempt 2

Probabilities

p_{ij} p_{ij} $p_{i\bullet}$
p_{ij} p_{ij} $p_{i\bullet}$
$p_{\bullet j}$ $p_{\bullet j}$ $p_{\bullet\bullet}$