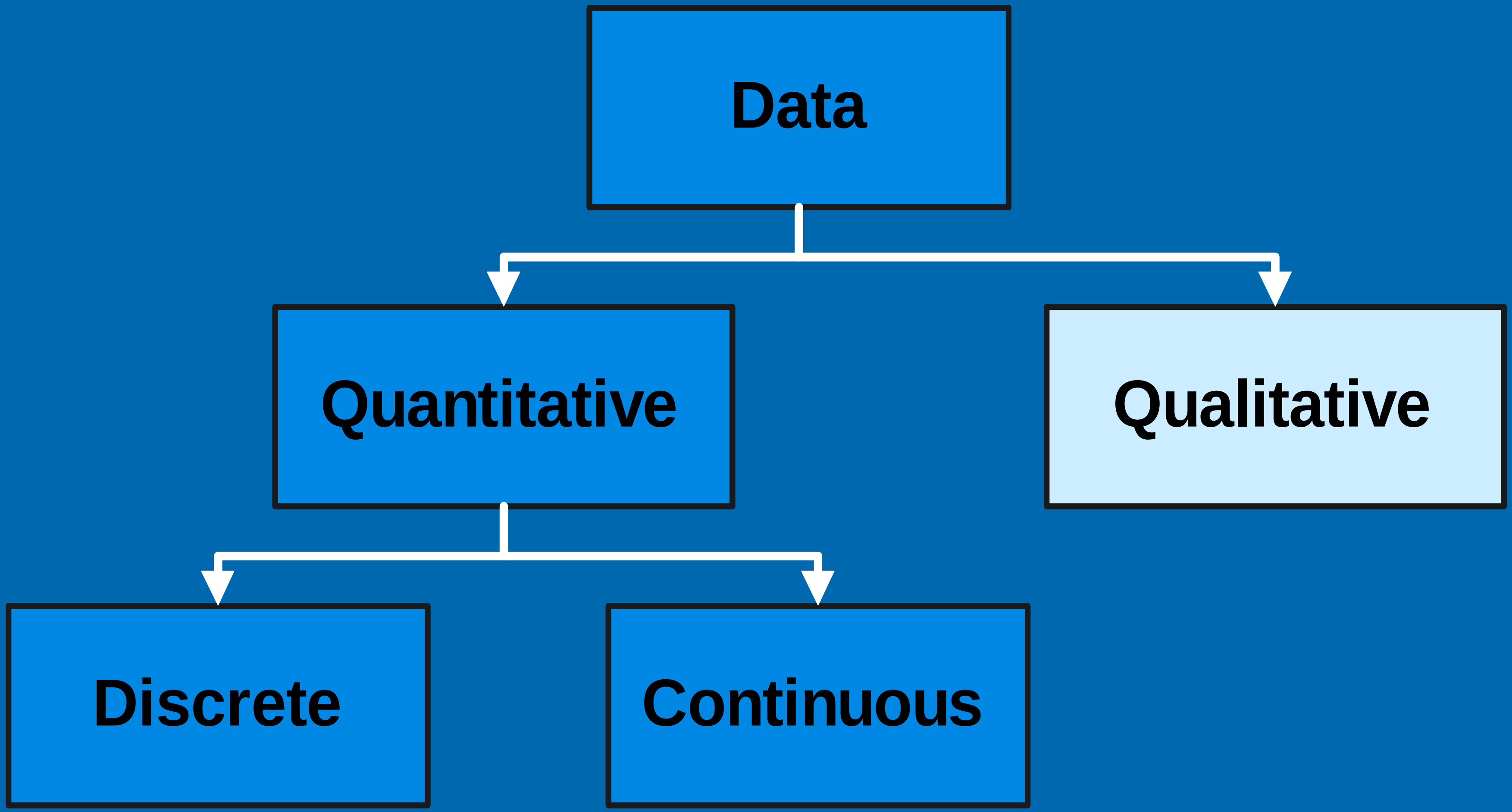
Lecture 10

Hypothesis testing: Categorical Data Analysis

Learning Objectives

1. Comparison of binomial proportion using Z and χ^2 Test.
2. Explain χ^2 Test for Independence of 2 variables
3. Explain The Fisher's test for independence
4. McNemar's tests for correlated data
5. Kappa Statistic
6. Use of Computer Program

Data Types



Data

Quantitative

Qualitative

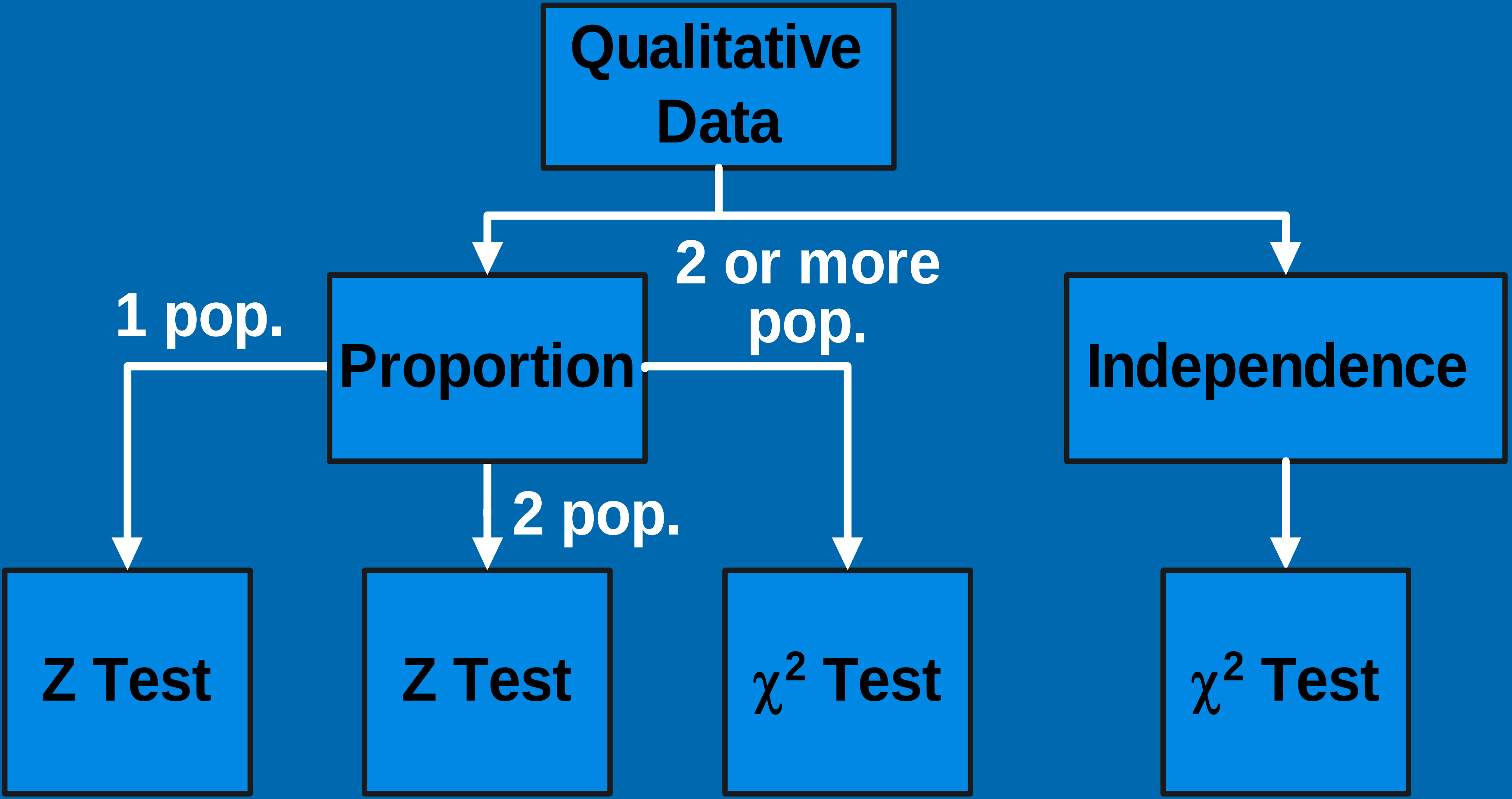
Discrete

Continuous

Qualitative Data

1. **Qualitative Random Variables Yield Responses That Can Be Put In Categories. Example: Sex (Male, Female)**
2. **Measurement or Count Reflect # in Category**
3. **Nominal (no order) or Ordinal Scale (order)**
4. **Data can be collected as continuous but recoded to categorical data. Example (Systolic Blood Pressure - Hypotension, Normal tension, hypertension)**

Hypothesis Tests Qualitative Data



Qualitative Data

Proportion

Independence

1 pop.

