Using pivots to construct confidence intervals

In Example 41 we used the fact that

$$Q(\bar{X}, \mu) = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \quad \text{for all } \mu.$$

We then said $|Q(\bar{X}, \mu)| \le z_{\alpha/2}$ with probability $1 - \alpha$, and converted this into a statement about μ.

Definition 21 Given a data vector $\boldsymbol{X}$, a random variable $Q(\boldsymbol{X}, \theta)$ is a *pivotal quantity* if the distribution of $Q(\boldsymbol{X}, \theta)$ is independent of all unknown parameters.

Suppose we want a $(1 - \alpha)100\%$ confidence interval for θ. Try to find a function of the data that also depends on θ but whose probability distribution *does not* depend on θ.

Ideally, the random variable $Q(\boldsymbol{X}, \theta)$ will depend on $\boldsymbol{X}$ only through a sufficient statistic.

Let α_1 and α_2 be positive and such that $\alpha_1 + \alpha_2 = \alpha$. Let q_1 and q_2 be the α_1 and $1 - \alpha_2$ quantiles of the distribution of $Q(\boldsymbol{X}, \theta)$. Then

$$P(q_1 < Q(\boldsymbol{X}, \theta) \leq q_2) = 1 - \alpha.$$

Now, try to find $C(\boldsymbol{X})$ such that

$$\theta \in C(\boldsymbol{X}) \quad \text{iff} \quad q_1 < Q(\boldsymbol{X}, \theta) \leq q_2.$$

It then follows that $C(\boldsymbol{X})$ is a $(1 - \alpha)100\%$ confidence set for θ.

Example 42 Let $X_1, \ldots, X_n$ be a random sample from the $U(0, \theta)$ distribution. Would like to construct a $(1-\alpha)100\%$ confidence interval for θ.

What kind of transformation of the data would make the distribution of the transformed data free of θ?

$$\frac{X_1}{\theta}, \ldots, \frac{X_n}{\theta} \quad \text{are i.i.d. } U(0,1).$$

It follows that $Q(\boldsymbol{X}, \theta) = \max(X_1/\theta, \ldots, X_n/\theta)$ has the same distribution as the maximum of a random sample from the $U(0,1)$ distribution. Therefore, $Q(\boldsymbol{X}, \theta)$ is a pivotal quantity. Note that

$$Q(\boldsymbol{X}, \theta) = \frac{X_{(n)}}{\theta},$$

and thus depends on the data through a sufficient statistic.

Choose a and b so that

$$P\left(a < \frac{X_{(n)}}{\theta} < b\right) = 1 - \alpha.$$

$$P\left(a < \frac{X_{(n)}}{\theta} < b\right) = \int_a^b nt^{n-1}\,dt = b^n - a^n.$$

Since

$$P\left(a < \frac{X_{(n)}}{\theta} < b\right) = P\left(\frac{X_{(n)}}{b} < \theta < \frac{X_{(n)}}{a}\right),$$

$[X_{(n)}/b, X_{(n)}/a]$ is a $(1-\alpha)100\%$ confidence interval for θ so long as $0 < a, b < 1$ and $b^n - a^n = 1 - \alpha$.

It would be desirable to find an interval of the above form that has the shortest possible length. The length is

$$X_{(n)}\left(\frac{1}{a} - \frac{1}{b}\right).$$

We have no control over $X_{(n)}$, but we could choose a and b to minimize $a^{-1} - b^{-1}$ subject to the constraint $b^n - a^n = 1 - \alpha$.

It is straightforward to show that the solution to the previous problem is

$$b = 1 \quad \text{and} \quad a = \alpha^{1/n}.$$

Therefore, among $(1-\alpha)100\%$ confidence intervals of the form $[X_{(n)}/b, X_{(n)}/a]$, the shortest one is $[X_{(n)}, X_{(n)}/\alpha^{1/n}]$.

Example 43 Let $X_1, \ldots, X_n$ be a random sample from a density of the form

$$\frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right),$$

where f is *known*. The parameter space is

$$\Theta = \{(\mu, \sigma) : -\infty < \mu < \infty, \sigma > 0\}.$$

The family of densities

$$\left\{\frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) : (\mu, \sigma) \in \Theta\right\}$$

is called a *location-scale family*.

