Lesson 12

Chi-square test

Fisher’s Exact test

McNemar’s Test

Binomial Test for Matched Pairs

Lesson 12 Overview

- Lesson 11 covered two inference methods for categorical data from 2 groups
  - Confidence Intervals for the difference of two proportions
  - Two sample z-test of equality of two proportions
- Lesson 12 covers three inference methods for categorical data
  - Chi-square test for comparisons between 2 categorical variables (Fisher’s exact test)
  - McNemar ‘sChi-square (Binomial Test) test for paired categorical data

Chi-square test

The Chi-square test can be used for two applications

- The Chi-square test can also be used to test for independence between two variables
  - The null hypothesis for this test is that the variables are independent (i.e. that there is no statistical association).
  - The alternative hypothesis is that there is a statistical relationship or association between the two variables.

- The Chi-square test can be used to test for equality of proportions between two or more groups.
  - The null hypothesis for this test is that the 2 proportions are equal.
  - The alternative hypothesis is that the proportions are not equal (test for a difference in either direction)

Contingency Tables

- Setting: Let X_1 and X_2 denote categorical variables, X_1 having **I** levels and X_2 having **J** levels. There are **IJ** possible combinations of classifications.

	Level=1	Level=2	………	Level=J
Level=1				
Level=2				
⋮				
Level=I				

- When the cells contain frequencies of outcomes, the table is called a contingency table.

Chi-square Test: Testing for Independence

Step 1: Hypothesis (always two-sided):

H_o: Independent

H_A: Not independent

Step 2: Calculate the test statistic:

$$X^2 = \sum \frac{(x_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2 \text{ with } \text{df} = (I-1)(J-1)$$

Step 3: Calculate the p-value

p-value = P($\mathcal{X}^2$ > X^2) <- value 2-sided

Step 4: Draw a conclusion

p-value<α reject independence

p-value>α do *not* reject independence

Chi-square Test: Testing for Independence

Racial differences and cardiac arrest.

An Example

- In a large mid-western city, the association in the incidence of cardiac arrest and subsequent survival was studied in 6117 cases of non-traumatic, out of hospital cardiac arrest.

- During a 12 month period, fewer than 1% of African-Americans survived an arrest-to-hospital discharge, compared to 2.6% of Caucasians.

Chi-square Test: Testing for Independence

Racial differences and cardiac arrest

Survival to Discharge

Race	YES	NO	Total
Caucasian	84	3123	3207
African-American	24	2886	2910
Total	108	6009	6117

Chi-square Test: Testing for Independence

Scientific Hypothesis:

An association exists between race (African-American/Caucasian) and survival to hospital discharge (Yes/No) in cases of non-traumatic out-of-hospital cardiac arrest.

Statistical Hypothesis:

H_o:

H_A:

Chi-square Test: Testing for Independence

1. Obtain a random sample of *n* independent observations (*the selection of one observation does not influence the selection of any other*).
2. Observations are classified subsequently according to cells formed by the intersection rows and columns in a contingency table.
    - Rows (r) consist of mutually exclusive categories of one variable.
    - Columns (c) consist of mutually exclusive categories of the other variable.
3. The frequency of observations in each cell is determined along with marginal totals.

Chi-square Test: Testing for Independence

4. Expected frequencies are calculated under the null hypothesis of independence (no association) and compared to observed frequencies.

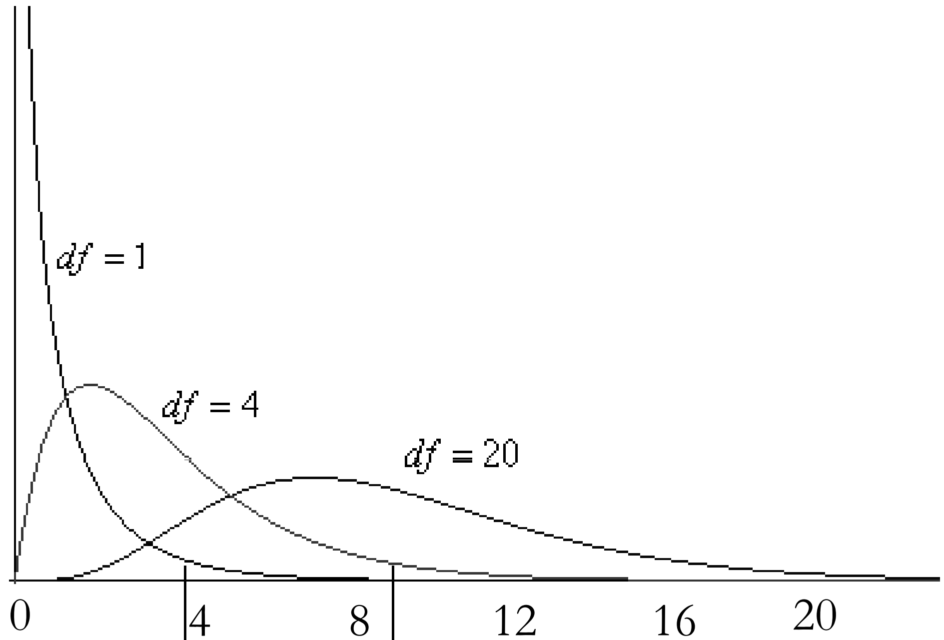
Recall: A and B are independent if:

5. Use the Chi-square (X^2) test statistic to observe the difference between the observed and expected frequencies.

χ^2 Distribution

- The probabilities associated with the chi-square distribution are in appendix D.
- The table is set up in the same way as the t-distribution.
- The chi-square distribution with 1 df is the same as the square of the Z distribution.
- Since the distribution only takes on positive values all the probability is in the right-tail.

Chi-square distributions and critical values for 1 df, 4 df and 20 df



For Chi-square with 20 df, the critical value (α = 0.05) = 31.4

Critical value for α = 0.05 and Chi-square with 1 df is 3.84

Critical value for α = 0.05 and Chi-square with 4 df is 9.49

Since the Chi-square distribution is always positive, the rejection region is only in the right tail

How to Identify the critical value

- The rejection region of the Chi-square test is the upper tail so there is only one critical value
- First calculate the df to identify the correct Chi-square distribution
  - For a 2 X 2 table, there are (2-1)*(2-1) = 1 df
- Use the CHIINV function to find the critical value
  - General formula: =CHIINV(alpha, df)

State the conclusion

The p-value for $P(\chi^2 > X^2) =$ CHIDIST(X^2,df)

- Reject the null hypothesis by either the rejection region method or the p-value method

  X^2 $\boxed{}$ Critical Value

  or

  Pvalue $\boxed{}$ α

Chi-squared Test: Testing for Independence

Calculating expected frequencies

Survival to Discharge

Race	YES	NO	Total
Caucasian	84 56.62	3123	3207
African-American	24	2886	2910
Total	108	6009	6117

Under the assumption of independence:

P(YES and Caucasian) = P(YES)*P(Caucasian)

= 108/6117 * 3207/6117 = 0.009256

Expected cell count =e_{ij} = 0.009256 * 6117 = 56.62

Chi-square Test: Testing for Independence

Calculating expected frequencies

Survival to Discharge

Race	YES	NO	Total
Caucasian	84 56.62	3123	3207
African-American	24	2886	2910
Total	108	6009	6117

Expected Cell Counts = (Marginal Row total * Marginal Column Total)/ n

Rule of Thumb: Check to see if expected frequencies are > 2

No more than 20% of cells with expected frequencies < 5

Chi-square Test: Testing for Independence

Step 1: Hypothesis (always two-sided):

H_o: Independent (Race/Survival)

H_A: Not independent

Step 2: Calculate the test statistic:

$$X^2 = \sum \frac{(x_{ij} - e_{ij})^2}{e_{ij}}$$

$$= \frac{(84 - 56.62)^2}{56.62} + \frac{(3123 - 3151.43)^2}{3151.43} + \frac{(24 - 51.38)^2}{51.38} + \frac{(2886 - 2854.82)^2}{2854.82}$$

$$= 13.24 + 0.26 + 14.59 + .34 = 28.42$$

Chi-square Test: Testing for Independence

Step 3: Calculate the p-value

p-value = P(X^2 >_______) =

Chidist(_____,_____) < 0.001

Step 4: Conclusion??

Chi-square Test: Testing for Equality or Homogeneity of Proportions

Testing for equality or homogeneity of proportions – examines differences between two or more independent proportions.

In chi-square test for independence, we examine the cross-classification of a ***single sample*** of observations on two qualitative variables.

The chi-square test can also be used for problems involving ***two or more*** independent populations.

