13.4 Chi-Squared Independence Test

Study Ch. 13.4, # 65 - 71, 75, 76

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

13.4 Chi-Squared Independence Test

At the 1% s.l., does an association exist between educational level and diabetic state?

Educ	Diabetic	Not Diabetic	Total
< HS	33	218	251
HS	25	389	414
< C	20	393	413
C	17	178	195
TOTAL	95	1178	1273

$$\chi^2_T = \sum\left[\frac{(O-E)^2}{E}\right]$$

OBSERVED values in Red
TOTALS in Green

How can we find the EXPECTED values?

Class Notes: Prof. G. Battaly, Westchester Community College, NY

Statistics Home Page Class Notes Homework

x^2

13.4 Chi-Squared Independence Test

OBSERVED values in Red
TOTALS in Green

Educ	Diabetic	Not Diabetic	Total
< HS	33	218	251
HS	25	389	414
< C	20	393	413
C	17	178	195
TOTAL	95	1178	1273

How many HS graduates would we EXPECT to be diabetic?

Start:
> Proportion of total that have HS education?

414 / 1273

> Find this proportion of the total number that are diabetic.
ie: Multiply by number with Diabetes:

95 (414 / 1273) = **30.896**

30.896

Therefore: $\chi^2_\tau = \sum\left[\frac{(O-E)^2}{E}\right]$

= 17.513

X^2

13.4 Chi-Squared Independence Test

OBSERVED values in Red
TOTALS in Green

Educ	Diabetic	Not Diabetic	Total
< HS	33	218	251
HS	25 **30.896**	389	414
< C	20	393	413
C	17	178	195
TOTAL	95	1178	1273

How many C graduates wouldwe EXPECT to be diabetic?

Start:
> Proportion of total that have HS education?

195 / 1273

> Find this proportion of the total number that are diabetic.
ie: Multiply by number with Diabetes:

95 (**195** / 1273) = **14.552**

14.552

Therefore: $\chi^2_\tau = \sum\left[\frac{(O-E)^2}{E}\right]$

= 17.513

X^2

13.4 Chi-Squared Independence Test

OBSERVED values in Red
TOTALS in Green

Educ	Diabetic	Not Diabetic	Total
< HS	33	218	251
HS	25 **30.896**	389	414
< C	20	393	413
C	17 **14.552**	178	195
TOTAL	95	1178	1273

How many C graduates wouldwe EXPECT to be diabetic?

Start:
> Proportion of total that have HS education?

195 / 1273

> Find this proportion of the total number that are diabetic.
ie: Multiply by number with Diabetes:

95 (**195** / 1273) = **14.552**

Therefore:

$$E = \frac{R\,C}{n}$$

$$\chi^2_\tau = \sum\left[\frac{(O-E)^2}{E}\right]$$

= 17.513

x^2

13.4 Chi-Squared Independence Test

OBSERVED values in Red
TOTALS in Green

Educ	Diabetic	Not Diabetic	Total
< HS	33	218	251
HS	25 **30.896**	389	414
< C	20	393	413
C	17 **14.552**	178	195
TOTAL	95	1178	1273

Find:
> The remaining Expected Values, using

Therefore:

$$E = \frac{R\,C}{n}$$

$$\chi^2_\tau = \sum\left[\frac{(O-E)^2}{E}\right]$$

= 17.513

Class Notes: Prof. G. Battaly, Westchester Community College, NY
Statistics Home Page ©Gertrude Battaly, 2014 Class Notes Homework

x^2

13.4 Chi-Squared Independence Test

Educ	Diabetic	Not Diabetic	Total
< HS	33 18.731	218 232.269	251
HS	25 30.846	389 383.104	414
< C	20 30.821	393 382.179	413
C	17 14.552	178 180.448	195
TOTAL	95	1178	1273

Therefore:

$$E = \frac{R\,C}{n}$$

Now that have Expected Values, can proceed to do the

χ^2 Hypothesis Test

χ^2

13.4 Chi-Squared Independence Test

Chi-Square Test for Independence

Assumptions: 1. All expected frequencies ≥ 1
2. At most 20% of the expected frequencies are less than 5
3. SRS

Step 2: Decide α

Step 1: H_0: The 2 variables are not associated. indep.
H_a: The 2 variables are associated. dependent

Step 3: Calculate the expected frequencies, $E = RC/n$
where n = sample size, R = row total, C = column total

Step 4: Compare values for E to assumptions to determine if can use this procedure.

