Annoocements

- Exams graded by some time on Monday.
- A feedback survey will be released today.
- Cut-offs finalization will be announced Monday.

Interval Estimation

Recall: Point estimators, $X_1, X_2, \dots, X_n$, random sample

Estimate θ by a point estimator

$$\hat{\Theta} = \hat{\Theta}(X_1, \dots, X_n).$$

How good is the estimate?

Interval estimator.

- Instead of estimating a point, we estimate an interval.

$[\hat{\Theta}_L, \hat{\Theta}_H]$ so that θ belongs to this interval with high probability.

Def: (Confidence interval). $X_1, X_2, \dots X_n$, θ.

An interval estimator is two estimators

$$\hat{\Theta}_L = \hat{\Theta}_L(X_1, \dots, X_n).$$
$$\hat{\Theta}_H = \hat{\Theta}_H(X_1, \dots, X_n).$$

such that

$$P\left(\theta \in [\hat{\Theta}_L, \hat{\Theta}_H]\right) \geq \underbrace{1-\alpha}_{\text{confidence level}}.$$

$$\hat{\Theta}_L \leq \theta \leq \hat{\Theta}_H$$

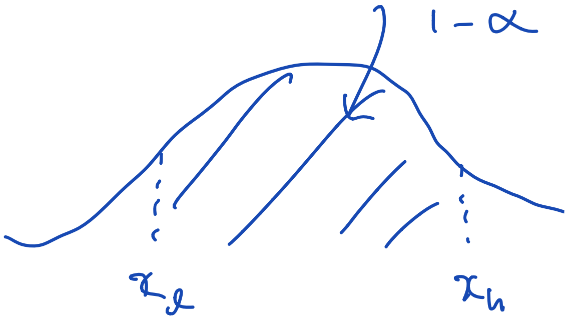
We say $[\hat{\Theta}_L, \hat{\Theta}_H]$ is a $(1-\alpha)\%$ confidence interval for θ.

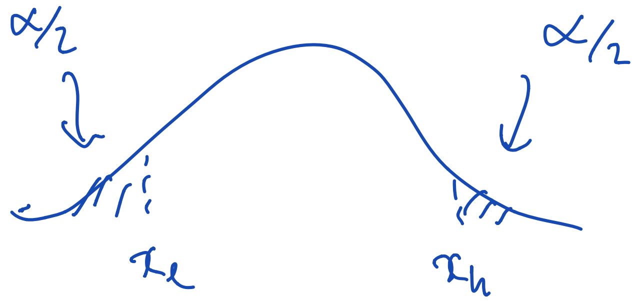
Consider first. X - continuous RV

What are x_l, x_h s.t.

$$P\left(x_l \leq X \leq x_h\right) = 1-\alpha. \ ?$$

$$\Rightarrow \quad P(x_l \geq X) + P(X \geq x_h) = \alpha.$$





Find x_ℓ, x_h by solving

$$P(X \le x_\ell) = \alpha/2, \quad P(X \ge x_h) = \alpha/2$$

use CDF

$$F_X(x_\ell) = \alpha/2, \quad F_X(x_h) = 1 - \alpha/2$$

$$\Rightarrow \quad x_\ell = F_X^{-1}(\alpha/2), \quad x_h = F_X^{-1}(1-\alpha/2)$$

Ex $Z \sim N(0,1)$ find x_ℓ, x_h

$$P(x_\ell \le Z \le x_h) = .95$$

$$\alpha = 1 - .95 = 0.05$$

$$x_\ell = \Phi^{-1}(\alpha/2) = \Phi^{-1}(0.025) = \boxed{-1.96}$$

$$x_h = \Phi^{-1}(1-\alpha/2) = \Phi^{-1}(1-0.025) = \boxed{1.96}$$

$x_h = -x_\ell$ by symmetry.

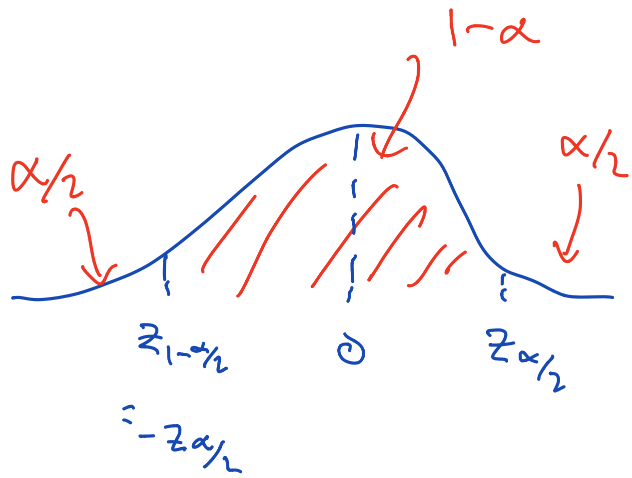
For general α.

<u>Def</u>

$$z_{\alpha/2} = \Phi^{-1}(1 - \alpha/2)$$

$$z_{1-\alpha/2} = \Phi^{-1}(\alpha/2) = -z_{\alpha/2}$$

symmetry



How to find a confidence interval.

- Start w/ a point estimator $\hat{\Theta}$ (MLE)
- Add and subtract numbers to get an interval.

