Test of Goodness of Fit

Lecture 41
Sections 14.1 - 14.3

Robb T. Koether

Hampden-Sydney College

Mon, Apr 7, 2008

Outline

Introduction

- The one-proportion test that we just studied allows us to test a hypothesis concerning one proportion, or two categories, e.g., “yes” and “no.”
- What if their are several categories, e.g., “excellent,” “good,” “fair,” “poor,” and “not sure?”
- Or the categories may be “Republican,” “Democrat,” and “Independent.”
- Depending on whether the data are univariate (one variable) or bivariate (two variables), we will perform a goodness-of-fit test or a test of homogeneity.

Count Data

Definition (Count data)

Data that counts the number of observations that fall into each of several categories.

- The data may be univariate or bivariate.
  - Univariate example - Observe a person's opinion of the mayor's performance (excellent, good, fair, poor, not sure).
  - Bivariate example - Observe a opinion of the mayor's performance and their political affiliation (Republican, Democrat, Independent).

Univariate Example

- Observe a person's opinion of the mayor's performance.

Excellent	Good	Fair	Poor	Not Sure
72	124	123	150	30

Bivariate Example

- Observe each person's opinion and political affiliation.

	Excellent	Good	Fair	Poor	Not Sure
Dem	38	70	59	86	16
Rep	21	31	33	23	7
Ind	13	23	31	41	7

The Two Basic Questions

Basic Question for Univariate Data

Do the data fit a specified distribution?

- For example, could these data have come from a uniform distribution?

Basic Question for Bivariate Data

Could the data in every row have come from the same distribution, whatever that distribution may be?

Observed and Expected Counts

Definition (Observed Counts)

The counts that were actually observed in the sample.

Definition (Expected Counts)

The counts that would be expected if the null hypothesis were true.

Tests of Goodness of Fit

Definition (Goodness-of-fit Test)

A hypothesis test that determines whether a set of data reasonably fits a specified distribution.

- The goodness-of-fit test applies only to univariate data.
- The null hypothesis specifies a discrete distribution for the population.
- We want to determine whether a sample from that population supports this hypothesis.

Example

- If we rolled a die 60 times, we expect 10 of each number.
- If we get frequencies 8, 10, 14, 12, 9, 7, does that indicate that the die is not fair?
- What is the distribution if the die were fair?

Example

- If we selected 20 people from a group that was 60% male and 40% female, we would expect to get 12 males and 8 females.
- If we got 15 males and 5 females, would that indicate that our selection procedure was not random (i.e., discriminatory)?
- What if we selected 100 people from the group and got 75 males and 25 females?

Example

- If we deal a 5-card poker hand, we should get
  - One pair 42.26% of the time,
  - Two pairs 9.51% of the time,
  - Three of a kind 2.11% of the time, and
  - Something else 46.12% of the time.
- Suppose we deal 1000 poker hands and get one pair 427 times, two pairs 102 times, three of a kind 26 times, and something else 445 times. Does this agree with the theory?

The Hypotheses

- The null hypothesis specifies the probability (or proportion) for each category.
- To test a die for fairness, the null hypothesis would be

$$H_0 : p_1 = \frac{1}{6}, p_2 = \frac{1}{6}, \ldots, p_6 = \frac{1}{6}.$$

- The alternative hypothesis will always be a simple negation of H_0:

$$H_1 : H_0 \text{is false.}$$

- Let $\alpha = 0.05$.

Expected Counts

- The test statistic will involve the observed and the expected counts.
- To find the expected counts, we apply the hypothetical probabilities to the sample size.
- For example, if the hypothetical probabilities are all $\frac{1}{6}$ and the sample size is 60, then the expected counts are

$$\left(\frac{1}{6}\right) \times 60 = 10.$$

The Test Statistic

- Make a chart showing both the observed and expected counts (in parentheses).

