The Chi-Square Test

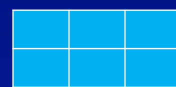
used when data are not scores that can be averaged, but instead are frequencies of observations that can only be counted – e.g., how many subjects fall into various categories

H0: population frequencies are as stated by the expected frequencies

H1: population frequencies are different than the stated expected frequencies

Test of Goodness of Fit: one row of cells

Test of Independence: cells in rows and columns
- a "contingency table" or "matrix"
- are the row variable and column variable related?



Test of Goodness of Fit: Are the expected frequencies a "good fit" to the observed frequencies? Rejecting the null hypothesis says no, they're not.

Test of Independence: Use expected frequencies that are based on the assumption that the row and column variables are independent, i.e., there's no relationship
- then rejecting H0 STILL means rejecting those expected frequencies as a "good fit" to the observed frequencies -- which means rejecting that assumption that the variables are independent, which means believing they're related!

The Chi-Square Test

ASSUMPTIONS OF THE CHI-SQUARE TEST

1) all observations are independent – not the variables, but the observations themselves (e.g. knowing the category one subject is in provides no information about another subject's category)
- e.g., in a study of helping behavior, one subject can't tell the next that they helped or not, which might influence their behavior

This also means each observation must fall into ONE, AND ONLY ONE, category – otherwise, an observation in category 1 would not be independent of the same subject observed in category 2
- e.g., a subject who intially did help but then did not help in another scenario cannot be included in both the "helped" and "didn't help" category counts

2) all expected frequencies (f_E) are at least 5 (all $f_E \geq 5$)
- this is due to the math of finding probabilities or p-values; with any of the $f_E < 5$, the probabilities are distorted and inaccurate

The Chi-Square Test

$\chi^2(df) = \Sigma\,[(f_O - f_E)^2 / f_E]$ or nicer looking, $\chi^2(df) = \Sigma \dfrac{(f_O - f_E)^2}{f_E}$

The tests only differ in how you find "df" and "expected frequencies"

Test of Goodness of Fit

df = (C-1) = (number of cells or categories or "columns" minus 1)

f_E for each cell = "proportion P (NOT a p-value!) of the population in that category" times "total number of subjects" = P*N; for equal probabilities, all f_E are same, and the formula simplifies to "total number of subjects / number of cells" = N/C

Test of Independence

df = (R-1)(C-1) = (number of rows minus 1) times (number of columns minus 1)

$f_E = (f_R * f_C) / N$ = (row margin frequency) times (column margin frequency), divided by total number of observations

The Chi-Square Test

WHY those df in each test?...

df = (C-1)... because if you know how many total observations there are, then knowing 2 of the 3 category frequencies, or 3 of the 4, etc., tells you what the last category frequency must be

- e.g. with 40 observations, if three of the frequencies are 7, 13, and 15, the last must be 5

7	13	15	?

df = (R-1)(C-1)... because if you know the margin frequencies, then knowing 1 of the 2x2 frequencies, or 4 of the 3x3 frequencies, etc., tells you what the other category frequencies must be

- e.g. if row margin 1 = 10 and 2 = 20, and column margin 1 = 12 and 2 = 18, knowing R1C1 frequency is 4 means R1C2 must be 6, and then R2C1 must be 8 and R2C2 must be 22

4	?	**10**
?	?	**20**
12	**18**	**30**

The Chi-Square Test

WHY those expected frequencies in the Test of Independence?...

$f_E = (f_R * f_C) / N$... because if the rows and columns are independent, then the proportion of observations in the overall column variable (which would be f_C/N) must be applicable to each separate row too

Example: applying that proportion, which for column 1 is f_C/N = 12/30, to the row 1 margin 10 gives $f_C / N * f_R$ = 12/30 * 10 = 4
- but $f_E = f_C/N * f_R$ can just be rearranged into $f_E = (f_R * f_C) / N$

	col 1	col 2	row marginal frequencies:
row 1	4		10
row 2			20
column marginal frequencies:	12	18	30

The Chi-Square Test

Finding a p-value to go with the value of χ^2 :
- use df to choose a row to read in the table
- read the "proportion in critical region" column under your p-value cutoff (or α, Greek letter "alpha") – usually .05
- the value for χ^2 is listed that has exactly that p-value for those df
- remember the general rule: a LARGER statistic has a SMALLER p
- if YOUR χ^2 value is LARGER than the χ^2 for p = .05, then YOUR p-value must be SMALLER than .05

For df = 4 (e.g,, from a 5-category Goodness of Fit test, or 3x3 Test of Independence) and for $\alpha = .05$, $\chi^2 = 9.49$

If your χ^2 is less than 9.49, you FAIL TO REJECT H0; if it's greater, you REJECT H0

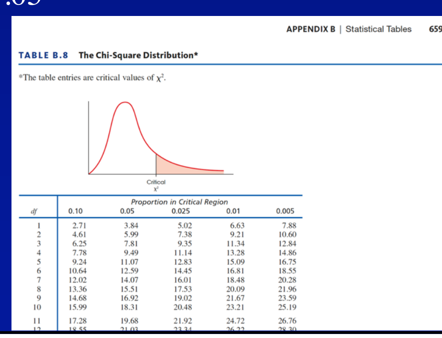

APPENDIX B | Statistical Tables 659

TABLE B.8 The Chi-Square Distribution*

*The table entries are critical values of χ^2.

