Quantitative Business Analysis

Estimation and confidence intervals

Point estimates, uncertainty, and a calculation you can interpret

Lectures 18–19 · Shi Qiu

Start the narrated lesson, then pause any time to discuss or calculate.

Parameter, estimator, estimate

A retailer wants the population mean daily demand.

Parameter

The fixed population quantity we want to learn.

$$\mu$$

Estimator

A rule that produces a number from a random sample.

$$\bar{X}$$

Estimate

The value of that rule for the sample we actually observed.

$$\bar{x} = 246$$

> The estimator varies across samples. This sample gives one estimate.

Point and interval estimates

Two ways to report the same sample's information.

Point estimate

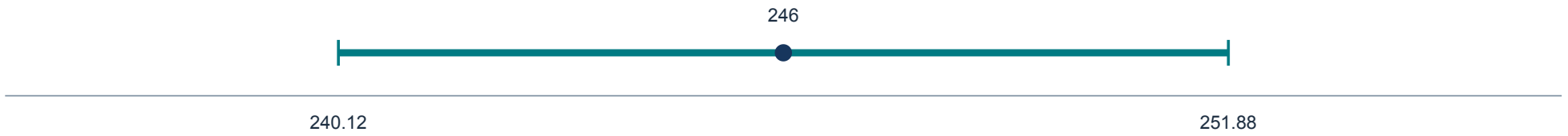
$$\bar{x} = 246$$

One estimated value for the population mean.

Interval estimate

$(240.12,\ 251.88)$

A range produced by a stated confidence procedure.



> The interval also communicates how precisely the sample estimates the mean.

Constructed example from the current lecture deck. We calculate it shortly.

Properties of an estimator

Unbiasedness	Across repeated random samples, the average estimate equals the target. $\mathrm{E}(\bar{X}) = \mu$
Consistency	As the sample grows, the estimator concentrates near the target. $\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n} \longrightarrow 0$
Efficiency	Among unbiased estimators of the same parameter under the same model and sample size, smaller variance gives greater precision.

For the sample-mean statements, assume independent draws from one population with finite variance. Shrinking variance alone does not establish consistency if bias persists.

Standard error and assumptions

Standard error measures the sampling variability of the mean.

$$\mathrm{SE}(\bar{X}) = \frac{\sigma}{\sqrt{n}}$$

σ describes individual observations.
SE describes sample means.

The model for this lecture

Independent random observations from one population.

The population standard deviation σ is known.

A normal population gives an exactly normal sample mean.

A sufficiently accurate central limit approximation supports an approximate interval.

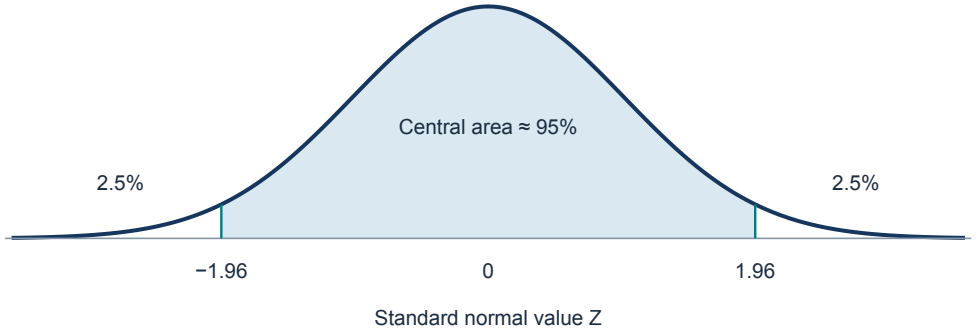
n denotes the sample size. The formula assumes independent observations.

Confidence level and critical value

The positive critical value z* leaves the requested central area under a standard normal curve.

$$P(-z^* \le Z \le z^*) = 0.95$$

Confidence	Critical value z*
90%	1.645
95%	**1.960**
99%	2.576



At 95% confidence, each tail has area 0.025.