Value	1	2	3	4	5	6
Observed (Expected)	8 (10)	10 (10)	14 (10)	12 (10)	9 (10)	7 (10)

- Denote the observed counts by O and the expected counts by E.
- Define the chi-square (χ^2) statistic to be

$$\chi^2 = \sum_{\text{all cells}} \frac{(O - E)^2}{E}.$$

The Value of the Test Statistic

- Clearly, if all of the deviations $O - E$ are small, then χ^2 will be small.
- But if even a few the deviations $O - E$ are large, then χ^2 will be large.
- Now calculate χ^2.

$$\begin{aligned}\chi^2 &= \frac{(8-10)^2}{10} + \frac{(10-10)^2}{10} + \frac{(14-10)^2}{10} \\ &\quad + \frac{(12-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(7-10)^2}{10} \\ &= 0.4 + 0.0 + 1.6 + 0.4 + 0.1 + 0.9 \\ &= 3.4\end{aligned}$$

Compute the p-Value

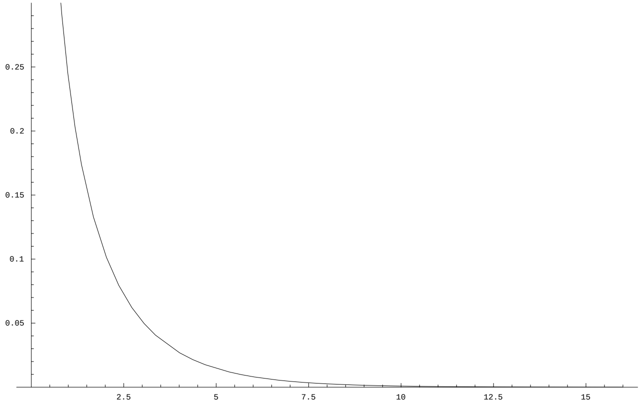
- To compute the p-value of the test statistic, we need to know more about the distribution of χ^2.
- How likely is it that we would observe a value as large as 3.4?

Chi-Square Degrees of Freedom

- The χ^2 distribution has an associated degrees of freedom, just like the *t* distribution.
- Each χ^2 distribution has a slightly different shape, depending on the number of degrees of freedom.
- In this test, *df* is one less than the number of cells.

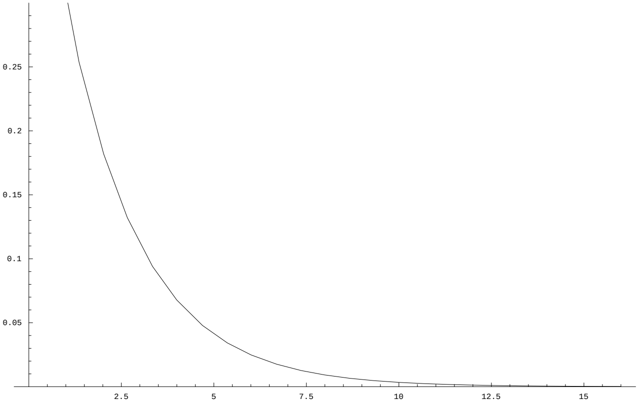
Chi-Square Degrees of Freedom

The Graph of χ^2_1 $(df = 1)$



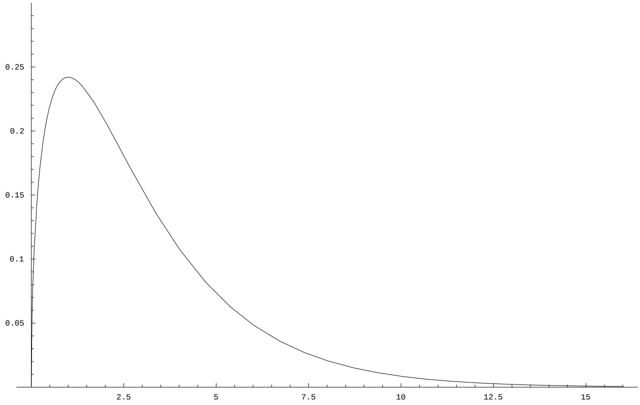
Chi-Square Degrees of Freedom

The Graph of χ^2_2 ($df = 2$)



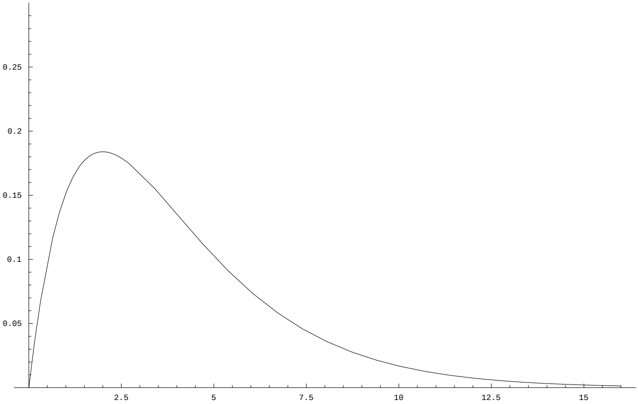
Chi-Square Degrees of Freedom

The Graph of χ^2_3 ($df = 3$)