2 pop.

2 or more pop.

Z Test

Z Test

χ^2 Test

χ^2 Test

Z Test for Differences in Two Proportions

Hypotheses for Two Proportions

Hypotheses for Two Proportions

Hypothesis	Research Questions		
	No Difference Any Difference	Pop 1 $\geq$ Pop 2 Pop 1 < Pop 2	Pop 1 $\leq$ Pop 2 Pop 1 > Pop 2
H_0			
H_a			

Hypotheses for Two Proportions

Hypothesis	Research Questions		
	No Difference Any Difference	Pop 1 ≥ Pop 2 Pop 1 < Pop 2	Pop 1 ≤ Pop 2 Pop 1 > Pop 2
H_0	$p_1 - p_2 = 0$		
H_a	$p_1 - p_2 \neq 0$		

Hypotheses for Two Proportions

Hypothesis	Research Questions		
	No Difference Any Difference	Pop 1 ≥ Pop 2 Pop 1 < Pop 2	Pop 1 ≤ Pop 2 Pop 1 > Pop 2
H_0	$p_1 - p_2 = 0$	$p_1 - p_2 \geq 0$	
H_a	$p_1 - p_2 \neq 0$	$p_1 - p_2 < 0$	

sawilopo@yahoo.com

Hypotheses for Two Proportions

Hypothesis	Research Questions		
	No Difference Any Difference	Pop 1 ≥ Pop 2 Pop 1 < Pop 2	Pop 1 ≤ Pop 2 Pop 1 > Pop 2
H_0	$p_1 - p_2 = 0$	$p_1 - p_2 \geq 0$	$p_1 - p_2 \leq 0$
H_a	$p_1 - p_2 \neq 0$	$p_1 - p_2 < 0$	$p_1 - p_2 > 0$

Hypotheses for Two Proportions

Hypothesis	Research Questions		
	No Difference Any Difference	Pop 1 ≥ Pop 2 Pop 1 < Pop 2	Pop 1 ≤ Pop 2 Pop 1 > Pop 2
H_0	$p_1 - p_2 = 0$	$p_1 - p_2 \geq 0$	$p_1 - p_2 \leq 0$
H_a	$p_1 - p_2 \neq 0$	$p_1 - p_2 < 0$	$p_1 - p_2 > 0$

Z Test for Difference in Two Proportions

1. Assumptions
   - Populations Are Independent
   - Populations Follow Binomial Distribution
   - Normal Approximation Can Be Used for large samples (All Expected Counts ≥ 5)
2. Z-Test Statistic for Two Proportions

$$Z \cong \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\hat{p} \cdot (1 - \hat{p}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \qquad \text{where } \hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

Sample Distribution for Difference Between Proportions

$$\overline{X_1} - \overline{X_2} \;\sim \mathrm{N}\left(\mu_1 - \mu_2; \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$$

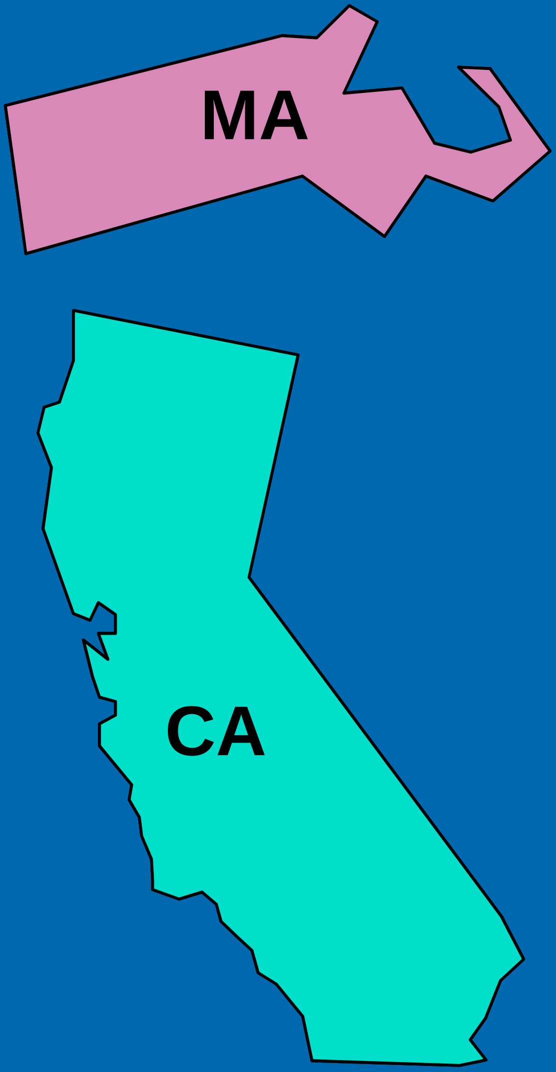
$$\overline{p_1} - \overline{p_2} \;\; \cong \mathrm{N}\left(p_1 - p_2; \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}\right)$$

$$\cong \mathrm{N}\left(0; \sqrt{\overline{p}\overline{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}\right) \quad under \;\; H_0 : \;\; p_1 = p_2$$

$$\overline{p} = \frac{x_1 + x_2}{n_1 + n_2},$$

Z Test for Two Proportions Thinking Challenge

- You're an epidemiologist for the Department of Health in your country. You're studying the prevalence of disease X in two provinces (MA and CA). In **MA**, **74** of **1500** people surveyed were diseased and in **CA**, **129** of **1500** were diseased. At **.05** level, does **MA** have a **lower** prevalence rate?



Z Test for Two Proportions Solution*

Z Test for Two Proportions Solution*

H_0:

H_a:

α =

n_{MA} = n_{CA} =

Critical Value(s):

Test Statistic:

Decision:

Conclusion:

Z Test for Two Proportions Solution*

H_0: $p_{MA} - p_{CA} = 0$

H_a: $p_{MA} - p_{CA} < 0$

$\alpha =$

$n_{MA} =$ $n_{CA} =$

Critical Value(s):

Test Statistic:

Decision:

Conclusion:

Z Test for Two Proportions Solution*

H_0: $p_{MA} - p_{CA} = 0$

H_a: $p_{MA} - p_{CA} < 0$

$\alpha = .05$

$n_{MA} = 1500$ $n_{CA} = 1500$

Critical Value(s):

Test Statistic:

Decision:

Conclusion:

Z Test for Two Proportions Solution*

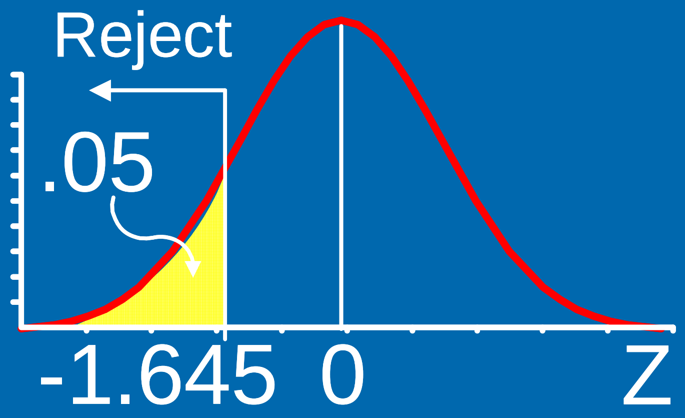
H0: $p_{MA} - p_{CA} = 0$

Ha: $p_{MA} - p_{CA} < 0$

$\alpha = .05$

$n_{MA} = 1500$ $n_{CA} = 1500$

Critical Value(s):



Test Statistic:

Decision:

Conclusion:

Z Test for Two Proportions Solution*

$$\hat{p}_{MA} = \frac{X_{MA}}{n_{MA}} = \frac{74}{1500} = .0493 \qquad \hat{p}_{CA} = \frac{X_{CA}}{n_{CA}} = \frac{129}{1500} = .0860$$

$$\hat{p} = \frac{X_{MA} + X_{CA}}{n_{MA} + n_{CA}} = \frac{74+129}{1500+1500} = .0677$$

$$Z \cong \frac{(.0493 - .0860) - (0)}{\sqrt{(.0677)\cdot(1-.0677)\cdot\left(\frac{1}{1500}+\frac{1}{1500}\right)}}$$

$$= -4.00$$

Z Test for Two Proportions Solution*

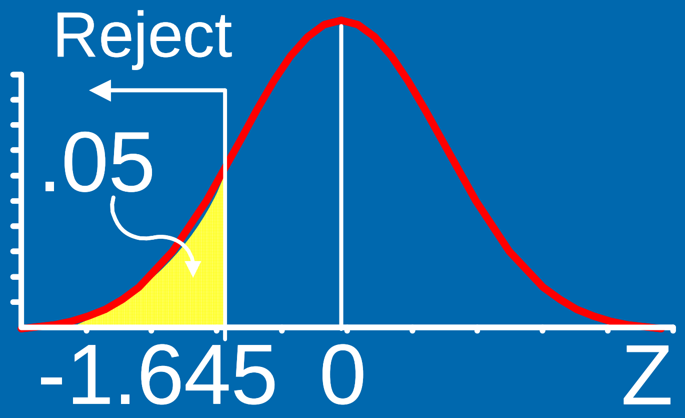
H_0: $p_{MA} - p_{CA} = 0$

H_a: $p_{MA} - p_{CA} < 0$

$\alpha = .05$

$n_{MA} = 1500 \quad n_{CA} = 1500$

Critical Value(s):



Test Statistic:

$Z = -4.00$

Decision:

Conclusion:

Z Test for Two Proportions Solution*

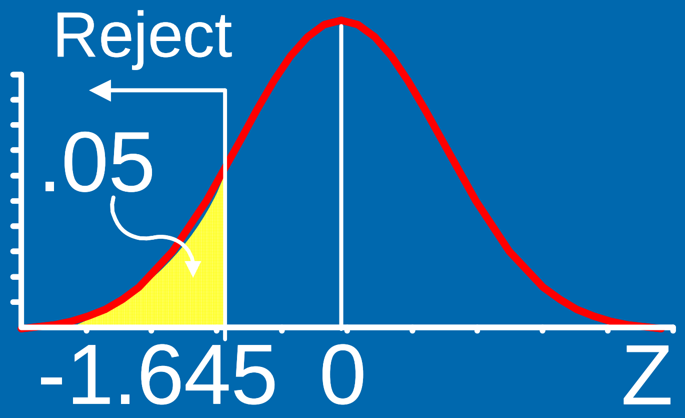
H_0: $p_{MA} - p_{CA} = 0$

H_a: $p_{MA} - p_{CA} < 0$

$\alpha = .05$

$n_{MA} = 1500$ $n_{CA} = 1500$

Critical Value(s):



Test Statistic:

$Z = -4.00$

Decision:

Reject at $\alpha = .05$

Conclusion:

Z Test for Two Proportions Solution*

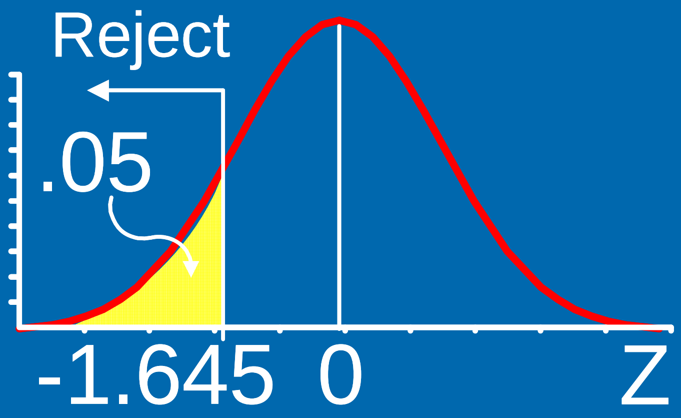
H0: $p_{MA} - p_{CA} = 0$

Ha: $p_{MA} - p_{CA} < 0$

$\alpha = .05$

$n_{MA} = 1500$ $n_{CA} = 1500$

Critical Value(s):



Test Statistic:

$Z = -4.00$

Decision:

