Stat 2911 Lecture Notes

Class 36, 2017

Uri Keich

CIs for: ratio of variances of normal distributions, Bernoulli(p), the normal mean with unknown variance

Confidence Intervals

Given a sample $\underline{x} = x_1, \ldots, x_n$ drawn from F_θ,
$I_\alpha(x_1, \ldots, x_n)$ is a $100(1-\alpha)\%$ CI for θ if for $\forall \theta$ and
$X_1, \ldots, X_n$ ind. F_θ-distributed RVs

$$P_\theta\left(\theta \in I_\alpha^{(n)}(X_1, \ldots, X_n)\right) \geq 1-\alpha.$$

Constructing CIs

Solve:

$$h(\theta_L; \hat\theta(\underline{x})) = P_{\theta_L}\left(\hat\Theta \geq \hat\theta(\underline{x})\right) = \alpha/2.$$

$$g(\theta_R; \hat\theta(\underline{x})) = P_{\theta_R}\left(\hat\Theta \leq \hat\theta(\underline{x})\right) = \alpha/2.$$

<u>Claim.</u> (θ_L, θ_R) is a $100(1-\alpha)\%$ CI for θ.

If $F = F_{\hat\Theta - \theta}$ is universal (θ is a location parameter) and F^{-1} exists. Then,

$$\theta_R = \hat\theta - F^{-1}(\alpha/2)$$
$$\theta_L = \hat\theta - F^{-1}(1-\alpha/2)$$

If $F = F_{\hat\Theta/\theta}$ is universal (θ is a scale parameter) and F^{-1} exists. Then,

$$\theta_R = \hat\theta / F^{-1}(\alpha/2)$$

$$\theta_L = \hat\theta / F^{-1}(1-\alpha/2)$$

4) $X_1, \dots, X_n$ are iid $N(\mu_x, \sigma_x^2)$ RVs, which are ind. of $Y_1, \dots, Y_m$ that are iid $N(\mu_y, \sigma_y^2)$.
$\mu_x, \mu_y \in \mathbb{R}$ and $\sigma_x, \sigma_y > 0$ are all unknown.

$$H_0 : \sigma_x^2 = \sigma_y^2 \quad \text{vs.} \quad H_1 : \sigma_x^2 \neq \sigma_y^2$$

If we can construct a 95% CI for $\Delta := \sigma_x^2 - \sigma_y^2$, or for $\theta := \sigma_x^2/\sigma_y^2$, then we can test H_0 by asking if 0 (Δ) or 1 (θ) is in the CI (general principle of duality between hypothesis testing and CIs).

Better to look at $\theta = \sigma_x^2/\sigma_y^2$.

Reasonable estimator:

$$\hat{\theta} = \hat{\sigma}_x^2/\hat{\sigma}_y^2 \quad \text{where} \quad \hat{\sigma}_x^2 = \frac{1}{n-1} s_x^2 = \frac{1}{n-1} \sum_1^n (x_i - \bar{x})^2$$

$$\hat{\sigma}_y^2 = \frac{1}{m-1} s_y^2 = \frac{1}{m-1} \sum_1^m (y_i - \bar{y})^2$$

Need to solve:

$$\frac{\alpha}{2} = P_{\theta_L}\left(\hat{\Theta} \geq \hat{\theta}\right) = P_{\theta_L}\left(\frac{1/(n-1)\, S_x'^2}{1/(m-1)\, S_y'^2} \geq \hat{\theta}\right)$$

Recall that $S_x'^2/\sigma_x^2 \sim \chi^2_{n-1}$ and $S_y'^2/\sigma_y^2 \sim \chi^2_{m-1}$

It follows that

$$\hat{\Theta}/\theta = \frac{1/(n-1)\, S_x'^2}{1/(m-1)\, S_y'^2} \cdot \frac{1}{\theta} = \frac{\frac{1}{n-1} S_x'^2/\sigma_x^2}{\frac{1}{m-1} S_y'^2/\sigma_y^2} \sim \frac{\frac{1}{n-1}\chi^2_{n-1}}{\frac{1}{m-1}\chi^2_{m-1}}.$$

In particular, this dist. is invariant of all unknown parameters. It has a name:

Def. The dist. of $\left(\frac{1}{n}X\right)/\left(\frac{1}{m}Y\right)$ where $X \sim \chi^2_n$ is ind. of $Y \sim \chi^2_m$ is called an F-dist. with n, m degrees of freedom ($F_{n,m}$). Let $f^{n,m}_\beta$ be its β-quantile.