df	Proportion in Critical Region 0.10	0.05	0.025	0.01	0.005
1	2.71	3.84	5.02	6.63	7.88
2	4.61	5.99	7.38	9.21	10.60
3	6.25	7.81	9.35	11.34	12.84
4	7.78	9.49	11.14	13.28	14.86
5	9.24	11.07	12.83	15.09	16.75
6	10.64	12.59	14.45	16.81	18.55
7	12.02	14.07	16.01	18.48	20.28
8	13.36	15.51	17.53	20.09	21.96
9	14.68	16.92	19.02	21.67	23.59
10	15.99	18.31	20.48	23.21	25.19
11	17.28	19.68	21.92	24.72	26.76
12	18.55	21.03	23.34	26.22	28.30

The Chi-Square Test

EFFECT SIZE

Test of Goodness of Fit (one row, any number of cells):
Cohen's w = $\sqrt{[\chi^2 / N]}$ (textbook's formula is equivalent)

- same as ϕ below, only w can't be interpreted as a correlation like ϕ can, because there's only one variable in the Goodness of Fit test

Test of Independence
when there are 2 rows and 2 columns (2x2 contingency table):
"phi coefficient" $\phi = \sqrt{[\chi^2 / N]}$ – which is actually a form of correlation
when there are more than 2 rows or columns (2x3, 3x2, 4x3, etc.):
"Cramér's V" = $\sqrt{[\chi^2 / N(df^*)]}$ where N is multiplied by df* (= whichever is SMALLER out of R-1 and C-1)

- notice that means you could always use V, because it's same as ϕ for 2x2 tables (try it!) – and incidentally also for 2x3 and 3x2 tables

The Chi-Square Test

EXAMPLE

$$\chi^2(df) = \Sigma \frac{(f_O - f_E)^2}{f_E}$$

Test of Goodness of Fit
df = (C-1) = 4 – 1 = 3
f_E = P*N = .25*40 (= N/C = 40/4) = 10, for each cell

(a)	(b)	(c)	(d)
7 (obs) (exp) 10	13 10	15 10	5 10

$\chi^2(3)$ = $(7 - 10)^2 / 10$ + $(13 - 10)^2 / 10$ + $(15 - 10)^2 / 10$ + $(5 - 10)^2 / 10$
= $(-3)^2 / 10$ + $(3)^2 / 10$ + $(5)^2 / 10$ + $(-5)^2 / 10$
= 9 / 10 + 9 / 10 + 25 / 10 + 25 / 10
= 0.9 + 0.9 + 2.5 + 2.5
= 6.8

value of chi-square on 3 df corresponding to α=.05 is $\chi^2(3)_{.05} = 7.81$
6.8 < 7.81, so p for 6.8 is > .05
not significant ("N.S."): fail to reject H0...
Instructor has no preference for using a particular answer choice

w = $\phi = \sqrt{[\chi^2 / N]} = \sqrt{[6.8 / 40]}$ = 0.41
(NOTE: if obs. were 70, 130, 150, 50, χ^2 = 68, p < .05
but w = ϕ = 0.41 still – not dependent on N)

df	0.10	0.05	0.025
1	2.71	3.84	5.02
2	4.61	5.99	7.38
3	6.25	7.81	9.35
4	7.78	9.49	11.14
5	9.24	11.07	12.83

The Chi-Square Test

EXAMPLE

$$\chi^2(df) = \Sigma \frac{(f_O - f_E)^2}{f_E}$$

Test of Independence

$df = (R-1)(C-1) = (2-1)(4-1) = (1)(3) = 3$

$f_E = (f_R * f_C) / N$

R1C1: $f_E = (50 * 100) / 200 = 25$
R1C2: $f_E = (50 * 20) / 200 = 5$
R1C3: $f_E = (50 * 40) / 200 = 10$
R1C4: $f_E = (50 * 40) / 200 = 10$
R2C1: $f_E = (150 * 100) / 200 = 75$
R2C2: $f_E = (150 * 20) / 200 = 15$
R2C3: $f_E = (150 * 40) / 200 = 30$
R2C4: $f_E = (150 * 40) / 200 = 30$

CAR COLOR: INT/EXT:	RED	YELLOW	GREEN	BLUE	
INTRO-VERTS	10 25	3 5	15 10	22 10	**50**
EXTRA-VERTS	90 75	17 15	25 30	18 30	**150**
	100	**20**	**40**	**40**	**200**

$\chi^2(3) = (10-25)^2 / 25 + (3-5)^2 / 5 + (15-10)^2 / 10 + (22-10)^2 / 10$
$+ (90-75)^2 / 75 + (17-15)^2 / 15 + (25-30)^2 / 30 + (18-30)^2 / 30 = 35.6$

$\chi^2(3)_{.05} = 7.81$

$35.6 > 7.81$, so p for 35.6 is $< .05$

significant: $\chi^2(3) = 35.6^*$, $p<.05$ so reject H0 (that the variables are independent)...

Introversion / Extraversion and Car Color Preference are related

Cramér's V = $\sqrt{[\chi^2 / N(df^*)]}$ with $df^* = 2-1 = 1$
[because df* is smaller of (2-1) or (4-1)]
$= \sqrt{[35.6 / 200(1)]} = \sqrt{[0.178]} = 0.422$

df	0.10	0.05	0.025
1	2.71	3.84	5.02
2	4.61	5.99	7.38
3	6.25	7.81	9.35
4	7.78	9.49	11.14
5	9.24	11.07	12.83