Interpreting Analysis Results

JAY BORTHEN

Our Test Statistics

- Chi-square
- Gamma
- Z and t Statistics
- Correlation coefficient and the R^2 value
- Significance and the p-value
- 1-sided and 2-sided tests

Variable Types

- Nominal, Ordinal, and Interval
  - <u>Nominal Example</u>: Type of Anesthetic (e.g. Local, General, and Regional)
  - <u>Ordinal Example</u>: Hospital Trauma Levels (e.g. Level-1, Level-2, Level-3)
  - <u>Interval Example</u>: Drug Dosage (e.g. mg of substance)

- There are multiple kinds of tests for each type of variable depending on assumptions being made and what is being tested for (e.g. means, variances, association, etc.)

Our focus is on <u>ordinal variables</u>
and testing for associations

χ^2 Test of Independence

- **Does not measure strength of an association**, but, rather, it indicates how much evidence there is that the variables are or are not independent
- The χ^2 statistic ranges from 0 to ∞ (always ≥ 0)
- The larger the χ^2 value, the more confident we can be that the variables are not independence

<u>Example</u>: Assume we have two ordinal variables (Hospital Size and the Relative Cost of the Laryngoscope being used):

H_0: There is no association between variables (χ^2 value is near zero)

H_1: There is an association between variables

Statistical Independence

Sample 1

	LARYNGOSCOPE USED			
HOSPITAL SIZE	Inexpensive	Moderately Expensive	Very Expensive	**Total**
Small (0 - 99 beds)	54 (45%)	48 (40%)	18 (15%)	120 (100%)
Medium (100 - 249 beds)	90 (45%)	80 (40%)	30 (15%)	200 (100%)
Large (250 or more beds)	135 (45%)	120 (40%)	45 (15%)	300 (100%)
Total	279	248	93	**620**

The χ^2 Test Statistic

Sample 2

	LARYNGOSCOPE USED			
HOSPITAL SIZE	Inexpensive	Moderately Expensive	Very Expensive	**Total**
Small (0 - 99 beds)	**53** 39 (33%)	**47** 74 (63%)	**18** 5 (4%)	118 (100%)
Medium (100 - 249 beds)	**77** 81 (47%)	**69** 84 (49%)	**26** 7 (4%)	172 (100%)
Large (250 or more beds)	**149** 159 (48%)	**132** 89 (27%)	**49** 82 (25%)	330 (100%)
Total	279 **(45%)**	248 **(40%)**	93 **(15%)**	**620**

"Observed Value"

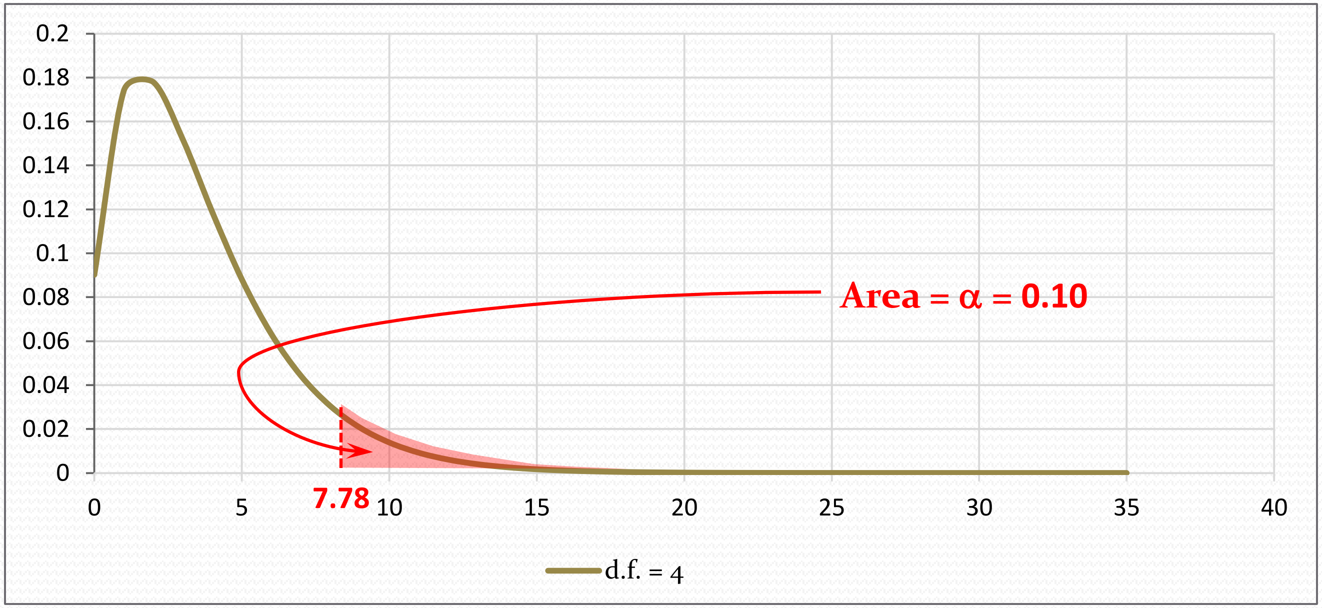
$$\chi^2 = \sum \frac{(O - E)^2}{E} = 83$$

