What is the difference between a "test of statistical significance" and a "measure of association?"

1. Chi Square is a test of statistical significance providing the level of probability that the null hypothesis is true. That is, the probability that the association found between two variables in our sample is due to sampling error and not a real association found in the whole population (the t test and z test are also tests of statistical significance)
2. Measures of association measure the "strength" of the relationship"
   (e.g., lambda, gamma)

So, if Chi Square is a test of statistical significance. What question does chi square answer?

Chi Square tests the null hypothesis of "no difference" or "no association between variables."

It answers the question: **what is the probability**, that we would have the results we have from our sample (e.g., the existence of an association between two variables), if there is actually no association between the two variables in the population.

How does Chi Square Work?

First, it creates a table of "no relationship". This is done by, first, creating a table that shows the two variables and their margin totals found from the sample (no numbers are placed in the cells of the table yet).

Next, it uses the marginal totals to determine what numbers will go in each cell of the table, assuming there is no association between the two variables.

Once this table is created, it compares this "table of no relationship"
to the
table displaying the actual data found from the sample

The more similar the
table of no relationship is to the
actual-data table the more likely/probable that there is NO association between the two variables in the population.

That is, whatever relationship is found in the sample, and subsequently shown in the actual-data table, is likely due to sampling error and not a true reflection of the population.

How do we interpret the Chi Square Statistic?

Answer: a Chi Square number is calculated by comparing the table of "no association" to the actual data.

This chi square number is then found on a chi square distribution table which provides the probability that the size of the association exists when there is actually no association.

If the probability is 5% (.05) or less, we reject the null hypothesis of no difference.

IBM-SPSS will provide the chi square number and apply the distribution table to determine the probability.

If the probability is 5% (.05) or less that we would find the existing association (found in our sample), when there is actually no association, then we reject the null hypothesis.

So, we look to see if the chi square significance level is .05 or less.

Limitations of Chi Square:

1. Chi Square is sensitive to sample size. The larger the sample size the larger the chi square. Consequently, the null hypothesis is more likely be rejected with a large sample.
2. Chi Square is sensitive to small expected frequencies. Each cell should include at least 5 cases to be sure that chi square is accurate.
3. While Chi Square shows us statistical significance it does not give us information about the strength of the relationship or substantive significance. (This is left for measures of association and interpretation of the data)

So, given the limitations, it is useful to revisit our earlier question:

Does it make sense to report (or even examine) the measures of association if the test of statistical significance tells us that we should not reject the hypothesis of "no association"?

Answer: typically, if the chi square is not significant, then the measure of association (such as lambda) should not be considered since we must accept the null hypothesis of no difference.

However, because chi square is affected by table cells that are small (i.e. less than 5 cases in a particular table cell), the chi square results could be wrong if this condition exists.

Therefore, it would be wise to examine the size of the measure of association (e.g., lambda) even if chi square is not significant when this condition exists.

Measures of Association: Overview

1. Examine the strength of the relationship between two variables (such as between an independent and dependent variable).

2. Some of the "measures of association" (such as gamma) show the direction of the relationship (either positive or negative).

3. Measures the ability (i.e., strength) of one variable (the independent variable) to predict another (the dependent variable). For example, if two variables have a positive relationship then, when one variable goes up, we can predict that the second variable will also go up.

Example:

We want to know (predict) the strength of the relationship between "sex" and "level of education" among Dallas residents.

First, we take a random sample of Dallas residents. We use chi square to determine the probability that we can reject the null hypothesis.

Second, if we can reject the null hypothesis, then we can assume that the relationship found in our sample also exists in the population but we don't know how strong the relationship is between the two variables.

Third, we use one of the measures of association to determine the strength of the relationship between the two variables.

How does a "measures of association" determine the strength of the relationship between two variables (i.e., how is it calculated)?

The various measures of association (e.g., lambda, gamma, tau c) each have there own method for calculating the strength of the relationship.

However, regardless of the particular measure, they use a similar logic.

First, the measures of association calculate one's ability to predict how each respondent will be scored on the dependent variable if there is NO independent variable to help us.

Second, the measures of association calculate the ability to predict the dependent variable if we are able to consider the independent variable before guessing (predicting) how the respondent would be scored on the dependent variable.

Finally, the measures of association compare our first calculation to our second and determines how much better (less wrong) we are at guessing the respondent's answer by considering the independent variable.

