STA 1013 : Statistics through Examples

Lecture 24: Interval Estimation

Hwiyoung Lee

November 4, 2019

Department of Statistics, Florida State University

Point Estimation

Estimating the population parameter

We want to know the population parameter

- In many instances, one cannot study an entire population.
- Main reasons are cost, time and effort involved in studying the entire population.

How can we estimate the population parameter ?

- Do Sampling !!!
- Use the sample Statistic to estimate the population Parameter

Examples

- Consider a manufacturing assembly line that produces thousands or millions of items of a product. To determine the quality of this product, is it necessary to inspect each item of the product?

- One finds 3 percent of the items in the sample as defective, the **conclusion is made that 3 percent of the items in the population is defective.**

Examples

- Goodyear tire manufacturer wants to know the mean life of its new brand of tires. One way is testing and wearing out each tire manufactured. Obviously, this does not make sense.

- Goodyear takes a sample of tires (1,000), tests and wears out each of these tires and then calculates the average life of the sampled tires.

- Suppose, the mean life is calculated as 42,000 miles. Based on this sample, **it is concluded that the mean life all new brand of tires (that is population) is 42,000 miles.**

Point Estimation

- In the previous two examples, we dealt with Mean (or average) and Proportion.
- We used the sample statistic to estimate population characteristics (mean, and proportion).
  1. Sample proportion (sample defective rate) $\rightarrow$ the population defective rate
  2. Sample mean $\rightarrow$ the population mean
- Using sample statistic (**single value**) to estimate a population mean or proportion is known as **Point Estimation**

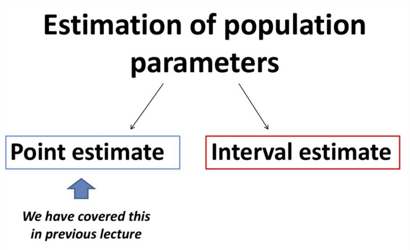
Examples of Point Estimations

We use the data from the sample to compute a single value of a sample statistic that serves as an estimate of a population parameter

	point estimator	Parameter
Mean	$\bar{x}$	μ
Proportion	p	π
Standard Deviation	s	σ

-

Two types of Estimations



- The Point estimation (ex : $\bar{X}$) is a single approximated value for true population parameter
- Interval estimate is a range of numbers around the point estimate within which the parameter is believed to fall

Interval Estimation

Interval Estimation

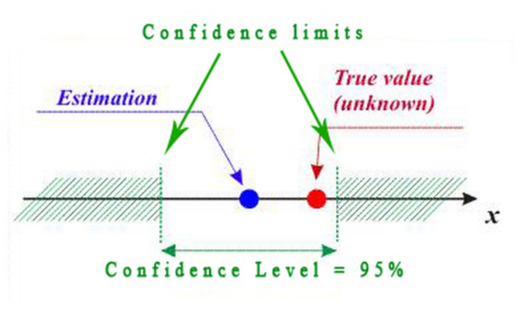
Suppose, someone asks you how long does it take to get from FSU to City Hall. What would be a more reliable estimate

1. 30 minutes
2. Between 25 and 40 minutes

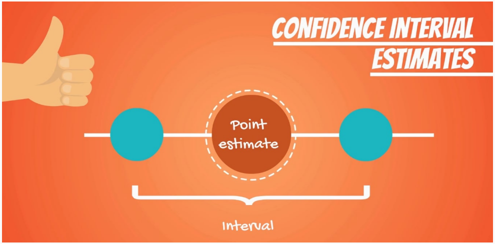
- If you use 30 minutes as estimate, you are using a point estimate. On the other hand, if you use between 25 and 40 minutes as estimate, you are using an interval estimation.

Interval Estimation

- To increase the level of confidence in estimation, we use a **range of values** (rather than a single value) as the estimate of a population parameter.



Interval Estimation



- The point estimation and the interval estimation are closely related.
- In fact, the point estimate is located exactly in the middle of the confidence interval.
- **However, confidence intervals provide much more information and are preferred when making inferences.**

Interval Estimation

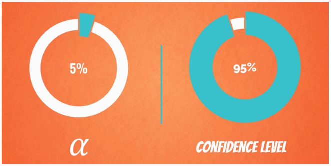
- Interval estimation uses a range of values.
- **The width of the range indicates the level of confidence.**
  1. **The Narrower the range, lower the confidence.**
  2. **The Wider the range, higher the confidence.**
- **Trade off** between confidence level and width of range
  - Want narrow range → but Low confidence
    : 3% sure average weight of male : (179lbs, 181lbs)
  - Want high confidence → but Wide range
    : 99.999% sure average weight of male : (100lbs, 250lbs)
- Confidence level has to be pre specified

The Level of Confidence $(1-\alpha) \times 100\%$

- $(1-\alpha)\times 100$ % is called **the confidence level of the interval.**
- α is a value between 0 and 1

The Level of Confidence

For example, if we want to be 95% confident that the parameter is inside the interval, α is 5%.



