Section 3.4 – Estimating a Population Standard Deviation

Chi-Square Distribution

In a normally distributed population with variance σ^2 assume that we randomly select independent samples of size *n* and, for each sample, compute the sample variance s^2 (which is the square of the sample standard deviations). The sample statistic χ^2 (pronounced chi-square) has a sampling distribution called the ***chi-square distribution***.

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

Where

n = sample size

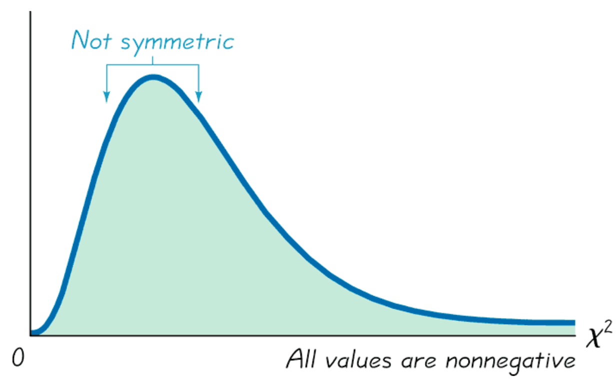
s^2 = sample variance

σ^2 = population variance

degrees of freedom = $n - 1$

Properties of the Distribution of the Chi-Square Statistic

1. The chi-square distribution is not symmetric, unlike the normal and Student *t* distributions. As the number of degrees of freedom increases, the distribution becomes more symmetric.



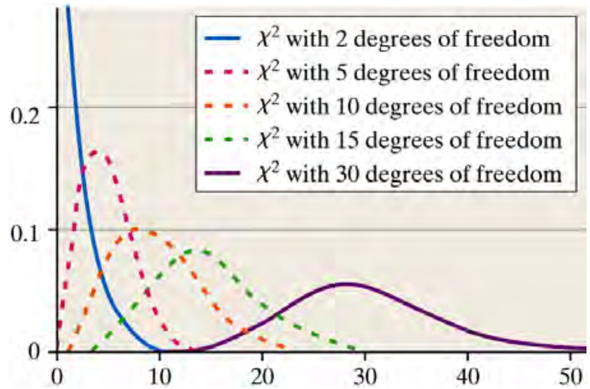
Chi-Square Distribution

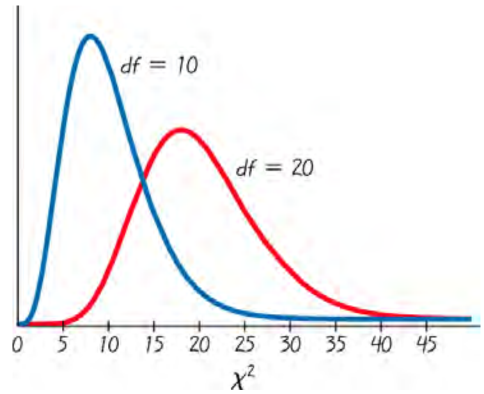
2. /The shape of the chi-square distribution depends on the degrees of freedom, just like the Student's *t*-distribution.

3. The values of chi-square can be zero or positive, but they cannot be negative.

4. The chi-square distribution is different for each number of degrees of freedom, which is df = $n - 1$. As the number of degrees of freedom increases, the chi-square distribution becomes more nearly symmetric.

In ***Table Chi-Square*** $\left(\chi^2\right)$ ***Distribution***, each critical value of χ^2 corresponds to an area given in the top row of the table, and that area represents the cumulative area located to the right of the critical value.

