Confidence intervals for normal parameters

18.650, Sept. 11, 2015

1 Introduction

This is a continuation of the handout "Normal distributions and sample statistics," to be called briefly "Normal samples" (normalsamples.pdf on the course website).

Suppose we observe $X_1, ..., X_n$ assumed to be i.i.d. $N(\mu, \sigma^2)$ for some unknown real μ and some $\sigma > 0$. We can estimate μ by the sample mean $\overline{X}$ and σ^2 by the sample variance s_X^2. Somewhat similarly as experimental scientists give "error bars" for their estimates, what will be done next is to give "confidence intervals" for μ and for σ.

2 Confidence intervals for σ and σ^2

Although in practice, μ may be more interesting than σ, it's simpler first to give confidence intervals for σ, as follows. Consider Theorem 3(b) of "Normal samples." Two paragraphs above the theorem, $\chi^2(d)$ distributions with d "degrees of freedom," abbreviated "df," were defined, and Theorem 3(b) says that

$$\frac{(n-1)s_X^2}{\sigma^2} = \frac{\sum_{j=1}^n (X_j - \overline{X})^2}{\sigma^2} \tag{1}$$

has a $\chi^2(n-1)$ distribution. Rice, Appendix B, p. A8, gives a "table of percentiles" of the χ^2 distribution. For $d = df = 1, 2, 3, ..., 16, 18, 24, 30, 40, 60, 120$ and $q = 0.005, 0.01, 0.025$, 0.05, 0.1, and $1 - q$ in place of q for these values of q, the table gives numbers $\chi_q^2(d)$ such that

$$\Pr\left(\chi^2(d) \leq \chi_q^2(d)\right) = q. \tag{2}$$

For our given observations $X = (X_1, ..., X_n)$, their sample variance s_X^2, and some small α such as $\alpha = 0.05$ or 0.01, we'd like to find an interval $[a, b] = [a(X), b(X)]$ with $0 < a < b < +\infty$ such that $\Pr(a \leq \sigma^2 \leq b) = 1 - \alpha$. For such a "two-sided" interval, it's usual to do this by making $\Pr(\sigma^2 < a) = \Pr(\sigma^2 > b) = \alpha/2$.

From (1) we have for $q = \alpha/2$ first by (2), then by multiplying by σ^2 and dividing by $\chi_q^2(n-1)$,

$$q = \Pr\left(\frac{(n-1)s_X^2}{\sigma^2} < \chi_q^2(n-1)\right) = \Pr\left(\sigma^2 > \frac{(n-1)s_X^2}{\chi_q^2(n-1)}\right), \tag{3}$$

and likewise

$$q = \Pr\left(\frac{(n-1)s_X^2}{\sigma^2} > \chi_{1-q}^2(n-1)\right) = \Pr\left(\sigma^2 < \frac{(n-1)s_X^2}{\chi_{1-q}^2(n-1)}\right). \tag{4}$$

Combining (3) and (4) gives that

$$\Pr\Big(a(X) \leq \sigma^2 \leq b(X)\Big) = 1 - 2q = 1 - \alpha \tag{5}$$

for $a(X) = (n-1)s_X^2/\chi_{1-q}^2(n-1)$ by (4) and $b(X) = (n-1)s_X^2/\chi_q^2(n-1)$ by (3). Then, $[a(X), b(X)]$ is said to be a $100(1-\alpha)\%$ confidence interval for σ^2, specifically a 95% confidence interval if $\alpha = 0.05$, so $q = 0.025$ and $1 - q = 0.975$, or a 99% confidence interval if $\alpha = 0.01$, so $q = 0.005$ and $1 - q = 0.995$ (respectively the smallest and largest values of q or $1 - q$ given in Rice's table).

To get corresponding confidence intervals for σ, we can just take square roots and use $[\sqrt{a(X)}, \sqrt{b(X)}]$. Note that these confidence intervals are not of "error bar" form: s_X^2 is not the midpoint of the confidence intervals for σ^2, nor is s_X the midpoint of those for σ.

3 Confidence intervals for the mean μ, and t distributions

To get a confidence interval for μ is a bit more complicated. We know from Theorem 2 of "Normal samples" that the sample mean $\overline{X}$ has a $N(\mu, \sigma^2/n)$

distribution. Thus $\sqrt{n}(\overline{X}-\mu)/\sigma$ has a $N(0,1)$ distribution. But σ is unknown. Suppose we substitute for it the sample standard deviation s_X, then what distribution does

$$\frac{\sqrt{n}(\overline{X}-\mu)}{s_X} = \frac{\sqrt{n}(\overline{X}-\mu)/\sigma}{s_X/\sigma} \tag{6}$$

have? The numerator always has a $N(0,1)$ distribution, and it is independent of the denominator by Theorem 3(a) of "Normal samples." By Theorem 3(b), the denominator has the distribution of the square root of a $\chi^2(n-1)$ variable, divided by its number $n-1$ of degrees of freedom. So the distribution of the quotient is of the following form, for $d = n-1$:

Definition. A random variable T has a t distribution with d degrees of freedom, or a $t(d)$ distribution, if it has the distribution of $Z/\sqrt{U/d}$ where Z has $N(0,1)$ distribution, U has a $\chi^2(d)$ distribution, and Z and U are independent.

The distribution of such a T is symmetric around 0, in other words T and $-T$ have the same distribution, because $-Z$ also has $N(0,1)$. Rice, Appendix B, p. A9, Table 4, is a table of the $t(d)$ distribution for $d = 1, 2, ..., 29, 30, 40, 60, 120, \infty$, giving $t_q(d)$ such that $\Pr(t(d) \le t_q(d)) = q$ or equivalently $\Pr(t(d) \ge t_q(d)) = 1-q$ and so $\Pr(|t(d)| \ge t_q(d)) = 2(1-q) = \alpha$ for $\alpha = 0.01, 0.02, 0.05, 0.1$, or some larger values. As $d \to +\infty$, $U/d \to 1$ by the law of large numbers, so the $t(d)$ distribution converges to the $N(0,1)$ distribution. To get a $100(1-\alpha)\%$ confidence interval for μ with endpoints $c(X,\alpha)$ and $d(X,\alpha)$, we have from (6) and the definition of t distribution that

$$1-\alpha = \Pr\left(\frac{\sqrt{n}|\overline{X}-\mu|}{s_X} \le t_{1-\alpha/2}(n-1)\right) = \Pr\left(|\overline{X}-\mu| \le \frac{t_{1-\alpha/2}(n-1)s_X}{\sqrt{n}}\right),$$

so the endpoints of the confidence interval for μ are

$$\overline{X} \pm \frac{t_{1-\alpha/2}(n-1)s_X}{\sqrt{n}}.$$

For μ, $\overline{X}$ is at the midpoint of the confidence intervals.