Slide 1

AP Statistics

Chapter 11:
Inference for Distribution of Categorical Data

Day 1

HW: Finish Lesson 1 Practice Worksheet

Slide 2

Inference for Categorical Data

The two we have learned so far are the one-sample and two-sample proportion z-tests.

○ These only work for categorical data of a single variable that is binomial (success or failure).

Chi-Square Goodness-of-Fit Test:

○ Allows us to determine whether a hypothesized distribution seems valid.

Slide 3

Examples of Chi-Square Goodness-of-Fit Tests

○ Are births evenly distributed across the days of the week?

○ Does background music influence customer purchases?

○ Is there an association between anger level and heart disease?

Slide 4

M & M's

The Mars Company produces M & M's and claims that 24% of the candies in each bag are blue. Jerry counted the number of each type of color in his bag of M & M's and came up with the following one-way table.

Color:	Blue	Orange	Green	Yellow	Red	Brown	Total
Count:	9	8	12	15	10	6	60

Is there evidence to conclude that the the proportion of blue M & M's in a bag differs from 24%?

$\hat{p} = \frac{9}{60} = .15$ 1-prop z-test

$p = .1026$

M & M's

What if we wanted to know if there was a difference in the overall distribution of the proportions of all the colors given to us by the Mars Company?

Color:	Blue	Orange	Green	Yellow	Red	Brown	Total
	.24	.20	.16	.14	.13	.13	
Count:	9	8	12	15	10	6	60

What would our hypotheses be?

H_0: The company's stated color distribution is correct. $p_{Blue}=.24$, $p_{orange}=.20$, etc ....

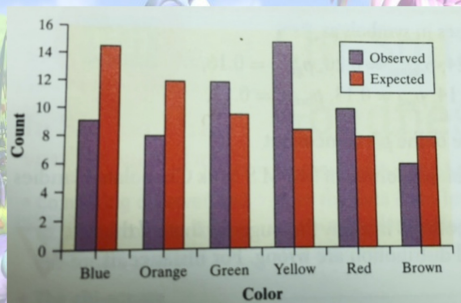H_a: The company's stated distribution is not correct.

Observed Counts Vs. Expected Counts

Expected counts are what expect to get if our H_0 is true.

Observed counts are what we actually found in our sample.

Color:	Blue	Orange	Green	Yellow	Red	Brown	Total
	.24	.20	.16	.14	.13	.13	
Observed Count:	9	8	12	15	10	6	60
Expected Count:	14.4	12.0	9.6	8.4	7.8	7.8	

Observed Counts Vs. Expected Counts

Chi-Square Statistic

○ **The measure of how far the observed counts are from the expected counts.**

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

Chi-Square Statistic

Color:	Blue	Orange	Green	Yellow	Red	Brown
Observed Count:	9	8	12	15	10	6
Expected Count:	14.40	12.00	9.60	8.40	7.80	7.80

$$\chi^2 = \frac{(9-14.40)^2}{14.4} + \frac{(8-12)^2}{12} + \frac{(12-9.6)^2}{9.6} + \frac{(15-8.4)^2}{8.4}$$

$$+ \frac{(10-7.8)^2}{7.8} + \frac{(6-7.80)^2}{7.8} = 10.18$$

$$\chi^2 = 10.18$$