Notation & Definitions

The **standard deviation** and the **variance** are measures of variability. They measure how much the values of a variable in a given sample or population vary above and below the mean. A population or sample with values that are closer to the mean would have a lower variance and a lower standard deviation than a population or sample with values more spread out around the mean.

|  |  |  |  |
| --- | --- | --- | --- |
| Population Parameters | | Sample Statistics | |
|  | The standard deviation of a population  The lowercase Greek letter “sigma” |  | The standard deviation of a sample |
|  | The variance of a population  “Sigma squared” |  | The variance of a sample |

The **standard deviation** is the **square root** of the **variance**:

Likewise, the **variance** is the **standard deviation squared**, exactly how it appears in the notation.

|  |
| --- |
| *Example.* If the standard deviation of a sample is 12, what is the variance of the sample?  Thus, the variance of the sample is 144 and is denoted with |

NOTE: problem descriptions could refer to either a variance OR a standard deviation. You must treat them accordingly in the equations for this chapter. If you notate the values properly, then you will know when a value needs to be squared and when it does not.

Each of these letters can be subscripted to refer to specific populations and samples, like this:

|  |  |  |  |
| --- | --- | --- | --- |
| Population Parameters with subscripts | | Sample Statistics with subscripts | |
|  | The standard deviation of population 1 |  | The standard deviation of sample 1 (i.e. the sample from population 1) |
|  | The standard deviation of population 2 |  | The standard deviation of sample 2 (i.e. the sample from population 2) |
|  | The variance of population 1 |  | The variance of sample 1 (i.e. the sample from population 1) |
|  | The variance of population 2 |  | The variance of sample 2 (i.e. the sample from population 2) |

This chapter will introduce two new test statistics, with corresponding sampling distributions and a host of critical values. The notation we will use is:

|  |  |
| --- | --- |
| Chi-Square | |
|  | This is the Greek letter “chi” squared.  Pronounced “kai square" |
|  | A chi-square test statistic |
|  | A critical value of chi-square at in the lower tail.  The critical value in a **lower-tail** test with a test statistic. |
|  | A critical value of chi-square at in the upper tail.  The critical value in an **upper-tail** test with a test statistic. |
|  | A critical value of chi-square at in the lower tail.  The **lower tail critical value** in a **two-tailed** test with a test statistic. |
|  | A critical value of chi-square at in the upper tail.  The **upper tail critical value** in a **two-tailed** test with a test statistic. |

|  |  |
| --- | --- |
|  | |
|  | An test statistic |
|  | An upper-tail critical value of at  The critical value in an **upper-tail** test with an test statistic. |
|  | A two-tailed critical value of at  The critical value in a **two-tailed** test with an test statistic. |

**Part One: Hypothesis Tests about a Single Population Variance,**

1. **Formulating the Hypotheses**: There are three possible forms of hypotheses. They each correspond to a different question you might want to ask about the true value of the population variance,

|  |  |  |
| --- | --- | --- |
| Hypotheses for Hypothesis Tests about a Single Population Variance, | | |
| **Lower Tail Test** | **Upper Tail Test** | **Two-Tailed Test** |
|  |  |  |
| Answers Questions About: | | |
| If the true population variance, is less than a given  number, | If the true population variance, is greater than a given number, | If the true population variance, is different from a given number, |

Upper Tail and Lower Tail Tests are called *one-tailed* or *directional* tests. Two-tailed tests are called *non-directional* tests.

1. **The Test Statistic:**

Under the assumption that is true as an equality, the sampling distribution of the sample variance, follows the distribution. Therefore, we standardize our sample against the distribution by calculating the following test statistic:

where

and the degrees of freedom are

NOTE: some problems will give you the variances (, in which case you simply plug the values into the equation as is. However, some problems will give you standard deviations ( in which case you would need to square those values in this equation.

1. **Deciding whether or not to Reject :**

There are two approaches to deciding whether or not to reject the Null:

* In the ***p-value approach***, we **compare** the **p-value** of our test statistic to the  **significance level** and reject if the p-value is less than or equal to the significance level.
* In the ***critical value approach***, we **compare** the **test statistic** to **a critical value** – this can be done with a diagram in which we use the critical value(s) to construct a rejection region or regions; if the test statistic is in a rejection region, we reject and accept Or, this step can be accomplished by following the mathematical rules given in the tables below.