"Expected Value"

Remember: This number does not reflect the strength of an association. It only provides evidence that an association actually exists.

The χ^2 Distribution

- Okay…so what does χ^2 = 83 tell us?
- With 4 degrees of freedom (d.f.) and an α = 0.10, our benchmark (critical value) for dependence is 7.78

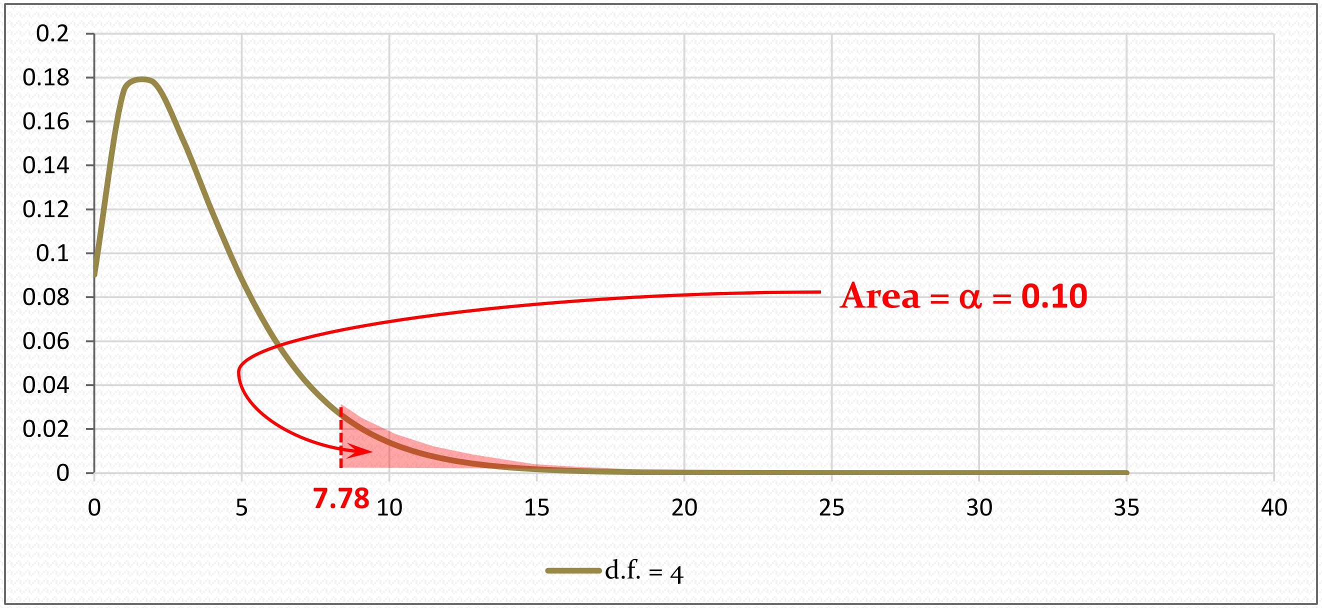


The χ^2 Distribution

H_o: There is no association between variables (χ^2 value is near zero)