Chi-square Test: Testing for Equality or Homogeneity of Proportions

Example

Patients with evolving myocardial infarction were assigned independently and randomly to one of four thrombolytic treatments, and then followed to determine 30 day mortality.

30 day outcome	Streptokinase and SC Heparin	Streptokinase and IV Heparin	Accelerated t-PA and IV Heparin	Accelerated t-PA Streptokinase with IV Heparin	Total
Survived	9091	9609	9692	9605	37997
Died	705	768	652	723	2848
Total	9796	10377	10344	10328	40845

Are these four treatment populations equal with respect to 30-day mortality?

Chi-square Test: Testing for Equality or Homogeneity of Proportions

Example

30 day outcome	Streptokinase and SC Heparin	Streptokinase and IV Heparin	Accelerated t-PA and IV Heparin	Accelerated t-PA Streptokinase with IV Heparin	Total
Survived	9091 9112.95	9609	9692	9605	37997
Died	705	768	652	723	2848
Total	9796	10377	10344	10328	40845

Under the assumption of independence:

P(Streptokinase and SC Heparin **and** Survival) = P(Streptokinase and SC Heparin)*P(Survival)

= 9796/40845 * 37997/40845 =0.223

Expected cell count =e_{ij} = 0.223 * 40845 = 9112.95

Chi-square Test: Testing for Equality or Homogeneity of Proportions

Example

30 day outcome	Streptokinase and SC Heparin	Streptokinase and IV Heparin	Accelerated t-PA and IV Heparin	Accelerated t-PA Streptokinase with IV Heparin	Total
Survived	9091 9112.95	9609 9653.44	9692 9622.74	9605 9607.86	37997
Died	705 683.05	768 723.56	652 721.26	723 720.14	2848
Total	9796	10377	10344	10328	40845

Under the assumption of independence:

Expected Cell Counts = (Marginal Row total * Marginal Column Total)/ n

Chi-square Test: Testing for Equality or Homogeneity of Proportions

Step 1: Hypothesis (always two-sided):

H_o: The four treatment options are homogeneous with respect to 30 day survival.

H_A: The four treatment options are not homogeneous with respect to 30 day survival.

Step 2: Calculate the test statistic:

$$X^2 = \sum \frac{(x_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2 \text{ with df } = (I-1)(J-1)$$

Step 3: Calculate the pvalue = P($\mathcal{X}^2$ > X^2)

Step 4: Draw a conclusion

p-value<α reject independence

p-value>α do *not* reject independence

Chi-square Test: Testing for Equality or Homogeneity of Proportions

Step 1: Hypothesis (always two-sided):

H_o: The four treatment options are homogeneous with respect to 30 day survival.

H_A: The four treatment options are not homogeneous with respect to 30 day survival.

Step 2: Calculate the test statistic:

$$X^2 = \sum \frac{(x_{ij} - e_{ij})^2}{e_{ij}} = 10.85 \text{ with df } = \_\_\_\_\_\_\_\_\_\_\_\_ \quad \_\_\_$$

Chi-square Test: Testing for Equality or Homogeneity of Proportions

Step 3: Calculate the p-value

p-value = P(X^2 >______)

=chidist(_____,_____) = 0.013

Step 4: Conclusion???

Chi-square test in EXCEL or online calculator

- There is no Excel function or Data Analysis Tool for the Chi-square test
  - For the Excel examples, you'll need to calculate the expected cell frequencies from the observed marginal totals and then calculate the statistic from the observed and expected frequencies.
- This website will calculate the Chi-square statistic and p-value for data in a 2 X 2 table. Enter the cell counts in the table. Choose the Chi-square test without Yate's correction to obtain the same results as in the example

www.graphpad.com/quickcalcs/contingency1.cfm

Chi-Square Testing: Rules of Thumb

- All expected frequencies should be equal to or greater than 2 (*observed frequencies can be less than 2*).

- No more than 20% of the cells should have expected frequencies of less than 5.

- What if these rules of thumb are violated?

Small Expected Frequencies

- Chi-square test is an approximate method.
- The chi-square distribution is an *idealized* mathematical model.
- In reality, the statistics used in the chi-square test are qualitative (have discrete values and not continuous).
- For 2 X 2 tables, use *Fisher's Exact Test (i.e. P(x=k) ~ B(n,p))* if your expected frequencies are less than 2. (Section 6.6)

Fisher's Exact Test: Description

- The Fisher's exact test calculates the *exact probability of the table* of observed cell frequencies given the following assumptions:
  - The null hypothesis of independence is true
  - The marginal totals of the observed table are fixed
- Calculation of the probability of the observed cell frequencies uses the factorial mathematical operation.
  - Factorial is notated by ! which means multiply the number by all integers smaller than the number
  - Example: 7! = 7*6*5*4*3*2*1 = 5040.

