MATH 450: Mathematical statistics

September 19th, 2019

Lecture 8: Method of moments

Topics

Week 2	*Chapter 6: Statistics and Sampling Distributions*
Week 4	Chapter 7: Point Estimation
Week 6	*Chapter 8: Confidence Intervals*
Week 9	*Chapter 9: Test of Hypothesis*
Week 11	Chapter 10: Two-sample inference
Week 12	Regression

Chapter 7: Overview

7.1 Point estimate

- unbiased estimator
- mean squared error

7.2 Methods of point estimation

- method of moments
- method of maximum likelihood.

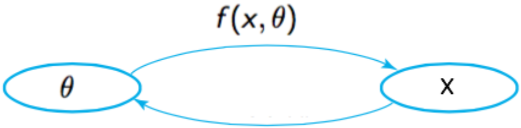
7.3 Sufficient statistic

7.4 Information and Efficiency

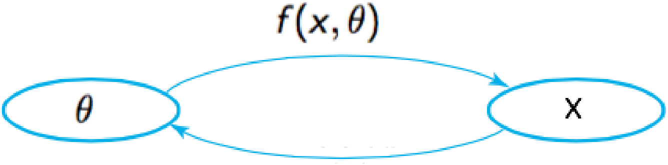
- Large sample properties of the maximum likelihood estimator

Question of this chapter

- Given a random sample $X_1, \ldots, X_n$ from a distribution with pmf/pdf $f(x, \theta)$ parameterized by a parameter θ
- Goal: Estimate θ



Point estimate



Definition

A point estimate $\hat{\theta}$ of a parameter θ is a single number that can be regarded as a sensible value for θ.

$$\begin{array}{ccccc} \text{population parameter} & \Longrightarrow & \textit{sample} & \Longrightarrow & \textit{estimate} \\ \theta & \Longrightarrow & X_1, X_2, \ldots, X_n & \Longrightarrow & \hat{\theta} \end{array}$$

Mean Squared Error

- Measuring error of estimation

$$|\hat{\theta} - \theta| \qquad \text{or} \qquad (\hat{\theta} - \theta)^2$$

- The error of estimation is random

Definition

The mean squared error of an estimator $\hat{\theta}$ is

$$E[(\hat{\theta} - \theta)^2]$$

Bias-variance decomposition

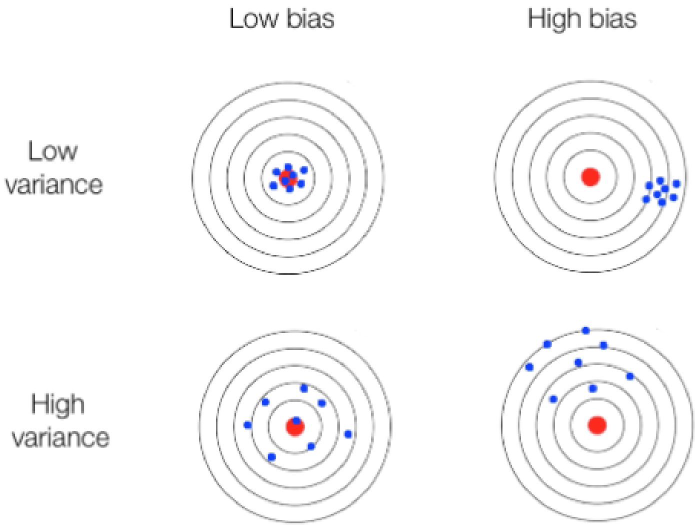
Theorem

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2] = V(\hat{\theta}) + \left(E(\hat{\theta}) - \theta\right)^2$$

Bias-variance decomposition

Mean squared error $=$ variance of estimator $+$ $(bias)^2$

Bias-variance decomposition



Unbiased estimators

Definition

A point estimator $\hat{\theta}$ is said to be an unbiased estimator of θ if

$$E(\hat{\theta}) = \theta$$

for every possible value of θ.

Unbiased estimator

$\Leftrightarrow$ Bias $= 0$

$\Leftrightarrow$ Mean squared error $=$ variance of estimator

Example: sample proportion

Problem

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a Bernoulli distribution with probability of success p

x	0	1
p(x)	1-p	p

Assume that we estimate p by using the sample mean

$$\bar{X} = \frac{X_1 + X_2 + \ldots X_n}{n}$$

What are the bias and the variance of this estimator?

Example: sample proportion

Problem

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a Bernoulli distribution with probability of success p

x	0	1
p(x)	1-p	p

Assume that we estimate p by using the sample mean

$$\bar{X} = \frac{X_1 + X_2 + \ldots X_n}{n}$$

Compute the MSE of this estimator.

Example: sample proportion

Problem

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a Bernoulli distribution with probability of success p

