STAT 4352 - Mathematical Statistics Notes

JaimeGoB

March 1, 2021

1 Chapter 11 - Interval Estimation

Point Estimators
θ is a unknown parameter (feature of a population)

- Ex: population mean μ
- **Fixed.**

$\hat{\theta}$ is a point estimator of θ (it is a numerical value)

- Ex: sample mean $\bar{x}$
- **Varies from sample to sample.**
- No guarantee of <u>accuracy</u>
- Must be *supplemented by* Var(θ)
  Standard Error SE($\hat{\theta}$) measures how much $\hat{\theta}$ varies from sample to sample.
  small SE $\implies$ low variance thus a more reliable estimate of θ

Interval Estimators

Def: Interval Estimate
Provides a range of values that best describe the population.
Let L = L(x) be the Lower Limit
U = U(x) be the Upper Limit
Both L,U are Random Variables because they are functions of sample data.

Def: Confidence Level / Confidence Coefficient
Is the probability that the **interval estimate** will include population parameter θ.

- Sample means will follow the <u>normal probability distribution</u> for large sample sizes (n $\geq$ 30)
- For small sample forces us to use the <u>t-distribution</u> probability distribution(n $<$ 30)
- <u>A confidence level of 95%</u> implies that **95% of all samples would give an interval that includes θ, and only 5%of all samples would yield an erroneous interval.**
- <u>The most frequently used confidence levels are 90%, 95%, and 99% with corresponding Z-scores 1.645, 1.96, 2.576.</u>
- The higher the confidence level, the more strongly we believe that the value of the parameter lies within the interval.

Def: Confidence Interval
Gives plausible values for the parameter θ being estimated where degree of plausibility specified by a confidence level.

To construct an interval estimator of unknown parameter θ. We must find two statistics **L** and **U** such that:

$$P\{\mathbf{L} \leq \theta \leq \mathbf{U}\} = 1 - \alpha$$

- $P\{\mathbf{L} \leq \theta \leq \mathbf{U}\}$ **Coverage Probability**, in repeated sampling, what percent of samples or Confident Intervals capture true θ.
- 100(1- α) **Confidence Interval** - for unknown fixed parameter θ.
- L,U - **Lower and Upper Bounds** - RVs because they are functions of sample data. Vary from sample to sample.
- 1-α **Confidence Level** (Probability) estimate will include population parameter θ.
- α **Level of Significance** Percent chance Confidence Interval will not contain population parameter θ.

Def: Coverage Probability
$P\{\mathbf{L} \leq \theta \leq \mathbf{U}\}$ Gives what % of samples or Confidence Intervals capture true θ.

Ex: Coverage Probability = 95%
Will capture θ, 95% of the time.
Will NOT capture θ, 5% of the time.

Properties of Confidence Intervals

- Confidence Intervals are not unique.
- Desirable to have E[Length of CI] to be small.
- A one-sided 100(1-α) lower-confidence interval on θ: L = -∞ $\implies$ $P\{L \leq \theta\} = 1-\alpha$
- A one-sided 100(1-α) upper-confidence interval on θ: $U = \infty$ $\implies$ $P\{\theta \leq U\} = 1-\alpha$
- If L,U are both finite, then we have a two sided interval.

Correctly Interpreting Confidence Intervals
Not Correct
There is 90% probability that the true population mean is within the interval.
Correct
There is a 90% probability that any given Confidence Interval from a random sample will contain the true population mean.

Theorem 11.1: Confidence Interval on the Mean of a Normal Distribution with known Variance

Let X be normal random variable with:

Unknown mean μ

Known variance σ^2

Suppose a random sample n, $(X_1, X_2, ..., X_n)$ is taken.

A 100(1-α)% confidence interval on μ can be obtained by considering sampling distribution of the sample mean $\bar{X}$.

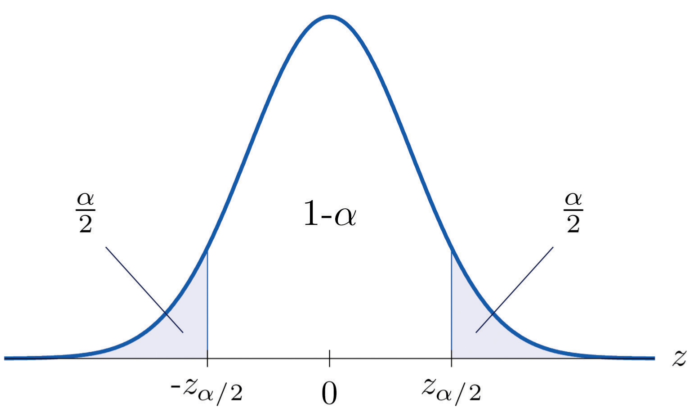
Central Limit Theorem:

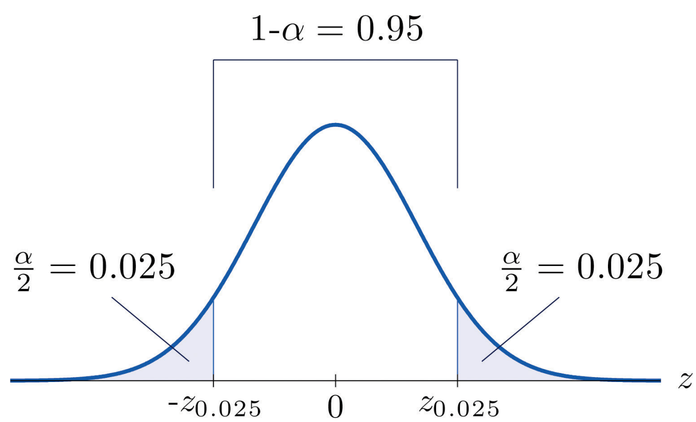
E($\bar{X}$) = μ and SD($\bar{X}$) = $\dfrac{\sigma}{\sqrt{n}}$, so $\bar{X} \sim \mathcal{N}(\mu, \dfrac{\sigma^2}{n})$ as n$\to \infty$

Let Z = Standardizing $\bar{X}$, Z will follow a Standard Normal Distribution

Let Z = $\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$

We can see from the image to the left: **Distribution of Z** and the image to the right: **Confidence Interval of 95%**





We can see that:

$$P\{-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}\} = 1 - \alpha$$

substituting Z into equation:

$$P\{-Z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq Z_{\alpha/2}\} = 1 - \alpha$$

isolating μ:

$$P\{\bar{X} - Z_{\alpha/2}(\sigma/\sqrt{n}) \leq \mu \leq \bar{X} + Z_{\alpha/2}(\sigma/\sqrt{n})\} = 1 - \alpha$$

Conclusion $\left[\bar{X} - Z_{\alpha/2}(\sigma/\sqrt{n}), \bar{X} + Z_{\alpha/2}(\sigma/\sqrt{n})\right]$ is a 100(1-α) CI for μ

How to Construct Confidence Interval Using Pivot Approach:
Suppose we have a random sample $X_1, X_2, \ldots, X_n$ from a population distribution and the parameter of interest is θ.

Given value $\alpha \in (0, 1)$. We would like to construct a 1-α Confidence Interval using a Pivot Approach:

1. **Find a variable Y, that is function of the parameter θ and data x.**
2. **The distribution of newly created variable Y is free of θ.**

<u>In many cases:</u>

$$Y = \frac{\hat{\theta} - \theta}{SE(\hat{\theta})} \text{ is a pivot and the distribution of Y is symmetric about 0.}$$

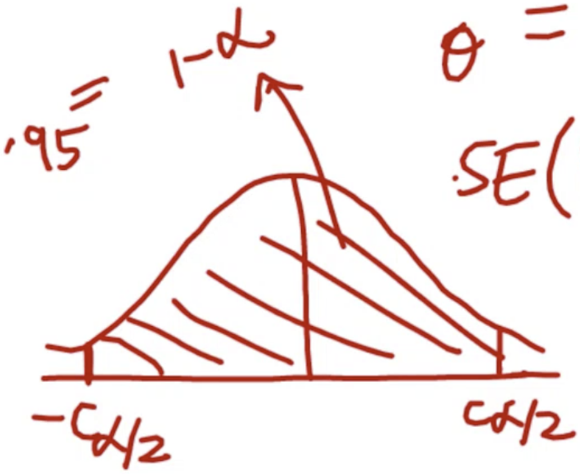
Using Pivot Approach for Two-Sided Intervals:
Find the cut/critical points(quantiles) $c_{\alpha/2}$ such that:

$$P\{-c_{\alpha/2} \leq Y \leq c_{\alpha/2}\} = 1 - \alpha$$

$c_{\alpha/2}$ is the upper (α / 2)100th percentile.

Quantiles- give you the area to the right of that point. They are the subscript of cut/critical points.

Visualizing elements from Pivot Approach:
Let μ be parameter of interest. We can construct CI using pivot approach.



$\hat{\theta} = \bar{X}$, $\theta = \mu$

$SE(\hat{\theta}) = SE(\bar{X}) = \frac{\sigma}{\sqrt{n}}$.

