PRINCIPLES OF POINT ESTIMATION

D. Uma

Computer Science and Engineering

STATISTICS FOR DATA SCIENCE

Point Estimation

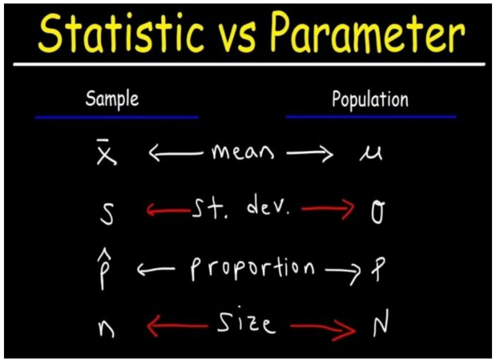
D. Uma

Computer Science and Engineering

Topics to be covered...

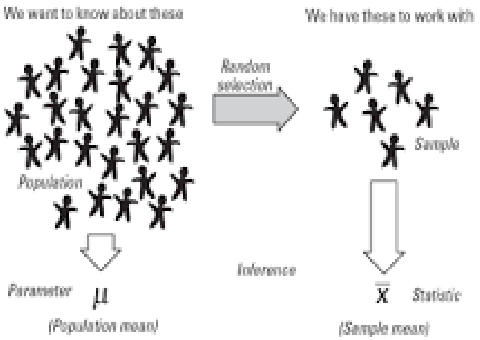
- ✓ Recap of Statistic and Parameter
- ✓ Point Estimator
- ✓ Measuring Goodness of an Estimator

Descriptive vs. Inferential Statistics



Descriptive Statistics: **To summarize the data.**

Inferential Statistics: **To make conclusions about population parameter.**



Statistic is associated with a **sample.**
Parameter is associated with a **population.**

Statistic is used **to estimate** the **value** of the **parameter.**

Source Images: reddit.com, youtube.com

$\hat{p}$

Sample Statistics & Population Parameters

Point Estimate & Point Estimator

Let **X ~ Bin(n, p)** where **p is not known** and let X = 9 and n=20

Then p_hat = 9/20 (**point Estimate**)

p_hat = X/n (**point Estimator**)

Point Estimate

$$\mu = \overline{x} \ \pm$$

Average marks(μ) of all students of a particular degree of a city.

x_bar = 7895/100=78.95 (**point Estimate**)

x_bar = ∑x/n (**point Estimator**)

Source Image: CQE Academy

point estimator

- How to determine goodness of point estimator?
- What are the methods available to construct a good point estimators?

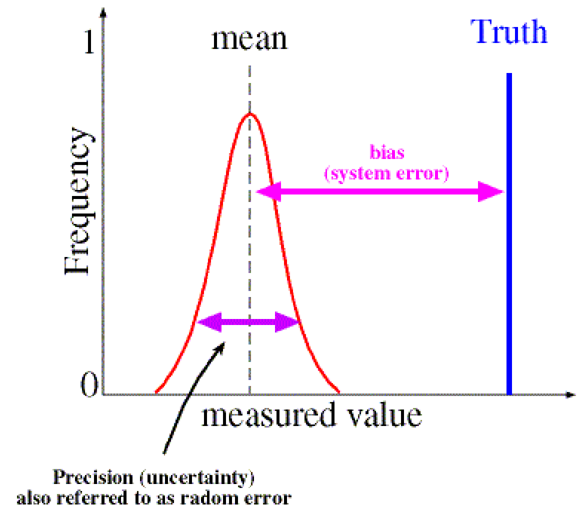
Measuring goodness of an Estimator

An **estimator** must be both **accurate**(measured by **bias**) and **precise**(measured by **uncertainty**).

Mean Squared Error(**MSE**) : quantity used to evaluate overall goodness of an estimator.

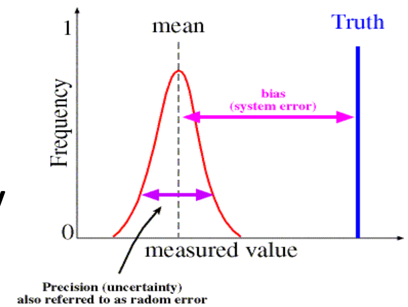
MSE = Bias^2 + Uncertainty^2

MSE = Bias^2 + Variance



Source: apollo.lsc.vsc.edu

Measuring goodness of an Estimator



An **estimator** must be both **accurate**(measured by **bias**) and **precise**(measured by **uncertainty**).