$$\Rightarrow \qquad \theta_L = \hat{\theta} / f^{n-1,m-1}_{1-\alpha/2} = \frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2} \frac{1}{f^{n-1,m-1}_{1-\alpha/2}} .$$

Similarly,

$$\theta_R = \frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2} \frac{1}{f^{n-1,m-1}_{\alpha/2}} .$$

Exercise. $f^{n,m}_\beta = \frac{1}{f^{m,n}_{1-\beta}}$ Hint! $\frac{1}{F_{n,m}} \sim F_{m,n}$

$$\Rightarrow \left(\frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2} \cdot f^{m-1,n-1}_{\alpha/2} , \frac{\hat{\sigma}_x^2}{\hat{\sigma}_y^2} \cdot f^{m-1,n-1}_{1-\alpha/2} \right)$$

is a $100(1-\alpha)\%$ CI for $\theta = \sigma_x^2/\sigma_y^2$.

5) $X_i \sim \text{Bernoulli}(\theta)$.

Clearly $\hat{\theta}(\underline{x}) = \bar{x}$ and $\hat{\Theta} = \bar{X} \sim \text{binom}(n,\theta)/n$.

Neither $\hat{\Theta}-\theta$ nor $\hat{\Theta}/\theta$ have a universal dist.

However, going back to the general procedure we note that

$$h(\theta;\bar{x}) = P_\theta(\bar{X} \geq \bar{x}) = \sum_{k=n\bar{x}}^{n} \binom{n}{k} \theta^k (1-\theta)^{n-k}$$

is cont. and strictly $\uparrow$, so θ_L is well-defined.

Moreover, θ_L can be readily found by numerically solving the equ. $\alpha/2 = h(\theta_L;\bar{x})$.

Similarly, $g(\theta;\bar{x}) = P_\theta(\bar{X} \leq \bar{x})$ is cont. and strictly $\downarrow$ so θ_R is well-defined and we can find it by numerically solving $\alpha/2 = g(\theta_R;\bar{x})$.

Thus obtaining $(\theta_L\ \theta_R)$, which is an exactly computed $100(1-\alpha)\%$ CI for θ. It is however a little different!

Exact CI: $P(\theta \in I_\alpha(X_1,\dots,X_n)) = 1-\alpha$ (non-bootstrap examples)

Approximate CI: $P(\theta \in I_\alpha(X_1,\dots,X_n)) \approx 1-\alpha$ (bootstrap)

Conservative CI: $P(\theta \in I_\alpha(X_1,\dots,X_n)) \geq 1-\alpha$ (Bernoulli θ)

(typical for discrete dist).

Alternatively, $\exists$ approximate CIs for Bernoulli(θ) based on asymptotic normality.

6) $X_i \sim N(\theta, \sigma^2)$ where σ^2 is unknown but of no interest.
$\hat{\theta}(\underline{x}) = \bar{x}$ as before but now $\hat{\Theta} = \hat{\theta}(\underline{X}) = \bar{X} \sim N(\theta, \sigma^2/n)$.
Following our general procedure we first try to solve

$$\alpha/2 = h(\theta; \hat{\theta}(\underline{x}) = \bar{x}) = P_\theta(\hat{\theta}(\underline{X}) \geq \hat{\theta}(\underline{x})) = P_\theta(\bar{X} \geq \bar{x}).$$

Unfortunately this is not possible since

$$P_\theta(\bar{X} \geq \bar{x}) = P\left(\frac{\bar{X}-\theta}{\sqrt{\sigma^2/n}} \geq \frac{\bar{x}-\theta}{\sqrt{\sigma^2/n}}\right) = 1 - \Phi\left(\frac{\bar{x}-\theta}{\sqrt{\sigma^2/n}}\right),$$

which clearly depends on the unknown σ^2 (as well as θ).

One recourse we have is to recall that when σ^2 is known

$$\left(\bar{x} - z_{1-\alpha/2} \cdot \sigma/\sqrt{n},\ \bar{x} + z_{1-\alpha/2} \cdot \sigma/\sqrt{n}\right)$$

is a $100(1-\alpha)\%$ CI for θ.

This suggests an approximate CI, replacing σ with $\hat{\sigma}$,
where $\hat{\sigma}^2 = \frac{1}{n-1} S^2 = \frac{1}{n-1} \sum_1^n (x_i - \bar{x})^2$.

$$\Rightarrow \quad \left(\bar{x} - z_{1-\alpha/2} \cdot \hat{\sigma}/\sqrt{n},\ \bar{x} + z_{1-\alpha/2} \hat{\sigma}/\sqrt{n}\right)$$

is an approximate $100(1-\alpha)\%$ CI for θ.