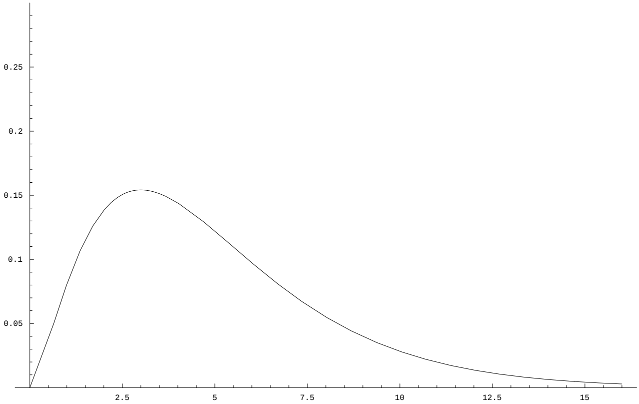
Chi-Square Degrees of Freedom

The Graph of χ^2_4 ($df = 4$)



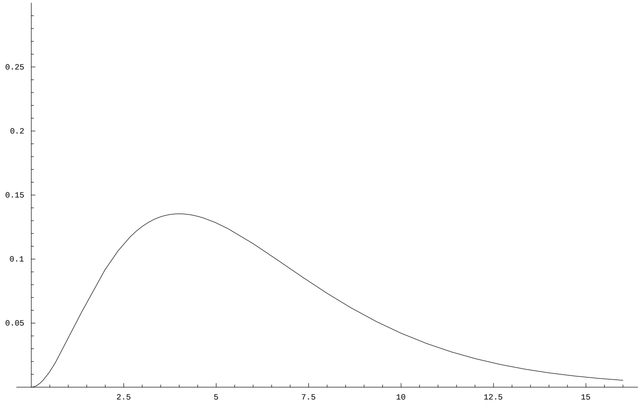
Chi-Square Degrees of Freedom

The Graph of χ^2_5 ($df = 5$)



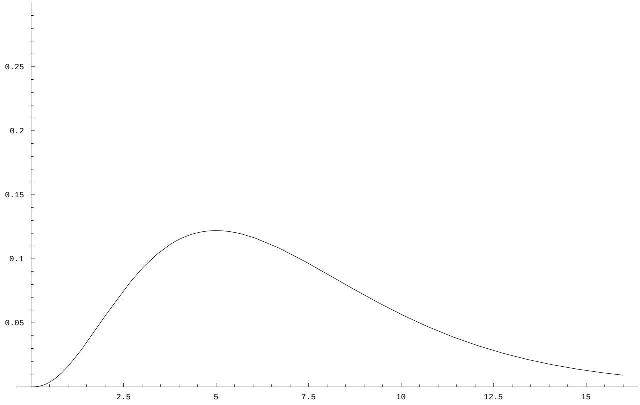
Chi-Square Degrees of Freedom

The Graph of χ^2_6 ($df = 6$)



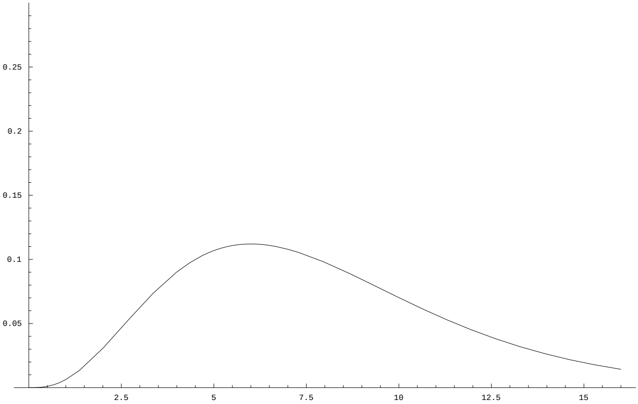
Chi-Square Degrees of Freedom

The Graph of χ^2_7 ($df = 7$)



Chi-Square Degrees of Freedom

The Graph of χ^2_8 ($df = 8$)