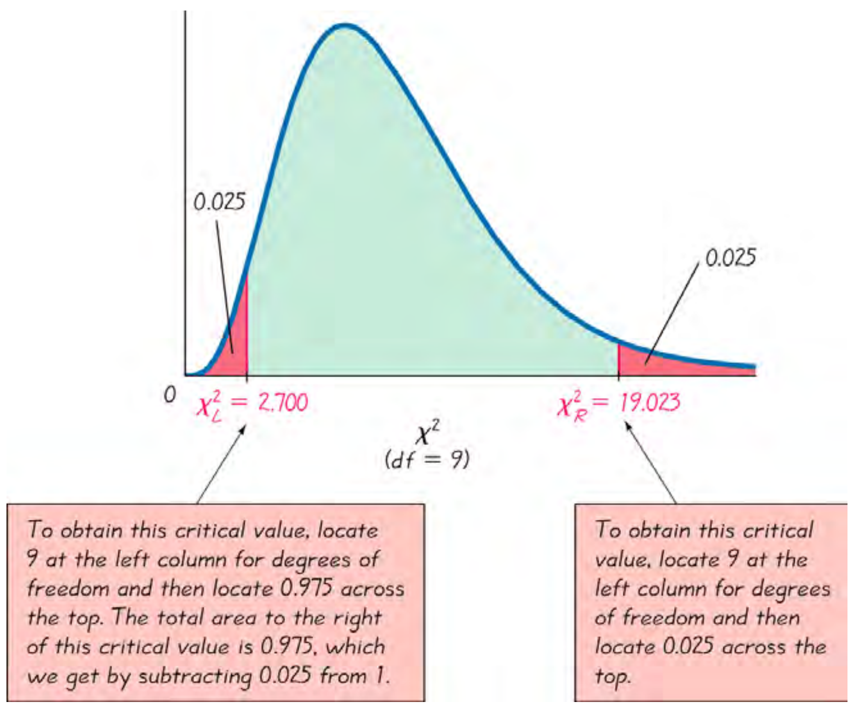




Chi-Square Distribution for df = 10 and df = 20

Example

A simple random sample of ten voltage levels is obtained. Construction of a confidence interval for the population standard deviation σ requires the left and right critical values of χ^2 corresponding to a confidence level of 95% and a sample size of $n = 10$. Find the critical value of χ^2 separating an area of 0.025 in the left tail, and find the critical value of χ^2 separating an area of 0.025 in the right tail.

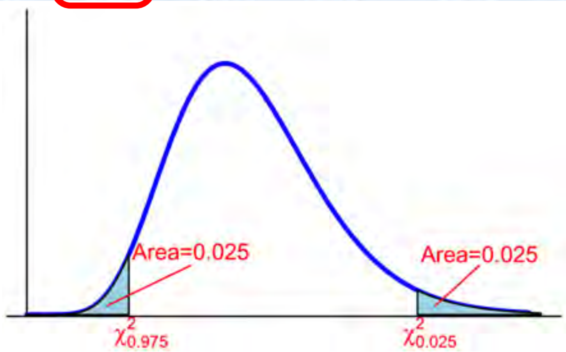


Example

Find the chi-square values that separate the middle 95% of the distribution from the 2.5% in each tail. Assume 18 degrees of freedom.

Solution

Degrees of Freedom	Chi-Square (χ^2) Distribution Area to the Right of Critical Value									
	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156



$\chi^2_{0.975} = 8.231$ $\quad\quad$ $\chi^2_{0.025} = 31.526$

Confidence Interval for Estimating a Population Standard Deviation or Variance

σ = population standard deviation
s = sample standard deviation
n = number of sample values
$\chi^2_L = \chi^2_{\alpha/2}$ = left-tailed critical value of χ^2

σ^2 = population variance
s^2 = sample variance
E = margin of error
$\chi^2_R = \chi^2_{1-\alpha/2}$ = right-tailed critical value of χ^2

Estimators of σ^2

The sample variance s^2 is the best point estimate of the population variance σ^2.

Estimators of σ

The sample standard deviation s is a commonly used point estimate of σ (even though it is a biased estimate).

Requirements:

1. The sample is a simple random sample.
2. The population must have normally distributed values (even if the sample is large).

Confidence Interval for the Population Variance σ^2

$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$

Confidence Interval for the Population Standard Deviation σ

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

Procedure for Constructing a Confidence Interval for σ or σ^2

1. Verify that the required assumptions are satisfied.
2. Using $n - 1$ degrees of freedom, refer to ***Table Chi-Square $\left(\chi^2\right)$ Distribution*** or use technology to find the critical values χ_R^2 and χ_L^2 that correspond to the desired confidence level.
3. Evaluate the upper and lower confidence interval limits using this format of the confidence interval:
$$\frac{(n-1)s^2}{\chi_R^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_L^2}$$
4. If a confidence interval estimate of σ is desired, take the square root of the upper and lower confidence interval limits and change σ^2 to σ.
5. Round the resulting confidence level limits. If using the original set of data to construct a confidence interval, round the confidence interval limits to one more decimal place than is used for the original set of data. If using the sample standard deviation or variance, round the confidence interval limits to the same number of decimals places.

Confidence Intervals for Comparing Data *Caution*

Confidence intervals can be used ***informally*** to compare the variation in different data sets, but *the overlapping of confidence intervals should not be used for making formal and final conclusions about equality of variances or standard deviations*.

Example

The proper operation of typical home appliances requires voltage levels that do not vary much. Listed below are ten voltage levels (in volts) recorded in the author's home on ten different days. These ten values have a standard deviation of $s = 0.15$ volt. Use the sample data to construct a 95% confidence interval estimate of the standard deviation of all voltage levels.