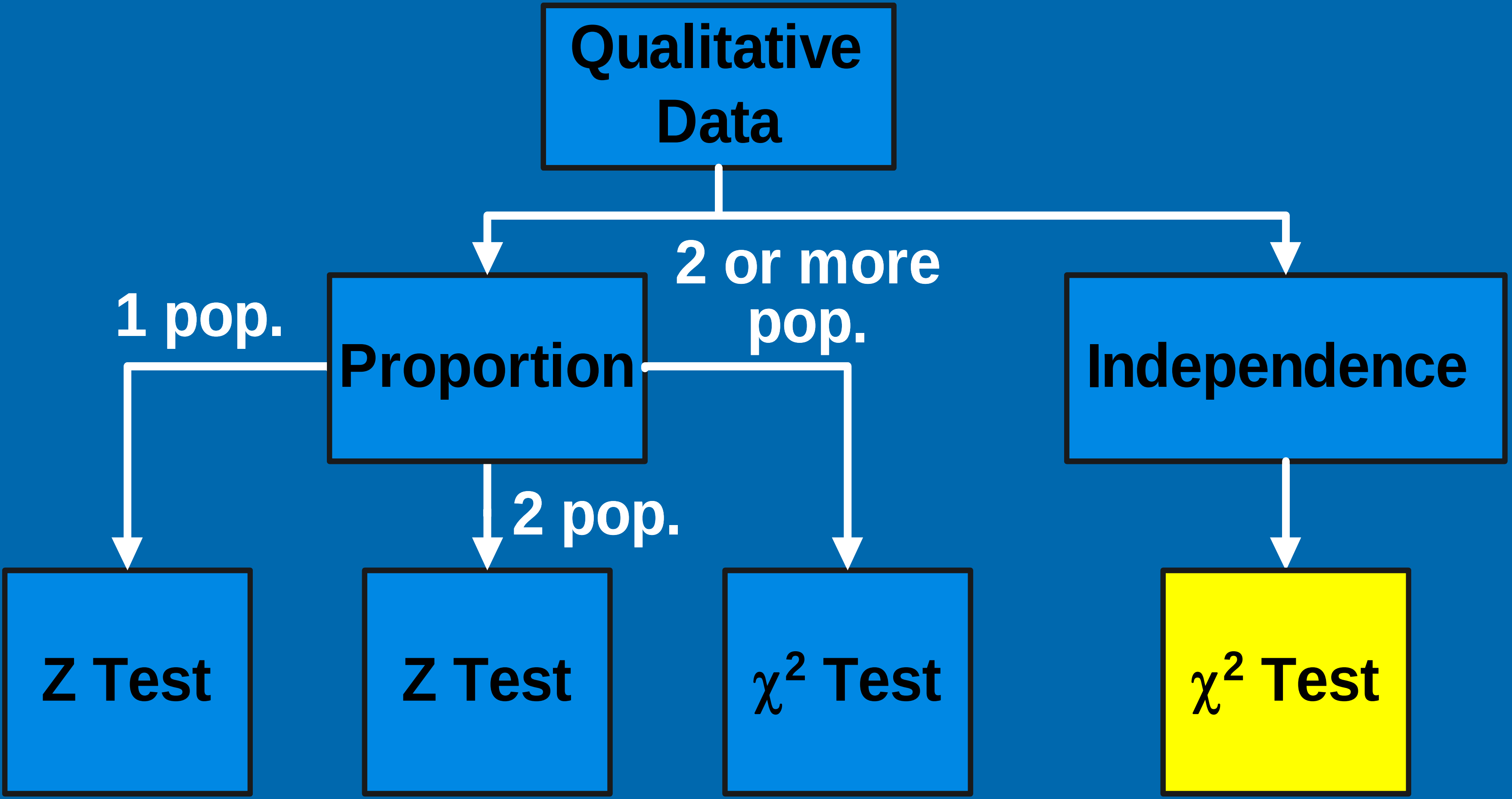
Reject at $\alpha = .05$

Conclusion:

There is evidence MA is less than CA

χ^2 Test of Independence Between 2 Categorical Variables

Hypothesis Tests Qualitative Data



Qualitative Data

Proportion

Independence

1 pop.

2 pop.

2 or more pop.

Z Test

Z Test

χ^2 Test

χ^2 Test

χ^2 Test of Independence

1. **Shows If a Relationship Exists Between 2 Qualitative Variables, but does Not Show Causality**
2. **Assumptions**
   - Multinomial Experiment
   - All Expected Counts ≥ 5
3. **Uses Two-Way Contingency Table**

χ^2 Test of Independence Contingency Table

- 1. Shows # Observations From 1 Sample Jointly in 2 Qualitative Variables

χ^2 Test of Independence Contingency Table

1. Shows # Observations From 1 Sample Jointly in 2 Qualitative Variables

Levels of variable 2

	Residence		
Disease Status	Urban	Rural	Total
Disease	63	49	112
No disease	15	33	48
Total	78	82	160

Levels of variable 1

χ^2 Test of Independence Hypotheses & Statistic

1. Hypotheses
   - H_0: Variables Are Independent
   - H_a: Variables Are Related (Dependent)

χ^2 Test of Independence Hypotheses & Statistic

1. **Hypotheses**

   H_0: Variables Are Independent

   H_a: Variables Are Related (Dependent)

2. **Test Statistic**

Observed count

$$\chi^2 = \sum_{\text{all cells}} \frac{\left[n_{ij} - \hat{E}(n_{ij})\right]^2}{\hat{E}(n_{ij})}$$

Expected count

χ^2 Test of Independence Hypotheses & Statistic

1. **Hypotheses**

   H_0: Variables Are Independent

   H_a: Variables Are Related (Dependent)

2. **Test Statistic**

Observed count

$$\chi^2 = \sum_{\text{all cells}} \frac{\left[n_{ij} - \hat{E}(n_{ij})\right]^2}{\hat{E}(n_{ij})}$$

Expected count

Rows Columns

Degrees of Freedom: (r - 1)(c - 1)

χ^2 Test of Independence Expected Counts

1. Statistical Independence Means Joint Probability Equals Product of Marginal Probabilities
2. Compute Marginal Probabilities & Multiply for Joint Probability
3. Expected Count Is Sample Size Times Joint Probability

Expected Count Example

Expected Count Example

Marginal probability = $\frac{112}{160}$

	Residence		
Disease Status	Urban Obs.	Rural Obs.	Total
Disease	63	49	112
No Disease	15	33	48
Total	78	82	160

Expected Count Example

Marginal probability = $\frac{112}{160}$

	Residence		
Disease Status	Urban Obs.	Rural Obs.	Total
Disease	63	49	112
No Disease	15	33	48
Total	78	82	160

Marginal probability = $\frac{78}{160}$

Expected Count Example

Joint probability = $\frac{112}{160}\frac{78}{160}$

Marginal probability = $\frac{112}{160}$

	Residence		
Disease Status	Urban Obs.	Rural Obs.	Total
Disease	63	49	112
No Disease	15	33	48
Total	78	82	160

Marginal probability = $\frac{78}{160}$

Expected Count Example

Joint probability = $\frac{112}{160}\,\frac{78}{160}$

Marginal probability = $\frac{112}{160}$

Disease Status	Residence Urban Obs.	 Rural Obs.	Total
Disease	63	49	112
No Disease	15	33	48
Total	78	82	160

Marginal probability = $\frac{78}{160}$

Expected count = $\cancel{160}\cdot\frac{112}{\cancel{160}}\,\frac{78}{160}$

= 54.6

Expected Count Calculation

Expected Count Calculation

$$\text{Expected count} = \frac{(\text{Row total})\cdot(\text{Column total})}{\text{Sample size}}$$

Expected Count Calculation

$$\text{Expected count} = \frac{(\text{Row total}) \cdot (\text{Column total})}{\text{Sample size}}$$

$\frac{112x78}{160}$

$\frac{112x82}{160}$

Disease Status	Residence				Total
	Urban		Rural		
	Obs.	Exp.	Obs.	Exp.	
Disease	63	54.6	49	57.4	112
No Disease	15	23.4	33	24.6	48
Total	78	78	82	82	160

$\frac{48x78}{160}$

$\frac{48x82}{160}$

χ^2 Test of Independence Example on HIV

- You randomly sample 286 sexually active individuals and collect information on their HIV status and History of STDs. At the .05 level, is there evidence of a relationship?

	HIV		
STDs Hx	No	Yes	Total
No	84	32	116
Yes	48	122	170
Total	132	154	286

χ^2 Test of Independence Solution

χ^2 Test of Independence Solution

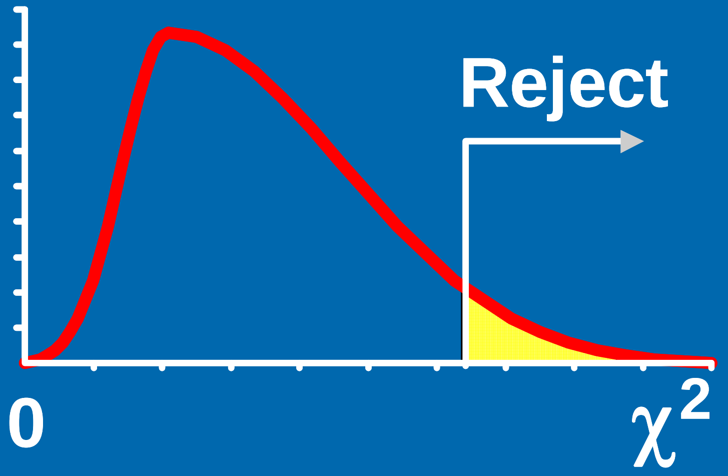
H_0:

H_a:

α =

df =

Critical Value(s):



Test Statistic:

Decision:

Conclusion:

