14.1 Test for Goodness of Fit

M:M activity

M:M claim:

Color	%
Brown	13
Red	13
Yellow	14
Green	16
Orange	20
Blue	24
	100%

① Complete chart (pg. 834)

	Br	R	Y	G	O	Bl	Total
O							
E							
$\frac{(O-E)^2}{E}$							χ^2

Expected amounts = (# in bag)(claimed %)

② Discuss questions in #5 pg. 835
③ Combine all counts for a class count
④ Complete a chart for the class

Dice and the Random # Generator

Use the calculator to simulate rolling a die 100 times. MATH PRB randInt (1, 6, 100)

(1.9)	X (face value)	1	2	3	4	5	6	
	observed	13	16	12	15	19	21	= 96

Expected value?

We expect a fair die to roll ea. value $\frac{1}{6}$ of time

Does the table of values from the random # generator represent a fair die?

L_1: observed L_2: expected STAT TEST χ^2GOF

$\chi^2 = 3.75$

$p = .586$

$.586 > .05$

not enough evidence to claim unfair

14.1 Notes (Review 4.2)

Chi-square (χ^2) test for goodness of fit:
determines if an observed sample distribution is significantly different in some way from the hypothesized (expected) population distribution

Chi-square test for homogeneity of populations
compares two or more population proportions

Chi-square test of association/independence
determines whether the distribution of one variable has been influenced by another variable

* Observed data only applies to count data, not proportions. Proportions may be stated in the hypothesis. Expected values may be fractional since they are averages but not proportions

Expected count - found by multiplying the proportion of each category by the sample size.

Chi-Square Statistic $(\chi^2) = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}} = \sum \frac{(O-E)^2}{E}$

* The larger the difference between observed and expected values, the larger χ^2 will be providing more evidence against H_0.

14.1 cont'd

Chi-Square GOF Process (Example 14.3 includes all steps)

① Hypotheses

H_0: $P_1 = P_2 = \ldots =$ ____ ; or state in context

H_a: At least one proportion is incorrect ; or state in context

② Procedure

χ^2 GOF test

③ Conditions

S - random samples

E - Expected values ≥ 1 and no more than 20% ≤ 5

* Observed values do not matter here, only expected

$\Sigma \frac{(O-E)^2}{E} \Rightarrow$ Do not want to divide by $\emptyset$

④ all expected values ≥ 5

④ Computation

$\chi^2 = \Sigma \frac{(O-E)^2}{E}$ $df =$ ____ p-value = ____

Exam questions: show first and last computations with Σ

not necessary to show each category.

⑤ Conclusion

Since the p-value = ____ is $\leq$ $\alpha = .05$ we reject / fail to reject H_0. There is / is not significant evidence for (context)

(If there is evidence to reject, identify the largest component - largest indiv. value $\frac{(O-E)^2}{E}$)

Chi-square density curve properties (pg. 841-842)

p-value - probability of observing a value of χ^2 at least as extreme as the one actually observed. Larger χ^2 values result in smaller p-values and more evidence to reject H_0.

Example 14.1 Auto accidents and cell phones

Day	Sun	Mon	Tue	Wed	Thur	Fri	Sat	Total
Number	20	133	126	159	136	113	12	699

We expect each day to have approximately the same values

Expected counts = $\frac{1}{7}(699) = 99.857$

Bar Graph pg 858 pg

Calculator:

L_1 = Observed Counts L_2 = Expected counts

$L_3 = (L_1 - L_2)^2 / L_2$ χ^2 = sum(list 3) or χ^2 GOF test

p-value = χ^2 GOF-test $df = n - 1$

$\chi^2 = 208.847$ p-value = 2.479×10^{-45} $df = 6$

Sum ⇒ [List]

χ^2 ⇒ [STAT] [TEST]

Solution:

① context: H_0: Motor vehicle accidents involving cell phone use are equally likely to occur on any day of the week

H_a: The probability of motor vehicle accidents involving cell phone use vary day to day.

alternative: H_0: $P_{sunday} = P_{monday} = \dots P_{saturday} = \frac{1}{7}$

H_a: At least one proportion differs day to day

② We will use the χ^2 GOF test

③ We will assume the 699 drivers were chosen at random and all expected values are ≥ 5. $\frac{1}{7}(699) = 99.857 \geq 5$

④ $\chi^2 = \frac{(20-99.857)^2}{99.857} + \dots \frac{(12-99.857)^2}{99.857} = 208.847$

p-value = 2.479×10^{-42} $df = 6$

141 cont'd

④ Since $2.479 \times 10^{-42} < \alpha = .05$ there is sufficient evidence to reject H_0 and conclude accidents involving cell phones are not equally likely to occur on each of the seven days of the week. The largest components of the sample are Saturday (77.30) and Sunday (63.86).