Chapter 8 - Statistical Modeling and Inference

STAT 251

Lecture 26

Examples: point estimates and confiaence intervals
Confidence Interval for μ when σ unknown
t-distribution
Confidence Intervals for the mean - Applet

Dr. Lasantha Premarathna

Chapter 8 - Learning Outcomes

- Point Estimation for μ ans σ
- Bias of an estimator
- Confidence Interval for μ
- Testing of Hypotheses about μ
- One sample problems
- Two sample problems

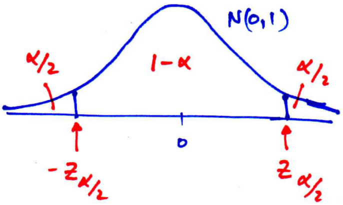
Re-cap: $(1-\alpha)100\%$ Confidence Interval (CI) for μ

In general, $(1-\alpha)100\%$ Confidence Interval (CI) for Population mean μ is

$$\left[\bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \;,\; \bar{x} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

or

$$\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$



$\bar{x}$ is the point estimate for μ and

$Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ is the margin of error

Example: 1

The average zinc concentration recovered from a sample of measurements taken in 36 different locations in a river is found to be 2.6 grams per milliliter. Assume that the population standard deviation is 0.3 grams per milliliter.

(a) Find the point estimate for μ, the true mean zinc concentration in the river.

(b) Find the 95% confidence interval for the mean zinc concentration in the river.

(c) Find the 99% confidence interval for the mean zinc concentration in the river.

(d) Suppose that the researchers would like to have the margin of error 0.05 grams per milliliter. Calculate the sample size they need to take, if they use 95% confidence level.

Example: 1 - Solution

$$\bar{x} = 2.6, \quad n = 36, \quad \sigma = 0.3$$

(a) point estimate for the true mean zinc

$$\text{concentration } (\mu) = \hat{\mu} = \bar{x} = 2.6 \ g/ml$$

Example: 1 - Solution

(b) Find the 95% confidence interval for the mean zinc concentration in the river.

<u>95% CI for μ</u>

$$\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \qquad\qquad Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$$

$$\Rightarrow \quad 2.6 \pm 1.96 \left(\frac{0.3}{\sqrt{36}}\right)$$

$$\Rightarrow \quad 2.6 \pm 0.093)$$

$$\Rightarrow \quad (2.6 - 0.093 \ , \ 2.6 + 0.093)$$

$$\Rightarrow \quad (2.507 \ , \ 2.693)$$

We are 95% confident that the true mean zinc concentration is between 2.507 g/ml and 2.693 g/ml.

Example: 1 - Solution

(c) Find the 99% confidence interval for the mean zinc concentration in the river.

<u>99% CI for μ</u>

$$\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \qquad\qquad Z_{\frac{0.01}{2}} = Z_{0.005} = 2.575$$

$$\Rightarrow \quad 2.6 \pm 2.575 \frac{0.3}{\sqrt{36}}$$

$$\Rightarrow \quad (2.471\ ,\ 2.729)$$

We are 99% confident that the true mean zinc concentration is between 2.471 g/ml and 2.729 g/ml.

Example: 1 - Solutions

(d) Suppose that the researchers would like to have the margin of error 0.05 grams per milliliter. Calculate the sample size they need to take, if they use 95% confidence level.

$$\bar{x} = 2.6, \quad \sigma = 0.3, \quad n = ?$$

$(1-\alpha)100\%$ CI for μ is $\quad \Rightarrow \quad \bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

margin of error (m.e.) $= Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$

$$\Rightarrow \quad m.e. = 0.05 = Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \qquad ; Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96$$

$$0.05 = (1.96)\frac{0.3}{\sqrt{n}}$$

$$n = \left((1.96)\frac{(0.3)}{0.05}\right)^2$$

$$n = 138.29 \quad \Rightarrow \mathbf{n \approx 139} \quad \text{(round up the number)}$$

Construct a Confidence Interval for Population mean μ

$(1-\alpha)100\%$ CI for μ is

$$\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$\bar{x}$ is the point estimate for μ and $Z_{\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}}$ is the margin of error

- The exact standard error of the sample mean is $\frac{\sigma}{\sqrt{n}}$
- In practice we don't know the population standard deviation σ
- In practice, we estimate σ by the sample standard deviation s
- substituting s for σ to get standard error $\frac{s}{\sqrt{n}}$, introduces extra error
- To account for this increased error, we replace the Z-score by slightly larger score, the t-score

Construct a Confidence Interval for Population mean μ

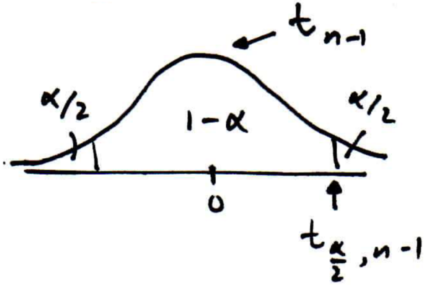
- for particular $\bar{x}$, $\Rightarrow$ $Z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$
- if σ unknown, $\dfrac{s}{\sqrt{n}}$ is used to estimate $\dfrac{\sigma}{\sqrt{n}}$
- The distribution of t is called the student's t-distribution

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t_{n-1}$$

where t_{n-1} denotes the t-distribution with $n - 1$ degrees of freedom.

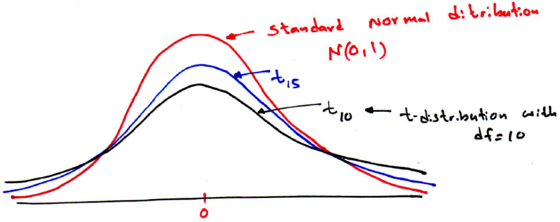
$(1-\alpha)100\%$ CI for μ (σ unknown)

$$\bar{x} \pm t_{\frac{\alpha}{2},\, n-1} \frac{s}{\sqrt{n}}$$



t - Distribution

- The $t-$distribution is bell shaped and symmetric about 0
- Probabilities depend on the degrees of freedom, $df = n - 1$
- the t_{n-1} has longer tails than the normal
- as $n \to \infty$, $t_{n-1} \to Z \sim N(0, 1)$
- $n \geq 30$, you can replace t_{n-1} with Z



- $t-$distribution has thicker tails and is more spread out than the standard normal distribution

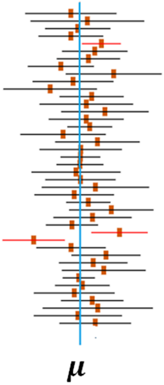
Conditions for Constructiong Confidence Intervals for μ

- Need a random sample from the population
- Normal Condition
  - A basic assumption of the CI using $t-$distribution is that the population distribution is normal
  - If the random sample is large, the population distribution need not to be normal because of the CLT

Interpreting a Confidence Interval for μ

Interpreting a CI for μ (for 95%)

If we repeatedly obtain samples of size n and construct the corresponding 95% confidence intervals for μ, on average, 95% of these intervals will include the value of μ (explain this using an applet)



Confidence Intervals for the mean - Applet

Simulating Confidence Intervals for the mean

Following applet shows the meaning of a confidence interval, calculating confidence intervals of the means of repeated samples.

`https://www.zoology.ubc.ca/~whitlock/Kingfisher/CIMean.htm`

Learning objectives:

- Understand the meaning of a confidence interval.
- Predict the effects of changing sample size on confidence intervals.

Before the next class ...

Visit the course website at canvas.ubc.ca

- Review Lecture 26 and related sections in the text book
- Topic of next class: **Chapter 8: Hypothesis Testing about the Mean**