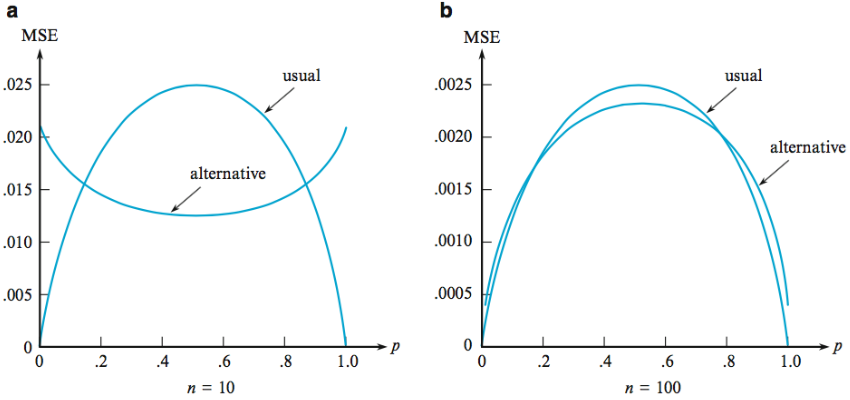
x	*0*	*1*
$p(x)$	*1-p*	p

Assume that we estimate p by using

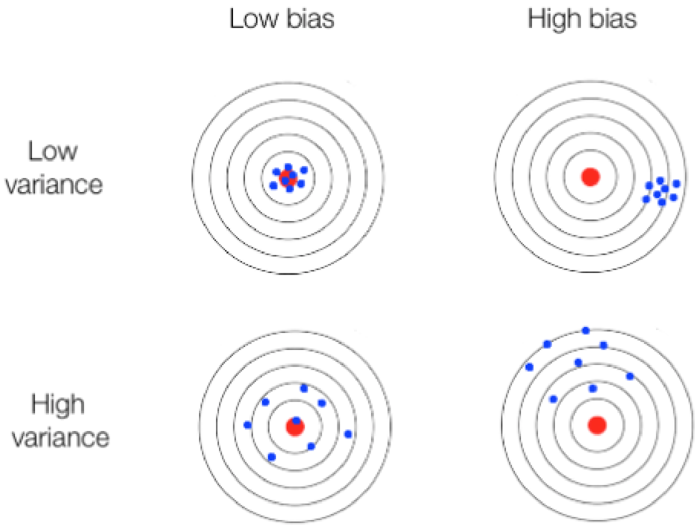
$$\tilde{p} = \frac{X_1 + X_2 + \ldots + X_n + 2}{n + 4}$$

Compute the MSE of this estimator.

Example 7.1 and 7.4



Bias-variance decomposition



Minimum variance unbiased estimator (MVUE)

Definition

Among all estimators of θ that are unbiased, choose the one that has minimum variance. The resulting $\hat{\theta}$ is called the minimum variance unbiased estimator (MVUE) of θ.

Recall:

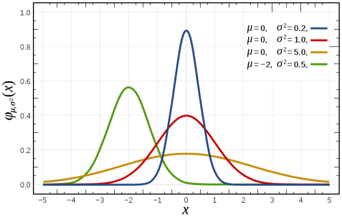
- Mean squared error $=$ variance of estimator $+$ $(bias)^2$
- unbiased estimator $\Rightarrow$ bias $=0$

$\Rightarrow$ MVUE has minimum mean squared error among unbiased estimators

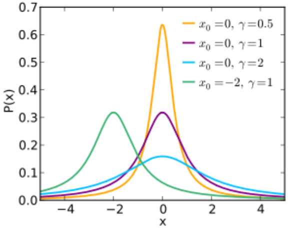
What is the best estimator of the mean?

Question: Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean μ. What is the best estimator of μ?

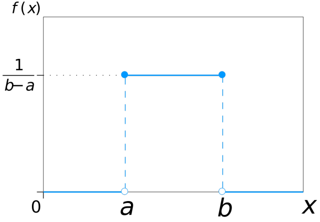
Example 7.8



Normal

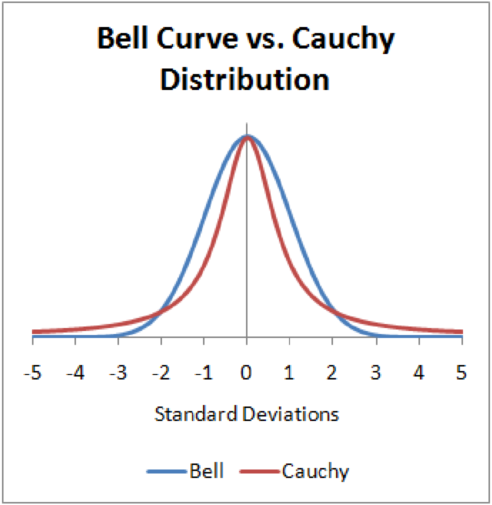


Cauchy



Uniform

Normal vs. Cauchy



What is the best estimator of the mean?

Question: Let $X_1, \ldots, X_n$ be a random sample from a distribution with mean μ. What is the best estimator of μ?

Answer: It depends.

- Normal distribution $\rightarrow$ sample mean $\bar{X}$
- Cauchy distribution $\rightarrow$ sample median $\tilde{X}$
- Uniform distribution $\rightarrow$

$$\hat{X}_e = \frac{\text{largest number} + \text{smaller number}}{2}$$

- In all cases, 10% trimmed mean performs pretty well

MVUE of normal distributions

Theorem

Let $X_1, \ldots, X_n$ be a random sample from a normal distribution with mean μ. Then the estimator $\hat{\mu} = \bar{X}$ is the MVUE for μ.

Method of moments

Example

Problem

Let $X_1, \ldots, X_{10}$ be a random sample from a distribution with pdf

$$f(x) = \begin{cases} (\theta + 1)x^\theta & \textit{if } 0 \le x \le 1 \\ 0 & \textit{otherwise} \end{cases}$$

A random sample of ten students yields data

$$x_1 = .92,\ x_2 = .79,\ x_3 = .90,\ x_4 = .65,\ x_5 = .86,$$

$$x_6 = .47,\ x_7 = .73,\ x_8 = .97,\ x_9 = .94,\ x_{10} = .77$$

Provide an estimator of θ.

Idea

- We can compute $E(X) \rightarrow$ the answer will be a function of θ
- For large n, we have

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

  is close to $E[X]$
- We can compute $\bar{x}$ from the data $\rightarrow$ approximate λ