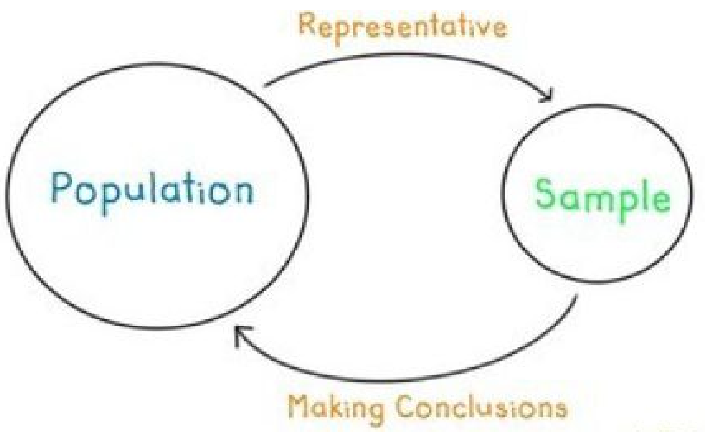
$Bias = \mu_{\hat{p}} - p \Rightarrow 0$

Mean Squared Error(**MSE**) : quantity used to evaluate overall goodness of an estimator.

$$\textbf{MSE} = \textbf{Bias}^2 + \textbf{Uncertainty}^2$$

Source: apollo.lsc.vsc.edu, webpages.uidaho.edu

Measuring goodness of an Estimator



- The sample may not provide an complete depiction of the population.
- There will always be an uncertainty when drawing conclusions about the population from the sample.

DATA SCIENCE

Example – Weather Forecasting

The forecasting application in smart phones, it shows the probability of spell expected tomorrow is 80%.

There are some uncertain situations where the prediction may fail also (standard errors).

Source Images: earth.com, demand-planning.com

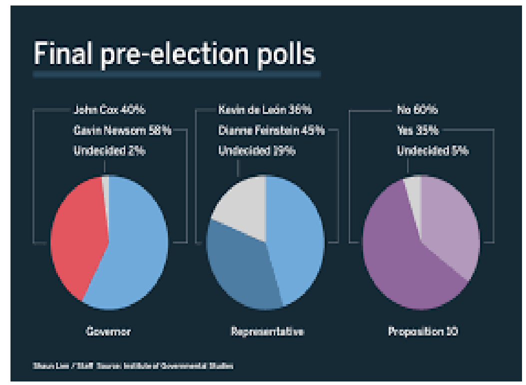
DATA SCIENCE

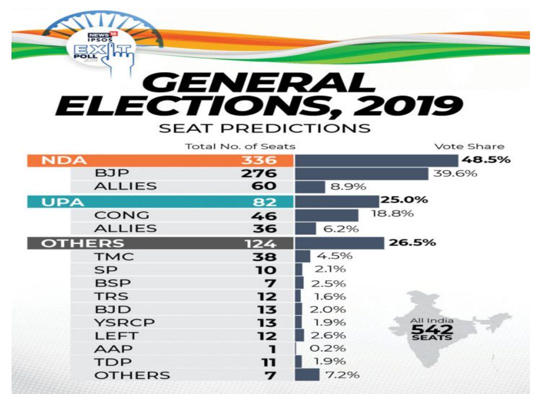
Example – Election Poll

Samples are selected and surveyed to which candidate they would vote to predict the results of the election.

Worst case, the some polled may change their decision at the time of voting in elections.

How to make the process a good guess by reducing the uncertainty?

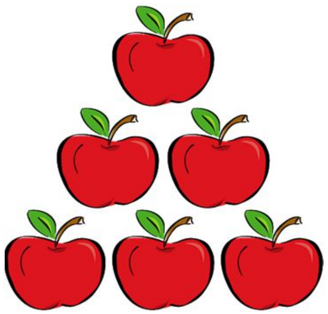




Source Images: dailycal.org, News18..com

Understanding Point Estimation

- Anna is interested in finding the mean weight of the apples that are imported from Kashmir. However a **survey claims** that the average **weight of an apple** is **around 90g**.



