MA204: Mathematical Statistics

Tutorial 7

T7.1 Confidence Interval (CI)

Let $X_1, \ldots, X_n$ be a random sample from a population with pdf $f(x;\theta)$. Define $\mathbf{x} = (X_1, \ldots, X_n)^\top$. Let $T_1(\mathbf{x})$ and $T_2(\mathbf{x})$ be two statistics such that $T_1 \leqslant T_2$ and

$$\Pr(T_1 \leqslant \theta \leqslant T_2) = 1 - \alpha.$$

Then the random interval $[T_1, T_2]$ is called a $100(1-\alpha)\%$ *confidence interval* (CI) for θ.

T7.2 Pivotal Quantity

Assume that $X_1, \ldots, X_n \sim f(x;\theta)$ and $T = T(\mathbf{x})$ is a sufficient statistic of θ. Let $P = P(T, \theta)$ be a function of T and θ. If the distribution of P does not depend on θ, then P is said to be a *pivotal quantity*.

Example T7.1 (A normal distribution with known variance). Let $X_1, \ldots, X_n \sim N(\mu, 3.3^2)$ with $n = 30$ and $\bar{x} = 27$. Construct a 90% CI for μ.

Solution: Since $\bar{X}$ is a sufficient statistic of μ, we have a pivotal quantity

$$\begin{aligned}
& P = \frac{\sqrt{n}(\bar{X} - \mu)}{3.3} \sim N(0,1), \\
\Rightarrow \quad & \Pr\left\{-z_{\alpha/2} \leqslant \frac{\sqrt{n}(\bar{X} - \mu)}{3.3} \leqslant z_{\alpha/2}\right\} = 1 - \alpha, \\
\Rightarrow \quad & \Pr\left(\bar{X} - z_{\alpha/2}\frac{3.3}{\sqrt{n}} \leqslant \mu \leqslant \bar{X} + z_{\alpha/2}\frac{3.3}{\sqrt{n}}\right) = 1 - \alpha,
\end{aligned}$$

$$\Rightarrow \quad \Pr\left(\bar{X} - z_{0.05}\frac{3.3}{\sqrt{n}} \leqslant \mu \leqslant \bar{X} + z_{0.05}\frac{3.3}{\sqrt{n}}\right) = 0.9.$$

Therefore, a 90% CI for μ is given by

$$\left[27 - 1.645\frac{3.3}{\sqrt{30}},\ 27 + 1.645\frac{3.3}{\sqrt{30}}\right] = [26.0089,\ 27.9911]\,. \qquad \|$$

Example T7.2 (A normal distribution with unknown variance). Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ with $n = 30$, $\bar{x} = 105$ and $s = 11$.

(a) Construct a 95% CI for μ.

(b) Construct a 95% CI for σ.

(c) When the sample size n is large, the sample variance S^2 can be approximated by a normal distribution with mean σ^2 and variance $2\sigma^4/(n-1)$. Based on this result, show that an approximate 95% CI for σ^2 is given by

$$\left[\frac{S^2}{1 + 1.96\sqrt{2/(n-1)}},\ \frac{S^2}{1 - 1.96\sqrt{2/(n-1)}}\right].$$

Hence, compute an approximate 95% CI for σ.

Solution: (a) Since $\bar{X}$ and S^2 are jointly sufficient for (μ, σ^2), we have a pivotal quantity

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1),$$

$$\Rightarrow \quad \Pr\left\{-t(\alpha/2, n-1) \leqslant \frac{\sqrt{n}(\bar{X} - \mu)}{S} \leqslant t(\alpha/2, n-1)\right\} = 1 - \alpha,$$

$$\Rightarrow \quad \Pr\left\{\bar{X} - t(\alpha/2, n-1)\frac{S}{\sqrt{n}} \leqslant \mu \leqslant \bar{X} + t(\alpha/2, n-1)\frac{S}{\sqrt{n}}\right\} = 1 - \alpha,$$

$$\Rightarrow \quad \Pr\left\{\bar{X} - t(0.025, n-1)\frac{S}{\sqrt{n}} \leqslant \mu \leqslant \bar{X} + t(0.025, n-1)\frac{S}{\sqrt{n}}\right\} = 0.95.$$

