

Point Estimation: Key Concepts

Maura Mezzetti

Department of Economics and Finance
Università Tor Vergata

Outline

1 Method of Moments

2 MLE

Method of Moments

The Method of Moments

consists of equating sample moments and population moments

Population moments

$$\mu_r(\theta) = E_{\theta}(X^r) = \int_{-\infty}^{\infty} x^r f(x; \theta) dx$$

Method of Moments

The Method of Moments

consists of equating sample moments and population moments

Population moments

$$\mu_r(\theta) = E_{\theta}(X^r) = \int_{-\infty}^{\infty} x^r f(x; \theta) dx$$

Sample moments

$$M_r = \frac{1}{n} \sum_i X_i^r$$

Method of Moments

The Method of Moments

consists of equating sample moments and population moments

Population moments

$$\mu_r(\theta) = E_{\theta}(X^r) = \int_{-\infty}^{\infty} x^r f(x; \theta) dx$$

Sample moments

$$M_r = \frac{1}{n} \sum_i X_i^r$$

Method of Moments

If a population has q parameters, the MOM consists of solving the system of q equations

1)

Equate the k -th population moment to the k -th sample moment

$$M_k = \mu_k \quad k = 1 \dots, q$$

Method of Moments

If a population has q parameters, the MOM consists of solving the system of q equations

1)

Equate the k -th population moment to the k -th sample moment

$$M_k = \mu_k \quad k = 1 \dots, q$$

2)

Solve the resulting system of equations for θ

Method of Moments

If a population has q parameters, the MOM consists of solving the system of q equations

1)

Equate the k -th population moment to the k -th sample moment

$$M_k = \mu_k \quad k = 1 \dots, q$$

2)

Solve the resulting system of equations for θ

$$X \sim N(\mu, \sigma^2)$$

$$\begin{aligned}\hat{\mu}_{MOM} &= \bar{X} \\ \hat{\sigma}_{MOM}^2 &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}\end{aligned}$$

MOM and unbiasedness

Are MOM estimators unbiased estimators?

$$X \sim N(\mu, \sigma^2)$$

$$\begin{aligned}\hat{\mu}_{MOM} &= \bar{X} \\ \hat{\sigma}_{MOM}^2 &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}\end{aligned}$$

MOM and unbiasedness

Are MOM estimators unbiased estimators?

$$X \sim N(\mu, \sigma^2)$$

$$\begin{aligned}\hat{\mu}_{MOM} &= \bar{X} \\ \hat{\sigma}_{MOM}^2 &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}\end{aligned}$$

MOM and unbiasedness

Are MOM estimators unbiased estimators?

$$X \sim U(0, \theta)$$

 $\hat{\theta}_{MOM}$

?

MOM and Sufficiency

Are MOM estimators a function of sufficient statistics?

$$X \sim U(0, \theta)$$

 $\hat{\theta}_{MOM}$

?

MOM and Sufficiency

Are MOM estimators a function of sufficient statistics?

$$X \sim U(0, \theta)$$

 $\hat{\theta}_{MOM}$

?

MOM and Sufficiency

Are MOM estimators a function of sufficient statistics?

$$X \sim U(0, \theta)$$

 $\hat{\theta}_{MOM}$

$$\hat{\theta}_{MOM} = 2\bar{X}$$

MOM and Sufficiency

Are MOM estimators a function of sufficient statistics?

$$X \sim U(0, \theta)$$

 $\hat{\theta}_{MOM}$

$$\hat{\theta}_{MOM} = 2\bar{X}$$

MOM and Sufficiency

Are MOM estimators a function of sufficient statistics?

$$X \sim U(0, \theta)$$

 $\hat{\theta}_{MOM}$

$$\hat{\theta}_{MOM} = 2\bar{X}$$

MOM and Sufficiency

Are MOM estimators a function of sufficient statistics?

Intuition

Suppose I tell you I have 100 cookies in my backpack. The cookies are of two types: chocolate chip cookies and fortune cookies. Moreover, I tell you that the number of fortune cookies is either 10 or 90. You draw a cookie out of my backpack at random and see that it is a fortune cookie.



Intuition

Based on this data, what is more likely: there are

- ① 10 fortune cookies and 90 chocolate chip cookies, or
- ② 90 fortune cookies and 10 chocolate chip cookies?

Based solely on one sample (fortune cookie), 2 is more likely.

Intuition

This is exactly the idea behind maximum likelihood estimation.

The method asks: what value of the parameter is most consistent with the data?