H_1: There is an association between variables

We reject H_o if our χ^2 value is > our "benchmark" value of 7.78



The χ^2 Distribution

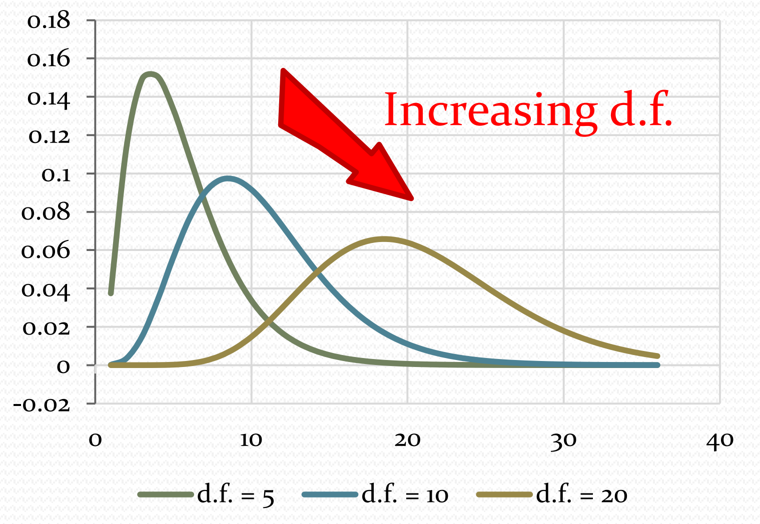
	PREFERRED LARYNGOSCOPE			
HOSPITAL SIZE	Macintosh	Glidescope	Other	**Total**
Small (0 - 99 beds)	39 (33%)	[illegible]4 (63%)	5 (4%)	118 (100%)
Medium (100 - 249 beds)	81 (47%)	[illegible]4 (49%)	7 (4%)	172 (100%)
Large (250 or more beds)	159 (48%)	89 (27%)	82 (25%)	330 (100%)
Total	279	248	93	**620**

3 x 3

Increasing d.f.

	PREFERRED LARYNGOSCOPE					
HOSPITAL SIZE	Macintosh	Macintosh	Glidescope	Macintosh	Other	**Total**
Small (0 - 99 beds)	39 (33%)	5 (4%)	74 (63%)	39 (33%)	5 (4%)	118 (100%)
Small Medium	39 (33%)	159 (48%)	5 (4%)	39 (33%)	159 (48%)	210 (100%)
Medium (100 - 249 beds)	81 (47%)	7 (4%)	[illegible]4 (49%)	81 (47%)	7 (4%)	172 (100%)
Medium Large	5 (4%)	89 (27%)	[illegible] (33%)	5 (4%)	89 (27%)	310 (100%)
Large (250 or more beds)	159 (48%)	82 (25%)	89 (27%)	159 (48%)	82 (25%)	330 (100%)
Total	279	144	248	169	93	**1240**

