Day 1: 11.1 Chi-square tests

Read 676-681

Emphasize that each of the variables in the introduction is categorical.

What is a one-way table? What is a chi-square goodness-of-fit test?

Test compares an observed distribution to a hypothesized one. Not about a parameter!

one sample / for categorical → i.e. % each color of m&m observed vs expected by Co. claim

What are the null and alternative hypotheses for a chi-square goodness-of-fit test?

Alternative is always two-sided. ≠ (at least one - does not have to be all categories)
Always in terms of the population! (not the sample)

How do you calculate the expected counts for a chi-square goodness-of-fit test? Should you round these to the nearest integer?

Don't round the expected counts. These represent the average number in each category in repeated samples. Take % given & multiply by sample size

What is the chi-square test statistic? Is it on the formula sheet? — yes

Must use counts, not proportions! $\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$

HW #1 page 692 (1-4)

Day 2: 11.1 Chi-Square Goodness-of-Fit Tests

Read 682-687

In a goodness-of-fit test, when does the chi-square test statistic follow a chi-square distribution? How do you calculate the degrees of freedom for a chi-square goodness-of-fit test?

When expected counts are at least 5.
Df = number of categories – 1
Note that df does not depend on sample size, just the number of categories.
Trace dotplot on page 682 (FIGURE 11.3)

Describe the shape, center, and spread of the chi-square distributions. How are these based on the degrees of freedom?

Sketch several distributions (see bottom of 682)
Skewed right, mean = df, peak = df – 2, SD = $\sqrt{2df}$

How do you calculate *p*-values using chi-square distributions? — χ^2 table gives area to right (mode)
or technology → χ^2cdf(lower, upper, df)

Tables and chi-square pdf

What are the conditions for conducting a chi-square goodness-of-fit test?

See page ~~684~~ 685

Expected counts not observed counts! Must show these.

Random —

Large Sample Size (replaces normal)

All expected counts at least 5

In the birth day example, emphasize nor accepting null.

reject or fail to reject H_0

never accept H_0

Indep — Indep + 10% rule if sampling w/o rep

Alternate Example: *Landline surveys*

According to the 2000 census, of all U.S. residents aged 20 and older, 19.1% are in their 20s, 21.5% are in their 30s, 21.1% are in their 40s, 15.5% are in their 50s, and 22.8% are 60 and older. The table below shows the age distribution for a sample of U.S. residents aged 20 and older. Members of the sample were chosen by randomly dialing landline telephone numbers. Do these data provide convincing evidence that the age distribution of people who answer landline telephone surveys is not the same as the age distribution of all U.S. residents?

Category	Count
20–29	141
30–39	186
40–49	224
50–59	211
60+	286
Total	1048

Read 687-690

χ^2 GOF-Test
TI-84 Plus
Put obs
Exp into lists

Can you use your calculator to conduct a chi-square goodness-of-fit test?

It depends. Newer calculators have it built in. Others have to use lists.

When should you do a follow-up analysis? How do you do a follow-up analysis?

In general, whenever the results are statistically significant. However, on AP exam, only do a follow-up analysis when asked.
Look at chi-square components—and direction (were there more observed than expected? Less?)

* After you run χ^2GOF-Test, list of Indv. Components is stored under CNTRB (list #)

Alternate Example: *Birthdays and hockey*
In his book *Outliers,* Malcolm Gladwell suggests that a hockey player's birth month has a big influence on his chance to make it to the highest levels of the game. Specifically, since January 1 is the cutoff date for youth leagues in Canada (where many National Hockey League players come from), players born in January will be competing against players up to 12 months younger. The older players tend to be bigger, stronger, and more coordinated and hence get more playing time and more coaching and have a better chance of being successful. To see if birth date is related to success (judged by whether a player makes it into the NHL), a random sample of 80 NHL players from the 2009–2010 season was selected and their birthdays were recorded. Overall, 32 were born in the first quarter of the year, 20 in the second quarter, 16 in the third quarter, and 12 in the fourth quarter.
(a) Do these data provide convincing evidence that the birthdays of NHL players are not uniformly distributed throughout the year?
(b) If the results are significant, do a follow-up analysis.

