

Lecture 9: Estimation

Lecturer: Jie Fu

EEL 3850

Motivating problem

- Estimating the average height of 3rd grade students in schools

Data sample 1:	53.49	51.59	53.94	56.57	51.30	51.30	Average = 51.51 inches
	56.74	54.30	50.59	53.63	50.61	50.60	
	52.73	46.26	46.83	50.31	48.96	52.94	
	49.28	47.76	56.40	51.32	52.20	47.73	
Data sample 2:	52.33	48.55	53.13	50.20	51.12	50.19	Average = 52.56 inches
	57.56	51.96	48.83	54.47	48.34	52.63	
	46.12	48.02	52.59	54.22	52.51	51.65	
	51.10	47.56	49.84	50.62	55.17	53.03	

Classical inference

- unknown parameter θ as a **deterministic** (not random!) but unknown quantity.
 - Average height.

- The observation X is random and its distribution
 - $p_X(x; \theta)$ if X is discrete
 - $f_X(x; \theta)$ if X is continuous
 - **depends on** the value **of** θ .

Classical inference

$$\theta \rightarrow f_X(x; \theta) \xrightarrow{x_1, x_2, \dots, x_n} \text{Estimator} \rightarrow \hat{\theta}$$

- Given observations $X = (X_1, \dots, X_n)$, an estimator $\hat{\theta} = g(X)$ is function of X .
- Thus, $\hat{\theta}$ is a _____.
- Let n be the number of observations, the mean and variance of $\hat{\theta}_n$ are denoted $E_{\theta}[\hat{\theta}_n]$ and $var_{\theta}[\hat{\theta}_n]$, respectively.

Terminology regarding estimators

- The underlying parameter θ to be estimated is a **constant**.
- **Estimation error**: $\tilde{\Theta}_n = \hat{\Theta}_n - \theta$.
- **Bias** of the estimator: $b_\theta(\hat{\Theta}_n) = E_\theta[\hat{\Theta}_n] - \theta$, is the expected value of the estimation error.

Estimation of the Mean

- Suppose that the observations X_1, \dots, X_n are i.i.d., with an **unknown** common mean μ_X .
- $\hat{\mu}_X = \frac{1}{n} \sum_{i=1}^n X_i$ is **unbiased estimator**
 - for any n , the expected value of the average is equal to the true mean.

```
heights= np.array([121.92, 132.64, 113.31, 97.20, 140.94, 139.04, 115.98, 128.27, 121.84, 97.73])
```

The average height estimate

```
average = np.mean(heights)= 122.185
```

Properties of the Estimator of the mean

- $\hat{\mu}_X = \frac{1}{n} \sum_{i=1}^n X_i$ is unbiased estimator
 - for any n , the expected value of the average is equal to the true mean.

Estimating the variance

Let σ_X^2 denote the variance of the random variables. Then there are two cases that should be considered for estimating the variance.

Known mean: If the mean of the random variables, μ_X , is known. Let the sample-variance estimator for this case be defined by

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \mu_X)^2.$$

Estimating the variance

- Let's first determine if the sample variance estimator is biased when the true mean is known: (*experiment validate*)

Estimating the variance

Known mean: If the mean of the random variables, μ_X , is known. Let the sample-variance estimator for this case be defined by

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \mu_X)^2.$$

Unknown mean: it is natural to replace μ_X with its sample estimate $\hat{\mu}_X$:

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \hat{\mu}_X)^2.$$

See the ADDITIONAL PDF for a proof why the variance estimator above is biased.

Estimating the variance

- Let's first determine if the sample variance estimator is biased when we replace the true mean with its sample estimate: (*experiment validate*)

Known mean: If the mean of the random variables, μ_X , is known. Let the sample-variance estimator for this case be defined by

$$\hat{\sigma}_X^2 = \frac{1}{N} \sum_{i=1}^N (X_i - \mu_X)^2.$$

Unknown mean: unbiased estimator:

$$\hat{\sigma}_X^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \hat{\mu}_X)^2.$$

The change in denominator is often referred to as a *degrees-of-freedom (dof) correction*.

Example

- Suppose we have a sample of student scores from an exam, and we want to estimate the population mean score.
- Sample data: `array([84, 78, 71, 74, 60, 76, 50, 86, 67, 82, 89, 93, 79, 72, 78, 76, 71, 85, 86, 52, 61, 63, 92, 71, 80, 60, 76, 81, 57, 88])`
- Total 30 samples
- Using the same sample data, we want to estimate the population variance.

Point estimate vs interval estimate

- Instead of estimating a single value, an interval estimate is also used:
- For an unknown parameter

$$\theta \rightarrow f_X(x; \theta) \xrightarrow{x_1, x_2, \dots, x_n} \text{Interval Estimator} \rightarrow [\hat{\theta}^-, \hat{\theta}^+]$$

An interval contain the unknown parameter with high probability

Confidence intervals (CIs)

- The value of an estimator may not be informative enough
- Let us first fix a desired confidence level, $1 - \alpha$, where α is typically a small number.
- We then replace the point estimator $\hat{\theta}_n$ by a lower estimator $\hat{\theta}_n^-$ and an upper estimator $\hat{\theta}_n^+$, s.t.

$$P(\hat{\theta}_n^- \leq \theta \leq \hat{\theta}_n^+) \geq 1 - \alpha$$

for every possible value of θ .

- We call $[\hat{\theta}_n^-, \hat{\theta}_n^+]$ a $(1 - \alpha)$ confidence interval.

Confidence intervals (CIs)

- $\hat{\mu}_X = \frac{1}{n} \sum_{i=1}^n X_i$
- Recall: the observations X_1, \dots, X_n are i.i.d., with an **unknown** common mean μ_X

$$\hat{\mu}_X \sim \mathcal{N}\left(\mu_X, \frac{\sigma^2}{n}\right)$$

- Recall CLT:

We call $[\hat{\mu}_X^-, \hat{\mu}_X^+]$ a $(1 - \alpha)$ confidence interval if

$$P(\hat{\mu}_X^- \leq \mu_X \leq \hat{\mu}_X^+) > 1 - \alpha$$

Confidence intervals (CIs)

- Suppose $\alpha = 0.05$
- Let's compute the 95% confidence interval about the mean of unknown RV using the samples.

<https://www.mathsisfun.com/data/standard-normal-distribution-table.html>

Confidence interval for mean estimate with unknown variance

- Recall if the variance is unknown, we have an unbiased estimate for the variance

Unknown mean: unbiased estimator:

$$\hat{\sigma}_X^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \hat{\mu}_X)^2.$$

$$\frac{\hat{\mu}_X - \mu_X}{\sigma_X / \sqrt{n}}$$

has a Student's t -distribution with $\nu = n - 1$ degrees of freedom (dof) t_ν .
(Like a Gaussian, but more spread out!)

1. Point Estimation for Mean with prior knowledge of the population variance

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

2. Standard Error of the Mean with known population variance (not an estimate but can be computed based on the property of variance.)

$$SE = \frac{\sigma}{\sqrt{n}}$$

3. Point Estimation for Mean without prior knowledge of the population variance

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i$$

4. Point Estimation for Variance without prior knowledge of the population variance

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

- Formula to compute confidence interval

For a population mean μ when the sample size is large ($n \geq 30$), the confidence interval is given by:

$$\left(\hat{\mu}_X - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad \hat{\mu}_X + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right) \quad (1)$$

where:

- $\hat{\mu}_X$ = sample mean
- σ = population standard deviation (or sample standard deviation if unknown)
- n = sample size
- $z_{\alpha/2}$ = critical value from the standard normal table for a given confidence level