123.3 123.5 123.7 123.4 123.6 123.5 123.5 123.4 123.6 123.8

Solution

Requirements are satisfied: simple random sample and normality

Construct the confidence interval: $n = 10$, $s = 0.15$

$$\chi^2_L = 2.700 \quad and \quad \chi^2_R = 19.023$$

$$\frac{(n-1)s^2}{\chi^2_R} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_L}$$

$$\frac{(10-1)(0.15)^2}{19.023} < \sigma^2 < \frac{(10-1)(0.15)^2}{2.70}$$

$$0.010645 < \sigma^2 < 0.0750$$

Finding the square root of each part (before rounding), then rounding to two decimal places, yields this 95% confidence interval estimate of the population standard deviation:

$$0.10\ volt < \sigma < 0.27\ volt$$

✓ Based on this result, we have 95% confidence that the limits of 0.10 volt and 0.27 volt contain the true value of σ.

Example

We want to estimate the standard deviation σ of all voltage levels in a home. We want to be 95% confident that our estimate is within 20% of the true value of σ. How large should the sample be? Assume that the population is normally distributed.

Solution

From the Table, we can see that 95% confidence and an error of 20% for σ correspond to a sample of size 48. We should obtain a simple random sample of 48 voltage levels form the population of voltage levels.

TI-83/84 PLUS The TI-83/84 Plus calculator does not provide confidence intervals for σ or σ^2 directly, but the program **S2INT** can be used. That program was written by Michael Lloyd of Henderson State University, and it can be downloaded from www.aw.com/triola. The program S2INT uses the program ZZINEWT, so that program must also be installed. After storing the programs on the calculator, press the **PRGM** key, select **S2INT,** and enter the sample variance s^2, the sample size n, and the confidence level (such as 0.95). Press the **ENTER** key, and wait a while for the display of the confidence interval limits for σ^2. Find the square root of the confidence interval limits if an estimate of σ is desired.

Exercises *Section* 3.4 – Estimating a Population Standard Deviation

1. Using the weights of the M&M candies. We use the standard deviation of the sample $(s = 0.05179\ g)$ to obtain the following 95% confidence interval estimate of the standard deviation of the weights of all M&Ms: $0.0455\ g < \sigma < 0.0602\ g$. Write a statement that correctly interprets that confidence interval.

2. Find χ_L^2 and χ_R^2 that corresponds to: 95%; $n = 9$

3. Find χ_L^2 and χ_R^2 that corresponds to: 99%; $n = 81$

4. Find χ_L^2 and χ_R^2 that corresponds to: 90%; $n = 51$

5. Find a confidence interval for the population standard deviation σ.
95% *confidence*; $n = 30$, $\bar{x} = 1533$, $s = 333$ (*Assume has a normal distribution*)

6. Find a confidence interval for the population standard deviation σ
95% *confidence*; $n = 25$, $\bar{x} = 81.0\ mi/h$, $s = 2.3\ mi/h$ (*Assume has a normal distribution*)

7. Find a confidence interval for the population standard deviation σ
99% *confidence*; $n = 7$, $\bar{x} = 7.106$, $s = 2.019$ (*Assume has a normal distribution*)

8. In a study of the effects of prenatal cocaine use on infants, the following sample data were obtained for weights at birth: $n = 190$, $\bar{x} = 2700\ g$, $s = 645\ g$. Use the sample data to construct a 95% confidence interval estimate of the standard deviation of all birth weights of infants born to mothers who used cocaine during pregnancy. Because from the Table, a maximum of 100 degrees of freedom while we require 189 degrees of freedom, use these critical values to obtained $\chi_L^2 = 152.8222$ and $\chi_R^2 = 228.9638$. Based on the result, does the standard deviation appear to be different from the standard deviation of 696*g* for birth weights of babies born to mothers who did not use cocaine during pregnancy?

9. In the course of designing theather seats, the sitting heights (in *mm*) of a simple random sample of adults women is obtained, and the results are

849 807 821 856 864 877 772 848 802 807 887 815

Use the sample data to construct a 95% confidence interval estimate of σ, the standard deviation of sitting heights of all women. Does the confidence contain the value of 35 *mm*, which is believed to be the standard deviation of sitting heights of women?

10. One way to measure the risk of a stock is through the standard deviation rate of return of the stock. The following data represent the weekly rate of return (in percent) of Microsoft for 15 randomly selected weeks. Compute the 90% confidence interval for the risk of Microsoft stock.

5.34	9.63	–2.38	3.54	–8.76	2.12	–1.95	0.27
0.15	5.84	–3.90	–3.80	2.85	–1.61	–3.31	