

6.5: Confidence Intervals for Variances - Optional Material

Learning Objectives

- Develop a second general methodology of constructing confidence intervals
- Find critical values in the χ^2 -distribution
- Construct confidence intervals for variances and standard deviations using sample data

Confidence Intervals for Variances

At this point in our development of confidence intervals, we introduce another methodology. The general interpretation of confidence intervals remains the same: a confidence interval built at an 85% confidence level catches the population parameter for 85% of all samples of that given size, or alternatively, catches the population parameter for 85% of the confidence intervals constructed if random sampling is repeatedly conducted. We, however, will no longer base the construction methodology around the idea that 85% of all sample statistics fall within the margin of error of the population parameter. This worked really well when the sampling distributions were approximately normal and hence symmetric about the population parameter, but the sampling distributions for variances are not symmetric. Recall that the sampling distributions of sample variances were not normal, but skewed right, and that we could transform them into χ^2 -distributions to determine probabilities. We studied the sampling distribution of sample variances only when the parent population was normal; we shall remain in this realm of normal parent populations throughout this section.

Previously when we were constructing confidence intervals, we routinely produced a figure split into three regions with known areas within a specific distribution: the left tail, the right tail, and the central region. Each tail had an area that was equal to $\frac{\alpha}{2}$, and the central region had an area equal to CL. We shall begin our development with such a figure within the context of randomly drawing a sample of size n from a population normally distributed with the intent to build a confidence interval for the population variance at the confidence level CL. We sketch such regions in the following figure of the χ^2 -distribution with $n - 1$ degrees of freedom.

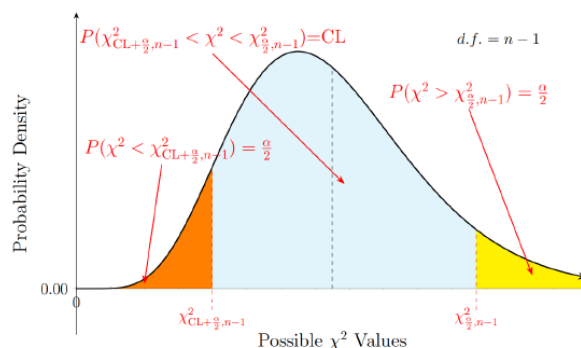


Figure 6.5.1: χ^2 -distribution

The boundary points of the central region are again called critical values; just like the critical values in the t -distribution, they depend on the confidence level and the degrees of freedom. There are, however, some significant differences. Notice that they are not the same distance away from the mean of the distribution (the black dashed line); the distribution is positively skewed, and we constructed the regions so that the tails each have an area of $\frac{\alpha}{2}$. Observe that both critical values are positive. To help us distinguish between the two critical values, we use the first part of the subscript to indicate the area to the left of the critical value. Note the smaller critical value has $CL + \frac{\alpha}{2}$ to the right of it; while, the larger critical value has $\frac{\alpha}{2}$. We label the critical values as seen in the figure. The smaller critical value is $\chi^2_{CL+\frac{\alpha}{2}, n-1}$, and the larger critical value is $\chi^2_{\frac{\alpha}{2}, n-1}$.

Now that we have an understanding of the figure and its labels, consider the probability statement at the top of the figure.

$$P\left(\chi^2_{CL+\frac{\alpha}{2}, n-1} < \chi^2 < \chi^2_{\frac{\alpha}{2}, n-1}\right) = CL$$

It says that the probability that the random variable χ^2 with $n - 1$ degrees of freedom falls between the two critical values is equal to the confidence level. In our context, this χ^2 variable with $n - 1$ degrees of freedom is related to the sampling distribution of

sample variances through the following transformation.

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

We can understand the probability statement about the random variable χ^2 in terms of the random variable s^2 , the sample variance.

$$P\left(\chi_{\text{CL}+\frac{\alpha}{2}, n-1}^2 < \frac{(n-1)}{\sigma^2} \cdot s^2 < \chi_{\frac{\alpha}{2}, n-1}^2\right) = \text{CL}$$

Recall that the underlying random experiment for the random variable s^2 is randomly selecting a sample of size n from a population that is normally distributed. We can think of this probability statement as follows: the probability of randomly selecting a sample of size n so that the sample variance scaled by $\frac{(n-1)}{\sigma^2}$ falls between the critical values is the confidence level. Recall that we are interested in constructing a confidence interval for σ^2 . We can use algebraic manipulation to get σ^2 isolated in the expression.

$$\begin{aligned} P\left(\chi_{\text{CL}+\frac{\alpha}{2}, n-1}^2 < \frac{(n-1)}{\sigma^2} \cdot s^2 < \chi_{\frac{\alpha}{2}, n-1}^2\right) &= \text{CL} \\ P\left(\frac{\chi_{\text{CL}+\frac{\alpha}{2}, n-1}^2}{(n-1)s^2} < \frac{1}{\sigma^2} < \frac{\chi_{\frac{\alpha}{2}, n-1}^2}{(n-1)s^2}\right) &= \text{CL} \\ P\left(\frac{(n-1)s^2}{\chi_{\text{CL}+\frac{\alpha}{2}, n-1}^2} > \sigma^2 > \frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}, n-1}^2}\right) &= \text{CL} \end{aligned}$$