χ^2 Test of Independence Solution

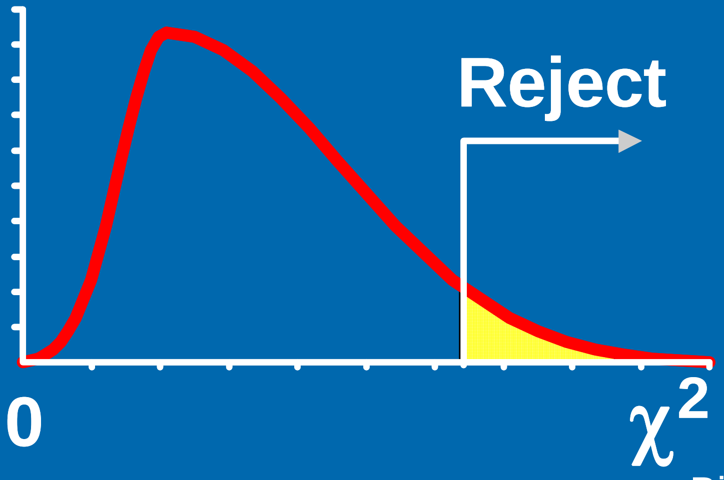
H_0: No Relationship

H_a: Relationship

α =

df =

Critical Value(s):



Test Statistic:

Decision:

Conclusion:

χ^2 Test of Independence Solution

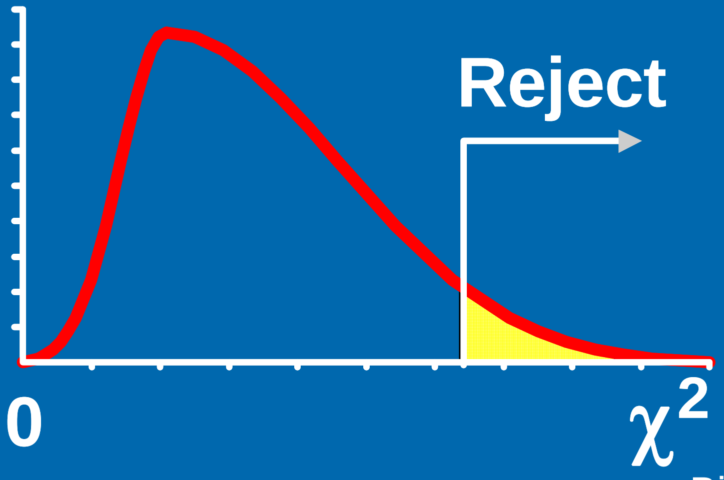
H_0: No Relationship

H_a: Relationship

$\alpha = .05$

df = (2 - 1)(2 - 1) = 1

Critical Value(s):



Test Statistic:

Decision:

Conclusion:

χ^2 Test of Independence Solution

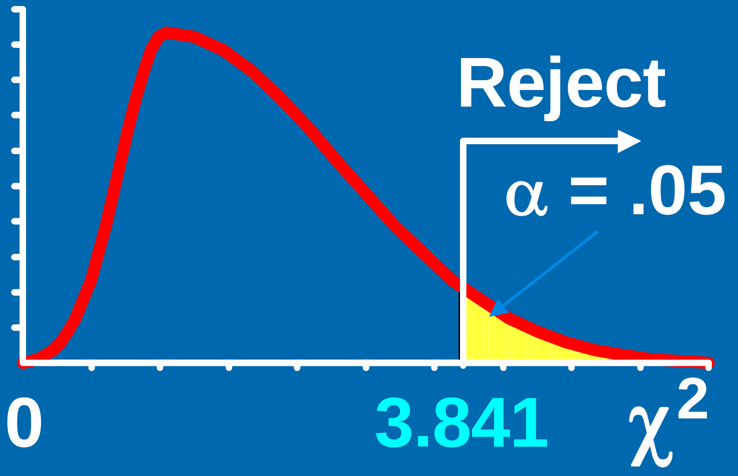
H_0: No Relationship

H_a: Relationship

$\alpha = .05$

df = (2 - 1)(2 - 1) = 1

Critical Value(s):



Test Statistic:

Decision:

Conclusion:

χ^2 Test of Independence Solution

✓ $E(n_{ij}) \geq 5$ in all cells

$\frac{116 \times 132}{286}$

$\frac{154 \times 116}{286}$

	HIV				
	No		Yes		
STDs HX	Obs.	Exp.	Obs.	Exp.	Total
No	84	53.5	32	62.5	116
Yes	48	78.5	122	91.5	170
Total	132	132	154	154	286

$\frac{170 \times 132}{286}$

$\frac{170 \times 154}{286}$

χ^2 Test of Independence Solution

$$\chi^2 = \sum_{\text{all cells}} \frac{\left[n_{ij} - \hat{E}\left(n_{ij}\right)\right]^2}{\hat{E}\left(n_{ij}\right)}$$

$$= \frac{\left[n_{11} - \hat{E}\left(n_{11}\right)\right]^2}{\hat{E}\left(n_{11}\right)} + \frac{\left[n_{12} - \hat{E}\left(n_{12}\right)\right]^2}{\hat{E}\left(n_{12}\right)} + \cdots + \frac{\left[n_{22} - \hat{E}\left(n_{22}\right)\right]^2}{\hat{E}\left(n_{22}\right)}$$

$$= \frac{\left[84 - 53.5\right]^2}{53.5} + \frac{\left[32 - 62.5\right]^2}{62.5} + \cdots + \frac{\left[122 - 91.5\right]^2}{91.5} = 54.29$$

χ^2 Test of Independence Solution

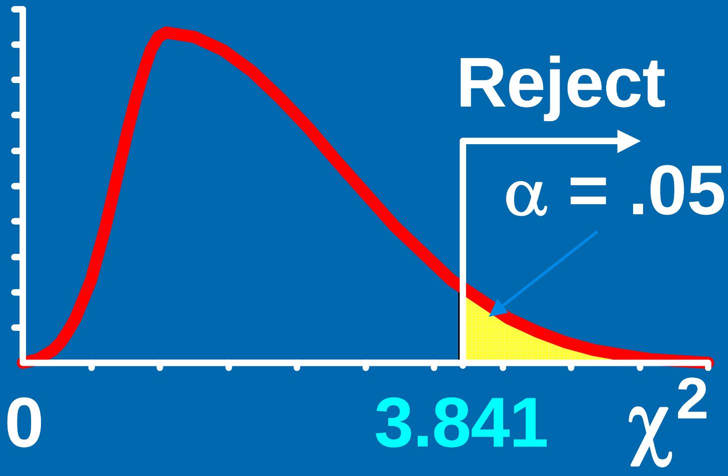
H_0: No Relationship

H_a: Relationship

$\alpha = .05$

df = (2 - 1)(2 - 1) = 1

Critical Value(s):



Test Statistic:

$\chi^2 = 54.29$

Decision:

Conclusion:

χ^2 Test of Independence Solution

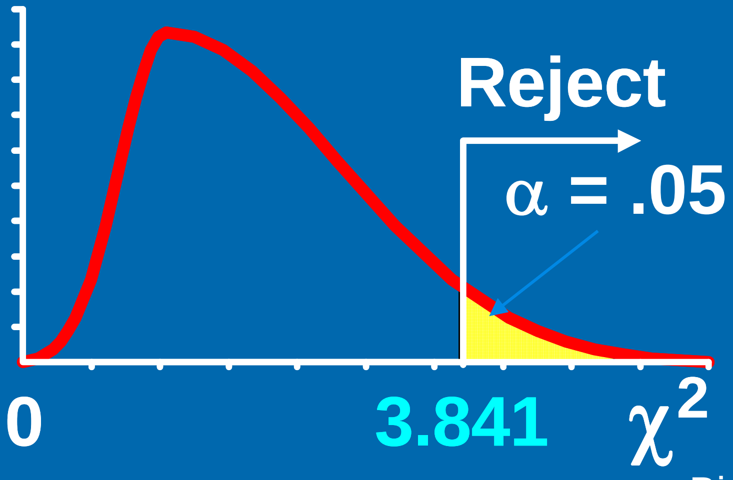
H_0: No Relationship

H_a: Relationship

α = .05

df = (2 - 1)(2 - 1) = 1

Critical Value(s):



