ST102 Week 17

Confidence Interval (continued)

For i.i.d. random sample $\{X_i\}_1^n$ from $F(x;\theta)$, and a statistic $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$ is the MLE of θ.
Recall that, under some regularity assumptions,

$$\hat{\theta} \underset{n \nearrow +\infty}{\overset{approx}{\sim}} N\left(\theta, \frac{1}{nI(\theta)}\right)$$

with $I(\theta)$ the Fisher Information.

Then the $(1-\alpha)$ confidence interval of θ is

$$\left(\hat{\theta} - z_{\frac{\alpha}{2}} \cdot [nI(\hat{\theta})]^{-\frac{1}{2}},\ \hat{\theta} + z_{\frac{\alpha}{2}} \cdot [nI(\hat{\theta})]^{-\frac{1}{2}}\right)$$

<u>For variances of Normals</u>

$\{X_i\}_1^n$ from Normal with known μ & $\sigma^2 > 0$.

Let $M = (n-1)S^2 \implies \frac{M}{\sigma^2} \sim \chi^2_{n-1}$ (proof in Lecture)

By the distribution of $Y \sim \chi^2_{n-1}$, we can find $0 < k_1 < k_2$:

$$\mathbb{P}(Y \le k_1) = \mathbb{P}(Y \ge k_2) = \frac{\alpha}{2}, \forall \alpha \in (0,1);$$

$$\implies 1-\alpha = \mathbb{P}\left(k_1 < \frac{M}{\sigma^2} < k_2\right)$$

$\implies$ An $(1-\alpha)$ confidence interval for σ^2 is

$$\left(\frac{M}{k_2}, \frac{M}{k_1}\right)$$

Hypothesis Testing

With i.i.d. random sample $\{X_i\}_1^n$ from some distribution with c.d.f. $F(x;\theta)$, we want to test

$$H_0: \theta = \theta_0 \qquad \text{null hypothesis}$$
$$H_1: \theta \in \Theta_1 \qquad \text{alternative hypothesis}$$

for some specific θ_0 and the set of alternative values Θ_1 with $\theta_0 \notin \Theta_1$.

α: significance level

We usually design some test statistic
$$T = T(X_1, \cdots, X_n)$$
For any given sample, T would correspondingly take a specific value $T = t$.

p-value: $p := \mathbb{P}_{H_0}(T = t \text{ or more "extreme" values})$

Decision rule: $\begin{cases} \text{reject } H_0 & \text{if } p \le \alpha \\ \text{don't reject } H_0 & \text{otherwise.} \end{cases}$

Alternatively, define critical value C_α with
$\mathbb{P}(|T|, \text{ under } H_0, \text{ takes value at least as extreme as } C_\alpha) = \alpha$

$\Rightarrow$ reject H_0 iff $|T| \ge C_\alpha$.

Two-sided test for normals

$\{X_i\}_1^n$ from $N(\mu, \sigma^2)$, $\sigma^2 > 0$ known.

$$H_0 : \mu = \underset{\substack{\downarrow \\ \text{given}}}{\mu_0} \quad \text{v.s.} \quad H_1 : \mu \neq \mu_0$$

One choice of test statistic is

$$T = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sigma} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \underset{\substack{\downarrow \\ \text{under } H_0}}{\sim} N(0, 1)$$

Reject H_0 if $|T|$ large.
In this case, $C_\alpha = z_{\frac{\alpha}{2}}$.

One-sided test for normals

$\{X_i\}_1^n$ from $N(\mu, \sigma^2)$, with $\sigma^2 > 0$ known. Test

$$H_0 : \mu = \mu_0 \quad \text{v.s.} \quad H_1 : \mu < \mu_0$$

Still we $T = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ as test statistic.

We can derive that $C_\alpha = -z_\alpha$ and reject H_0 if $T \le C_\alpha$.

Tests for normals with unknown variance

$\{X_i\}_1^n$ from $N(\mu, \sigma^2)$, both μ & σ^2 unknown.

$$H_0: \mu = \mu_0 \quad \text{v.s.} \quad H_1: \mu < \mu_0$$

Test statistic $T = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \underset{H_0}{\sim} t_{n-1}$

Then we know $C_\alpha = -t_{n-1,\alpha}$ and we reject H_0 if $T \le C_\alpha$.

The general formulation

$\{X_i\}_1^n$ from $F(x;\theta)$. Test (assuming $\Theta_0 \cap \Theta_1 = \phi$)

$$H_0: \theta \in \Theta_0 \quad \text{v.s.} \quad H_1: \theta \in \Theta_1,$$

with significance level α.

Step 1. Design test statistic $T = T(X_1, \cdots, X_n)$, especially we need to know the distribution of T under H_0.

Step 2. According to such distribution & Θ_0 & Θ_1, identify the critical region $\mathbb{C}$ s.t.

$$\mathbb{P}_{H_0}(T \in \mathbb{C}) = \alpha$$

Step 3. Calculate T on data, reject H_0 if $T \in \mathbb{C}$.