Day 3: 11.2 Chi-Square Tests for Homogeneity

Read 696-702

How is section 11.2 different than section 11.1?

Comparing the distribution of a categorical variable in 2 or more populations or treatments rather than comparing the distribution of a categorical variable in 1 population to a hypothesized distribution. Example: comparing the color distributions of plain and peanut M&M's. Two-way tables vs. one-way tables

What are the two explanations for the differences in the distributions of wine purchases? (example p. 696-697)

Music has an effect on purchases.
Music has no effect and the differences are due to the chance variation in the random assignment.

How do you state hypotheses for a test of homogeneity?

H_0: There is no diff. in distr of wine purchases at this store when no music, French music or Italian music is played.

H_a: There is a difference . . . (same as above)

What is the problem of multiple comparisons?

The problem of how to do many comparisons at once w/ an overall measure of confidence

1. overall test – to see if there is good evidence of any differences
2. Follow-up analysis – to decide which parameters differ & to est. how large diff. are

How do you calculate the expected counts for a test that compares the distribution of a categorical variable in multiple groups or populations?

Assuming null is true, the proportion in each category should be the same in each population or group.
Reminder: Don't round!
Formula is not on the formula sheet

$$\text{expected count} = \frac{(\text{row total}) \cdot (\text{column total})}{\text{table total}}$$

What is the formula for the chi-square test statistic? Is it on the formula sheet?

$$\chi^2 = \sum \frac{(\text{Obs.} - \text{Exp})^2}{\text{Exp}}$$

$$df = (\#\text{rows} - 1)(\#\text{columns} - 1)$$ ← without totals!

Alternate Example: *Saint-John's-wort and depression*
An article in the *Journal of the American Medical Association* (vol. 287, no. 14, April 10, 2002) reports the results of a study designed to see if the herb Saint-John's-wort is effective in treating moderately severe cases of depression. The study involved 338 subjects who were being treated for major depression. The subjects were randomly assigned to receive one of three treatments—Saint-John's-wort, Zoloft (a prescription drug), or a placebo—for an eight-week period. The table below summarizes the results of the experiment.

	Saint-John's-wort	Zoloft	Placebo	Total
Full response	27	27	37	91
Partial response	16	26	13	55
No response	70	56	66	192
Total	113	109	116	338

(a) Calculate the conditional distribution (in proportions) of the type of response for each treatment.
(b) Make an appropriate graph for comparing the conditional distributions in part (a).
(c) Compare the distributions of response for each treatment.
(d) State the hypotheses we are interested in testing.
(e) Calculate the expected counts and the chi-square statistic.

Read 703-706

How do you calculate the degrees of freedom for a chi-square test of homogeneity?

Illustrate by filling in cells on a blank table (with totals)

$df = (\#\text{ rows w/o totals} - 1)(\#\text{ cols w/o totals} - 1)$

What are the conditions for a chi-square test of homogeneity?

Random — random sample or randomized experiment

Must show calc → Large Sample Size — expected counts at least 5 of each category

Indep — Indep / 10% cond.

Alternate Example: *Saint-John's-wort and depression*

(a) Verify that the conditions for this test are satisfied.

(b) Calculate the *P*-value for this test.

(c) Interpret the *P*-value in context.

(d) What is your conclusion?

Can you use your calculators to do a chi-square test of homogeneity? p. 705–706 instructs

Remember to state the expected counts! Good idea to show at least 1 term of the sum.