Z denotes a standard normal variable. Displayed critical values are rounded.

A confidence interval for a mean

With known σ, estimate the population mean μ using:

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$$

Margin of error

$$E = z^* \frac{\sigma}{\sqrt{n}}$$

Full interval width

$$\text{Width} = 2E$$

> Before sampling, the endpoints vary with the sample mean. After sampling, insert the observed value x̄.

The information in the question

Constructed inventory example. Assume a valid normal model or an accurate CLT approximation.

“Daily demand observations are independent.”	The standard error formula applies to independent observations.
“A sample of 100 days has mean demand of 246 units.”	$n = 100, \quad \bar{x} = 246$
“The population standard deviation is known to be 30 units.”	$\sigma = 30$
“Find a two-sided 95% confidence interval for mean demand.”	$z^* \approx 1.96$

The calculation

Sample mean 246 units · Sample size 100 · Known population SD 30 units

1 **Standard error**

$$\mathrm{SE} = \frac{30}{\sqrt{100}} = 3$$

2 **Margin of error**

$$E = 1.96(3) = 5.88$$

3 **Endpoints**

$$246 \pm 5.88 = (240.12,\ \ 251.88)$$

> The margin is 5.88 units. The full width is 11.76 units.

Interpretation

> “We are 95% confident that population mean daily demand is between **240.12 and 251.88 units**.”

The target	The population mean daily demand. Individual days can fall outside this interval.
The 95%	The long-run coverage of the procedure across repeated samples.

After observing a sample, both the calculated limits and the population mean are fixed. The frequentist confidence level belongs to the procedure.

Repeated intervals, fixed population mean

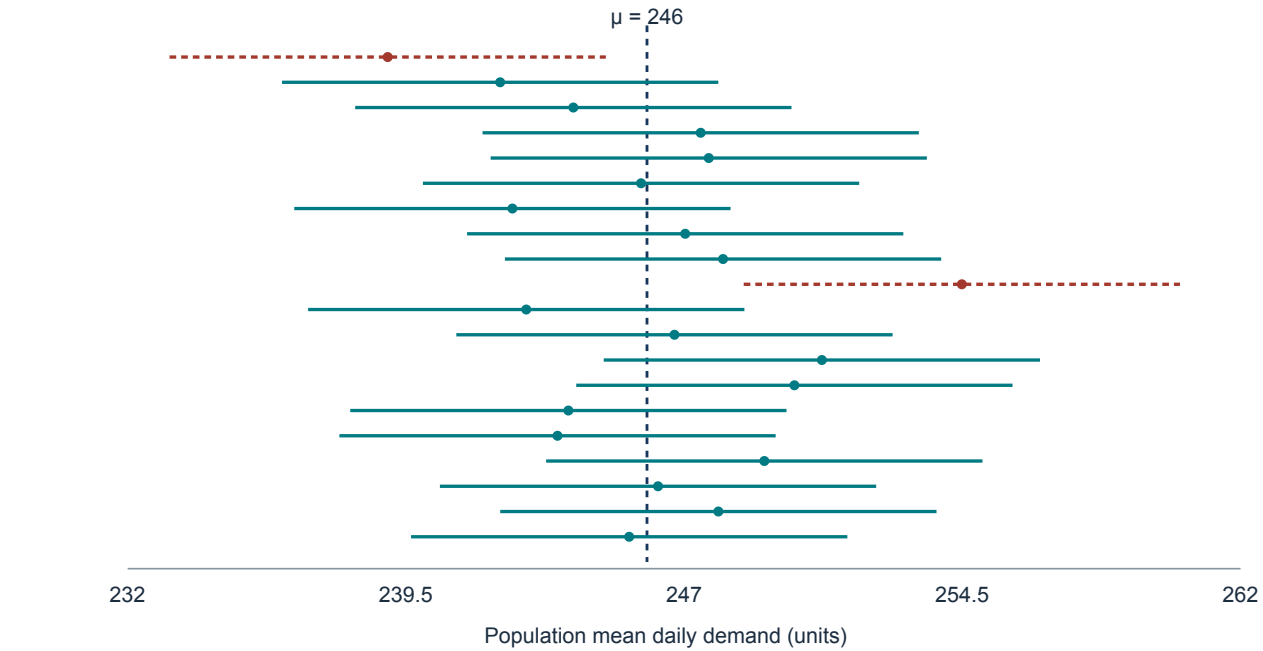
Each row uses a new sample of 100 observations from the same normal model.

