NOVA SCHOOL OF BUSINESS & ECONOMICS

MACROECONOMETRICS

Master in Economics

Maximum Likelihood Estimation

Maximum Likelihood Estimation

A Really Simple Example

- Model: y_t i.i.d. N ($\mu = ?$, $\sigma^2 = 1$)
- There is one unknown parameter: $\mu = ?$
- Observed Sample: y_1 (yes, just one observation!)
- Objective: Use this sample to obtain an estimate of μ
- Let's use the Maximum Likelihood Estimator

Maximum Likelihood Estimation

Wikipedia

(March 23, 2021)

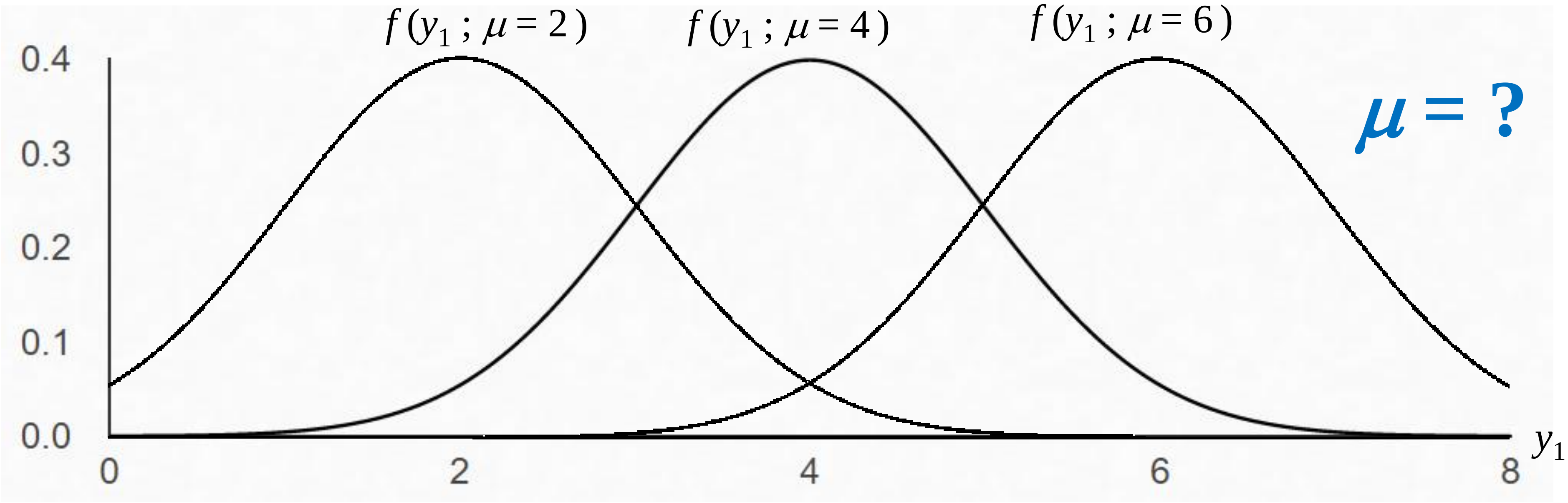
- ✓ *In statistics, maximum likelihood estimation (MLE) is a method of* ***estimating the parameters*** *of a probability distribution by maximizing a likelihood function,* ***so that*** *under the assumed statistical model* ***the observed data is most probable****.*

- ✓ *The point in the parameter space that maximizes the likelihood function is called the maximum likelihood estimate.*