If the independent variable is unable to provide us with some help in predicting the dependent variable, then we conclude the association is very weak.

In other words:

First, we determine how often we would be wrong, if we tried to guess (i.e., predict) each respondent's score (e.g., answer) on a dependent variable.

Second, we determine how often we would be wrong, if we used the independent variable (i.e., extra information) to help us guess or predict each respondent's score (e.g., answer) on a dependent variable.

Third, we determine how "less wrong" we are by using the independent variable to predict the dependent variable (i.e., each respondent's score on the dependent variable). The more "less wrong" we are, the stronger the relationship between the dependent and independent variables.

Here's another way to look at it: Take your best guess?

If you know nothing else about a person, except that he or she has a job and I asked you to guess the prestige score for his or her occupation, what would you guess?

The <u>mean</u> prestige score for occupations.

Take your best guess?

Now if I tell you that this person has a PhD, would you change your guess?

With quantitative analyses we are generally **trying to predict** or take our best guess at the value of the dependent variable.

One way to assess the strength of a relationship between two variables is to consider the degree to which the **extra information** of the second variable **makes your guess better**.

Measures of Association

"All good measures of association use a proportionate reduction in error (PRE) approach"

(Here we are talking about the number of errors made when trying to predict a specific score for each case in our sample)

What is a PRE measure?

Summary

PRE measures are derived by comparing:

1. the number of errors made when predicting the dependent variable (DV) while ignoring all independent variables (IV)

to:

2. the number of errors made when predicting the DV while using information about an IV.

How to Calculate a PRE Measure

$$PRE = \frac{E1 - E2}{E1}$$

E1 = errors of prediction made when the independent variable is ignored

E2 = errors of prediction made when the IV is used to make the predictions

Proportional Reduction of Error (PRE)

- If the DV is related to the IV, then the IV will allow us to make a better prediction (fewer errors) than the prediction we would make without considering the IV.
- The better the ability of the IV to help us predict the DV, the "stronger" the relationship between the DV and the IV.

Two PRE Measures: Lambda & Gamma

		Appropriate for…
Lambda	λ	Nominal variables
Gamma	γ	Ordinal & Dichotomous Nominal (has two values such as sex)

Measure of Association: Lambda λ

- Provides us with an indication of the **strength of an association** between the independent and dependent variables.
- Suitable for use with **nominal** variables
- Ranges from **0.0 to 1.0**
- A lower value represents a weaker association, while a higher value is indicative of a stronger association between the DV & IV

- **Lambda** is an asymmetrical measure of association.
  - A measure whose value may vary depending on which variable is considered the independent variable and which the dependent variable.
  - In this case, when performing a statistical analysis with SPSS, place the independent variable in the column and the dependent variable in the row.

The size of the Lambda is generally interpreted as follows:

.00 to .19	"little to no relationship"
.20 to .39	"weak relationship"
.40 to .59	"moderate relationship"
.60 to 1.00	"strong relationship"

Measure of Association: Gamma γ

- Gamma provides us with an indication of the **strength** and **direction** of the association between the variables (ranges from 0.0 to ±1.0).

- Appropriate for **ordinal** variables or with **dichotomous nominal** variables (dichotomous variables have only two values such as female/male).

- Gamma γ is symmetrical.

That is, a measure whose value will be the same when either variable is considered the independent variable or the dependent variable.

The size of the Gamma is generally interpreted as follows:

.00 to .19	"little to no relationship"
.20 to .39	"weak relationship"
.40 to .59	"moderate relationship"
.60 to 1.00	"strong relationship"

Four other measures include:

Yule's Q	a PRE symmetric measure used with a 2 x 2 table; gives misleading information when one of the four cells has a zero frequency.
Phi	a PRE symmetric measure used with a 2 x 2 table; doesn't reach a maximum or minimum 1 so a "phi adjusted" has been developed.

Four other measures include:

Tau C — a non-PRE symmetric measure used with two discrete ordered variables; due to how it calculates the size of an association it is always a smaller value than gamma (neither statistic is preferred to the other).

Somer's D — a PRE asymmetric measure. Due to how it calculates the size of an association, it is always a smaller value than gamma; because it is asymmetric the DV should always be specified when reporting the statistic.

