11.2b

χ^2 test for Association / Independence

- Use when you have a 2-way table from one, single SRS and each individual is classified by both categories.

Step 1:

H_0: There is no association between the 2 categorical variables in the population (independent)

H_a: There is an association between the two variables (not independent)

*Same 3 conditions as other χ^2 tests
Same steps 2,3,4 as χ^2 test for Homogeneity

To explore the direction/nature of the association, we must examine the row/column %'s

ex: Allergies/gender

Step 1 $\alpha = 0.05$

H_0: There is no association between gender and having allergies for U.S. H.S. students who filled out the survey.

H_a: there is an association between gender and having allergies for U.S. H.S. students who filled out the survey.

Step 2: Use a χ^2 test for Association

Conditions: ① Random - random sample of 40 H.S. students

② Large Sample Size

Expected counts	Female	Male
Allergies	10.35	7.65
No Allerg.	12.65	9.35

③ Independent - more than 400 H.S. students in the U.S.

Step 3

$\chi^2 = 0.051$ $df = 1$

p-val = .822

Step 4 Since our p-val (.822) is greater than $\alpha = .05$ we fail to reject H_0. We can't conclude that there is an association between gender and having allergies..

χ^2 test for GOF

- Compare sample distribution to theoretical distrib.
- Expected = p · n
- df = # of categories − 1
- χ^2 GOF test in calc or by hand

χ^2 test for Homogeneity

- 2 or more samples, compare distributions
- Expected = $\frac{\text{row tot} \cdot \text{col tot}}{\text{table tot}}$
- df = (r−1)(c−1)
- χ^2 test (matrix)

χ^2 test for Association/Independ.

- 1 sample, 2 variables

All χ^2 tests

- Same 4 general steps
- Same 3 conditions ← SRS, exp. counts > 5
- Categorical variable
- dealing w/ counts (not means, ... p used in GOF test)
- $\chi^2 = \sum \frac{(obs - exp)^2}{exp}$
- χ^2 cdf to find p-val