This last step might require further explanation. We want an expression with σ^2 not with $\frac{1}{\sigma^2}$. Notice that these two expressions are reciprocals of each other. So, we reciprocate each term in the string of inequalities and figure out what happens with the inequalities. Consider the very simple string of inequalities $2 < 3 < 4$. The reciprocals of each of the terms in our inequality are $\frac{1}{2}$, $\frac{1}{3}$, and $\frac{1}{4}$. Notice that $\frac{1}{2} > \frac{1}{3} > \frac{1}{4}$. We are in a similar situation in our probability statements. We have a string of inequalities with positive values in each term; so, when we reciprocate, we must flip the inequality signs. We generally have lower bounds on the left and upper bounds on the right; so, we reorder this last line to arrive at a final probability statement.

$$P\left(\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}, n-1}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{\text{CL}+\frac{\alpha}{2}, n-1}^2}\right) = \text{CL}$$

We understand this last probability statement to say that the probability of randomly selecting a sample of size n from the normally distributed parent population so that $\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}, n-1}^2}$ is less than σ^2 and $\frac{(n-1)s^2}{\chi_{\text{CL}+\frac{\alpha}{2}, n-1}^2}$ is greater than σ^2 is the confidence level. In other words, we have constructed an interval so that the population variance falls within that interval the confidence level percent of the time. Thus, this construction of confidence intervals for variances yields confidence intervals of the following form.

$$\left(\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}, n-1}^2}, \frac{(n-1)s^2}{\chi_{\text{CL}+\frac{\alpha}{2}, n-1}^2}\right)$$

Constructing Confidence Intervals for Variances

We now have a method of constructing confidence intervals for variances that produces confidence intervals of a specific form; this form is quite different from the forms for means and proportions. The sample statistic is no longer the center of the interval. Critical values still play an important role in the construction of the interval, and computing these critical values is the last aspect of construction that we need to hammer out. In the section on sampling distributions of sample variances, we introduced the CHISQ.DIST accumulation function. In order to calculate critical values, we need an inverse accumulation function. We introduce CHISQ.INV which works similarly to the inverse accumulation functions that have been previously introduced. Given an area and the degrees of freedom of the distribution, CHISQ.INV returns the point such that that area is to the left of the point in that distribution.

$$a = \text{CHISQ.INV}(\text{area to the left of } a, d. f.) = \text{CHISQ.INV}(P(\chi^2 < a), d. f.)$$

? Text Exercise 6.5.1

Within the context of constructing confidence intervals for variances from populations that are normally distributed. Determine the two critical values for the given confidence level and sample size by roughly sketching the χ^2 -distribution and using technology.

1. CL = 0.9 and $n = 8$

Answer

We begin by sketching a χ^2 -distribution with 7 degrees of freedom. It is a skewed right distribution starting at 0, 0 with its peak at 5 (in general, at $(n - 3)$). We then draw our three regions and label them with the appropriate areas.

