Point estimator brief recall

Recall:

- We say that $\bar{X}$ s 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?
- By the CLT, for the moderately large (>30) sample $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- Let's look at the two boundary points, $\pm z_{\alpha/2}$:

$$\pm z_{\alpha/2} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \to \mu = \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

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

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.

Confidence intervals

Estimating population mean, μ

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,

Margins of errors

One of possible ways to write:

$$E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}.$$

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_α
- Find out the Error margin E, $E = Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$
- Make confidence interval

Example 1

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.

Example 2

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).

Example 3

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

Estimating population mean, μ

Population variance – unknown

Error margin:

$$E_t = t_{\alpha/2,n-1}\frac{S}{\sqrt{n}},$$

where S is a sample standard deviation, $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

Then, CI will be of the form:

$$\bar{X} - E_t < \mu < \bar{X} + E_t.$$

Example 4

Problem C.2.7. A random sample of 5 observations from a normal distribution with mean μ and variance σ^2 gives a sample mean 100. An independent random sample of size 10 from the same population has sample variance 9. Find a 90% confidence interval for the population mean.

Estimating population proportion, p

Error margin:

$$E_p = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

where $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ is called a standard error of $\hat{p}$.

Then, CI will be of the form:

$$\hat{p} - E_p < p < \hat{p} + E_p.$$

Example 5

Problem C.2.16. A random sample was taken of 189 National Basketball Association games in which the score was not tied after one quarter. In 132 of these games, the team leading after one quarter won the game.

(a) Find a 90% confidence interval for the population proportion of all occasions on which the team leading after one quarter wins the game.

(b) Without doing the calculations, state whether a 95% confidence interval for the population mean would be wider than or narrower than that found in (a).