8.5: Chi Squared Distribution

The χ^2 (chi squared) distribution is a consequence of a random process based on the normal distribution. It is derived from the normal distribution as the result of the following stochastic process :

1. Suppose you have a population that has variance σ^2 and is normally distributed.
2. Take a sample of size n from the population and compute $x_1 = \frac{(n-1)s_1^2}{\sigma^2}$ using the sample standard deviation s_1 from that sample.
3. Put the sample back into the population.
4. Take another sample of size n from the population and compute $x_2 = \frac{(n-1)s_2^2}{\sigma}$ using the sample standard deviation s_2 from that sample.
5. etc.
6. The distribution of the values of $x_i = \frac{(n-1)s_i^2}{\sigma^2}$ values will be a χ^2 distribution with $\nu = n-1$ degrees of freedom.

Like the t-distributions, the χ^2 distributions are a family, see Figure 8.10.

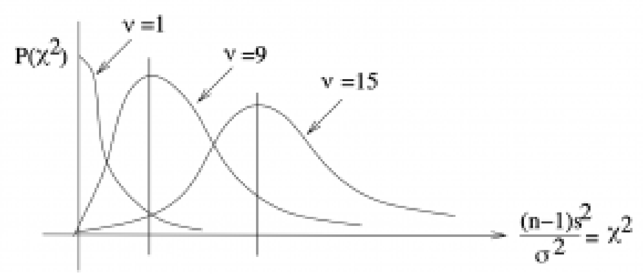

Figure 8.10 : The χ^2 distributions are enumerated by degrees of freedom.

The χ^2 distribution underlies why s is the best estimate for σ. It mean, or expected value is $\nu = n-1$ so the expected value of s is σ. The expected value of $\sum(x-\bar{x})/n$ in a random sample of size n is not σ.

Confidence Intervals on σ and σ^2

The χ^2 distribution is already normalized in its definition through including s in its definition. Therefore no z-transforms are needed and we can work directly with a table that gives right tail areas under the χ^2 distribution. That table is the **Chi-squared Distribution Table**, in the Appendix, and it gives values of χ^2 for given values of area to the right of χ^2, see Figure 8.11.

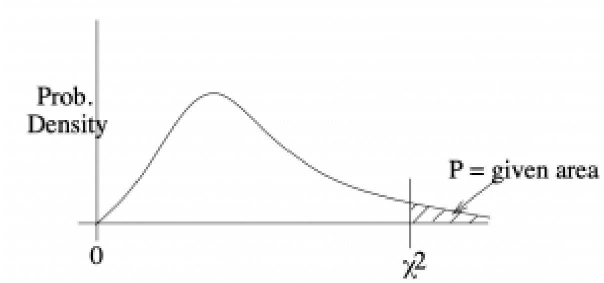

Figure 8.11 : The Chi-squared Distribution Table gives χ^2 associated with given right tail areas.

We'll need χ^2_{left} and χ^2_{right} such that the tail areas are equal and such that the area between them is C, see Figure 8.12.

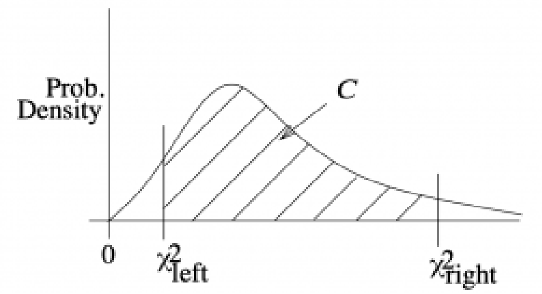

Figure 8.12 : The values χ^2_{left} and χ^2_{right} define the confidence region $\mathcal{C}$.

Notation : Let's call the α in the **Chi-squared Distribution Table** α_T and let $\chi^2(\alpha_T)$ be the table value that corresponds to α_T. In other words $\chi^2(\alpha_T)$ is the χ^2 value that corresponds to a right tail area of α_T.

So given $\mathcal{C}$, the appropriate χ^2_{left} and χ^2_{right} are the following values from the **Chi-squared Distribution Table**:

$$\chi^2_{\text{right}} = \chi^2\left(\frac{1-\mathcal{C}}{2}\right) \tag{8.5.1}$$

$$\chi^2_{\text{left}} = \chi^2\left(1-\left[\frac{1-\mathcal{C}}{2}\right]\right). \tag{8.5.2}$$

Note the symmetry of the **Chi-squared Distribution Table**. If χ^2_{right} comes from the column 3 columns from the right edge of the table then χ^2_{left} comes from a column 3 columns from the left edge of the table. Only small and large areas appear in the table, there are no intermediate values.

Finally, the confidence interval for σ^2 is given by

$$\frac{(n-1)s^2}{\chi^2_{\text{right}}} < \sigma^2 < \frac{(n-1)s^2}{\chi^{-2}_{\text{left}}} \tag{8.5.3}$$

and for σ by:

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{\text{right}}}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_{\text{left}}}} \tag{8.5.4}$$

Where the χ^2 distribution with $\nu = n-1$ degrees of freedom (giving the line to use in the **Chi-squared Distribution Table**) is used.

Example 8.5 : Find the 90% confidence interval on σ and σ^2 for the following data

$$59, 54, 53, 52, 51, 39, 49, 46, 49, 48 \tag{8.5.5}$$

Solution : Compute, using your calculator :

$$s^2 = 28.2 \tag{8.5.6}$$

$$\nu = n-1 = 9. \tag{8.5.7}$$

From the **Chi-squared Distribution Table**, in the $\nu = 9$ line, find :

$$\chi^2_{\text{right}} = \chi^2\left(\frac{1-0.90}{2}\right) = \chi^2(0.05) = 16.919 \tag{8.5.8}$$

and

$$\chi^2_{\text{left}} = \chi^2(1-0.05) = \chi^2(0.95) = 3.325 \tag{8.5.9}$$

So

$$\begin{aligned}\frac{(n-1)s^2}{\chi^2_{\text{right}}} &< \sigma^2 < \frac{(n-1)s^2}{\chi^2_{\text{left}}} \\ \frac{9\cdot 28.2}{16.919} &< \sigma^2 < \frac{9\cdot 28.2}{3.325} \\ 15.0 &< \sigma^2 < 76.3 \qquad\qquad \text{with } 90\backslash\% \text{ confidence.}\end{aligned}$$

Taking square roots:

$$3.87 < \sigma < 8.73 \qquad\qquad \text{with } 90\backslash\% \text{ confidence.} \tag{8.5.10}$$

□