Therefore, a 95% CI for μ is given by

$$\begin{aligned} & \left[\bar{x} - t(0.025, 29)\frac{s}{\sqrt{n}},\ \bar{x} + t(0.025, 29)\frac{s}{\sqrt{n}}\right] \\ = & \left[105 - 2.045\frac{11}{\sqrt{30}},\ 105 + 2.045\frac{11}{\sqrt{30}}\right] = [100.8930,\ 109.1070]\,. \end{aligned}$$

(b) Since

$$\begin{aligned} & \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1), \\ \Rightarrow\ & \Pr\left\{\chi^2(1-\alpha/2, n-1) \leqslant \frac{(n-1)S^2}{\sigma^2} \leqslant \chi^2(\alpha/2, n-1)\right\} = 1-\alpha, \\ \Rightarrow\ & \Pr\left\{\frac{(n-1)S^2}{\chi^2(\alpha/2, n-1)} \leqslant \sigma^2 \leqslant \frac{(n-1)S^2}{\chi^2(1-\alpha/2, n-1)}\right\} = 1-\alpha, \\ \Rightarrow\ & \Pr\left\{\sqrt{\frac{(n-1)S^2}{\chi^2(0.025, n-1)}} \leqslant \sigma \leqslant \sqrt{\frac{(n-1)S^2}{\chi^2(0.975, n-1)}}\right\} = 0.95. \end{aligned}$$

Therefore, a 95% CI for σ is given by

$$\left[\sqrt{\frac{(n-1)s^2}{\chi^2(0.025, 29)}},\ \sqrt{\frac{(n-1)s^2}{\chi^2(0.975, 29)}}\right] = \left[\sqrt{\frac{29 \times 11^2}{45.722}},\ \sqrt{\frac{29 \times 11^2}{16.047}}\right] = [8.7605,\ 14.7875]\,.$$

(c) Based on the result

$$S^2 \mathrel{\dot\sim} N\left(\sigma^2, \frac{2\sigma^4}{n-1}\right) \quad \Rightarrow \quad \frac{S^2 - \sigma^2}{\sqrt{2\sigma^4/(n-1)}} = \sqrt{\frac{n-1}{2}}\left(\frac{S^2}{\sigma^2} - 1\right) \mathrel{\dot\sim} N(0,1),$$

we have

$$\begin{aligned} & \Pr\left\{-z_{\alpha/2} \leqslant \sqrt{\frac{n-1}{2}}\left(\frac{S^2}{\sigma^2} - 1\right) \leqslant z_{\alpha/2}\right\} = 1-\alpha, \\ \Rightarrow\ & \Pr\left\{\frac{S^2}{1 + z_{\alpha/2}\sqrt{2/(n-1)}} \leqslant \sigma^2 \leqslant \frac{S^2}{1 - z_{\alpha/2}\sqrt{2/(n-1)}}\right\} = 1-\alpha, \\ \Rightarrow\ & \Pr\left\{\frac{S^2}{1 + z_{0.025}\sqrt{2/(n-1)}} \leqslant \sigma^2 \leqslant \frac{S^2}{1 - z_{0.025}\sqrt{2/(n-1)}}\right\} = 0.95. \end{aligned}$$

Therefore, an approximate 95% CI for σ^2 is given by

$$\left[\frac{S^2}{1+1.96\sqrt{2/(n-1)}},\ \frac{S^2}{1-1.96\sqrt{2/(n-1)}}\right].$$

Hence, an approximate 95% CI for σ is given by

$$\left[\sqrt{\frac{11^2}{1+1.96\sqrt{2/(30-1)}}},\ \sqrt{\frac{11^2}{1-1.96\sqrt{2/(30-1)}}}\right] = [8.9377, 15.7905]\,. \qquad \|$$

Example T7.3 (Two normal distributions). Let $X_1, \ldots, X_{n_1} \sim N(\mu_1, \sigma_1^2)$ and $Y_1, \ldots, Y_{n_2} \sim N(\mu_2, \sigma_2^2)$ with $n_1 = 18$, $\bar{x} = 13.5$, $s_1 = 5$ and $n_2 = 12$, $\bar{y} = 9.5$, $s_2 = 6$.