Test Statistic:

$$\chi^2 = 54.29$$

Decision:

Reject at α = .05

Conclusion:

χ^2 Test of Independence Solution

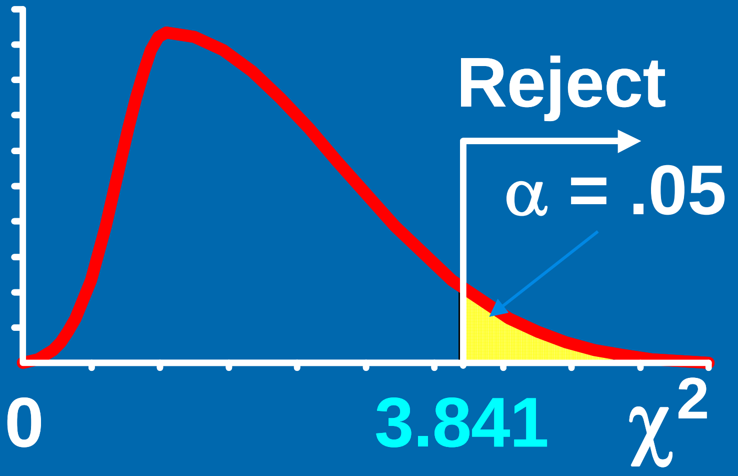
H_0: No Relationship

H_a: Relationship

$\alpha = .05$

df = (2 - 1)(2 - 1) = 1

Critical Value(s):



Test Statistic:

$\chi^2 = 54.29$

Decision:

Reject at $\alpha = .05$

Conclusion:

There is evidence of a relationship

STATA

tabi 30 18 38 \ 13 7 22, chi2

```
           |                col
       row |         1          2          3 |     Total
-----------+---------------------------------+----------
         1 |        30         18         38 |        86
         2 |        13          7         22 |        42
-----------+---------------------------------+----------
     Total |        43         25         60 |       128

          Pearson chi2(2) =   0.7967   Pr = 0.671
```

χ^2 Test of Independence

SAS CODES

```
Data dis;
input STDs HIV count;
cards;
1 1  84
1 2  32
2 1  48
2 2 122
;
run;

Proc freq data=dis order=data;
     weight Count;
     tables STDs*HIV/chisq;
run;
```

sawilopo@yahoo.com

χ^2 Test of Independence
SAS OUTPUT

```
      Statistics for Table of STDs by HIV

Statistic                     DF       Value      Prob
------------------------------------------------------
 Chi-Square                    1     54.1502    <.0001
Likelihood Ratio Chi-Square    1     55.7826    <.0001
Continuity Adj. Chi-Square     1     52.3871    <.0001
Mantel-Haenszel Chi-Square     1     53.9609    <.0001
Phi Coefficient                       0.4351
Contingency Coefficient               0.3990
Cramer's V                            0.4351
```

Fisher's Exact Test

- Fisher's Exact Test is a test for independence in a 2 X 2 table. It is most useful when the total sample size and the expected values are small. The test holds the marginal totals fixed and computes the hypergeometric probability that n_{11} is at least as large as the observed value
- Useful when E(cell counts) < 5.

Hypergeometric distribution

- Example: 2x2 table with cell counts a, b, c, d. Assuming marginal totals are fixed:

  M1 = a+b, M2 = c+d, N1 = a+c, N2 = b+d.

  for convenience assume N1<N2, M1<M2.

  possible value of a are: 0, 1, ...min(M1,N1).
- Probability distribution of cell count a follows a hypergeometric distribution:

  N = a + b + c + d = N1 + N2 = M1 + M2
  - Pr (x=a) = N1!N2!M1!M2! / [N!a!b!c!d!]
  - Mean (x) = M1N1/ N
  - Var (x) = M1M2N1N2 / [N^2(N-1)]
- Fisher exact test is based on this hypergeometric distr.

Fisher's Exact Test Example

HIV Infection

Hx of STDs

	yes	no	total
yes	3	7	10
no	5	10	15
total	8	17	25

- **Is HIV Infection related to Hx of STDs in Sub Saharan African Countries? Test at 5% level.**

Hypergeometric prob.

- Probability of observing this specific table given fixed marginal totals is

  Pr (3,7, 5, 10) = 10!15!8!17!/[25!3!7!5!10!]

  = 0.3332
- Note the above is not the p-value. Why?
- Not the accumulative probability, or not the tail probability.
- Tail prob = sum of all values (a = 3, 2, 1, 0).

Hypergeometric prob

Pr (2, 8, 6, 9) = 10!15!8!17!/[25!2!8!6!9!]
= 0.2082

Pr (1, 9, 7, 8) = 10!15!8!17!/[25!1!9!7!8!]
= 0.0595

Pr (0,10, 8, 7) = 10!15!8!17!/[25!0!10!8!7!]
= 0.0059

Tail prob = .3332+.2082+.0595+.0059 = .6068

. tabi 3 7 \ 5 10, row replace chi exact

```
+-------------------+
| Key               |
|-------------------|
|     frequency     |
|  row percentage   |
+-------------------+

           |          col
       row |         1          2 |     Total
-----------+----------------------+----------
         1 |         3          7 |        10 
           |     30.00      70.00 |    100.00 
-----------+----------------------+----------
         2 |         5         10 |        15 
           |     33.33      66.67 |    100.00 
-----------+----------------------+----------
     Total |         8         17 |        25 
           |     32.00      68.00 |    100.00 

          Pearson chi2(1) =   0.0306   Pr = 0.861
           Fisher's exact =                 1.000
   1-sided Fisher's exact =                 0.607
```

Fisher's Exact Test
SAS Codes

```
Data dis;
input STDs $ HIV $ count;
cards;
no no   10
No Yes   5
yes no   7
yes yes  3
;
run;

proc freq data=dis order=data;
      weight Count;
      tables STDs*HIV/chisq fisher;
run;
```

Pearson Chi-squares test
Yates correction

- **Pearson Chi-squares test**

  $\chi^2 = \sum_i (O_i\text{-}E_i)^2/E_i$ follows a chi-squares distribution with df = (r-1)(c-1)

  if $E_i \geq 5$.

- Yates correction for more accurate p-value

  $\chi^2 = \sum_i (|O_i\text{-}E_i| - 0.5)^2/E_i$

  when O_i and E_i are close to each other.

Fisher's Exact Test SAS Output

```
 Statistics for Table of STDs by HIV

Statistic                     DF       Value      Prob
------------------------------------------------------
Chi-Square                     1      0.0306    0.8611
Likelihood Ratio Chi-Square    1      0.0308    0.8608
Continuity Adj. Chi-Square     1      0.0000    1.0000
Mantel-Haenszel Chi-Square     1      0.0294    0.8638
Phi Coefficient                      -0.0350
Contingency Coefficient               0.0350
Cramer's V                           -0.0350

 WARNING: 50% of the cells have expected counts less
          than 5. Chi-Square may not be a valid test.

                Fisher's Exact Test
          ----------------------------------
          Cell (1,1) Frequency (F)        10
          Left-sided Pr <= F          0.6069
          Right-sided Pr >= F         0.7263

          Table Probability (P)       0.3332
          Two-sided Pr <= P           1.0000
```

Fisher's Exact Test

- The output consists of three p-values:
- Left: Use this when the alternative to independence is that there is negative association between the variables. That is, the observations tend to lie in lower left and upper right.
- Right: Use this when the alternative to independence is that there is positive association between the variables. That is, the observations tend to lie in upper left and lower right.
- 2-Tail: Use this when there is no prior alternative.

