

POPULATION VARIANCES

© Nick Lee and Mike Peters 2016.

QUESTION 1.

You and a friend regularly eat cereal for breakfast. The brand that you buy states the average weight of the contents is 368 grams per box. You decide to check whether the manufacturer is telling the truth, so you decide to record the weight of the contents of each box you buy. After you have collected the data from 25 boxes you need to perform a statistical analysis. In order to do this you ask the manufacturer to let you know what they consider to be an acceptable level of variation between the boxes. They tell you they expect the standard deviation to be 15 grams.

Construct the null and alternative hypotheses.

$$H_0: \sigma^2 = \underline{\hspace{2cm}}$$

$$H_1: \sigma^2 \neq \underline{\hspace{2cm}}$$

The number of degrees of freedom, $df = \underline{\hspace{2cm}}$

At a level of significance of 0.5, the lower critical value is $\underline{\hspace{2cm}}$ and the upper critical value is $\underline{\hspace{2cm}}$.

Complete the following:

Reject H_0 if $\chi_{stat}^2 < \underline{\hspace{2cm}}$ or if $\chi_{stat}^2 > \underline{\hspace{2cm}}$.

$$\chi_{stat}^2 = \frac{(n-1)S^2}{\sigma^2} = \underline{\hspace{2cm}} = \underline{\hspace{2cm}}$$

Based on the statistical evidence, you can/cannot (delete as appropriate) reject the null hypothesis. Your conclusion is that there is sufficient/insufficient (delete as appropriate) evidence that the $\underline{\hspace{2cm}}$ $\underline{\hspace{2cm}}$ is different from $\underline{\hspace{2cm}}$ grams.

QUESTION 2.

A car manufacturer is interested in finding out if the reliability of its cars has changed within the past year. In order to do this, they have asked you to survey 25 of its customers asking them how much money they spent on car repairs in the previous 12 months. The company has supplied you with the results from the same survey carried out last year. The estimated standard deviation from last year's population was £200 and the standard deviation of this year's sample was £237.52.

Construct the null and alternative hypotheses.

$$H_0: \sigma = \underline{\hspace{2cm}}$$

$$H_1: \sigma = \underline{\hspace{2cm}}$$

$$df = \underline{\hspace{2cm}}$$

State the conditions for rejecting H_0 : $\underline{\hspace{2cm}}$

$$\chi^2 = \underline{\hspace{2cm}}$$

Is there sufficient evidence to suggest the population standard deviation is different from £200 at the 0.05 level of significance?

Complete the following:

In order to use the χ^2 test of a population variance the data is assumed to be $\underline{\hspace{2cm}}$.

The p -value is $\underline{\hspace{2cm}}$ which means there is a $\underline{\hspace{2cm}}$ % probability of obtaining a result greater than the sample value of £237.52 when the null/alternative (delete as appropriate) is true.

QUESTION 3.

You have been employed by a computer training company to assess the anxiety levels of employees working for a private care company. You use a rating scale from 20 (no anxiety) to 100 (highest anxiety) and administer the test to 172 of the care company's employees. The results of the survey are shown in the table below:

	Males	Females
\bar{X}	40.26	36.85
S	13.35	9.42
n	100	72

Construct the null and alternative hypotheses.

$$H_0: \underline{\hspace{2cm}}$$

$$H_1: \underline{\hspace{2cm}}$$

To do this analysis an F test can be done assuming both populations are $\underline{\hspace{2cm}}$ distributed.

Complete the following:

If $F_{stat} > \underline{\hspace{2cm}}$ then the $\underline{\hspace{2cm}}$ can be rejected.

The test statistic $F_{stat} = \frac{S_1^2}{S_2^2} = \frac{\quad}{\quad} = \frac{\quad}{\quad} = \quad$

Because the $F_{stat} = \quad$ which is greater than \quad , there is significant evidence to reject the null hypothesis.

This suggests there is evidence that the two population variances are different.

Because of this result, you decide to test whether there is a significant difference in mean computer anxiety for male and female employees.

Which test would you use to do this?

Question 4.

You have been asked by your university's Learning and Teaching committee to analyse some data about two ways of teaching mathematics. The traditional method did not use Matlab (a computer mathematics program) whereas the experimental method did. Exam results from two consecutive years were analysed and from these results 25 random samples were selected. The variances are given below.

Year 1: variance = 25.3

Year 2: variance = 33.5

Construct the null and alternative hypotheses.

H_0 : \quad

H_1 : \quad

Complete the following:

The test statistic $F_{stat} = \quad$

The p -value = \quad

The statistical evidence suggests that the variances are/are not (delete as appropriate) equal. Therefore the alternative hypothesis can/cannot be accepted.

MINI PROJECT

The government department for which you work is interested in using email to communicate with the general population. The section manager has to come up with a target time for responding to email enquiries. He wants to know if people of different age groups expect different response times. In order to answer his question you have decided to look at two distinct age groups: the over 60s and those aged between 18 and 60. The results from a survey of 1000 people over the age of 60 suggested that 70.7% of this age group expected a response within one working day whereas 53.6% of 1000 in the 18 to 60 age group expected an answer within one working day.

Using a 0.01 level of significance, decide if there is evidence to suggest a significant difference between the two age groups' expectations of a response within one working day.

Note: In order to complete this mini project you will need to investigate a technique not covered in the text associated with this workbook. To get you started put the data into a table as shown below:

	age > 60	18 < age < 60
1 day	707	536
3 days	293	464

You then need to calculate the estimated overall proportion for both groups:

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{X}{n}$$

Then you can use the formula below to calculate the test statistic:

$$\chi_{stat}^2 = \sum_{allcells} \frac{(f_o - f_e)^2}{f_e}$$

where

f_o is the observed frequency for one entry in the table.

f_e is the expected frequency for one entry in the table if the null hypothesis is true.

The final bit is to calculate the number of degrees of freedom.

This value is equal to the number of rows in the table minus 1 multiplied by the number of columns minus 1.

And finally...

If my half of an apple pie is bigger than your half, whose is the biggest proportion of the whole pie?