Steps for conducting a crosstabulation in order to determine chi square significance and strength of association:

1. open data set and click "analyze"
2. click "descriptive statistics"
3. click "crosstabs"
4. move the 2 variables of interest into the row and column boxes (put dependent in column)
5. click "statistics"
6. click "chi square"
7. click your statistics of interest (e.g., lambda, gamma)
8. click "continue"
9. click "cells"
10. click "expected" (so you can see what is expected if there were no association)

Continued: Steps for conducting a crosstabulation in order to determine chi square significance:

11. click "row percentages"
12. click "continue
13. click "display bar charts"
14. click "OK"

Example of IBM-SPSS Output:

```
CROSSTABS
  /TABLES=avar69 BY avar9
  /FORMAT=AVALUE TABLES
  /STATISTICS=CHISQ GAMMA
  /CELLS=COUNT EXPECTED ROW
  /COUNT ROUND CELL
  /BARCHART.
```

Example of IBM-SPSS Output:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
69-While at work I make many decisions on my own or with other CNAs * 9-I feel burned out from my work	391	95.4%	19	4.6%	410	100.0%

69-While at work I make many decisions on my own or with other CNAs * 9-I feel burned out from my work Crosstabulation

			9-I feel burned out from my work					
			Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Total
69-While at work I make many decisions on my own or with other CNAs	Strongly disagree	Count	12	15	6	11	1	45
		Expected Count	7.6	14.8	11.9	7.4	3.3	45.0
		% within 69-While at work I make many decisions on my own or with other CNAs	26.7%	33.3%	13.3%	24.4%	2.2%	100.0%
	Disagree	Count	13	37	29	9	6	94
		Expected Count	15.9	31.0	24.8	15.4	7.0	94.0
		% within 69-While at work I make many decisions on my own or with other CNAs	13.8%	39.4%	30.9%	9.6%	6.4%	100.0%
	Neutral	Count	10	26	36	22	9	103
		Expected Count	17.4	34.0	27.1	16.9	7.6	103.0
		% within 69-While at work I make many decisions on my own or with other CNAs	9.7%	25.2%	35.0%	21.4%	8.7%	100.0%
	Agree	Count	21	46	28	17	9	121
		Expected Count	20.4	39.9	31.9	19.8	9.0	121.0
		% within 69-While at work I make many decisions on my own or with other CNAs	17.4%	38.0%	23.1%	14.0%	7.4%	100.0%
	Strongly agree	Count	10	5	4	5	4	28
		Expected Count	4.7	9.2	7.4	4.6	2.1	28.0
		% within 69-While at work I make many decisions on my own or with other CNAs	35.7%	17.9%	14.3%	17.9%	14.3%	100.0%
Total		Count	66	129	103	64	29	391
		Expected Count	66.0	129.0	103.0	64.0	29.0	391.0
		% within 69-While at work I make many decisions on my own or with other CNAs	16.9%	33.0%	26.3%	16.4%	7.4%	100.0%

Some Notes from the previous table:

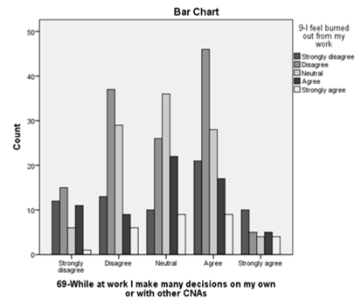
1. 54% of <u>those who feel empowered</u> (strongly agree + agree) are not burned out, while 32% do feel burned out.

2. 60% of <u>those who do not feel empowered</u> do feel burned out, while 27% do not feel burned out.

3. When considering "expected" verses the actual data: of those who felt empowered, we would expect 15 of them to report not being burned out (strongly agree or agree), the actual data found 15 people. this suggests no difference between the "table of no difference" and the actual data.

4. of those who are not empowered, we would expect 22 of them to not feel burned out, while the actual data shows that 27 did not feel burned out.

Example of IBM-SPSS Output:



Example of IBM-SPSS Output:

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	36.785[a]	16	.002
Likelihood Ratio	37.027	16	.002
Linear-by-Linear Association	.309	1	.578
N of Valid Cases	391		

a. 4 cells (16.0%) have expected count less than 5. The minimum expected count is 2.08.

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Ordinal by Ordinal	Gamma	.013	.059	.225	.822
N of Valid Cases		391			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Less than a .002 percent chance that the null hypothesis is true, i.e., we would get this association between the two variables when there is actually no association.

Gamma suggests that the association found is very weak.

再見

(see you later)