↳ enter observed counts in matrix [A]

Read 706-709 2nd x^{-1} (matrix), edit, enter dimensions (rows x c

How do you conduct a follow-up analysis for a test of homogeneity?

look at printout on p. 709

★ show expected counts for Large Sample size → Stat → tests → χ^2-Test

Obs [A]

Read 709-711

Alternate Example: *Ibuprofen or acetaminophen?*
In a study reported by the *Annals of Emergency Medicine* (March 2009), researchers conducted a randomized, double-blind clinical trial to compare the effects of ibuprofen and acetaminophen plus codeine as a pain reliever for children recovering from arm fractures. There were many response variables recorded, including the presence of any adverse effect, such as nausea, dizziness, and drowsiness. Here are the results:

	Ibuprofen	Acetaminophen plus codeine	Total
Adverse effects	36	57	93
No adverse effects	86	55	141
Total	122	112	234

(a) Is the difference between the two groups statistically significant? Conduct an appropriate chi-square test to find out.
(b) Show that the results of a two-sample z test for a difference in proportions are equivalent.

When should you use a chi-square test and when should you use a two-sample z test?

- *The chi-square test is always two-sided. That is, it only tests for a difference in the two proportions. If you want to test whether one proportion is larger than the other, use the two-sample z test.*
- *If you want to estimate the difference between two proportions, use a two-sample z interval. There are no confidence intervals that correspond to chi-square tests.*
- *If you are comparing more than two treatments or the response variable has more than two categories, you must use a chi-square test.*
- *You can also use a chi-square goodness-of-fit test in place of a one-sample z test for a proportion if the alternative hypothesis is two-sided. The chi-square test will use two categories (success and failure) and have df = 2 – 1 = 1.*

HW #3 page 694 (19-22), page 724 (27, 29, 31, 33, 35, 43)

(b) 2 Sample z test for Prop

$z = -3.34$

$X_1 = 36 \quad X_2 = 57$

$n_1 = 122 \quad n_2 = 112 \quad p_1 \neq p_2$

pvalue = 2(normalcdf(-99

7

Day 4: 11.2 Chi-Square Tests of Association/Independence

Read page 713-718

What does it mean if two variables have an association? What does it mean if two variables are independent?

How is a test of association/independence different than a test of homogeneity?

It's all about the way the data are gathered. If there is one sample and two variables, we are testing for an association. If there are 2+ samples or groups and one variable, we are testing for homogeneity.

How do you state hypotheses for a test of association/independence?

H_0: There is no association between the 2 categorical variables

H_a: There is an association btwn the variables

How do you calculate expected counts for a test of association/independence?

Remember not to round! expected count $= \frac{(\text{row total})(\text{column total})}{(\text{table total})}$ OR

H_0: 2 cat. var are independ

$\chi^2 = \sum \frac{(Obs - Exp)^2}{Exp}$ $df = (\#\text{rows}-1)(\#\text{col}-1)$

H_a: 2 cat. var. are not indep.

What are the conditions for a test of association/independence?

Random – data comes from random sample or randomized experiment

Large Sample Size – All expected counts are at least 5

Independent – Indiv. observ. are indep. when sampling w/o replacement check 10% rule

Alternate Example: *Allergies*

In an Alternate Example from Chapter 5, we investigated the relationship between gender and having allergies for a random sample of 40 students who completed a CensusAtSchool survey. Here is a two-way table that summarizes the data:

	Female	Male	Total
Allergies	10	8	18
No allergies	13	9	22
Total	23	17	40

(a) In Chapter 5, we concluded that the events "female" and "allergies" were not independent for the members of the *sample*. Calculate appropriate conditional distribution to verify that this is true.

(b) Do the data provide convincing evidence of an association between gender and having allergies for the *population* of U.S. high school students who filled out the CensusAtSchool survey?

Read 718-721

Alternate Example: *Online social networking*
An article in the *Arizona Daily Star* (April 9, 2009) included the following table:

Age (years):	18–24	25–34	35–44	45–54	55–64	65+	Total
Use online social networks:	137	126	61	38	15	9	386
Do not use online social networks:	46	95	143	160	130	124	698
Total:	183	221	204	198	145	133	1084

Suppose that you decide to analyze these data using a chi-square test. However, without any additional information about how the data were collected, it isn't possible to know which chi-square test is appropriate.
(a) Explain how you know that a goodness-of-fit test is not appropriate for analyzing these data.
(b) Describe how these data could have been collected so that a test for homogeneity is appropriate.
(c) Describe how these data could have been collected so that a test for association/independence is appropriate.