#### **QUANTITATIVE ANALYSIS**

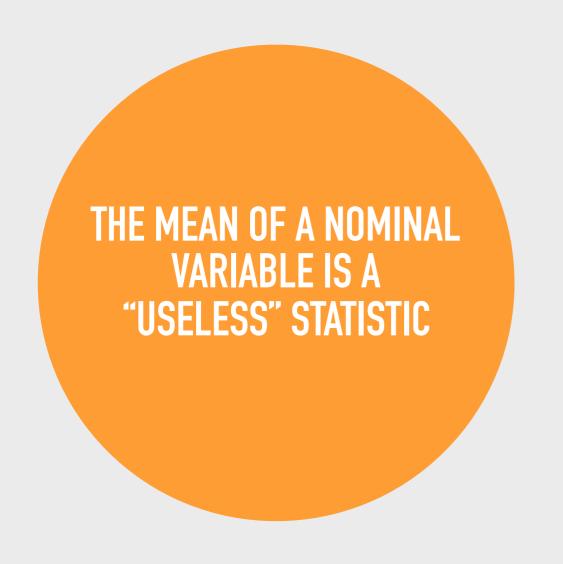
# CHI-SQUARED

#### **AGENDA**

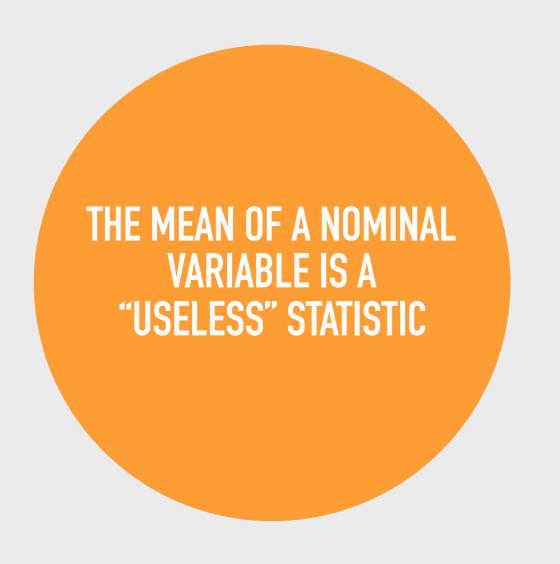
- 1. Follow-up
- 2. Chi-squared Logic
- 3. Cross-tabulations
- 4. Chi-squared Test in Stata