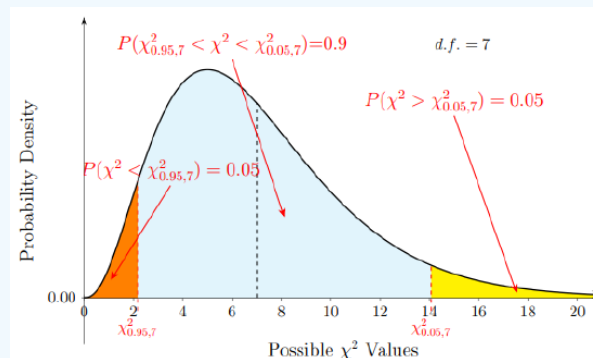


Figure 6.5.2 χ^2 -distribution

$$\chi^2_{0.95,7} = \text{CHISQ.INV}(0.95, 7) \approx 2.6174$$

$$\chi^2_{0.05,7} = \text{CHISQ.INV}(0.05, 7) \approx 14.0671$$

2. CL = 0.95 and $n = 24$

Answer

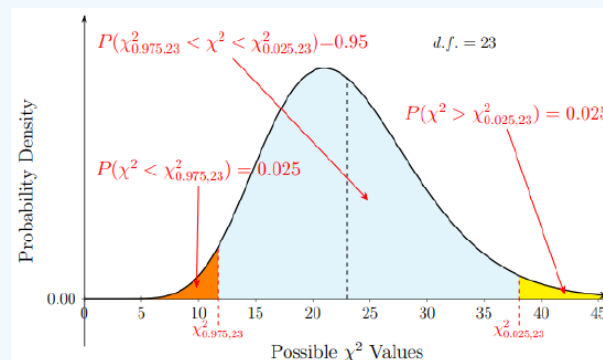


Figure 6.5.3 χ^2 -distribution

$$\chi^2_{0.975,23} = \text{CHISQ.INV}(0.975, 23) \approx 11.6886$$

$$\chi^2_{0.025,23} = \text{CHISQ.INV}(0.025, 23) \approx 38.0756$$

3. CL = 0.99 and $n = 17$

Answer

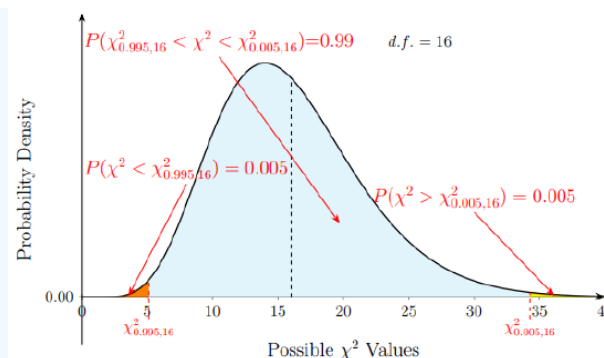


Figure 6.5.4 χ^2 -distribution

$$\chi^2_{0.995,16} = \text{CHISQ.INV}(0.995, 16) \approx 5.1422$$

$$\chi^2_{0.005,16} = \text{CHISQ.INV}(0.005, 16) \approx 34.2672$$

? Text Exercise 6.5.2

Let us return to the topic of the heights of adult males. In the previous section, we used a random sample of 10 adult males to construct a confidence interval for the population mean of adult males. Use the same sample data, provided again below, to construct a 90% confidence interval for the population variance.

64, 66, 67.5, 68, 69, 69.75, 71.25, 72, 73, 73.25

Answer

In order to construct a confidence interval for variances using the methods developed above, we need the parent population to be normally distributed and to use a randomly selected sample. The heights of adult males are normally distributed so we may construct our confidence interval.

To construct the confidence interval, we need the sample standard variance, sample size, and the two critical values. The sample variance of our particular sample is approximately 9.3924 square inches. We sampled 10 adult males. To calculate the critical values, we must find α . Since $CL = 0.90$, $\alpha = 0.1$. Our critical values come from the χ^2 -distribution with 9 degrees of freedom. We sketch the distribution to help us compute the critical values.

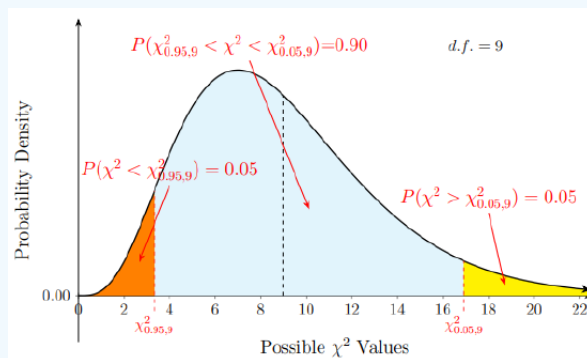


Figure 6.5.5 χ^2 -distribution

$$\chi^2_{0.95,9} = \text{CHISQ.INV}(0.95, 9) \approx 3.3251$$

$$\chi^2_{0.05,9} = \text{CHISQ.INV}(0.05, 9) \approx 16.919$$

$$\left(\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}, n-1}}, \frac{(n-1)s^2}{\chi^2_{CL+\frac{\alpha}{2}, n-1}} \right) \approx \left(\frac{9 \cdot 9.3924}{16.919}, \frac{9 \cdot 9.3924}{3.3251} \right) \approx (4.9962, 25.4221)$$

At the 90% confidence level, the population variance σ^2 of adult male heights is between 4.9962 square inches and 25.4221 square inches.

Constructing Confidence Intervals for Standard Deviations

We are often interested in the standard deviation of a population. Most of the theoretical work for sampling distributions and confidence intervals occurs within the realm of variance because the sample variance is an unbiased estimator of the population variance; while, the sample standard deviation is not an unbiased estimator of population standard deviation. But, we can use a confidence interval for variances to speak about standard deviations rather simply because the standard deviation is the square root of the variance, and \sqrt{x} is an increasing function meaning that it preserves order. Let us return to the context of our last text exercise about adult male height. We are 90% confident that the population variance is between 4.9962 and 25.4221 square inches and can, therefore, be 90% confident that the population standard deviation is between $\sqrt{4.9962} \approx 2.2352$ and $\sqrt{25.4221} \approx 5.0420$ inches. This range of values is right where we might expect the population standard deviation of adult male heights given that the population standard deviation of adult female heights is about 2.5 inches.

The section concludes with a general formulation. If a parent population is normally distributed, we may, by randomly selecting a sample of size n from the population and setting a confidence level CL, construct a confidence interval for standard deviations of the form

$$\left(\sqrt{\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2}, n-1}^2}}, \sqrt{\frac{(n-1)s^2}{\chi_{CL+\frac{\alpha}{2}, n-1}^2}} \right)$$

where the critical values come from the χ^2 -distribution with $n-1$ degrees of freedom.

6.5: Confidence Intervals for Variances - Optional Material is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by The Math Department at Fort Hays State University.

- **3.3: Measures of Central Tendency** by [David Lane](#) is licensed [Public Domain](#). Original source: <https://onlinestatbook.com>.