The simulation sets the true mean to 246 units and the population SD to 30.

18 of 20 intervals cover the mean

New simulated batch

A batch can capture more or less than 95%. Long-run coverage concerns many repetitions.



Dots mark sample means. Dashed intervals miss the true mean.

Confidence and precision

Keep the sample mean at 246 and the known population SD at 30 units.

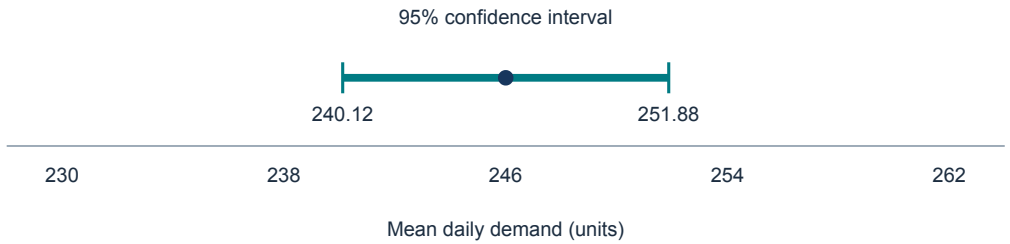
Sample size n **100**

Confidence level

95%

$$E = z^* \frac{\sigma}{\sqrt{n}}$$

Adjusting a control pauses autoplay.



Standard error **3.00**

Margin of error **5.88**

> Four times the sample size halves the margin at fixed confidence and population SD.

Minimum sample size

Mean order value (AED) · 95% confidence · Known σ = 75 · Desired margin at most 16

$$n \geq \left(\frac{z^{*}\sigma}{E_0}\right)^2, \quad E_0 = \text{target margin}$$

$$n \geq \left(\frac{1.96(75)}{16}\right)^2 = 84.41015625$$

$$n_{\min} = 85$$ The required size is the next whole number.

84 observations: margin 16.039, above the target.

85 observations: margin 15.944, within the target.

Quick check

A service company samples 64 independent repairs, with mean duration 260 minutes. The population is normal with known SD 40 minutes.

Find and interpret a two-sided 95% confidence interval for the population mean repair duration.

$$\mathrm{SE} = 40/\sqrt{64} = 5, \quad E = 1.96(5) = 9.80$$

$$260 \pm 9.80 = (250.20,\ 269.80)$$

> We are 95% confident that population mean repair duration is between 250.20 and 269.80 minutes.

Constructed practice example. Pause before revealing the solution.

A population proportion

Let x count successes among n independent observations. The population proportion is p.

$$\hat{p} = x/n$$

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Check random sampling and suitable independence. When ignoring a finite-population adjustment, the population should be large relative to the sample.

> For this course’s approximation, both observed counts must be at least 5.

This usual normal-based interval is approximate and can perform poorly near probability boundaries.

Proportion calculation and interpretation

Of 300 randomly sampled store customers, 168 choose home delivery. Assume independent responses and a much larger customer population.

$$\hat{p} = 168/300 = 0.56, \quad n - x = 132$$

$$\widehat{\mathrm{SE}} = \sqrt{\frac{0.56(0.44)}{300}} \approx 0.028659$$

$$E \approx 0.05617, \quad 95\%\ \mathrm{CI} = (0.5038,\ 0.6162)$$

> We are 95% confident that 50.38% to 61.62% of the store’s customer population choose home delivery.

Constructed customer-delivery example.