# 1 FOLLOW-UP

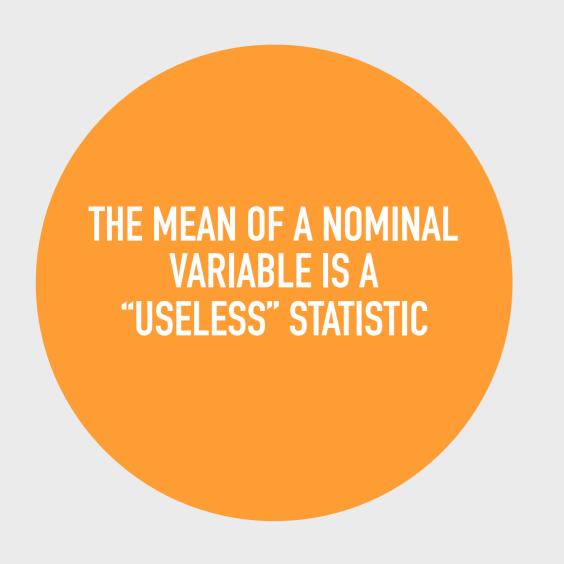
# 2 CHI-SQUARED LOGIC



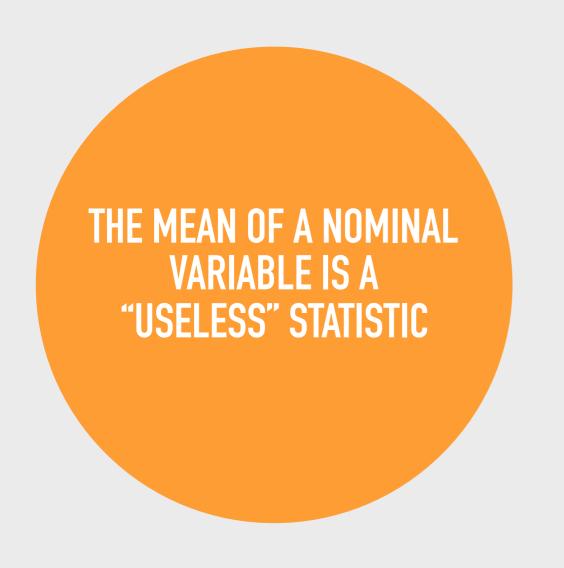
value	label	frequency
1	White	86
2	African American	43
3	Hispanic	24
4	Asian	18
mean		79



value	label	frequency
4	White	86
1	African American	43
2	Hispanic	24
3	Asian	18
mean		122.25



value	label	frequency
3	White	86
4	African American	43
1	Hispanic	24
2	Asian	18
mean		122.5



value	label	frequency
2	White	86
3	African American	43
4	Hispanic	24
1	Asian	18
mean		103.75



ALL OF THE STATISTICS WE'VE COVERED (S<sup>2</sup>, S, COV, R, T) RELY ON DEVIANCE (X<sub>1</sub> – X-BAR) AS A FOUNDATIONAL ELEMENT

CHI<sup>2</sup> (X<sup>2</sup>) ALLOWS FOR THE COMPARISON OF NOMINAL DATA WITHOUT RELYING ON MEANS TO UNDERSTAND THE RELATIONSHIP

CHI² COMPARES THE OBSERVED FREQUENCIES  $(F_0)$  To the expected frequencies  $(F_E)$  of the categories in both the X and Y variables

# 3 CROSS-TABULATIONS

# CROSS TABS IN STATA

tabulate xvar yvar [, row col cell]

tabulate foreign rep78

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

## CROSS TABS IN STATA

tabulate xvar yvar [, row col cell]

tabulate rep78 foreign

Densir

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

## CROSS TABS IN STATA

tabulate xvar yvar [, row col cell]

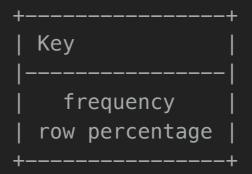
tabulate rep78 foreign

Densir

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

. tabulate rep78 foreign, row

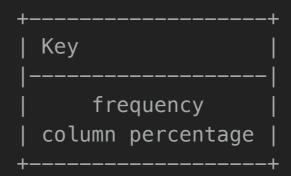
#### CROSS TABS IN STATA



Repair Record 1978	•	type Foreign	Total
			+
1	2   100.00	0 0.00	•
2	8   100.00	0 0.00	•
3	27   90 <b>.</b> 00	3 10.00	30   100.00
4	9   50 <b>.</b> 00	9 50 <b>.</b> 00	
5	2   18.18	9 81 <b>.</b> 82	•
Total	48   69 <b>.</b> 57	21 30.43	69   100.00

. tabulate rep78 foreign, col

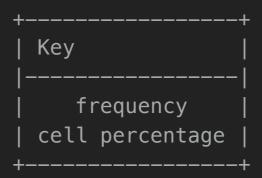
## CROSS TABS IN STATA



Repair	I		
Record	Car	type	
1978	Domestic +	Foreign	Total
1		0	2
	4.17 +	0.00 	2 <b>.</b> 90
2	8	0	8
	16.67 +	0.00 	11 <b>.</b> 59
3	27	3	30
	56 <b>.</b> 25 +	14 <b>.</b> 29	43.48
4	9	9	18
	18.75 +	42 <b>.</b> 86	26 <b>.</b> 09
5		9	11
	4 <b>.</b> 17	42 <b>.</b> 86	15 <b>.</b> 94
Total	48	21	69
	100.00	100.00	100.00

