

Section 10.1 Notes
Goodness of Fit Test

A _____ determines whether it is reasonable to assume that your sample came from a population in which, for each category, the proportion of the population that falls into the category is equal to some hypothesized proportion.

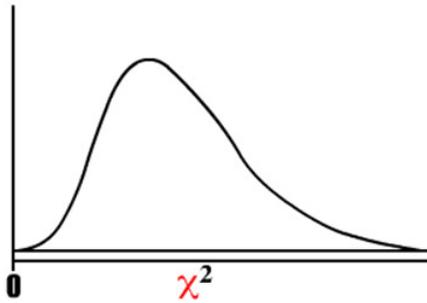
Test Statistic for Chi-Squared Tests: _____

Determine if a die is unfair.

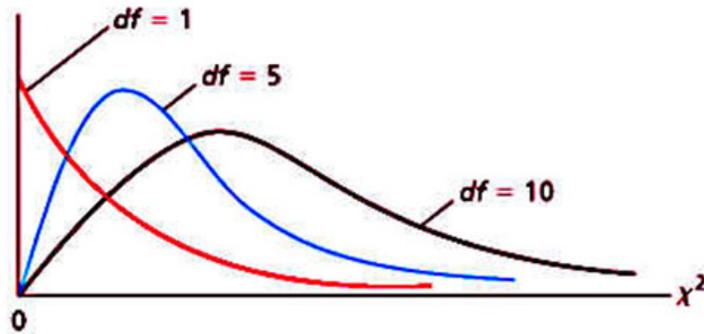
Outcome	Observed Frequency, O	Expected Frequency, E	$O - E$	$\frac{(O - E)^2}{E}$
1				
2				
3				
4				
5				
6				
Total	60	60		

Properties of the Distribution of the Chi-Square Statistic

- 1) The chi-square statistic is _____, unlike the normal and student t distributions. (As the number of degrees of freedom _____, the distribution becomes more symmetric.)



- 2) The value for chi-square can be _____ or _____, but they cannot be _____.
- 3) The chi-square distribution is _____ for each number of degrees of freedom.



Chi-Squared Goodness-of-Fit (GOF) Conditions

- Each outcome of your population falls into _____ of a fixed number of categories.
- You have a model that gives the hypothesized proportion of outcomes in the population that falls into each category.
- You have a _____ from your population.
- The expected frequency in each category is _____.

State the hypotheses:

- H_0 : The proportion in the population are equal to the proportions in the model.
- H_a : At least one proportion in the population is not equal to the corresponding proportion in your model.

df = _____

Example:

Acme Toy Company prints baseball cards. The company claims that 30% of the cards are rookies, 60% veterans, and 10% are All-Stars. The cards are sold in packages of 100.

Suppose a randomly selected package of cards has 50 rookies, 45 veterans, and 5 All-Stars. Is this consistent with Acme's claim? Use a 0.05 level of significance.

Section 10.2 Notes
The Chi-Square Test of Homogeneity

A chi-square test of homogeneity tests whether it is reasonable to believe that when several populations are sorted into the same _____, they will have the same _____ of members in each category.

Components of the test...

CONDITIONS:

1. _____ simple random samples of _____ (but not necessarily equal) sizes are taken from two or more large populations (or two or more treatments are randomly assigned to subjects who give a categorical responses.)
2. Each outcome falls into _____ of several categories, with the categories being the same in all populations.
3. The expected frequency in each cell is _____.

State the hypotheses:

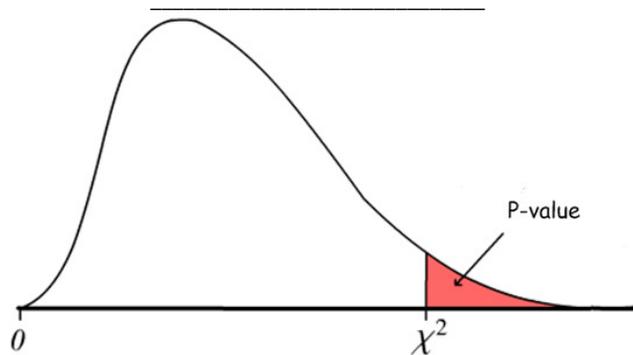
- H_0 : The proportion that falls into each category is the same for every population.
- H_a : For at least one category, it is not the case that each population has the same proportion in that category.

Calculations:

Test Statistic:

E =

df = _____



Conclusion:

- Linked to your computations and in the context of the problem.

Example: In a study of the television viewing habits of children, a developmental psychologist selects a random sample of 300 first graders - 100 boys and 200 girls. Each child is asked which of the following TV programs they like best: The Lone Ranger, Sesame Street, or The Simpsons. Results are shown in the contingency table below. Do the boys' preference for TV programs differ significantly from the girls' preferences? Use a 0.05 level of significance.

Viewing Preferences				
	Lone Ranger	Sesame Street	The Simpson	Row total
Boys	50	30	20	100
Girls	50	80	70	200
Column total	100	110	90	300

Section 10.3 Notes
Chi-Square Test for Independence

Conditions:

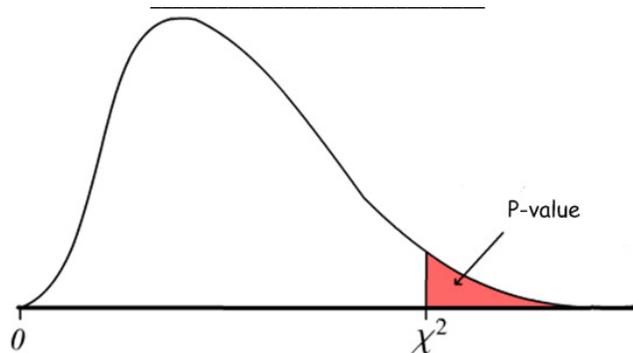
1. A simple random sample is taken from _____ population.
2. Each outcome is classified into one cell according to its _____ on one variable and its _____ on a second variable.
3. The expected frequency in each cell is _____.

Calculations:

Test Statistic:

E =

df = _____



Homogeneity vs. Independence

The distinguishing factor between the test for homogeneity and test for association/independence is how the data were produced. If the data came from independent random samples or independent groups in a randomized comparative experiment, then a test for homogeneity is called for. If the data came from a single random sample, with each individual categorized according to two categorical variables, then a test for independence is called for.

Example: A public opinion poll surveyed a simple random sample of 1000 voters. Respondents were classified by gender and by voting preference. Results are shown in the contingency table below. Is there a gender gap? Do the men's voting preference differ significantly from the women's preferences? Use a 0.05 level of significance.

Voting Preferences				
	Republican	Democrat	Independent	Row total
Male	200	150	50	400
Female	250	300	50	600
Column total	450	450	100	1000