What can we do now?

- It is well know that we cannot weigh every apple (population). So, we need to take samples.

- From those samples we can **make inferences** about the entire population.

- There are also chances that the samples what we examine will have **some errors**.

Image Source: http://clipart-library.com/

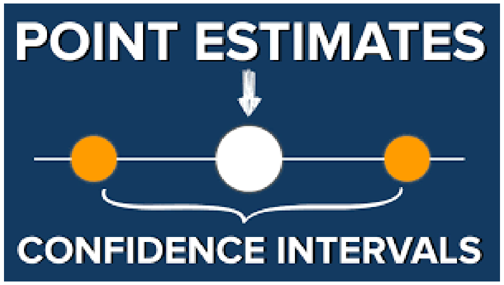
- Consider four samples of size 20 each.

$$Sample\,1: \ \bar{x} = 88g \qquad Sample\,2: \ \bar{x} = 89.5g$$

$$Sample\,3: \ \bar{x} = 91g \qquad Sample\,4: \ \bar{x} = 89.9g$$

What is Point Estimate?

- There is a claim that the average weight of an apple is around **90g**.
- From samples we understand that Sample 4 having 89.9g is pretty close to 90g.
- On the other hand, **89.9g** is close enough to **90g** and it is plausible to accept. This is referred as **Point Estimate.**



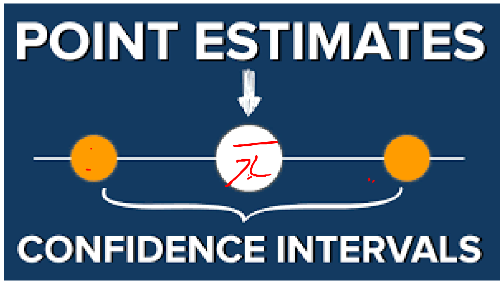
Source Image: 365datascience.com

Point Estimate

- It is a **single numeric value** specified for the data(sample) which is also referred as **sample statistic.**

- It is used to estimate an **unknown constant**, or **parameter**, is called a **point estimator.**

- Point estimate infers about the **population parameters.**



Source Image: 365datascience.com

Inferences from Point Estimator

Point Estimators	
Sample Mean	($\bar{X}$)
Sample Standard Deviation	(s)
Sample Proportion	($\hat{p}$)

Inference $\rightarrow$

Unknown Population Parameters	
Population Mean	(μ)
Population Standard Deviation	(σ)
Population Proportion	($\cancel{\hat{p}}$)

p

Point estimate is used to make an estimation of unknown population parameters including population mean, standard deviation and proportion .

Example – Point Estimation of Population Mean

- Point Estimate of Population mean μ is Sample mean $\bar{X}$.

- **Example:** Sample of heights of 34 male students in a class was obtained.

185 161 174 175 202 178 202 139 177
170 151 176 197 214 283 184 189 168
188 170 207 180 167 177 166 231 176
184 179 155 148 180 194 176

$$\bar{X} = 182.44$$

- This can be inferred as the single numeric point estimate for the population mean (true mean) of all the male students of that class.

Example – Point Estimation of Population Proportion

Example: A sample of 100 people were selected in a particular locality to estimate the proportion of them who go for walking in the park everyday. In this sample 40 of them go for walking everyday.

$$\hat{p} = \frac{X}{n}$$

$$\hat{p} = \frac{40}{100} = 0.4$$

- For a large population, when sampling technique is used, it is not going to be perfect always. There will always be some **uncertainty** in **estimation**.

Property 1 : Bias

- When the **expected value of an estimator** is different from the **value of the parameter** that is being estimated.

- When they are equal, we call it as unbiased.

Property 2 : Consistency

- This portrays **how close** the **point estimator** can be to the **true value of the parameter** even if it increases in size.

- The consistency and accuracy of point estimator can be achieved by using large samples.

