7 Statistical Intervals Based on a Single Sample

7.4 Confidence Intervals for the Variance and Standard Deviation of a Normal Population

Although inferences concerning a population variance σ^2 or standard deviation σ are usually of less interest than those about a mean or proportion, there are occasions when such procedures are needed.

In the case of a normal population distribution, inferences are based on the following result concerning the sample variance S^2.

Theorem

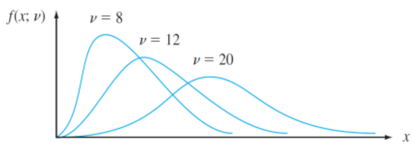
Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with parameters μ and σ^2. Then the rv

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum(X_i - \overline{X})^2}{\sigma^2}$$

has a chi-squared (χ^2) probability distribution with $n - 1$ df.

We know that the chi-squared distribution is a continuous probability distribution with a single parameter v, called the number of degrees of freedom, with possible values 1, 2, 3, ...

The graphs of several χ^2 probability density functions (pdf's) are illustrated in Figure 7.10.



Graphs of chi-squared density functions

Figure 7.10

Each pdf $f(x; v)$ is positive only for $x > 0$, and each has a positive skew (long upper tail), though the distribution moves rightward and becomes more symmetric as v increases.

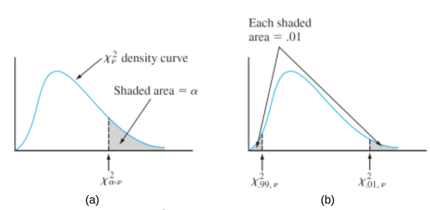
To specify inferential procedures that use the chi-squared distribution, we need notation analogous to that for a t critical value $t_{\alpha,v}$.

Notation

Let $\chi^2_{\alpha,\nu}$, called a **chi-squared critical value,** denote the number on the horizontal axis such that α of the area under the chi-squared curve with v df lies to the right of $\chi^2_{\alpha,\nu}$.

Symmetry of t distributions made it necessary to tabulate only upper-tailed t critical values ($t_{\alpha,v}$ for small values of α).

The chi-squared distribution is not symmetric, so Appendix Table A.7 contains values of $\chi^2_{\alpha,\nu}$ both for α near 0 and near 1, as illustrated in Figure 7.11(b).



$\chi^2_{\alpha,\nu}$ notation illustrated

Figure 7.11

Confidence Intervals for the Variance and Standard Deviation of a Normal Population

For example, $\chi^2_{.025,14}$ = 26.119, and $\chi^2_{.95,20}$ (the 5th percentile) = 10.851.

The rv $(n-1)S^2/\sigma^2$ satisfies the two properties on which the general method for obtaining a CI is based: It is a function of the parameter of interest σ^2, yet its probability distribution (chi-squared) does not depend on this parameter.

The area under a chi-squared curve with v df to the right of $\chi^2_{\alpha/2,v}$ is $\alpha/2$, as is the area to the left of $\chi^2_{1-\alpha/2,v}$.

Thus the area captured between these two critical values is $1-\alpha$.

Confidence Intervals for the Variance and Standard Deviation of a Normal Population

As a consequence of this and the theorem just stated,

$$P\left(\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2,n-1}\right) = 1-\alpha \qquad \textbf{(7.17)}$$

The inequalities in (7.17) are equivalent to

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}$$

Substituting the computed value s^2 into the limits gives a CI for σ^2, and taking square roots gives an interval for σ.

Confidence Intervals for the Variance and Standard Deviation of a Normal Population

A **100(1 – α)% confidence interval for the variance σ^2 of a normal population** has lower limit

$$(n-1)s^2/\chi^2_{\alpha/2,n-1}$$

and upper limit

$$(n-1)s^2/\chi^2_{1-\alpha/2,n-1}$$

A **confidence interval for** σ has lower and upper limits that are the square roots of the corresponding limits in the interval for σ^2. An upper or a lower confidence bound results from replacing $\alpha/2$ with α in the corresponding limit of the CI.

Example 15

The accompanying data on breakdown voltage of electrically stressed circuits was read from a normal probability plot that appeared in the article "Damage of Flexible Printed Wiring Boards Associated with Lightning-Induced Voltage Surges" (*IEEE Transactions on Components, Hybrids*, *and Manuf*. *Tech*., 1985: 214–220).

The straightness of the plot gave strong support to the assumption that breakdown voltage is approximately normally distributed.

1470	1510	1690	1740	1900	2000	2030	2100	2190
2200	2290	2380	2390	2480	2500	2580	2700	

Example 15

cont'd

Let σ^2 denote the variance of the breakdown voltage distribution. The computed value of the sample variance is s^2 = 137, 324.3, the point estimate of σ^2.

With df = n – 1 = 16, a 95% CI requires $\chi^2_{.975,16}$ = 6.908 and $\chi^2_{.025,16}$ = 28.845.

The interval is

$$\left(\frac{16(137{,}324.3)}{28.845}, \frac{16(137{,}324.3)}{6.908}\right) = (76{,}172.3,\ 318{,}064.4)$$

13

Example 15

cont'd

Taking the square root of each endpoint yields (276.0, 564.0) as the 95% CI for σ.

These intervals are quite wide, reflecting substantial variability in breakdown voltage in combination with a small sample size.

14