Intuition

This is exactly the idea behind maximum likelihood estimation.

The method asks: what value of the parameter is most consistent with the data?

In other words, what value of a parameter makes the data most likely?

Intuition

This is exactly the idea behind maximum likelihood estimation.

The method asks: what value of the parameter is most consistent with the data?

In other words, what value of a parameter makes the data most likely?

Bernoulli

$$X \sim \text{Bernoulli}(\theta)$$

Observations: 3 successes and 2 failures

Which of the following values for θ is the most likely:

0.2 0.4 0.6 0.8

Bernoulli

$X \sim \text{Bernoulli}(\theta)$

Observations: 3 successes and 2 failures

Which of the following values for θ is the most likely:

0.2 0.4 0.6 0.8

Likelihood?

$$L(\theta|x) = \prod_i \theta^{x_i} (1 - \theta)^{1-x_i}$$

$$L(\theta|x) = \theta^{\sum_i x_i} (1 - \theta)^{n - \sum_i x_i}$$

Bernoulli

$X \sim \text{Bernoulli}(\theta)$

Observations: 3 successes and 2 failures

Which of the following values for θ is the most likely:

0.2 0.4 0.6 0.8

Likelihood?

$$L(\theta|x) = \prod_i \theta^{x_i} (1 - \theta)^{1-x_i}$$

$$L(\theta|x) = \theta^{\sum_i x_i} (1 - \theta)^{n - \sum_i x_i}$$

Bernoulli

$$X \sim \text{Bernoulli}(\theta)$$

Observations: 3 successes and 2 failures

Likelihood

$$L(0.2|x) = 0.0051$$

$$L(0.4|x) = 0.023$$

$$L(0.6|x) = 0.0346$$

$$L(0.8|x) = 0.0205$$

Bernoulli

$$X \sim \text{Bernoulli}(\theta)$$

Observations: 3 successes and 2 failures

Likelihood

$$L(0.2|x) = 0.0051$$

$$L(0.4|x) = 0.023$$

$$L(0.6|x) = 0.0346$$

$$L(0.8|x) = 0.0205$$

More likely value?

Bernoulli

$$X \sim \text{Bernoulli}(\theta)$$

Observations: 3 successes and 2 failures

Likelihood

$$L(0.2|x) = 0.0051$$

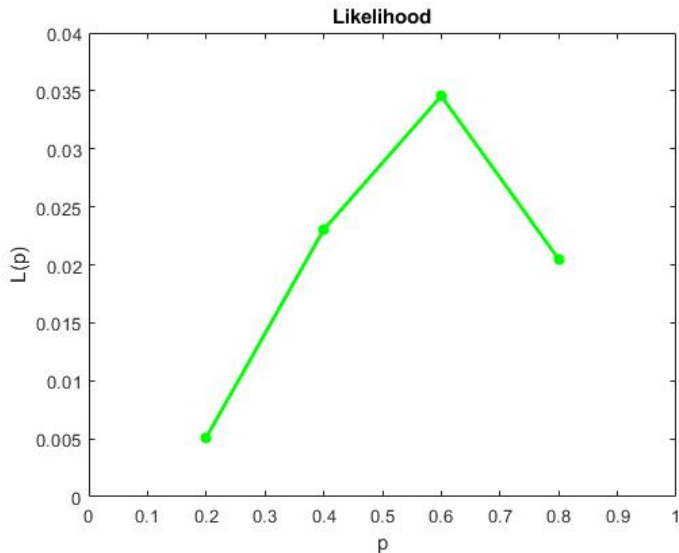
$$L(0.4|x) = 0.023$$

$$L(0.6|x) = 0.0346$$

$$L(0.8|x) = 0.0205$$

More likely value?