Sometimes you have to settle for approximate CIs, but not in this case.
The problem here is that the dist. of our estimator $\hat{\Theta} = \bar{X}$ depends on both unknown parameters: scale σ and location θ.

It helps to rephrase our strategy for constructing CIs.

Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ and for a fixed observed sample $\underline{x}$ define

$$h(\theta;\underline{x}) = P_\theta(\varphi(\underline{X}) \geq \varphi(\underline{x}))$$

$$g(\theta;\underline{x}) = P_\theta(\varphi(\underline{X}) \leq \varphi(\underline{x})).$$

Then assuming $\forall \underline{x} \in \mathbb{R}^n$, $h\uparrow$, $g\downarrow$ and both are cont. $\exists \theta_L, \theta_R$ s.t.

$$h(\theta_L;\underline{x})) = P_{\theta_L}(\varphi(\underline{X}) \geq \varphi(\underline{x})) = \alpha/2$$

$$g(\theta_R;\underline{x})) = P_{\theta_R}(\varphi(\underline{X}) \leq \varphi(\underline{x})) = \alpha/2.$$

and then (θ_L, θ_R) is a $100(1-\alpha)\%$ CI for θ.

Previously we used $\varphi_1(\underline{X}) = \hat{\Theta} - \theta$, ($N(\theta,1)$ and $N(\theta,\sigma^2)$)

and $\varphi_2(\underline{X}) = \hat{\Theta}/\theta$. ($N(\mu,\sigma^2)$ as well as ratio of variances)

Both can be considered as $d(\hat{\Theta},\theta)$.

In the same vein we can consider

$$\varphi(\underline{X}) = \frac{\sqrt{n}(\bar{X}-\theta)}{\sqrt{\frac{1}{n-1} S^2}} = \frac{\sqrt{n}(\bar{X}-\theta)}{\sqrt{\frac{1}{n-1}\sum_1 (X_i - \bar{X})^2}}.$$

In the previous examples $\varphi_i(\underline{X})$ had a universal dist.

Let $Z_i = (X_i - \theta)/\sigma$ then $Z_i \sim N(0,1)$ and recall that

(i) $\frac{\bar{X}-\theta}{\sigma} = \bar{Z} \sim N(0, 1/n)$

(ii) $S^2/\sigma^2 = \sum_1^n \frac{(X_i - \bar{X})^2}{\sigma^2} = \sum_1^n (Z_i - \bar{Z})^2 \sim \chi^2_{n-1}$

Moreover, $\bar{\underline{X}}$ ($\bar{\underline{Z}}$) is ind. of $\frac{1}{n-1}S^2$ (a non-trivial fact that is a property of the normal dist.) and therefore (tutorial 10):

$$\varphi(\underline{X}) = \frac{\sqrt{n}(\bar{\underline{X}}-\theta)}{\left[\frac{1}{n-1}\sum_i(\underline{X}_i-\bar{\underline{X}})^2\right]^{1/2}} \cdot \frac{1/\sigma}{1/\sigma}$$

$$= \frac{\sqrt{n}\,\bar{\underline{Z}}}{\left(\frac{1}{n-1}S^2/\sigma^2\right)^{1/2}} \sim t_{n-1}.$$

That is, $\varphi(\underline{X})$ has a t-dist. with $n-1$ degrees of freedom <u>invariantly</u> of <u>both</u> unknown parameters θ and σ^2.

$$\Rightarrow \quad h(\theta;\underline{x}) = P_\theta\left(\varphi(\underline{X}) \geq \varphi(\underline{x})\right)$$

$$= 1 - F_{t_{n-1}}(\varphi(\underline{x})) \quad \text{(where is } \theta \text{ now?)}$$

$$= 1 - F_{t_{n-1}}\left(\frac{\sqrt{n}(\bar{x}-\theta)}{\hat{\sigma}}\right)$$

Similarly,

$$g(\theta;\underline{x}) = P(\varphi(\underline{X}) \leq \varphi(\underline{x})) = F_{t_{n-1}}\left(\frac{\sqrt{n}(\bar{x}-\theta)}{\hat{\sigma}}\right).$$

It is not difficult to verify that h is strictly $\uparrow$, g is strictly $\downarrow$ and both are cont., so we can solve

$$\frac{\alpha}{2} = h(\theta_L;\varphi(\underline{x})) \quad \text{and} \quad \frac{\alpha}{2} = g(\theta_R;\varphi(\underline{x}))$$

and then (θ_L,θ_R) is a $100(1-\alpha)\%$ CI for θ.