(a) Construct a 95% CI for σ_1/σ_2.

(b) By making a further assumption based on the result in (a), construct a 95% CI for $\mu_1 - \mu_2$.

Solution: (a) Define $\upsilon_i = n_i - 1$, $i = 1, 2$. Since $f(1-\alpha/2, \upsilon_1, \upsilon_2) = f^{-1}(\alpha/2, \upsilon_2, \upsilon_1)$ and

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(\upsilon_1, \upsilon_2),$$

$$\Rightarrow \ \Pr\left\{f(1-\alpha/2, \upsilon_1, \upsilon_2) \leqslant \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \leqslant f(\alpha/2, \upsilon_1, \upsilon_2)\right\} = 1-\alpha$$

$$\Rightarrow \ \Pr\left\{\frac{S_1^2}{S_2^2}\cdot f^{-1}(\alpha/2, \upsilon_1, \upsilon_2) \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^2}{S_2^2}\cdot f^{-1}(1-\alpha/2, \upsilon_1, \upsilon_2)\right\} = 1-\alpha$$

$$\Rightarrow \ \Pr\left\{\frac{S_1^2}{S_2^2}\cdot f^{-1}(\alpha/2, \upsilon_1, \upsilon_2) \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^2}{S_2^2}\cdot f(\alpha/2, \upsilon_2, \upsilon_1)\right\} = 1-\alpha$$

$$\Rightarrow \ \Pr\left\{\sqrt{\frac{S_1^2}{S_2^2}\cdot f^{-1}(0.025, \upsilon_1, \upsilon_2)} \leqslant \frac{\sigma_1}{\sigma_2} \leqslant \sqrt{\frac{S_1^2}{S_2^2}\cdot f(0.025, \upsilon_2, \upsilon_1)}\right\} = 0.95.$$

Therefore, a 95% CI for σ_1/σ_2 is given by

$$\left[\sqrt{\frac{s_1^2}{s_2^2}\cdot f^{-1}(0.025, 17, 11)},\ \sqrt{\frac{s_1^2}{s_2^2}\cdot f(0.025, 11, 17)}\right]$$

$$= \left[\sqrt{\frac{5^2}{6^2} \cdot 3.2816^{-1}},\ \sqrt{\frac{5^2}{6^2} \cdot 2.8696} \right] = [0.4600,\ 1.4117]\,.$$

(b) From (a), the 95% CI for σ_1/σ_2 includes 1, thus we may assume that $\sigma_1/\sigma_2 = 1$, i.e., $\sigma_1 = \sigma_2$. Let $n_p = n_1 n_2/(n_1 + n_2)$, $v_p = n_1 + n_2 - 2$ and

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{v_p},$$

then we have

$$\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p/\sqrt{n_p}} \sim t(v_p),$$

$$\Rightarrow \ \Pr\left\{ -t(\alpha/2, v_p) \leqslant \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p/\sqrt{n_p}} \leqslant t(\alpha/2, v_p) \right\} = 1 - \alpha,$$

$$\Rightarrow \ \Pr\left\{ \bar{X} - \bar{Y} - t(\alpha/2, v_p)\frac{S_p}{\sqrt{n_p}} \leqslant \mu_1 - \mu_2 \leqslant \bar{X} - \bar{Y} + t(\alpha/2, v_p)\frac{S_p}{\sqrt{n_p}} \right\} = 1 - \alpha,$$

$$\Rightarrow \ \Pr\left\{ \bar{X} - \bar{Y} - t(0.025, v_p)\frac{S_p}{\sqrt{n_p}} \leqslant \mu_1 - \mu_2 \leqslant \bar{X} - \bar{Y} + t(0.025, v_p)\frac{S_p}{\sqrt{n_p}} \right\} = 0.95.$$

Therefore, a 95% CI for $\mu_1 - \mu_2$ is given by

$$\left[\bar{x} - \bar{y} \mp t(0.025, 28)\frac{s_p}{\sqrt{n_p}} \right] = \left[13.5 - 9.5 \mp 2.0484\frac{5.4149}{\sqrt{7.2}} \right] = [-0.1337,\ 8.1337]\,. \qquad \|$$