Bernoulli



Bernoulli

$$X \sim \text{Bernoulli}(\theta)$$

Observations: 3 successes and 2 failures

Likelihood

$$L(0.2|x) = 0.0051$$

$$L(0.4|x) = 0.023$$

$$L(0.6|x) = 0.0346$$

$$L(0.8|x) = 0.0205$$

Bernoulli

$$X \sim \text{Bernoulli}(\theta)$$

Observations: 3 successes and 2 failures

Likelihood

$$L(0.2|x) = 0.0051$$

$$L(0.4|x) = 0.023$$

$$L(0.6|x) = 0.0346$$

$$L(0.8|x) = 0.0205$$

More likely value? $\hat{p} = 0.6$

Bernoulli

$$X \sim \text{Bernoulli}(\theta)$$

Observations: 3 successes and 2 failures

Likelihood

$$L(0.2|x) = 0.0051$$

$$L(0.4|x) = 0.023$$

$$L(0.6|x) = 0.0346$$

$$L(0.8|x) = 0.0205$$

More likely value? $\hat{p} = 0.6$

Bernoulli

$X \sim \text{Bernoulli}(\theta)$

Observations: 3 successes and 2 failures

Likelihood

$$L(\theta|x) = \theta^{\sum_i x_i} (1 - \theta)^{n - \sum_i x_i}$$

$$\hat{\theta}_{MLE} = \hat{\theta}(x) \quad \text{such that} \quad L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta)$$

Bernoulli

$X \sim \text{Bernoulli}(\theta)$

Observations: 3 successes and 2 failures

Likelihood

$$L(\theta|x) = \theta^{\sum_i x_i} (1 - \theta)^{n - \sum_i x_i}$$

Most likely value

$$\hat{\theta}_{MLE} = \hat{\theta}(x) \quad \text{such that} \quad L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta)$$

Bernoulli

$X \sim \text{Bernoulli}(\theta)$

Observations: 3 successes and 2 failures

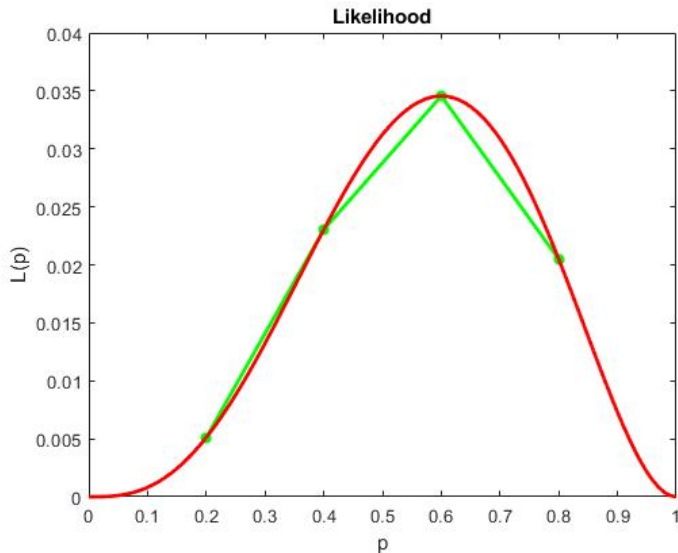
Likelihood

$$L(\theta|x) = \theta^{\sum_i x_i} (1 - \theta)^{n - \sum_i x_i}$$

Most likely value

$$\hat{\theta}_{MLE} = \hat{\theta}(x) \quad \text{such that} \quad L(\hat{\theta}) = \sup_{\theta \in \Theta} L(\theta)$$

Bernoulli



Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$L(\mu|x) = \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(x_i - \mu)^2}{2\sigma^2} \right) \right)$$

$$L(\mu|x) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n/2} \exp \left(-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right)$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$L(\mu|x) = \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(x_i - \mu)^2}{2\sigma^2} \right) \right)$$

$$L(\mu|x) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n/2} \exp \left(-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right)$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$L(\mu|x) = \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp \left(-\frac{(x_i - \mu)^2}{2\sigma^2} \right) \right)$$

$$L(\mu|x) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{n/2} \exp \left(-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} \right)$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$L(164|x) = 3.7 \times 10^{-8}$$

$$L(168|x) = 5.74 \times 10^{-8}$$

$$L(172|x) = 4.01 \times 10^{-8}$$

$$L(166|x) = 5.09 \times 10^{-8}$$

$$L(170|x) = 5.3 \times 10^{-8}$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$L(164|x) = 3.7 \times 10^{-8}$$

$$L(168|x) = 5.74 \times 10^{-8}$$

$$L(172|x) = 4.01 \times 10^{-8}$$

$$L(166|x) = 5.09 \times 10^{-8}$$

$$L(170|x) = 5.3 \times 10^{-8}$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$\log(L(164|x)) = -17.113$$

$$\log(L(166|x)) = -16.793$$

$$\log(L(168|x)) = -16.673$$

$$\log(L(170|x)) = -16.753$$

$$\log(L(172|x)) = -17.033$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$\log(L(164|x)) = -17.113$$

$$\log(L(166|x)) = -16.793$$

$$\log(L(168|x)) = -16.673$$

$$\log(L(170|x)) = -16.753$$

$$\log(L(172|x)) = -17.033$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$L(164|x) = 3.7 \times 10^{-8}$$

