

Point estimator brief recall

Recall:

- We say that \bar{X} is a point estimate of the population mean
- We want our estimator to be unbiased (or low bias). Estimator \bar{X} is unbiased provided our sample is random.
- We want our estimator to have a small variance (standard error)
- An efficient estimator is one that has lower MSE than another estimator

Errors in our estimate of the mean

- How large might $\bar{X} - \mu$ get? $(\bar{X} - \mu)$
- By the CLT, for the moderately large (>30) sample $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$ $CLT = \text{Central Limit Theorem}$
- Let's look at the two boundary points, $\pm z_{\alpha/2}$ $n = \text{sample length}$
 $\sigma = \text{std in population}$

$$! \quad \pm z_{\alpha/2} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \rightarrow \mu = \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \quad \mu \in (\bar{x} - E; \bar{x} + E)$$

- The probability that the true mean will fall within this range is $1-\alpha$

Confidence intervals

- A plausible range of values for the population parameter is called a confidence interval.
- Using only a point estimate to estimate a parameter is like fishing in a murky lake with a spear, and using a confidence interval is like fishing with a net.
- We can throw a spear where we saw a fish but we are more likely to miss. If we toss a net in that area, we have a better chance of catching the fish.
- If we report a point estimate, we probably will not hit the exact population parameter. If we report a range of plausible values – a confidence interval – we have a good shot at capturing the parameter.

$$\hat{\theta} \in \Theta$$

Confidence intervals

Estimating population mean, μ

$\sigma =$ Population variance -- known

What do you need?

- Random Sample
- Population variance, σ^2 , -- is known (!)
- $n > 30$ - CLT works fine, if not - assumption that population is normally distributed, $\text{if global distr. is normal}$

Point estimate:

\bar{X} (sample mean) is an unbiased point estimator for population mean

Margins of errors

One of possible ways to write:

$$E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = \text{error margin}$$

Can be thought of as Max difference between \bar{x} and μ .

$$\bar{X} - E < \mu < \bar{X} + E$$

Construction of CI

- Check requirements
- Calculate critical value $Z_{\alpha/2}$; $Z_{\alpha/2}$
- Find out the Error margin E , $E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$
- Make confidence interval

$$Z_{\alpha/2}: P(Z \in (Z_{\alpha/2}, +\infty)) = \frac{\alpha}{2}$$

$$\sigma^2 = \text{var.}$$

$$\sigma = \sqrt{\sigma^2} = \text{std. d.}$$

Example 1

$$\mu = 76.3 \pm 2.57 \frac{12.5}{\sqrt{40}}$$

Random sample of 40 students. The average resting heart-rate for the sample was 76.3 bpm. Assume the population std is 12.5 bpm. Construct a 99% CI for the average resting heart-rate of the population.

$$(1-\alpha) = 0.99 \rightarrow \alpha = 0.01 \rightarrow \alpha/2 = 0.005$$

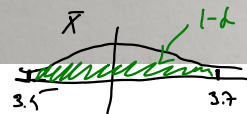
Example 2



$$Z_{\alpha/2}: P(0 < Z < Z_{\alpha/2}) = 0.5 - \frac{\alpha/2}{2} = 0.493$$

Problem C.2.1. Manager of a restaurant wants to estimate the mean amount m that a visitor spends for a lunch. A sample contains 36 visitors. Sample mean is $\bar{x} = \$3.60$. Manager knows that the standard deviation for one visitor is \$0.72. Find the confidence level corresponding to the interval (\$3.5; \$3.7).

$$\alpha/2 = 1/2 - P(0 < Z < Z_{\alpha/2})$$



$$3.5 = \bar{x} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$3.7 = \bar{x} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Example 3

$$\alpha/2 = 0.5 - 0.2969 = 0.2031 \rightarrow \alpha = 0.406$$

$$(1-\alpha) \cdot 100\% = 59.4\%$$

A college admission officer for an MBA program has determined that historically candidates have undergraduate grade point averages that are normally distributed with std 0.45. A random sample of 25 applications from the current year is taken, yielding a sample mean grade average of 2.90.

- Find a 95% CI for the populatin mean
- Based on these sample results, a statistician computes for the populatin mean a CI running from 2.81 to 2.99. Find the probability content associated with this interval.