<u>Ex</u> $X_1, X_2, \ldots, X_n \sim N(\theta, 1)$

<u>What is a</u> $1-\alpha$ % confidence interval for θ?

$$\hat{\Theta} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \sim N(\theta, 1/n)$$

Standardize (Pivot).

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

We know

$$P\left(-z_{\alpha/2} \le Z \le z_{\alpha/2}\right) = 1-\alpha.$$

don't depend on θ.

Use algebra

$$-z_{\alpha/2} \le \sqrt{n}(\bar{X} - \theta) \le z_{\alpha/2}$$

$$\Updownarrow$$

$$-z_{\alpha/2} \le \sqrt{n}(\theta - \bar{X}) \le z_{\alpha/2}.$$

$$\Updownarrow$$

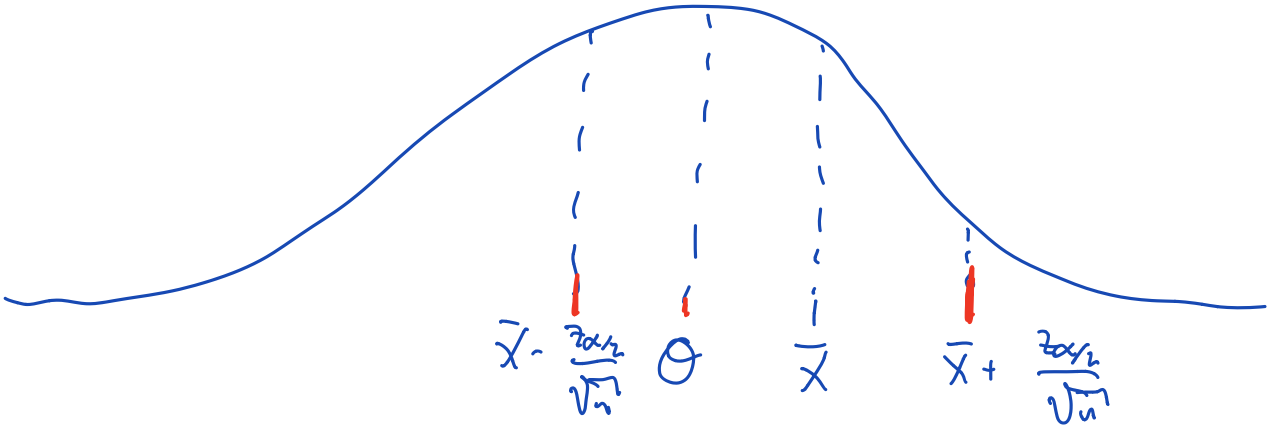
$$\bar{X} - \frac{z_{\alpha/2}}{\sqrt{n}} \le \theta \le \bar{X} + \frac{z_{\alpha/2}}{\sqrt{n}}$$

$$P\left(\bar{X} - \frac{z_{\alpha/2}}{\sqrt{n}} \le \theta \le \bar{X} + \frac{z_{\alpha/2}}{\sqrt{n}}\right) = 1-\alpha.$$

Therefore

$$[\hat{\theta}_L, \hat{\theta}_H] = \left[\bar{X} - \frac{z_{\alpha/2}}{\sqrt{n}}, \bar{X} + \frac{z_{\alpha/2}}{\sqrt{n}}\right]$$

is a $(1-\alpha)\%$ confidence interval for θ.



What is the main idea?

$$\bar{Z} = \frac{\bar{X} - \theta}{1/\sqrt{n}}$$ pivotal quantity.

the distribution doesn't depend on θ.

<u>Def</u> A pivotal quantity

$$Q = Q(X_1, X_2, \dots, X_n, \theta)$$

such that the probability distribution doesn't depend on θ.

<u>General Method</u>

① Find the pivotal quantity

② Find q_L, q_u $\quad P(q_L \le Q \le q_u) = 1 - \alpha$.

q_L, q_u don't depend on θ.

③ Use algebra to turn into

$$P(\hat{\Theta}_l \le \Theta \le \hat{\Theta}_u) = 1-\alpha.$$

What if X_i are not normal?

<u>Ex</u> Large sample size. (CLT Approach)

$X_1, X_2, \dots, X_n$, random sample

$\mathbb{E}X_i = \Theta$, $Var(X_i) = \sigma^2$.

unknown known.

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

Pivot $Q = \dfrac{\bar{X} - \Theta}{\sigma/\sqrt{n}} \approx N(0,1)$

n large CLT.

$$P(-z_{\alpha/2} \le Q \le z_{\alpha/2}) \approx 1-\alpha.$$

$\Downarrow$

$$P\left(\bar{X} - \frac{z_{\alpha/2}\,\sigma}{\sqrt{n}} \le \Theta \le \bar{X} + \frac{z_{\alpha/2}\,\sigma}{\sqrt{n}}\right) \approx 1-\alpha.$$

$$\left[\bar{X} - \frac{z_{\alpha/2}\,\sigma}{\sqrt{n}},\ \bar{X} + \frac{z_{\alpha/2}\,\sigma}{\sqrt{n}}\right]$$ is $(1-\alpha)\%$ confidence interval

What if you don't know σ?

① Find σ_{max} s.t. $\sigma \le \sigma_{max}$.

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