$$L(168|x) = 5.74 \times 10^{-8}$$

$$L(172|x) = 4.01 \times 10^{-8}$$

$$L(166|x) = 5.09 \times 10^{-8}$$

$$L(170|x) = 5.3 \times 10^{-8}$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$L(164|x) = 3.7 \times 10^{-8}$$

$$L(168|x) = 5.74 \times 10^{-8}$$

$$L(172|x) = 4.01 \times 10^{-8}$$

$$L(166|x) = 5.09 \times 10^{-8}$$

$$L(170|x) = 5.3 \times 10^{-8}$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$\log(L(164|x)) = -17.113$$

$$\log(L(168|x)) = -16.673$$

$$\log(L(172|x)) = -17.033$$

$$\log(L(166|x)) = -16.793$$

$$\log(L(170|x)) = -16.753$$

Gaussian

$$X \sim N(\mu, \sigma = 10)$$

Observations:

160 170 166 173 172

Likelihood

$$\log(L(164|x)) = -17.113$$

$$\log(L(168|x)) = -16.673$$

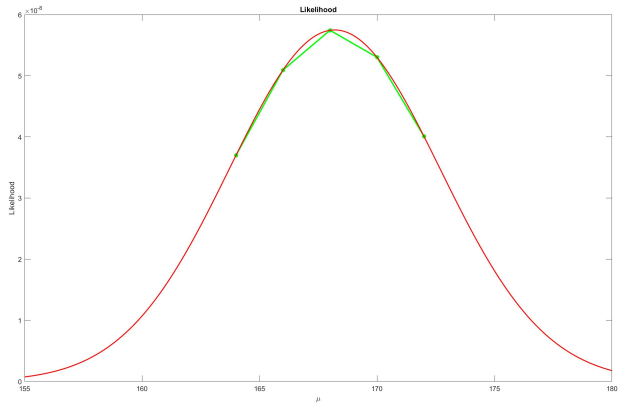
$$\log(L(172|x)) =$$

$$\log(L(166|x)) = -16.793$$

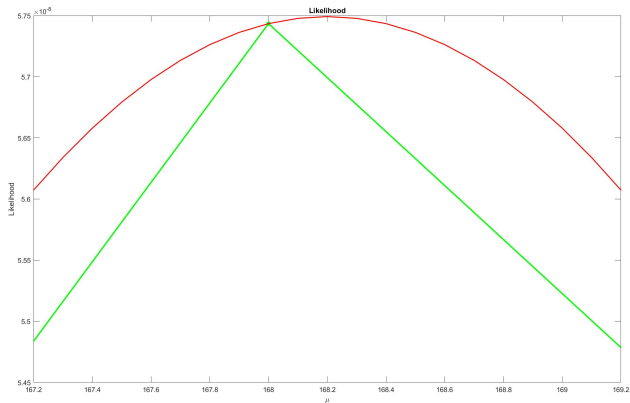
$$\log(L(170|x)) = -16.753$$

$$-17.033$$

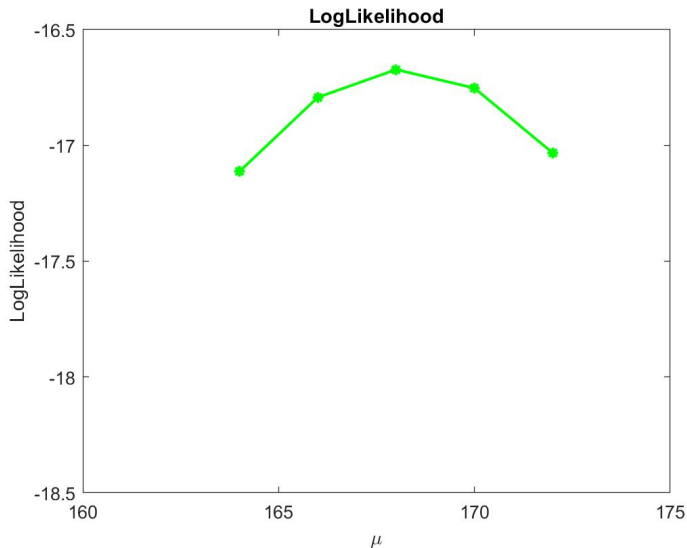
Gaussian-Likelihood



Gaussian-Likelihood



Gaussian-LogLikelihood



Gaussian-LogLikelihood