REMINDER: you can never *accept* the null hypothesis. Remember the swans.

|  |  |  |  |
| --- | --- | --- | --- |
| **When to Reject for Test Statistics:** | | | |
|  | **For a Lower Tail Test:** | **For an Upper Tail Test:** | **For a Two-Tailed Test:** |
| **p-value approach:** | Calculate the lower tail of  If the then reject and accept  If the then do not reject . is unsupported. | Calculate the upper tail of  If the then reject and accept  If the then do not reject . is unsupported. | The two-tailed is two times the one-tailed p-value of  If the then reject and accept  If the then do not reject . is unsupported. |
| **Critical Value: Approach** | If then reject and accept  If , then do not reject is unsupported. | If then reject and accept  If , then do not reject is unsupported. | If  , then reject and accept .  If then do not reject is unsupported. |
| NOTES:   1. is a Test Statistic 2. are Critical Values 3. is based on degrees of freedom. Be sure to calculate the appropriate degrees of freedom for the test you are performing. | | | |

1. **Interpreting the test:**

(Note: This explanation of interpretation holds for ALL hypothesis tests.) We start every hypothesis test with a question about the parameter of interest, so we must end every hypothesis test with the answer to that question. In other words, we must *interpret* the conclusion of our test in terms of the original question.

Remember: in hypothesis testing you can never prove the null hypothesis. You can only prove the alternative hypothesis: when you reject the null and accept the alternative, then at your given level of significance you may conclude that is true. If you do not reject then you must conclude that is unsupported by the evidence. This gives us a clear guideline for how to *interpret* hypothesis tests: *always look at the alternative hypothesis!*

In all that follows, you would substitute the actual words and numbers from your hypothesis test for the symbols. Notice that each interpretation simply states the alternative hypothesis in words, and says either that it is true or that we cannot conclude that it is true.

|  |  |  |  |
| --- | --- | --- | --- |
| **How to Interpret a Hypothesis Test about a Single Population Variance, :** | | | |
| **When you:** | **For a Lower Tail Test:** | **For an Upper Tail Test:** | **For a Two-Tailed Test:** |
| **Reject** | At the significance level, we can conclude that is less than . | At the significance level, we can conclude that is greater than . | At the significance level, we can conclude that is different than . |
| **Do not reject** | At the significance level, we cannot conclude that is less than . | At the significance level, we cannot conclude that is greater than . | At the significance level, we cannot conclude that is different than . |
| NOTES:   1. is the true value of the population variance 2. is a number that is the hypothesized value of | | | |

**Assumptions Underlying Hypothesis Tests about a Single Population Variance**

All hypothesis tests use sampling distributions to determine the probability of sample statistics. In order for us to be confident that our choice of sampling distribution for any given test really is the way the sample statistic is distributed, certain assumptions must be met. If the assumptions are not met – that is, if any given assumption is not true – then we cannot rely on the results of the hypothesis tests. They may mislead us, give us the wrong answers, and cause us to draw the wrong conclusions.

When making inferences about a single population variance using :

1. The underlying population must be normally distributed. If it is not normal, then the sampling distribution is unknown, so we cannot use as the sampling distribution.
2. The sample must be random. If it is not random, then the sampling distribution is unknown, so we cannot use as the sampling distribution.

**Part Two: Hypothesis Tests about Two Population Variances,**

1. **Formulating the Hypotheses**: There are two possible forms of hypotheses: upper tail and two-tailed. Whenever a directional question is asked (that is, a less-than or greater-than question), we will formulate the hypothesis test as an upper tail test. The two forms correspond to different questions you might want to ask about the true values of two population variances,

|  |  |
| --- | --- |
| Hypotheses for Hypothesis Tests about Two Population Variances, | |
| **Upper Tail Test** | **Two-Tailed Test** |
|  |  |
| Answers questions about: | |
| If the true population variance of population 1, is **greater than** the true population variance of population 2,  NOTE: This test also tells you if the variance of population 2, , is **less than** the variance of population 1, . | If the true population variance of population 1, is **different from** the true population variance of population 2, |

The Upper Tail Test is a *one-tailed* or *directional* test. Two-tailed tests are called *non-directional* tests.

1. **The Test Statistic:**

**To properly formulate the test statistic, you MUST use the larger sample variance as and the smaller sample variance as . You can use this information to label the populations: the sample with the larger sample variance is sample 1 from Population 1, and the sample with the smaller variance is sample 2 from Population 2. By doing this, you will ensure that the test statistic is in the upper tail of the F distribution, and the critical values on the F table will work properly.**

For hypothesis testing about two population variances, the sample statistics of interest are the sample variances, Under the assumption that is true as an equality, the sampling distribution of the ratio of the sample variances follows the F distribution. Therefore, we standardize our samples against the F distribution by calculating the following F test statistic:

where

NOTE: some problems will give you the variances (, in which case you simply plug the values into the equation as is. However, some problems will give you standard deviations ( in which case you would need to square those values in this equation.

1. **Deciding whether or not to Reject :**

There are two approaches to deciding whether or not to reject the Null:

* In the ***p-value approach***, we **compare** the **p-value** of our test statistic to the  **significance level** and reject if the p-value is less than or equal to the significance level.
* In the ***critical value approach***, we **compare** the **test statistic** to **a critical value** – this can be done with a diagram in which we use the critical value(s) to construct a rejection region or regions; if the test statistic is in a rejection region, we reject and accept Or, this step can be accomplished by following the mathematical rules given in the tables below.