Useful Measures of Association - Nominal Data

- **Cohen’s Kappa (κ)**
  - **Also referred to as Cohen’s General Index of Agreement.**
  - **It was originally developed to assess the degree of agreement between two judges or raters assessing n items on the basis of a nominal classification for 2 categories.**
  - **Subsequent work by Fleiss and Light presented extensions of this statistic to more than 2 categories.**

sawilopo@yahoo.com

Useful Measures of Association - Nominal Data

- Cohen's Kappa (κ)

Inspector A

		Good	Bad	
Inspector 'B'	Good	.33 (a)	.07 (b)	**.40 (p_B)**
	Bad	.13 (c)	.47 (d)	**.60 (q_B)**
		.46 (p_A)	**.54 (q_A)**	**1.00**

Useful Measures of Association - Nominal Data

- **Cohen's Kappa (κ)**
  - **Cohen's κ requires that we calculate two values:**
    - p_o : the proportion of cases in which agreement occurs. In our example, this value equals 0.80.
    - Pe : the proportion of cases in which agreement would have been expected due purely to chance, based upon the marginal frequencies; where

$$p_e = p_A p_B + q_A q_B = 0.508 \text{ for our data}$$

Useful Measures of Association - Nominal Data

- **Cohen's Kappa (κ)**
  - **Then, Cohen's κ measures the agreement between two variables and is defined by**

$$\kappa = \frac{p_o - p_e}{1 - p_e} = 0.593$$

Useful Measures of Association - Nominal Data

- **Cohen's Kappa (κ)**
  - To test the Null Hypothesis that the true kappa $\kappa = 0$, we use the Standard Error:

$$\sigma_\kappa = \frac{1}{(1 - p_e)\sqrt{n}} \sqrt{p_e + p_e^2 - \sum_{i=1}^{k} p_{i.}p_{.i}\,(p_{i.} + p_{.i})}$$

  - then $z = \kappa/\sigma_\kappa \sim N(0,1)$

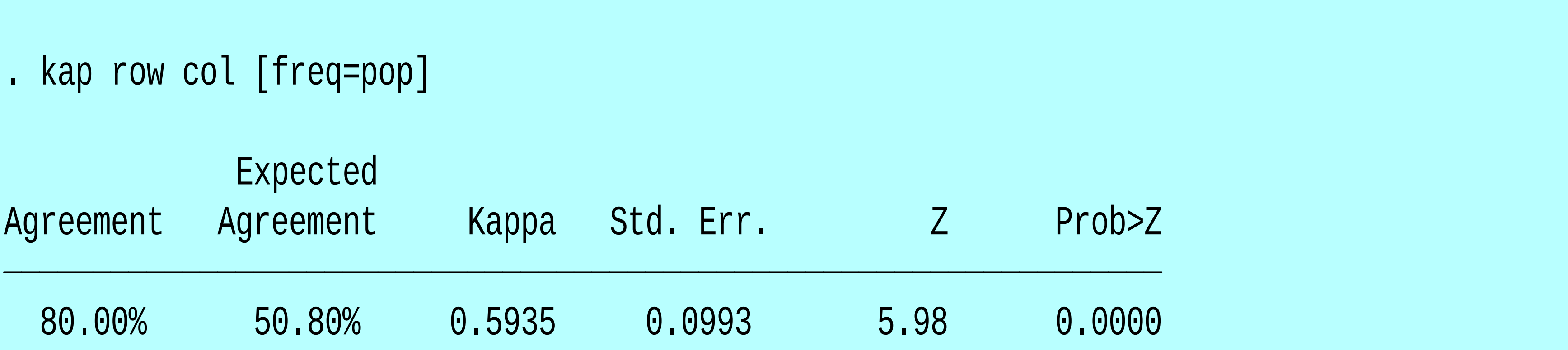
where $p_{i.}$ & $p_{.i}$ refer to row and column proportions (in textbook, $a_i = p_{i.}$ & $b_i = p_{.i}$)

. tabi 33 7 \ 13 47 , replace chi exact

```
. tabi 33   7   \   13   47 , replace chi exact

           |          col
       row |         1          2 |     Total
-----------+----------------------+----------
         1 |        33          7 |        40 
         2 |        13         47 |        60 
-----------+----------------------+----------
     Total |        46         54 |       100 

          Pearson chi2(1) =  35.7555   Pr = 0.000
           Fisher's exact =                 0.000
   1-sided Fisher's exact =                 0.000
```

```
. lis in 1/4


     +-----------------+
     | row   col   pop |
     |-----------------|
  1. |   1     1    33 |
  2. |   1     2     7 |
```

kap row col [freq=pop]



```
. kap row col [freq=pop]

             Expected
Agreement   Agreement     Kappa   Std. Err.         Z      Prob>Z
-----------------------------------------------------------------
  80.00%      50.80%     0.5935     0.0993       5.98      0.0000
```

Useful Measures of Association - Nominal Data- SAS CODES

```
Data kap;
input B $ A $ prob;
n=100;
count=prob*n;
cards;
Good Good .33
Good Bad .07
Bad Good .13
Bad Bad .47
;
run;

proc freq data=kap order=data;
     weight Count;
     tables B*A/chisq;
    test kappa;
run;
```

Useful Measures of Association - Nominal Data- SAS OUTPUT

```
      The FREQ Procedure

 Statistics for Table of B by A

    Simple Kappa Coefficient
--------------------------------
Kappa                     0.5935
ASE                       0.0806
95% Lower Conf Limit      0.4356
95% Upper Conf Limit      0.7514

     Test of H0: Kappa = 0

ASE under H0              0.0993
Z                         5.9796
One-sided Pr >  Z         <.0001
Two-sided Pr > |Z|        <.0001

        Sample Size = 100
```

McNemar's Test for Correlated (Dependent) Proportions

McNemar's Test for Correlated (Dependent) Proportions

Basis / Rationale for the Test

- **The approximate test previously presented for assessing a difference in proportions is based upon the assumption that the two samples are independent.**
- **Suppose, however, that we are faced with a situation where this is not true. Suppose we randomly-select 100 people, and find that 20% of them have flu. Then, imagine that we apply some type of treatment to all sampled peoples; and on a post-test, we find that 20% have flu.**

McNemar's Test for Correlated (Dependent) Proportions

- **We might be tempted to suppose that no hypothesis test is required under these conditions, in that the 'Before' and 'After' *p* values are identical, and would surely result in a test statistic value of 0.00.**
- **The problem with this thinking, however, is that the two sample *p* values are dependent, in that each person was assessed twice. It *is* possible that the 20 people that had flu originally still had flu. It is also possible that the 20 people that had flu on the second test *were a completely different set* of 20 people!**

McNemar's Test for Correlated (Dependent) Proportions

- It is for precisely this type of situation that McNemar's Test for Correlated (Dependent) Proportions is applicable.

- McNemar's Test employs two unique features for testing the two proportions:

  * a special fourfold contingency table; with a

  * special-purpose chi-square (X^2) test statistic (the approximate test).