5 x 5

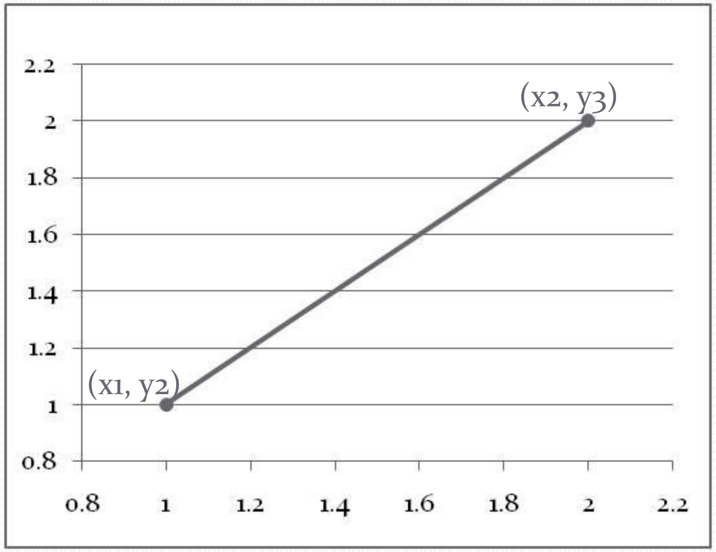


Degrees of Freedom = (# of rows – 1) x (# of columns – 1)

So, how do we determine the **strength and direction** of an established association between variables?

Concordant

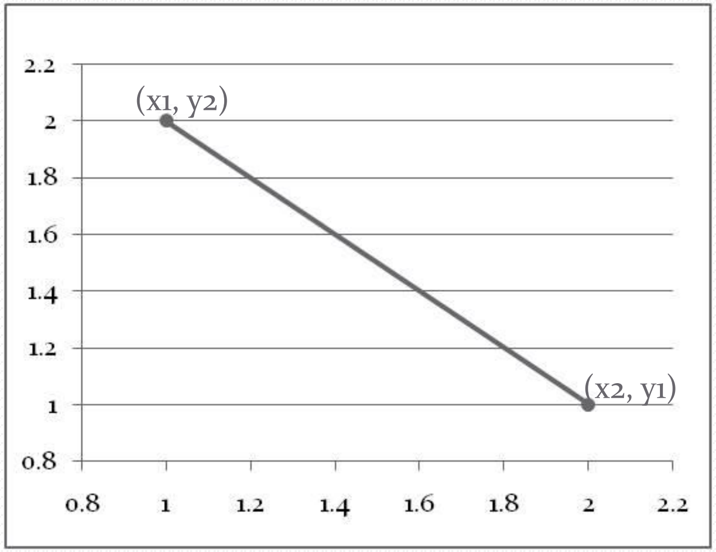
	Y			
	LARYNGOSCOPE USED			
HOSPITAL SIZE	Inexpensive	Moderately Expensive	Very Expensive	Total
Small (0 - 99 beds)	(X1, Y1)	(X1, Y2)	(X1, Y3)	
X Medium (100 - 249 beds)	(X2, Y1)	(X2, Y2)	(X2, Y3)	
Large (250 or more beds)	(X3, Y1)	(X3, Y2)	(X3, Y3)	
Total				



Discordant

Y

	LARYNGOSCOPE USED			
HOSPITAL SIZE	Inexpensive	Moderately Expensive	Very Expensive	**Total**
Small (0 - 99 beds)	(X1, Y1)	(X1, Y2)	(X1, Y3)	
X Medium (100 - 249 beds)	(X2, Y1)	(X2, Y2)	(X2, Y3)	
Large (250 or more beds)	(X3, Y1)	(X3, Y2)	(X3, Y3)	
Total				



The Gamma Value (γ)

- Appropriate for ordinal-by-ordinal data
- Generally more powerful test statistic than the χ^2 test statistic (i.e. we are able to detect significant differences with greater ease)
- Because ordinal variables can be ordered, we can talk about the direction of association between them (i.e. positive or negative)

Measure of Association

$$\gamma = \frac{C - D}{C + D}$$

C = Concordant Pairs

D = Discordant Pairs

γ lies between -1 and 1, where a value of 0 means that there is no relation between the two ordinal variables and $|\gamma| = 1$ represents the strongest associations.