REMINDER: you can never *accept* the null hypothesis. Remember the swans.

The Hypothesis Testing rejection rules for test statistics are:

|  |  |  |
| --- | --- | --- |
| **When to Reject if the Test Statistic is** | | |
|  | **For an Upper Tail Test:** | **For a Two-Tailed Test:** |
| **p-value approach:** | Calculate the upper tail of  If the then reject and accept  If the then do not reject is unsupported. | The two-tailed is  2 x upper tail p-value of  If the then reject and accept  If the then do not reject is unsupported. |
| **Critical Value Approach:** | If then reject and accept  If then do not reject is unsupported. | If then reject and accept  If then do not reject is unsupported. |
| NOTES:   1. is a Test Statistic 2. and are Critical Values. 3. The degrees of freedom are and | | |

1. **Interpreting the test:**

(Note: This explanation of interpretation holds for ALL hypothesis tests.) We start every hypothesis test with a question about the parameter of interest, so we must end every hypothesis test with the answer to that question. In other words, we must *interpret* the conclusion of our test in terms of the original question.

Remember: in hypothesis testing you can never prove the null hypothesis. You can only prove the alternative hypothesis: when you reject the null and accept the alternative, then at your given level of significance you may conclude that is true. If you do not reject then you must conclude that is unsupported by the evidence. This gives us a clear guideline for how to *interpret* hypothesis tests: *always look at the alternative hypothesis!*

In all that follows, you would substitute the actual words and numbers from your hypothesis test for the symbols. Notice that each interpretation simply states the alternative hypothesis in words, and says either that it is true or that we cannot conclude that it is true.

|  |  |  |
| --- | --- | --- |
| **How to Interpret a Hypothesis Test about Two Population Variances, :** | | |
| **When you:** | **For an Upper Tail Test:** | **For a Two-Tailed Test:** |
| **Reject** | This result can be interpreted two ways. Choose the interpretation that is appropriate in the context of the problem:   1. At the significance level, we can conclude that is **greater** than . 2. At the significance level, we can conclude that is **less** than . | At the significance level, we can conclude that is different than . |
| **Do not reject** | This result can be interpreted two ways. Choose the interpretation that is appropriate in the context of the problem:   1. At the significance level, we cannot conclude that is **greater** than . 2. At the significance level, we cannot conclude that is **less** than . | At the significance level, we cannot conclude that is different than . |
| NOTES:   1. is the true variance of population 1; is the true variance of population 2 | | |

**Assumptions Underlying Hypothesis Tests about Two Population Variances**

All hypothesis tests use sampling distributions to determine the probability of sample statistics. In order for us to be confident that our choice of sampling distribution for any given test really is the way the sample statistic is distributed, certain assumptions must be met. If the assumptions are not met – that is, if any given assumption is not true – then we cannot rely on the results of the hypothesis tests. They may mislead us, give us the wrong answers, and cause us to draw the wrong conclusions.

When making inferences about two population variances using :

1. The underlying populations must both be normally distributed. If they are not normal, then the sampling distribution is unknown, so you cannot use F as the sampling distribution.
2. Both samples must be random, and they must be independent of one another. If they are not random and independent, then the sampling distribution is unknown, so you cannot use F as the sampling distribution.

**The Chi-Square ( Distribution**

The test statistics and sampling distributions are used for inferences about a single population variance (Ch 11), goodness of fit tests (Ch 12), and tests of independence (Ch 12).

The **Chi-Square distribution**:

* is a probability distribution, so the area under the curve is 1
* has degrees of freedom (df). The calculation for degrees of freedom will depend on the type of hypothesis test or confidence interval.
* is ALWAYS positive and is NOT symmetrical
  + Implication: Critical Values in the upper tail **and** lower tail will be **two different positive numbers**
* The mean of a chi-square distribution is equal to its degrees of freedom. How is this relevant?
  + If , then the test statistic is in the **upper tail**
  + If , then the test statistic is in the **lower tail**

The Chi-Square table contains Critical Values. Use the column headings marked “Area in Lower Tail” for LT critical values, and “Area in Upper Tail” for UT critical values. Excel is used to calculate p-values for Chi-Square test statistics.

The Hypothesis Testing rejection rules for test statistics are listed in Ch 11: Handout #2, Part 1.

**Practice using Chi-Square: Exercises**

***Exercise 1*:**

a) If df = 20 and , what is the critical value of Chi-square for an upper tail test?

b) Suppose the test-statistic in this upper tail test is . Use the critical value approach to decide whether or not to reject the null hypothesis.

c) Suppose the test-statistic in this upper tail test is . Use the p-value approach to decide whether or not to reject the null hypothesis.