Maximum Likelihood Estimation

A Really Simple Example

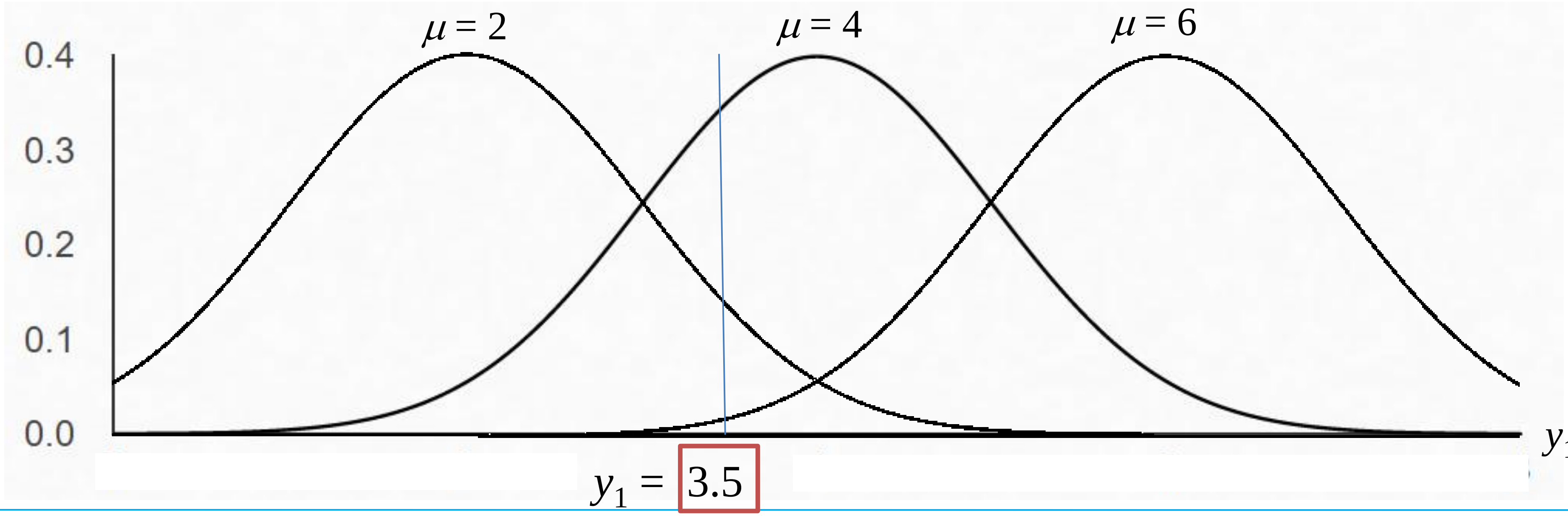
- Model: $y_1 \sim \mathrm{N}(\mu, \sigma^2 = 1)$
- Let $f(y_1 ; \mu)$ denote the probability density function (pdf) of y_1, which depends on the unknown value of μ



Maximum Likelihood Estimation

A Really Simple Example

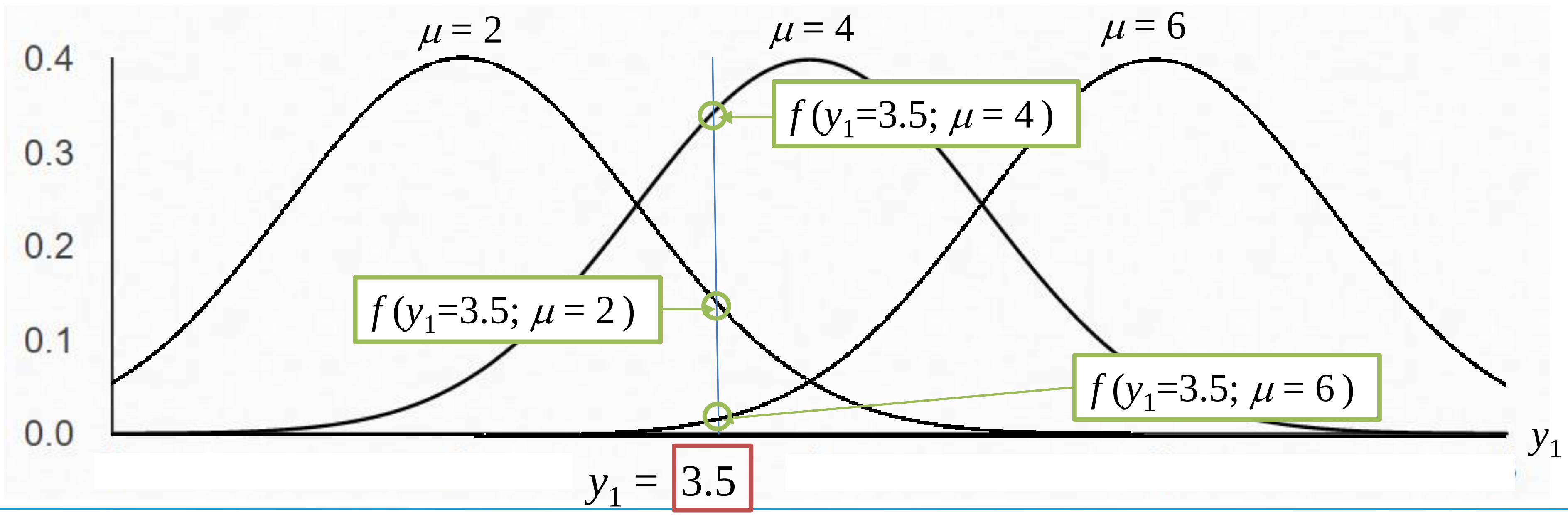
- Maximum Likelihood Estimation: Find μ so that the observed data is most probable
- Suppose $y_1 = 3.5$ is the observed data !