Fisher's Exact Test: Calculation

a	b	a+b
c	d	c+d
a+c	b+d	n

If margins of a table are fixed, the exact probability of a table with cells a,b,c,d and marginal totals (a+b), (c+d), (a+c), (b+d) =

$$\frac{(a+b)! * (c+d)! * (a+c)! * (b+d)!}{n! * a! * b! * c! * d!}$$

Fisher's Exact Test: Calculation Example

1	8	9
4	5	9
5	13	18

The exact probability of this table =

$$\frac{9! \ * \ 9! \ * \ 13! \ * \ 5!}{18! \ * \ 1! \ * \ 8! \ * \ 4! \ * \ 5!} = \frac{136080}{1028160} = 0.132$$

Probability for all possible tables with the same marginal totals

- The following slide shows the 6 possible tables for the observed marginal totals: 9, 9, 5, 13. The probability of each table is also given.
- The observed table is Table II
- The p-value for the Fisher's exact test is calculated by summing all probabilities less than or equal to the probability of the observed table.
- The probability is smallest for the tables (tables I and VI) that are least likely to occur by chance if the null hypothesis of independence is true.

Set of 6 possible tables with marginal totals: 9,9,5,13

I

0	9	9
5	4	9
5	13	18

Pr = 0.0147

II

1	8	9
4	5	9
5	13	18

Pr = 0.132

III

2	7	9
3	6	9
5	13	18

Pr = 0.353

IV

3	6	9
2	7	9
5	13	18

Pr = 0.353

V

4	5	9
1	8	9
5	13	18

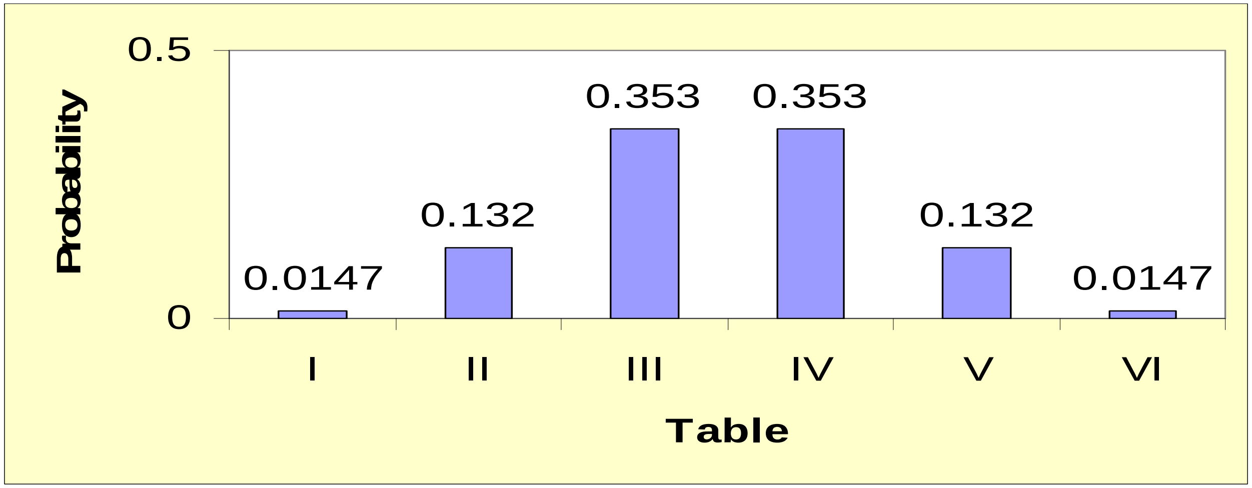
Pr = 0.132

VI

5	4	9
0	9	9
5	13	18

Pr = 0.0147

Fisher's Exact Test: p-value



The observed table (Table II) has probability = 0.132

P-value for the Fisher's exact test =
Pr (Table II) + Pr (Table V) + Pr (Table I) + Pr (Table VI)

= 0.132 + 0.132 + 0.0147 + 0.0147 = 0.293

Conclusion of Fisher's Exact test

- At significance level 0.05, the null hypothesis of independence is not rejected because the p-value of $0.294 > 0.05$.
- Looking back at the probabilities for each of the 6 tables, only Tables I and VI would result in a significant Fisher's exact test result:

  $p = 2*0.0147 = 0.0294$ for either of these tables.