McNemar's Test for Correlated (Dependent) Proportions

Nomenclature for the Fourfold (2 x 2) Contingency Table

A	B	(A + B)
C	D	(C + D)
(A + C)	(B + D)	n

where
(A+B) + (C+D) =
(A+C) + (B+D) =
n = number of units evaluated

and where

df = 1

McNemar's Test for Correlated (Dependent) Proportions

Underlying Assumptions of the Test

1. Construct a 2x2 table where the paired observations are the sampling units.
2. Each observation must represent a single joint event possibility; that is, classifiable in only one cell of the contingency table.
3. In it's Exact form, this test may be conducted as a One Sample Binomial for the B & C cells

McNemar's Test for Correlated (Dependent) Proportions

Underlying Assumptions of the Test

- **4. The expected frequency (f_e) for the B and C cells on the contingency table must be equal to or greater than 5; where**

$$f_e = (B + C) / 2$$

from the Fourfold table

McNemar's Test for Correlated (Dependent) Proportions

Sample Problem

A randomly selected group of 120 students taking a standardized test for entrance into college exhibits a failure rate of 50%. A company which specializes in coaching students on this type of test has indicated that it can significantly reduce failure rates through a four-hour seminar. The students are exposed to this coaching session, and re-take the test a few weeks later. The school board is wondering if the results justify paying this firm to coach all of the students in the high school. Should they? Test at the 5% level.

McNemar's Test for Correlated (Dependent) Proportions

Sample Problem

The summary data for this study appear as follows:

Number of Students	Status Before Session	Status After Session
4	Fail	Fail
4	Pass	Fail
56	Fail	Pass
56	Pass	Pass

McNemar's Test for Correlated (Dependent) Proportions

The data are then entered into the Fourfold Contingency table:

		Before		
		Pass	Fail	
After	Pass	56	56	112
	Fail	4	4	8
		60	60	120

McNemar's Test for Correlated (Dependent) Proportions

- **Step I : State the Null & Research Hypotheses**

$$H_0 : \pi_1 = \pi_2$$

$$H_1 : \pi_1 \neq \pi_2$$

where π_1 and π_2 relate to the proportion of observations reflecting changes in status (the B & C cells in the table)

- Step II : $\alpha = 0.05$

McNemar’s Test for Correlated (Dependent) Proportions

- **Step III : State the Associated Test Statistic**

$$X^2 = \frac{\{\text{ ABS (B - C) - 1 }\}^2}{\text{B + C}}$$

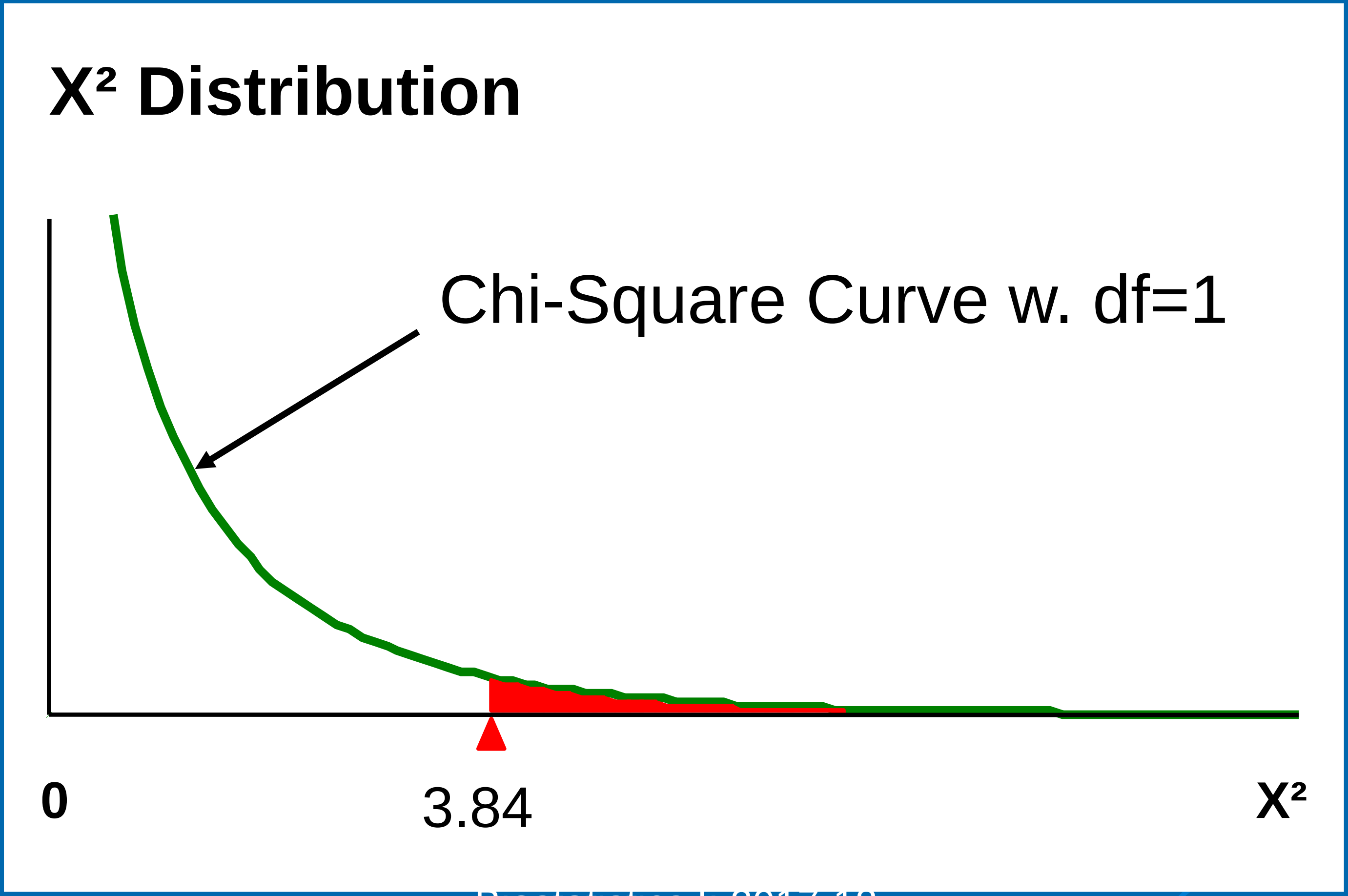
McNemar's Test for Correlated (Dependent) Proportions

- **Step IV : State the distribution of the Test Statistic When H_o is True**

$X^2 \overset{d}{=} X^2$ **with 1 *df* when H_o is True**

McNemar's Test for Correlated (Dependent) Proportions

Step V : Reject H_o if ABS (X^2) > 3.84



McNemar's Test for Correlated (Dependent) Proportions

- **Step VI : Calculate the Value of the Test Statistic**

$$X^2 = \frac{\{\text{ ABS } (56 - 4) - 1\ \}^2}{56 + 4} = 43.35$$

McNemar's Test for Correlated (Dependent) Proportions-SAS Codes

```
Data test;
input Before $ After $ count;
cards;
pass pass 56
pass fail 56
fail pass 4
fail fail 4;
run;

proc freq data=test order=data;
     weight Count;
     tables Before*After/agree;
run;
```

McNemar's Test for Correlated (Dependent) Proportions-SAS Output

```
Statistics for Table of Before by After

          McNemar's Test
      ________________________
      Statistic (S)    45.0667
      DF                     1
      Pr > S             <.0001


         Sample Size = 120
```

Without the correction

Conclusion: What we have learned

1. Comparison of binomial proportion using Z and χ^2 Test.
2. Explain χ^2 Test for Independence of 2 variables
3. Explain The Fisher's test for independence
4. McNemar's tests for correlated data
5. Kappa Statistic
6. Use of SAS Proc FREQ

Conclusion: Further readings

Read textbook for

1. Power and sample size calculation
2. Tests for trends