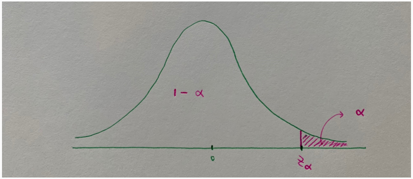
If we want a higher confidence level of, say, 99%, α will be 1%.

Confidence Interval for population mean μ

Recall : Critical value & Central Limit Theorem

Critical value $\mathbf{z}_\alpha$: the z score with an area of α to its right.

$$P(Z \geqslant z_\alpha) = \alpha$$



CLT : Suppose X follows a distribution with mean μ, standard deviation σ, then for large n, the distribution of the sample mean $\bar{X}$ follows

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Confidence Interval for population mean μ

Once a confidence level is specified, the interval estimate can be calculated using formulas;

- $(1-\alpha)\%$ Confidence interval estimate of μ

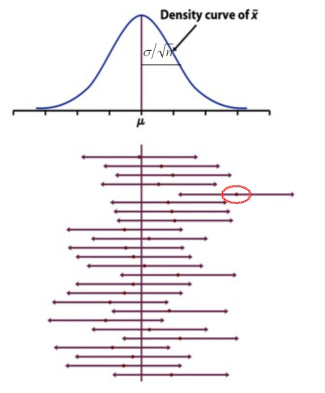
$$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

- which gives us the interval :

$$\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\ ,\ \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$$

- **Interpretation** : **We are $(1-\alpha)\times 100\%$ confident that the interval includes the population mean μ**

Meaning of the Confidence Interval



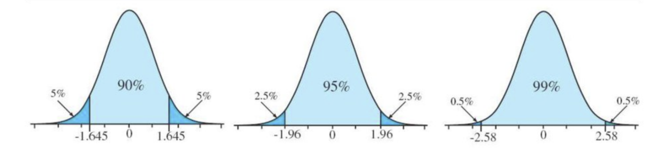
95% Confidence Interval

- If we repeat experiments like we just did (constructing the Interval) 100 times
- Then approximately 95 intervals out of 100 created intervals will contain the true unknown population mean μ
- There is a 1-in-20 chance (5%) that our Confidence Interval does NOT include true mean.

Commonly used confidence levels

Values of $z_{\alpha/2}$ that most commonly used for Confidence Interval

Confidence Level	α	$\alpha/2$	$z_{\alpha/2}$
90%	.10	.05	1.645
95%	.05	.025	1.960
99%	.01	.005	2.576



Confidence Intervals for population mean μ

- 90% Confidence Interval :

$$\left(\bar{X} \pm z_{0.1/2}\frac{\sigma}{\sqrt{n}}\right) = \left(\bar{X} - 1.645\frac{\sigma}{\sqrt{n}}\ ,\ \bar{X} + 1.645\frac{\sigma}{\sqrt{n}}\right)$$

- 95% Confidence Interval :

$$\left(\bar{X} \pm z_{0.05/2}\frac{\sigma}{\sqrt{n}}\right) = \left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}\ ,\ \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right)$$

- 99% Confidence Interval :

$$\left(\bar{X} \pm z_{0.01/2}\frac{\sigma}{\sqrt{n}}\right) = \left(\bar{X} - 2.576\frac{\sigma}{\sqrt{n}}\ ,\ \bar{X} + 2.576\frac{\sigma}{\sqrt{n}}\right)$$

Examples

A simple random sample of 50 items from a population with $\sigma = 6$ resulted in a sample mean of 32.

- Provide a 90% confidence interval for the population mean.

- Provide a 95% confidence interval for the population mean.

- Provide a 99% confidence interval for the population mean.

Examples

A simple random sample of 60 items resulted in a sample mean of 80. The population standard deviation is $\sigma = 15$

1. Compute the 95% confidence interval for the population mean.

2. Assume that the same sample mean was obtained from a sample of 120 items. Provide a 95% confidence interval for the population mean. Interpret the confidence interval.

Examples

The undergraduate grade point average (GPA) for students admitted to the top graduate business schools was 3.37. Assume this estimate was based on a sample of 120 students admitted to the top schools. Using past years' data, the population standard deviation can be assumed known with $\sigma = 0.28$.

- What is the 95% confidence interval estimate of the mean undergraduate GPA for students admitted to the top graduate business schools?