Chapter 6
Confidence Intervals

ACSTA 101 - Professor MG

6.1

Confidence Intervals For the Mean (σ known)

Definition

A point estimate is a single value estimate for a population parameter.

One example of this is using a sample mean to estimate a population mean.

Example

19	25	15	21	22	20	20	22
22	21	21	23	22	16	21	18
25	23	23	21	22	24	18	19
23	20	19	19	24	25	17	21
21	25	23	18	22	20	21	21

A researcher is collecting data about a college athletic conference and its student-athletes. A random sample of 40 student-athletes is selected and their numbers of hours spent on required athletic activities for one week are recorded (see table at left). Find a point estimate for the population mean m, the mean number of hours spent on required athletic activities by all student-athletes in the conference.

Answer: 21.1 hours approximates the population mean

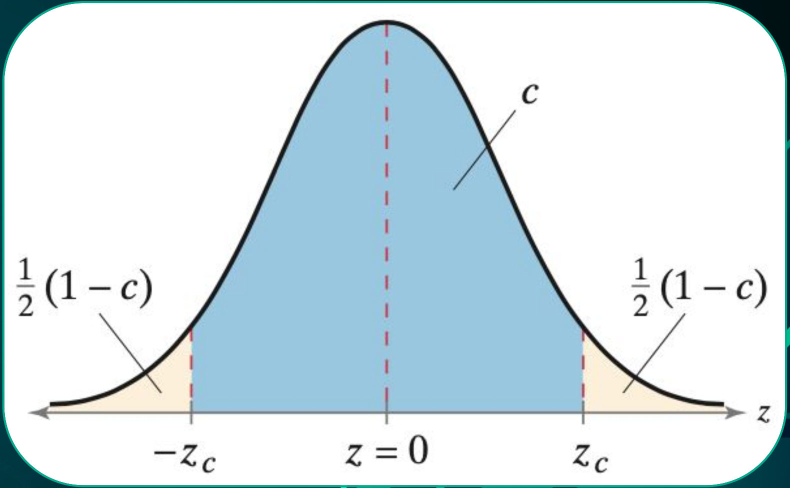
Definition

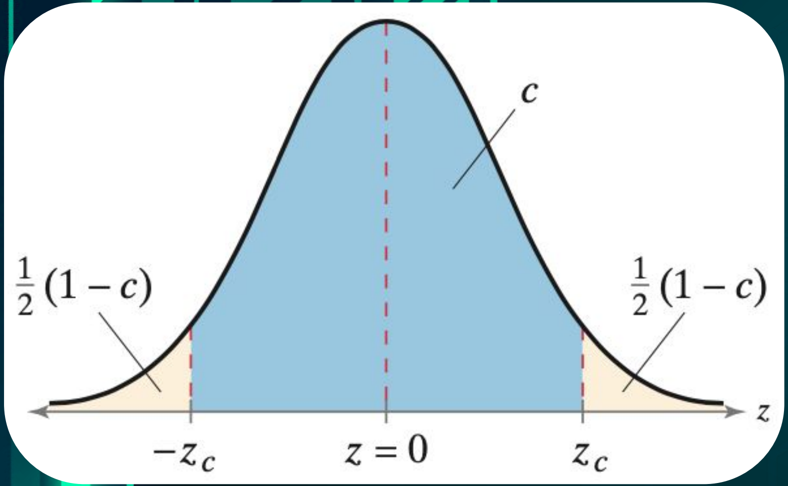
A point estimate is extremely unlikely to give an accurate representation of the actual value of the population parameter.

To improve on this, we instead use an interval estimate, which is an interval (i.e. a range of values) that the population parameter is likely to fall in.

The level of confidence (denoted c) is the probability that the interval estimate contains the population parameter. For normally distributed data, the level of confidence is the area under the confidence interval.

Usually, we use confidence levels c = 90 , 95, and 99.





Example: Find the critical values for a confidence level of 90%

$$\text{Area to left of } -z_c = \frac{1}{2}(1-0.9) = 0.05$$

$$-z_c = -1.645,\ z_c = 1.645$$

Definitions

Sampling Error

- The difference between a point estimate and the actual population.
  - Equal to (sample mean)-(population mean) when talking about mean estimates
  - Since the sample statistic changes from sample to sample, we instead look at the maximum sampling error, called the margin of error.

Margin of Error

- Given a level of confidence, the margin of error (denoted E) is calculated by:

$$E = z_c \sigma_{\bar{x}} = z_c \cdot \frac{\sigma}{\sqrt{n}}$$

When:
  - The sample taken is random
  - The population is approximately normal (i.e. normally distributed or $n \geq 30$)
- Sometimes called the maximum error of estimate or error tolerance

Example

For the previous sports data: $z_c = 1.96$ and $E = 0.7$

Definition

A confidence interval for a population with mean μ is given by

$$\bar{x} - E < \mu < \bar{x} + E \text{ or } (\bar{x} - E, \bar{x} + E)$$

For level of confidence c, this indicates that you are c% confident that the population mean lies in that interval.

Example

A college admissions director wishes to estimate the mean age of all students currently enrolled. In a random sample of 20 students, the mean age is found to be 22.9 years. From past studies, the standard deviation is known to be 1.5 years, and the population is normally distributed. Construct a 90% confidence interval for the population mean age.

$$n = 20$$
$$\bar{x} = 22.9$$
$$\sigma = 1.5$$
$$z_c = 1.645$$

$$E = 1.645 \cdot \frac{1.5}{\sqrt{20}} \approx 0.6$$

This gives a confidence interval of (22.3, 23.5)

You can improve the precision of your estimate by increasing your sample size.

Given a c confidence level and a margin of error E, the minimum sample size, n, needed to estimate the population mean is given by

$$n = \left(\frac{z_c \sigma}{E}\right)^2$$

Round up to the nearest whole number, if necessary

Example

Using the information from the previous example on student athletes, how many students should be included in a sample to be 95% confident that the sample mean is within 0.5 hour of the population mean?

$$c = 0.95,\ z_c = 1.96, \sigma = 2.3, E = 0.5$$

$$n = \left(\frac{1.96 \cdot 2.3}{0.5}\right)^2 \approx 81.29$$

At least 82 people would need to be randomly selected.