$P(-c_{\alpha/2} \leq Y \leq c_{\alpha/2}) = 1 - \alpha$

$c_{\alpha/2} = z_{.025}$

Symmetric Two-sided CI: Theorem
$\hat{\theta} \pm c_{\alpha/2}(SE(\hat{\theta}))$ is a $100(1-\alpha)\%$ confidence interval for θ

Proof:

$$1-\alpha = P\{-c_{\alpha/2} \leq Y \leq c_{\alpha/2}\}$$

$$= P\{-c_{\alpha/2} \leq \frac{\hat{\theta}-\theta}{SE(\hat{\theta})} \leq c_{\alpha/2}\}$$

$$= P\{\hat{\theta} - c_{\alpha/2}(SE(\hat{\theta})) \leq Y \leq \hat{\theta} + c_{\alpha/2}(SE(\hat{\theta}))\}$$

$\implies$ $\hat{\theta}$ is **within** $c_{\alpha/2}(SE(\hat{\theta}))$ of θ with **probability** 1-α

$c_{\alpha/2}(SE(\hat{\theta}))$ is known as *Margin of Error* (size of error in estimation)
Ex: In polls you might hear accurate with 0.02 (this is margin of error)

Asymmetric Two-sided CI(Non-symmetric distributions):
$[\hat{\theta} - c_{\alpha/2}(SE(\hat{\theta})), \hat{\theta} - c_{1-\alpha/2}(SE(\hat{\theta}))]$ is a $100(1-\alpha)\%$ confidence interval for θ

Proof:

$$1-\alpha = P\{c_{1-\alpha/2} \leq Y \leq c_{\alpha/2}\}$$

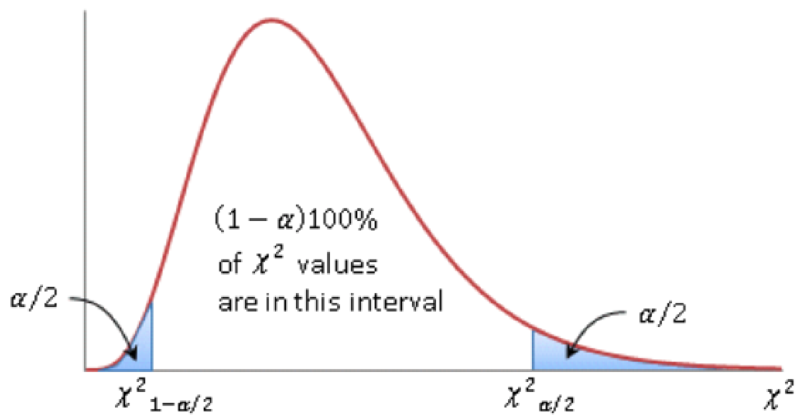
$$= P\{c_{1-\alpha/2} \leq \frac{\hat{\theta}-\theta}{SE(\hat{\theta})} \leq c_{\alpha/2}\}$$

$$= P\{\hat{\theta} - c_{\alpha/2}(SE(\hat{\theta})) \leq \theta \leq \hat{\theta} - c_{1-\alpha/2}(SE(\hat{\theta}))\}$$

Constructing Confidence Intervals about σ^2 and σ

Now that we have the basics of the distribution of the variable X^2, we can work on constructing a formula for the confidence interval.

From the distribution shape on the previous page, we know that $(1-\alpha)100\%$ of the X^2 values will be between the two critical values shown below.



This gives us the following inequality:

$$\chi^2_{1-\alpha/2} < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{\alpha/2}$$

If we solve the inequality for σ^2, we get the formula for the confidence interval:

A **(1-α)100% confidence interval for σ^2** is

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$$

Note: The sample *must* be taken from a normally distributed population.

Note #2: If a confidence interval for σ is desired, we can take the square root of each part.

Both variance and standard deviation are non-negative numbers. The Chi-Square distribution allows us to construct confidence intervals for the variance and the standard deviation (when the original population of data is normally distributed).

Since the distribution is not symmetrical(asymmetrical) you cannot just take plus,minus cut points. We can see from the image above that the **area to the right of cut off points** are:

- Lower bound critical point : 1-$\alpha/2$
- Upper bound critical point: $\alpha/2$

One-sided CI:
Upper Bound: U = $\hat{\theta} - c_{1-\alpha}(SE(\hat{\theta}))$ when L = -∞

Lower Bound: L = $\hat{\theta} - c_{\alpha}(SE(\hat{\theta}))$ when U = ∞

Proof(Upper Bound):
Coverage probability is 1 - α.
$1 - \alpha = P\{Y \geq c_{1-\alpha}\}$

$$= P\{\frac{\hat{\theta} - \theta}{SE(\hat{\theta})} \geq c_{1-\alpha}\}$$
$$= P\{\theta \leq \hat{\theta} - c_{1-\alpha}(SE(\hat{\theta}))\}$$

$\implies$ U= $\hat{\theta} - c_{1-\alpha}(SE(\hat{\theta}))$

The lower bound can be computed in the same manner.

2