Define

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

and

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

The random variable $(\bar{X} - \mu)/S$ is a pivotal quantity, a fact which leads to a confidence interval for μ.

Proof: Consider

$$\begin{aligned} X_i &= (X_i - \mu) + \mu \\ &= \sigma(X_i - \mu)/\sigma + \mu \\ &= \sigma Z_i + \mu \end{aligned}$$

We then have $\bar{X} - \mu = \sigma\bar{Z}$, where

$$\bar{Z} = n^{-1}\sum_{i=1}^{n} Z_i.$$

$$
\begin{aligned}
S^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(\sigma Z_i - \sigma \bar{Z})^2 \\
&= \sigma^2 S_Z^2.
\end{aligned}
$$

It follows that

$$\frac{(\bar{X}-\mu)}{S} = \frac{\bar{Z}}{S_Z}.$$

It is easy to verify that Z_i has density f, which is free of any unknown parameters, and the result follows.

The $N(\mu, \sigma^2)$ family of distributions is an example of a location-scale family, with

$$f(y) = \frac{1}{\sqrt{2\pi}} e^{-y^2/2}.$$

In this case

$$T = \frac{(\bar{X}-\mu)}{S/\sqrt{n}}$$

is a pivotal quantity having the Student's t-distribution with $n-1$ degrees of freedom.

Let $t_{p,n-1}$ be the $(1-p)$ quantile of Student's t-distribution with $n-1$ degrees of freedom. A $(1-\alpha)100\%$ confidence interval for μ is

$$\left[\bar{X} - t_{\alpha_2,n-1}\frac{S}{\sqrt{n}}, \bar{X} - t_{1-\alpha_1,n-1}\frac{S}{\sqrt{n}}\right],$$

where $\alpha_1 + \alpha_2 = \alpha$.

For any location-scale or scale family, S^2/σ^2 is a pivotal quantity. This fact leads to confidence intervals for σ^2 (or σ).

When $X_1, \ldots, X_n$ are i.i.d. $N(\mu, \sigma^2)$,

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Let a and b be such that $P(a < \chi^2_{n-1} < b) = 1-\alpha$. Then

$$\left[\frac{(n-1)S^2}{b}, \frac{(n-1)S^2}{a}\right]$$

is a $(1-\alpha)100\%$ confidence interval for σ^2.

Shortest length confidence intervals

Subject to the confidence coefficient being $1-\alpha$, we would like our confidence interval to be of shortest length.

Example 43 (continued) $X_1, \ldots, X_n$ i.i.d. $N(\mu, \sigma^2)$. The length of the previously derived confidence interval for μ is

$$\frac{S}{\sqrt{n}}(-t_{1-\alpha_1,n-1} + t_{\alpha_2,n-1}).$$

Subject to the constraint $\alpha_1 + \alpha_2 = \alpha$, choose α_1 and α_2 to minimize

$$t_{\alpha_2,n-1} - t_{1-\alpha_1,n-1}.$$

Theorem 12 Let f be a unimodal pdf. If the interval $[a, b]$ satisfies

(i) $\int_a^b f(x)\,dx = 1 - \alpha$

(ii) $f(a) = f(b) > 0$

(iii) $a \le x^* \le b$ where x^* is the mode of f,

then $[a, b]$ is the shortest of all intervals that satisfy (i).

Proof: See Casella and Berger, p. 442.

Continuing Example 43, since the t-distribution is unimodal, we may apply Theorem 13. Now, the t-distribution is symmetric, so $t_{\alpha_2,n-1} - t_{1-\alpha_1,n-1}$ is minimized subject to $\alpha_1 + \alpha_2 = \alpha$ when $\alpha_1 = \alpha_2 = \alpha/2$.

The shortest interval has the form

$$\left[\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right].$$

In some situations we can minimize length directly.

Example 44 In Example 42 we had $X_1, \ldots, X_n$ i.i.d. $U(0, \theta)$, and found that $[X_{(n)}/b, X_{(n)}/a]$ is a $(1-\alpha)100\%$ confidence interval for θ for each (a, b) such that $b^n - a^n = 1 - \alpha$. Here, we can't apply Theorem 12, but we could minimize the length directly.