- This is can be exercised by **mean** and the **variance.**

- To be more consistent the mean of the sample should move towards the true value of the population parameter.

Property 3 : Efficiency

- A very **efficient point estimator** should have the following,

  a) **smallest variance**.
  b) **unbiased observation.**
  c) **consistent**.

- All these parameters can be achieved from a **normally distributed population.**

How can we measure the goodness of an estimator?

- Given a point estimator, how do we determine how good it is?

  Goodness Measure - **Mean Squared Error(MSE)**

- What methods can be used to construct good point estimators?

  Good Method to construct Point Estimator - **Maximum Likelihood Estimate (MLE)**

Measuring Goodness of Estimator – Mean Squared Error (MSE)

- A good estimator should be both accurate and precise.
- Accuracy of an estimator is measured by bias.
- Precision is measured by standard deviation or uncertainty.
- It can be measured by a quantity called Mean Squared Error (MSE).
- MSE combines both bias and uncertainty.

Mean Squared Error (MSE) - Bias & Uncertainty

- The **bias** of the estimator is denoted by,

$$\hat{\theta} = \mu_{\hat{\theta}} - \theta$$

- The difference between the mean of the estimator and true value.

- θ is the unknown parameter

- $\hat{\theta}$ denote the estimator of θ

- The **uncertainty** is the standard deviation $\sigma_{\hat{\theta}}$, defined as the standard error of the estimator.

$\theta \Rightarrow$ parameter (of a population)

$\hat{\theta} \Rightarrow$ Estimator

Bias of $\hat{\theta} = \mu_{\hat{\theta}} - \theta$

$\hat{\theta} = ?$

$\hat{p} = \frac{X}{n}$

Bias of $\hat{p} = \mu_{\hat{p}} - p$

$\hat{p}$

unknown $= \mu_{\frac{X}{n}} - p$

$X \sim Bin(n, p)$

$0 \le p - p$

$= \frac{1}{n}\mu_X - p$

$= \frac{1}{n}(np) - p$

Mean Squared Error (MSE)

- **MSE** is found by adding the variance to square of the bias.

$$MSE_{\hat{\theta}} = Var(\hat{\theta}) + (Bias\ of\ \hat{\theta})^2$$

- By definition,

Let θ be a parameter, and $\hat{\theta}$ be the estimator of θ.

The mean squared error (MSE) of $\hat{\theta}$ is

$$MSE_{\hat{\theta}} = (\mu_{\hat{\theta}} - \theta)^2 + \sigma_{\hat{\theta}}{}^2$$

Bias

- An equivalent expression is,

$$MSE_{\hat{\theta}} = \mu_{(\hat{\theta} - \theta)^2}$$

Note: $(\hat{\theta} - \theta)^2$ is the difference between estimated value and true value and it is called as error.

Problem

Let $X \sim Bin(n, p)$ where p is unknown. Find the MSE of $\hat{p} = X/n$

Solution: As we learnt in the Sample Proportion of Binomial Distribution,

bias $= \mu_{\hat{p}} - p$

We know that, $\hat{p} = X/n$

$$\mu_{\hat{p}} = \mu_{X/n} = \frac{\mu_X}{n} = \frac{np}{n} = p$$

Since, $\mu_{\hat{p}} = p$, $\hat{p}$ is unbiased,

bias $= 0$.

The uncertainty is the standard deviation $\sigma_{\hat{p}}$.

The standard deviation of X is $\sigma_X = \sqrt{np(1-p)}$.

We know that, $\hat{p} = X/n$

$$\sigma_{\hat{p}} = \sigma_{X/n} = \frac{\sigma_X}{n} = \frac{\sqrt{np(1-p)}}{n} = \sqrt{\frac{p(1-p)}{n}}$$

$$\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}$$

$$MSE_{\hat{\theta}} = (Bias\ of\ \hat{\theta})^2 + Var(\hat{\theta}) = 0 + p(1-p)/n$$

$$MSE_{\hat{\theta}} = p(1-p)/n$$

Note: When bias is 0, MSE will be equal to variance.

THANK YOU

D. Uma

Computer Science and Engineering

umaprabha@pes.edu

+91 99 7251 5335