. tabulate rep78 foreign, cell

## CROSS TABS IN STATA



Repair	l .		
Record	Car	type	
1978	Domestic	Foreign	Total
1		0	2
	2 <b>.</b> 90 +	0.00 	2 <b>.</b> 90
2	8	0	8
	11 <b>.</b> 59 +	0.00 	11 <b>.</b> 59
3	27	3	30
	39 <b>.</b> 13 +	4 <b>.</b> 35 	43.48
4	9	9	18
	13.04 +	13 <b>.</b> 04 	26 <b>.</b> 09
5	2	9	11
	2.90 +	13.04 	15 <b>.</b> 94
Total	48	21	69
	69.57	30.43	100.00

# 4 CHI-SQUARED TEST IN STATA

#### HYPOTHESES

 $\blacktriangleright$  H0 = there is no difference in the variation between x and y

 $\blacktriangleright$  H1 = there is a difference in the variation between x and y

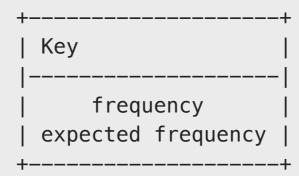
#### **ASSUMPTIONS**

- nominal data
- independence
- sufficient sample size
- less than 20% of cells can have an expected count of less than 5 cases

tabulate xvar yvar [, expected chi2 exact V]

. tabulate rep78 foreign, expected

#### **EXPECTED COUNTS**



Repair			
Record	Car	type	
1978	Domestic	Foreign	Total
1	2	0	2
	1.4 +	0.6 	2.0
2	8	0	8
	5.6	2 <b>.</b> 4	8.0
3	27	3	30
	20.9	9.1	30.0
4	9	9	18
	12.5	5.5	18.0
5	2	9	11
	7.7	3.3	11.0
Total	48	21	69
	48.0	21.0	69.0

# CALCULATING CHI<sup>2</sup>

tabulate rep78 foreign, chi2

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

Pearson chi2(4) = 27.2640 Pr = 0.000

#### CALCULATING CHI<sup>2</sup>

tabulate rep78 foreign, chi2

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

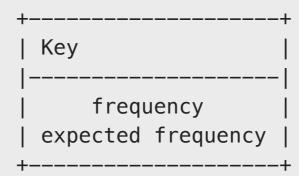
The results of the Chi2 test ( $chi^2 = 27.264$ , p < .001) suggest that there is a relationship between origin and repair record. Foreign cars are expected to have higher repair record values than domestic cars.

Pearson chi2(4) = 27.2640 Pr = 0.000

tabulate xvar yvar [, expected chi2 exact V]

. tabulate rep78 foreign, expected

#### **EXPECTED COUNTS**



Repair Record 1978			Total
1	2	0	2
	1.4	0.6	2.0
2	8	0	8
	5.6	2.4	8.0
3	27   20 <b>.</b> 9	3 9 <b>.</b> 1	30.0
4	9	9	18
	12.5	5 <b>.</b> 5	18.0
5	2   7 <b>.</b> 7	9 3.3	11   11.0
Total	48	21	69
	48.0	21.0	69.0

#### FISCHER'S EXACT TEST

. tabulate rep78 foreign, exact

```
Enumerating sample-space combinations:
stage 5: enumerations = 1
stage 4: enumerations = 3
stage 3: enumerations = 24
stage 2: enumerations = 203
stage 1: enumerations = 0
```

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

Fisher's exact =

0.000

#### **CRAMER'S V**

tabulate rep78 foreign, chi2 exact V

[output omitted]

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

Pearson chi2(4) = 27.2640 Pr = 0.000 Cramér's V = 0.6286 Fisher's exact = 0.000

SMALL = 0.1 MEDIUM = 0.3 LARGE = 0.5

#### **DOCUMENT DETAILS**

Document produced by <u>Christopher Prener, Ph.D</u> for the Saint Louis University course SOC 5050: QUANTITATIVE ANALYSIS - APPLIED INFERENTIAL STATISTICS. See the <u>course wiki</u> and the repository <u>README.md</u> file for additional details.



This work is licensed under a <u>Creative Commons Attribution 4.0 International License</u>.