Testing Independence with γ

H_0: There is no association between variables (γ value is near zero)

H_1: There is an association between variables

γ follows a normal distribution such that our *z* statistic becomes:

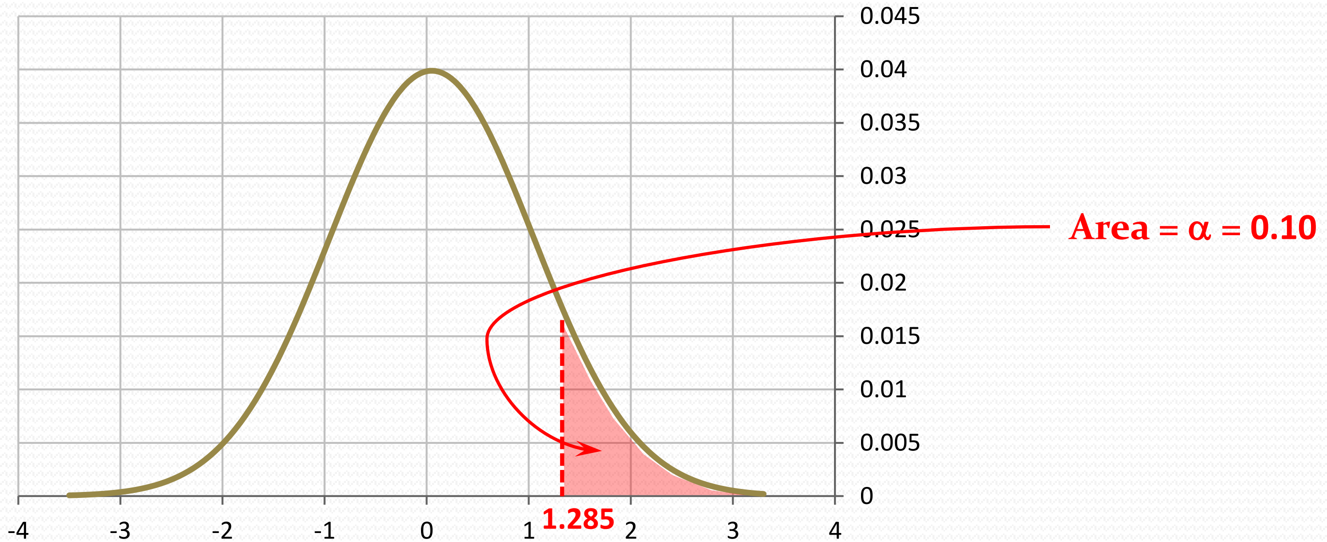
$$z = \frac{\gamma}{\text{s.e.}}$$

Comparing our *z* statistic with α, we can determine if the variables are statistically independent

Standard Error (s.e.) = sample standard deviation / $\sqrt{n}$

Testing Independence with γ

Assuming an α of 0.10 then our benchmark is $z_0 = 1.285$



If $z > z_0 = 1.285$ then reject H_0

Rejection of the H_0 means that the variables are not independent (see previous slide)

Using γ to test for independence does not rely on the degrees of freedom

Why do we not just use γ and forget about χ^2 since γ is a more powerful statistic?

χ^2 and γ

- An ordinal measure of association (i.e. γ) may equal 0 when the variables are actually statistically **dependent**

	LARYNGOSCOPE USED			
HOSPITAL SIZE	Inexpensive	Moderately Expensive	Very Expensive	**Total**
Small (0 - 99 beds)	59 (50%)	0 (0%)	59 (50%)	118 (100%)
Medium (100 - 249 beds)	20 (12%)	132 (76%)	20 (12%)	172 (100%)
Large (250 or more beds)	165 (50%)	0 (0%)	165 (50%)	330 (100%)
Total	244	132	244	**620**

$$\chi^2 = 437, \gamma = 0$$

Association vs. Causation

If

then

?

Hemline Theory = Association without Causation

Association vs. Causation

- Causation can often be obscure or counterintuitive so how do we tell the difference from Association?
  - Performing controlled comparisons
  - Increasing the variable resolutions (i.e. the number of data points)
  - Sequence Analysis (time becomes a variable)