  - This makes sense, intuitively, because these tables are least likely to occur by chance if the null hypothesis is true.

Fisher’s Exact test Calculator for a 2x2 table

This website will calculate the Fisher’s exact test p-value after you enter the cell counts for a 2 x 2 contingency table.

Use the p-value for the two-sided test.

www.graphpad.com/quickcalcs/contingency1.cfm

Tests for Categorical Data

- To compare proportions between two groups or to test for independence between two categorical variables, use the Chi-square test
- If more than 20% of the expected cell frequencies < 5, use the Fisher's exact test
- When categorical data are paired, the McNemar test is the appropriate test.

Comparing Proportions with Paired data

- When data are paired and the outcome of interest is a proportion, the ***McNemar Test*** is used to evaluate hypotheses about the data.
  - Developed by Quinn McNemar in 1947
  - Sometimes called the McNemar Chi-square test because the test statistic has a Chi-square distribution
- The McNemar test is only used for paired nominal data.
  - Use the Chi-square test for independence when nominal data are collected from independent groups.

Examples of Paired Data for Proportions

- Pair-Matched data can come from
  - Case-control studies where each case has a matching control (matched on age, gender, race, etc.)
  - Twins studies – the matched pairs are twins.
- Before - After data
  - the outcome is presence (+) or absence (-) of some characteristic measured on the same individual at two time points.

Summarizing the Data

- Like the Chi-square test, data need to be arranged in a contingency table before calculating the McNemar statistic
- The table will always be 2 X 2 but the cell frequencies are numbers of ‘pairs’ not numbers of individuals
- Examples for setting up the tables are in the following slides for
  - Case – Control paired data
  - Twins paired data: one exposed and one unexposed
  - Before – After paired data

Pair-Matched Data for Case-Control Study: outcome is exposure to some risk factor

		Control	
		Exposed	Unexposed
Case	Exposed	a	b
	Unexposed	c	d

The counts in the table for a case-control study are numbers of **pairs** not numbers of individuals.

Paired Data for Before-After Counts

	After treatment		
Before treatment		+	-
	+	a	b
	-	c	d

The counts in the table for a before-after study are numbers of **pairs** and number of individuals.

Null hypotheses for Paired Nominal data

- The null hypothesis for case-control pair matched data is that the proportion of subjects exposed to the risk factor is equal for cases and controls.
- The null hypothesis for twin paired data is that the proportions with the event are equal for exposed and unexposed twins
- The null hypothesis for before-after data is that the proportion of subjects with the characteristic (or event) is the same before and after treatment.

McNemar's test

- For any of the paired data Null Hypotheses the following are true if the null hypothesis is true:

  Ho: b = c

  Ho: b/(b+c) =0.5

- Cells 'b' and 'c' are called the *discordant* cells because they represent pairs with a difference

- Cells 'a' and 'd' are the *concordant* cells. These cells do not contribute any information about a difference between pairs or over time so they aren't used to calculate the test statistic.

McNemar Statistic

- The McNemar's Chi-square statistic is calculated using the counts in the 'b' and 'c' cells of the table:

$$\chi^2 = \frac{(b-c)^2}{b+c}$$

- Rule of thumb: b + c >=20
- If the null hypothesis is true the McNemar's Chi-square statistic = 0.

McNemar statistic distribution

- The sampling distribution of the McNemar statistic is a Chi-square distribution.

- For a test with alpha = 0.05, the critical value for the McNemar statistic = 3.84.

  - The null hypothesis is not rejected if the McNemar's statistic ___ 3.84.
  - The null hypothesis is rejected if the McNemar's statistic ____ 3.84.

P-value for McNemar statistic

- You can find the p-value for the McNemar statistic using the CHIDIST function in Excel
- Enter = CHIDIST(test statistic, 1) to obtain the p-value.
- If the test statistic is > 3.84, the p-value will be < 0.05 and the null hypothesis of equal proportions between pairs or over time will be rejected.

McNemar test Example

- Breast cancer patients receiving mastectomy followed by chemotherapy were matched to each other on age and cancer stage.