***Exercise 2:***

a) If df = 45 and , what is the critical value of Chi-Square for a lower tail test?

b) Suppose that the test-statistic in this lower tail test is . Use the critical value approach to decide whether or not to reject the null hypothesis.

c) Suppose that the test-statistic in this lower tail test is . Use the p-value approach to decide whether or not to reject the null hypothesis.

***Exercise 3:***

a) If df = 16 and , what are the critical values of Chi-square for a two tailed test?

b) Suppose that the test-statistic in this two-tailed test is Use the critical value approach to decide whether or not to reject the null hypothesis.

c) Suppose that the test-statistic in this two-tailed test is Use the p-value approach to decide whether or not to reject the null hypothesis.

**Hypothesis Tests about a Single Population Variance, : Exercises**

*Example 1:* A cranberry juice manufacturing company has designed its bottle filling process with a variance of 0.0225 ounces2. It conducts regular quality control tests to see whether the process is running according to this design specification. To that end, a quality control analyst takes a random sample of 25 bottles of cranberry juice and finds this sample has a variance of 0.04 ounces2. The population of bottle contents is assumed to be normal.

Conduct and interpret a hypothesis test at the significance level to determine whether the population variance of the contents of the bottles is different from 0.0225. Is there evidence that the process is out of adjustment?

*Example 2:* Kombucha is a sparkling, fermented tea which naturally contains some alcohol. In order to prevent kombucha from being regulated as an alcoholic beverage, makers must keep their tea from exceeding the legal limit for alcohol percentage. In order to stay legal, the manufacturer needs the population variance of their kombucha to be less than 0.00985. The population is normally distributed. The manufacturer takes a random sample of 66 bottles of kombucha and calculates a sample variance of 0.00624.

Conduct and interpret a hypothesis test to determine whether the population variance of alcohol percentage in the kombucha is less than 0.00985 at an significance level. Is the kombucha brewer staying within the legal limit?

**The F distribution**

The F test statistic and F sampling distribution are used for inferences about two population variances (Ch 11), analysis of variance tests (Ch 13), and regression analysis (Ch 14-15).

The distribution:

* A probability distribution, so the area under the curve is 1
* Has TWO separate degrees of freedom
  + The *numerator degrees of freedom ,* which is based on the numerator in the test statistic.
  + The *denominator degrees of freedom* , which is based on the denominator of the test statistic.
* is ALWAYS positive and is NOT symmetrical
  + In hypothesis tests about two population variances, we will construct our test statistic to be in the upper tail, so there will be one critical value for the directional test (upper tail) and one critical value for the two-tailed test.

The F table contains a selection of Critical Values. Excel is used to calculate p-values for F test statistics.

The Hypothesis Testing rejection rules for test statistics are listed in Ch 11: Handout #2, Part 2.

**Practice using F: Exercises**

*Exercise 1:* Consider an upper tail test at an significance level, with

1. What is the critical value?
2. Suppose the test statistic is . By the Critical Value approach, decide whether or not to reject the null hypothesis.
3. Suppose the test statistic is . By the p-value approach, decide whether or not to reject the null hypothesis.

*Exercise 2:* Consider a two-tailed test at an significance level, with

1. What is the critical value?
2. Suppose that the test statistic is . By the Critical Value approach, decide whether or not to reject the null hypothesis.
3. Suppose that the test statistic is By the p-value approach, decide whether or not to reject the null hypothesis.

**Hypothesis Tests about Two Population Variances, : Exercises**

*Example 1:* Our transport company would like to prove that we have lower variability in delivery times than our competitor. An analyst at our company obtains two random, independent samples. A random sample of 41 of our deliveries had a variance in delivery time of 2.56. A random sample of 61 of our competitor’s deliveries had a variance in delivery time of 4.15. Delivery times for both companies are normally distributed.

Conduct and interpret a hypothesis test to determine whether we have a lower population variance than our competitor at an significance level.

*Example 2:* A manufacturer of bathroom scales makes two models, Model A and Model B. Reliability, which is the extent to which a measurement is consistent (i.e. the same) when repeated under similar conditions, is an important characteristic of a bathroom scale. A scale that is more reliable can command a higher price. (HINT: Consider what reliability means in terms of the variance)

The manufacturer conducted a series of tests in which it weighed a standard 150lb weight on the two scales. It then sampled the weights displayed on the scales. These were independent, random samples, and the populations are assumed to be normally distributed. The sample size for Model A was 21 weigh-ins and the sample variance was 0.16. The sample size for Model B was 26 weigh-ins and the sample variance was 0.0625.

Conduct and interpret a hypothesis test to determine whether the population variances for these two scales differ at the significance level. Based on your analysis, which scale should have a higher price?