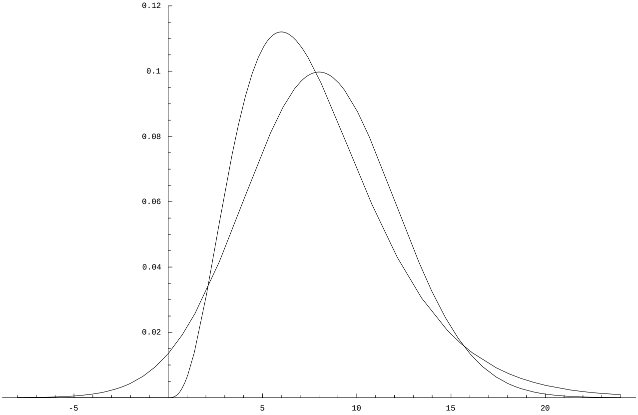


Properties of χ^2

- The chi-square distribution with *df* degrees of freedom has the following properties.
  - $\chi^2 \geq 0$.
  - It is unimodal.
  - It is skewed right (not symmetric!)
  - $\mu_{\chi^2} = df$.
  - $\sigma_{\chi^2} = \sqrt{2df}$.
  - If *df* is large, then χ^2_{df} is approximately normal with mean *df* and standard deviation $\sqrt{2df}$.

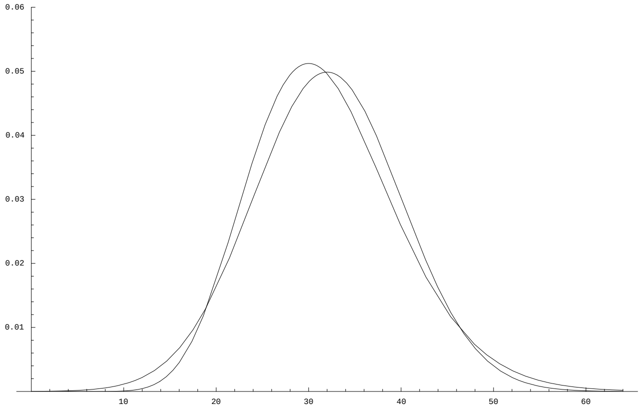
Chi-Square vs. Normal

The graph of χ^2_8 vs. $N(8, 4)$



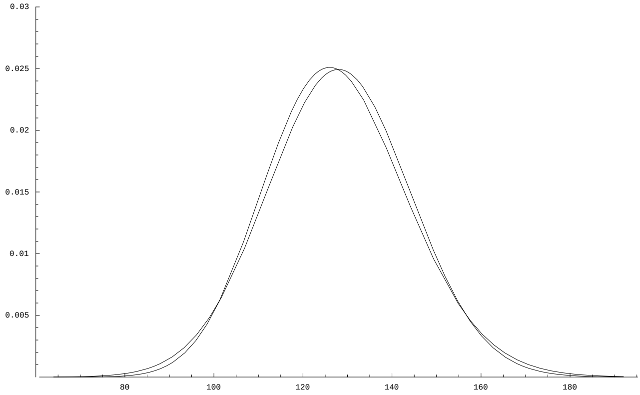
Chi-Square vs. Normal

The graph of χ^2_{32} vs. $N(32, 8)$



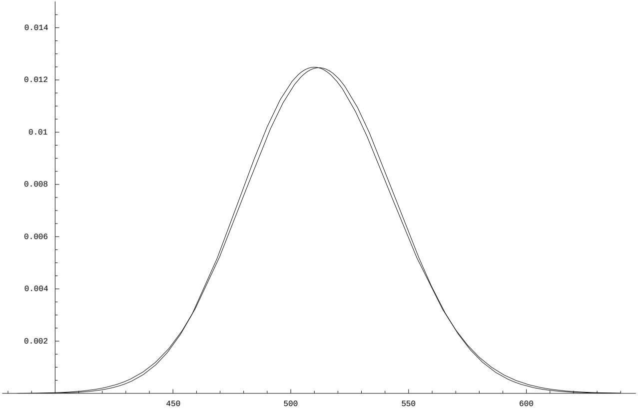
Chi-Square vs. Normal

The graph of χ^2_{128} vs. $N(128, 16)$



Chi-Square vs. Normal

The graph of χ^2_{512} vs. $N(512, 32)$



Summary

Summary

- The goodness-of-fit test will determine whether the data fit a specified distribution.
- The chi-square statistic is defined as

$$\chi^2 = \sum_{\text{all cells}} \frac{(O - E)^2}{E}.$$

- The closer the data match the distribution, the smaller the value of χ^2.