- By random assignment, one patient in each matched pair received chemo perioperatively and for an additional 6 months, while the other patient in each matched pair received chemo perioperatively only.

Data for Twin study in a table

	Periop. only		
		Survived 5 years	Died within 5 years
Periop. + 6 Months	Survived 5 years	510	17
	Died within 5 years	5	90

McNemar test hypotheses

Scientific Question: Does survival to 5 years differ by treatment group?

H_O:

H_A:

McNemar's test

- Check: b + c=
- Critical value for Chi-square distribution with 1 df = 3.84
- Calculate the test statistic

$$\chi^2 = \frac{(b-c)^2}{b+c} =$$

- P-value = P(χ^2>_____) = 0.01 =CHIDIST(___, __)

Decision and Conclusion

- Decision:
  - By the rejection region method:
  - By the p-value method:
- Conclusion:

McNemar test in EXCEL and online calculator

There is no EXCEL function or Data Analysis Tool for the McNemar Chi-square test.

This website will calculate the McMemar test statistic and p-value

http://www.graphpad.com/quickcalcs/McNemar1.cfm

What if b+c<20

- We know under the null

Ho: $b \sim B(b+c, 0.5)$

So, p-value for this test (one-sided)

$$p - value = P(b <= obs \mid b \sim B(b+c, 0.5))$$

Binomial test Example (Matched pairs)

- **Smoking status was examined over a two year period.**
- **In 2007, a sample of 212 adults, ages 18-22 years, was asked to identify themselves as smokers or nonsmokers.**
- **In 2009, the *same individuals* were again asked whether they were currently smokers or nonsmokers.**
- **Researchers want to know if a different number of participants switched from being smokers to nonsmokers and from nonsmokers to smokers.**

Changes in smoking status– The data

	2009	
2007	Smoker	Not Smoker
Smoker	62	10
Not Smoker	8	132

Hypotheses

Scientific Question: Did a different number of participants switch from being smokers to nonsmokers and from nonsmokers to smokers?

H_O:

H_A:

Binomial test (Matched pairs)

- Check: b + c =
- Calculate the p-value

$$p - value =$$

- P-value = 2*BINOMDIST(x, n, pi , cum?) = = 2*BINOMDIST(__,___, ___ , ___) = 2*0.41=0.82

Decision and Conclusion

- Decision:
- Conclusion:

Yogurt

Participants in a nutrition survey each answered *two* questions:

"Does whole fat yogurt taste good or bad to you?"

"Does non-fat yogurt taste good or bad to you?"

Yogurt

	Non-fat	
Whole Fat	Good	Bad
Good	55	12
Bad	5	28

How many participants were involved in this study?

1. 17
2. 83
3. 50
4. 100
5. 200

How many questions were asked?

1. 17
2. 83
3. 50
4. 100
5. 200

How many pairs of answers?

1. 17
2. 83
3. 50
4. 100
5. 200

How many discordant pairs?

1. 17
2. 83
3. 50
4. 100
5. 200

What is the shape of the test statistic under the null of no difference in response?

1. Chi-square with 1 degree of freedom
2. McNemar's Chi-Square
3. Binomial(17, ½)
4. Normal (μ, σ)
5. Cannot be determined.

Suppose you were given this table for the Yogurt study.

	Yogurt Type	
Taste	Whole Fat	Non-fat
Good	67	60
Bad	33	40

Using the table on the previous slide, can you answer the scientific question appropriately?

1. No
2. Yes

Death certificate accuracy

- Across two different hospitals, the results of 575 autopsies were compared to the cause of death listed on the death certificates.
- Based on this comparison, the death certificates were classified as accurate or inaccurate.
- One hospital was a community based hospital, the other a university hospital.
- Is type of hospital independent of accuracy of cause of death listed on the death certificate?

Death certificate accuracy – The data

	Accurate Death Certificate	Inaccurate Death Certificate
Community Hospital	157	72
University Hospital	268	78

What is the shape of the test statistic under the null of no difference in response?

1. Chi-square with 1 degree of freedom
2. McNemar's Chi-Square
3. Binomial(17, ½)
4. Normal (μ, σ)
5. Cannot be determined.

Readings and Assignments

- Reading
  - Chapter 6 pgs 149 – 153
  - Chapter 5 pgs 119-121
- Work through the Lesson 12 Practice Exercises
- Lesson 12 Excel Module