Step 5: Compute the test statistic, using a table of values:

$$\chi^2_T = \sum\left[\frac{(O-E)^2}{E}\right]$$

Step 6 : Find CV(s) **using df = (r - 1)(c-1) and Table VII.**

Step 7: Decide: reject H_0 or not?
Reject if test statistic is in rejection region (tail).

Step 5: p-value from calculator
p = 1- χ^2cdf (0, χ^2_T,df)

Step 6: Decide: reject H_0 or not?
Reject if $p < \alpha$

Step 8: Verbal interpretation

Class Notes: Prof. G. Battaly, Westchester Community College, NY
Statistics Home Page ©Gertrude Battaly, 2014 Class Notes Homework

13.4 Chi-Squared Independence Test

At the 1% s.l., does an association exist between educational level and diabetic state?

Educ	Diabetic	Not Diabetic	Total
< HS	33 (10.870) (18.731)	218 (0.877) (232.269)	251
HS	25 (7.125) (30.876)	389 (0.091) (383.104)	414
< C	20 (3.799) (30.821)	393 (0.306) (382.179)	413
C	17 (0.4118) (14.552)	178 (0.032) (180.448)	195
TOTAL	95	1178	1273

χ^2 Hypothesis Test

1. All E >1 and >5, srs assumptions met
H_0: Education level and diabetic status are NOT associated.
H_a: Education level and diabetic status ARE associated.
2. $\alpha = 0.01$, df = (r-1)(c-1)=3(1)=3
3. $\chi^2_\tau = \sum\left[\frac{(O-E)^2}{E}\right] = 12.5118$
$\frac{(33-18.731)^2}{18.731}$
4. p = 1 - χ^2cdf(0, 12.5118, 3) = ____
5. p ? α Reject H_0???
6. Conclude?

Therefore:
$E = \frac{R\,C}{n}$

χ^2

13.4 Chi-Squared Independence Test

Educ	Diabetic	Not Diabetic	Total
< HS	33	218	251
HS	25	389	414
< C	20	393	413
C	17	178	195
TOTAL	95	1178	1273

Therefore:
$E = \frac{R\,C}{n}$

Calculator:
1. 2nd Matrix/ Edit / [A]
2. Enter # of rows, # columns
3. Enter observations
   - the red numbers in the table
4. STAT/ TESTS/ χ^2-Test
5. Hit Enter for [A] and for [B]
6. Select Calculate
   Answer will include test statistic, p-value, and df.
   For expected values:
   2nd Matrix/ Edit / [B]

$\chi^2_\tau = \sum\left[\frac{(O-E)^2}{E}\right]$

$\chi^2 = 12.5118$

$p = 5.5545(10^{-4})$
$= .00055545$

13.4 Chi-Squared Independence Test

Com	Male	Female	Total
Mail	58	26	
email	151	86	
Both	72	40	
N/A	76	50	
TOTAL			

Therefore:

$$E = \frac{RC}{n}$$

On-line χ^2 Computation
Kristopher J. Preacher

13.4 Chi-Squared Independence Test

Com	Male	Female	Total
Mail	58 53.6 0.361	26 30.4 0.637	84
email	151 151.4 0.001	86 85.6 0.002	237
Both	72 71.5 0.003	40 40.5 0.006	112
N/A	76 80.5 0.252	50 45.5 0.445	126
TOTAL	357	202	559

H_0: gender & how indep.

H_a: gender & how depend.

$\alpha = 0.05$

$$\chi^2_T = \sum\left[\frac{(O-E)^2}{E}\right] \qquad df = (c-1)(r-1)$$

$\chi^2_T = 1.6757$ $P = 1 - \chi^2cdf(0, 1.707, 3)$

$= 0.635 > \alpha$ do not reject H_0

Therefore:

$$E = \frac{RC}{n}$$