

JMP INTRO® Lab Activities

Lab Activity – χ^2 Goodness-of-Fit Test

Data Set: Denim

As you may remember from the Lab Activity, Exploring Categorical Data, Denim is a data set from an experiment done by a company that manufactures blue jeans. Denim fabric naturally contains starch, creating stiffness in the fabric. The denim fabric the company uses is washed using various treatments to make it feel “worn”. Although the feel of the fabric is important, the company is also concerned about the strength of the treated fabric.

Open the data file **DENIM** from the Sample Data folder. **Thread Wear** is an ordinal variable, created by Taking the numerical values in **Thread Wear Measured** and dividing them into categories created the data in the **Thread Wear** column. Look at the data table and describe in your report how this appears to have been done.

Use the **Analyze** → **Distribution** command to make a bar graph (histogram) of **Thread Wear**. Your display should show a histogram, a mosaic plot and a frequencies table that shows both counts and probabilities. Add a count axis to your histogram and copy the two graphs into your report.

Use the red diamond menu next to **Thread Wear** to select **Test Probabilities**. A Test Probabilities table should appear under the frequencies table with blanks to fill in the hypothesized probabilities. Suppose that the thread wear counts 1 through 10 are equally likely. Since the Low category was created by combining the thread wear ratings of 1 to 4 out of a scale of 10, it might be reasonable to assign a hypothesized value of 0.4 (4 of the possible 10 values) to the category Low. Also fill in the hypothesized values for the categories Moderate and Severe using this presumed relationship.

Before you have JMP INTRO run a chi-square goodness-of-fit test, figure out the test statistic yourself. You’ll need to figure the expected values by multiplying your hypothesized probabilities by the total number of trials. Then, for each value in the counts column (your observed values)

- 1) subtract: observed value – the expected value (the expected count you created)
- 2) square this difference
- 3) divide by the expected value.

Find the sum of all these values. This is called the **chi-square statistic**, with formal notation χ^2 (the Greek letter chi).

This is the formula: $\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$. Show your work in your report.

Now you are ready to have JMP INTRO do the goodness-of-fit test. Click on the **Done** button. JMP INTRO will find two chi-square test statistics and the associated p-values.

Copy the **Frequencies** table and the **Test Probabilities** table into your report. Include only the Pearson Chi-Square Test results. Did these results agree with the values you calculated by hand?

To be statistically significant at a significance level of 0.05, you should have a p-value or probability less than 0.05. What was your p-value? Were your results statistically significant? Remember to report your answers in the context of the problem.

Print a copy of your lab report. Proofread your copy, make any needed changes, close JMP INTRO and the word processing program, and log off.

JMP INTRO® Lab Activities

Teacher Notes

Lab Activity – X^2 Goodness-of-Fit Test

Objectives:

- Use JMP INTRO to figure expected values in a contingency table and run a chi-square goodness-of-fit test.
- Create a word processing document, incorporating graphs and tables from JMP INTRO.

Time Required: 45 minutes

Materials:

- X^2 Goodness-of-Fit Test student activity directions
- Denim data set

Prerequisites:

- Students should have experience producing numerical summaries and graphical displays of categorical data by hand and/or with graphing calculators.
- Students should have basic knowledge of how to use JMP INTRO.
- Students should have completed Lab Activity Exploring Categorical Data.

JMPINTRO Notes:

- The **Denim** data file is included in the Sample Data folder as part of JMP INTRO.
- The chi-square statistic that is calculated by the formula shown in the student directions and found in introductory textbooks is the Pearson test statistic.
- Additional data sets are available for download from the JMP website: www.jmpdiscovery.com. These can be used to extend this activity.

JMP INTRO® Lab Activities Answer Key

The following are suggested answers for χ^2 Goodness-of-Fit Test:

To create the ordinal variable Thread Wear, the values 1 – 4 were combined as Low, 5 – 9 were combined as Moderate, and 10 was classified as Severe.

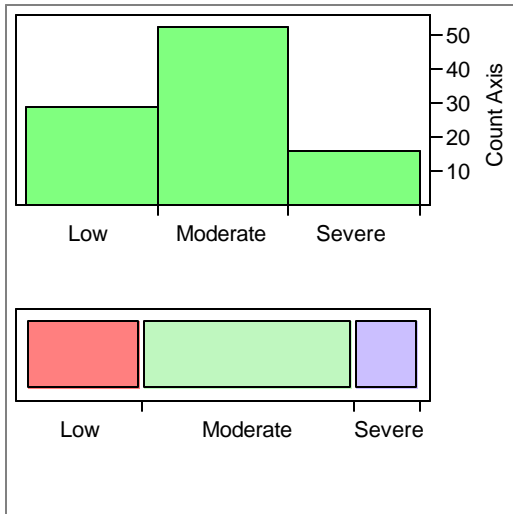
Observed	Expected
29	$0.4(98) = 39.2$
53	$0.5(98) = 49$
16	$0.1(98) = 9.8$

Chi-square calculation:

$$\chi^2 = \frac{(29 - 39.2)^2}{39.2} + \frac{(53 - 49)^2}{49} + \frac{(16 - 9.8)^2}{9.8}$$

$$\chi^2 = 2.65408 + 0.32653 + 3.92245$$

$$\chi^2 = 6.90306$$



Frequencies

Level	Count	Prob
Low	29	0.29592
Moderate	53	0.54082
Severe	16	0.16327
Total	98	1.00000

Test Probabilities

Level	Estim Prob	Hypoth Prob
Low	0.29592	0.40000
Moderate	0.54082	0.50000
Severe	0.16327	0.10000

Test	ChiSquare	DF	Prob>Chisq
Pearson	6.9031	2	0.0317

The chi-square value reported by JMP INTRO (6.9031) is essentially the same as the hand calculated value (6.90306) with rounding differences.

The reported p-value of 0.0317 is significant at the 0.05 level. There is evidence that the counts observed are different from what would be expected if the distribution were distributed as hypothesized. You would expect more fabric samples to have had low thread wear and fewer to have had moderate and severe thread wear than what was observed.