Explicitly, let t_p^n be the quantiles of F_{t_n}. Then

$$\frac{\alpha}{2} = g(\theta_R; \underline{x}) = F_{t_{n-1}}\left(\frac{\sqrt{n}(\bar{x}-\theta_R)}{\hat{\sigma}}\right)$$

$$\Rightarrow \theta_R = \bar{x} - t^{n-1}_{\alpha/2} \cdot \hat{\sigma}/\sqrt{n}$$

$$\overset{\text{sym}}{=} \bar{x} + t^{n-1}_{1-\alpha/2} \cdot \hat{\sigma}/\sqrt{n}$$

Similarly,

$$\frac{\alpha}{2} = h(\theta_L; \underline{x}) = 1 - F_{t_{n-1}}\left(\frac{\sqrt{n}(\bar{x}-\theta_L)}{\hat{\sigma}}\right)$$

$$\Rightarrow \theta_L = \bar{x} - t^{n-1}_{1-\alpha/2} \hat{\sigma}/\sqrt{n}$$

$$\Rightarrow \left(\bar{x} - t^{n-1}_{1-\alpha/2} \hat{\sigma}/\sqrt{n},\ \bar{x} + t^{n-1}_{1-\alpha/2} \hat{\sigma}/\sqrt{n}\right)$$

is a $100(1-\alpha)\%$ CI for θ.

Compare with approximate CI!

$$\left(\bar{x} - z_{1-\alpha/2} \cdot \hat{\sigma}/\sqrt{n},\ \bar{x} + z_{1-\alpha/2} \hat{\sigma}/\sqrt{n}\right)$$

Fact: $t_n \xrightarrow[n]{} N(0,1)$.

← ended here →

Extra material (not needed for final!)

Let
$$h(\theta;\underline{x}) = P_\theta(\varphi(\underline{\underline{X}}) \geq \varphi(\underline{x}))$$
$$g(\theta;\underline{x}) = P_\theta(\varphi(\underline{\underline{X}}) \leq \varphi(\underline{x})).$$

Then assuming $\forall \underline{x} \in \mathbb{R}^n$, $h\uparrow$, $g\downarrow$ and both are cont. $\exists \theta_L, \theta_R$ s.t.
$$h(\theta_L;\underline{x})) = P_{\theta_L}(\varphi(\underline{\underline{X}}) \geq \varphi(\underline{x})) = \alpha/2$$
$$g(\theta_R;\underline{x})) = P_{\theta_R}(\varphi(\underline{\underline{X}}) \leq \varphi(\underline{x})) = \alpha/2.$$

Claim. (θ_L, θ_R) is a $100(1-\alpha)\%$ CI for θ.

Proof. Let $\Theta_L = \theta_L(X_1, \dots, X_n)$, $\Theta_R = \theta_R(X_1, \dots, X_n)$

$$P_\theta(\theta \notin (\Theta_L, \Theta_R)) = P_\theta(\theta \leq \Theta_L) + P_\theta(\theta \geq \Theta_R)$$

Since g is strictly $\downarrow$, $\forall \underline{x} \in \mathbb{R}^n$

$$\theta \geq \theta_R(\underline{x}) \iff \alpha/2 \geq g(\theta;\underline{x})) = P_\theta(\varphi(\underline{\underline{X}}) \leq \varphi(\underline{x}))$$
$$\iff \alpha/2 \geq F^\theta_{\varphi(\underline{X})}(\varphi(\underline{x}))$$

$$\Rightarrow P_\theta(\theta \geq \theta_R(\underline{\underline{X}})) = P_\theta(F^\theta_{\varphi(\underline{X})}(\varphi(\underline{\underline{X}})) \leq \alpha/2)$$
$$\leq \alpha/2$$ (tutorial 9, problem 6)

Similarly, since h is strictly $\uparrow$, $\forall \underline{x} \in \mathbb{R}^n$

$$\theta \leq \theta_L(\underline{x}) \iff \alpha/2 \geq h(\theta; \underline{x}) = \underbrace{P_\theta\left(\varphi(\underline{X}) \geq \varphi(\underline{x})\right)}_{G^\theta_{\varphi(\underline{X})}(\varphi(\underline{x}))}$$

$$\Rightarrow P_\theta\left(\theta \leq \theta_L(\underline{X})\right) = P_\theta\left(G^\theta_{\varphi(\underline{X})}(\varphi(\underline{X})) \leq \alpha/2\right)$$

$$\leq \alpha/2.$$ (tutorial 9, problem 7)

$\square$