Maximum Likelihood Estimation

A Really Simple Example

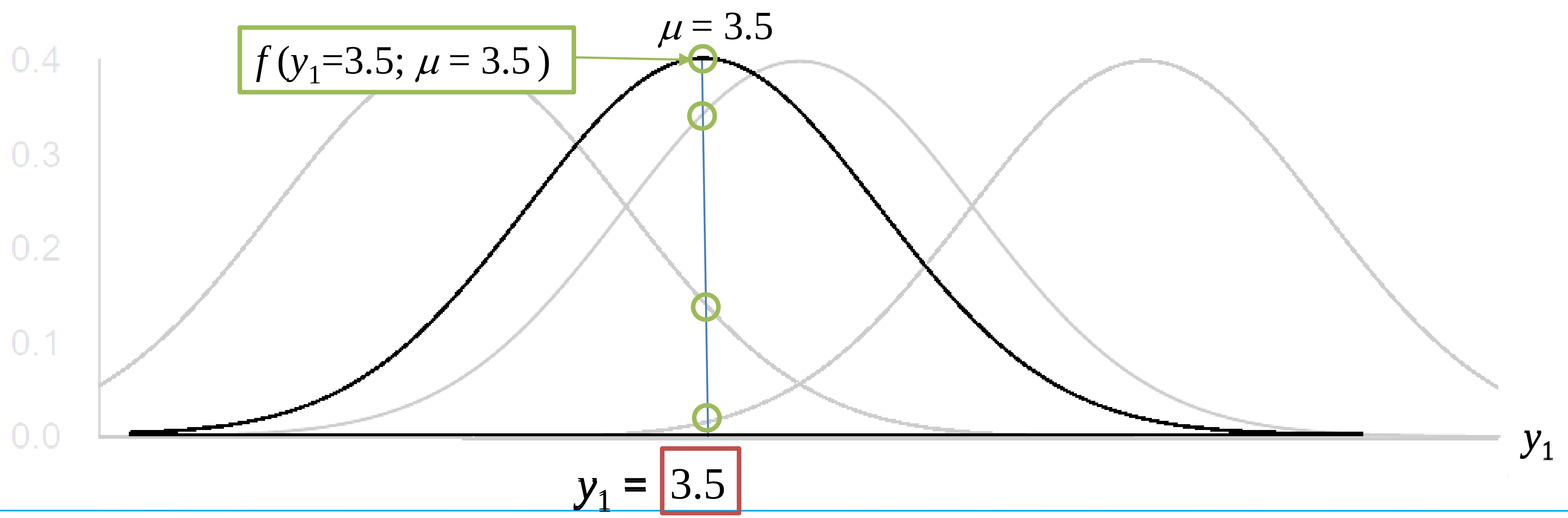
- Maximum Likelihood Estimation: Find μ so that the observed data is most probable
- Suppose $y_1 = 3.5$ is the observed data !



Maximum Likelihood Estimation

A Really Simple Example

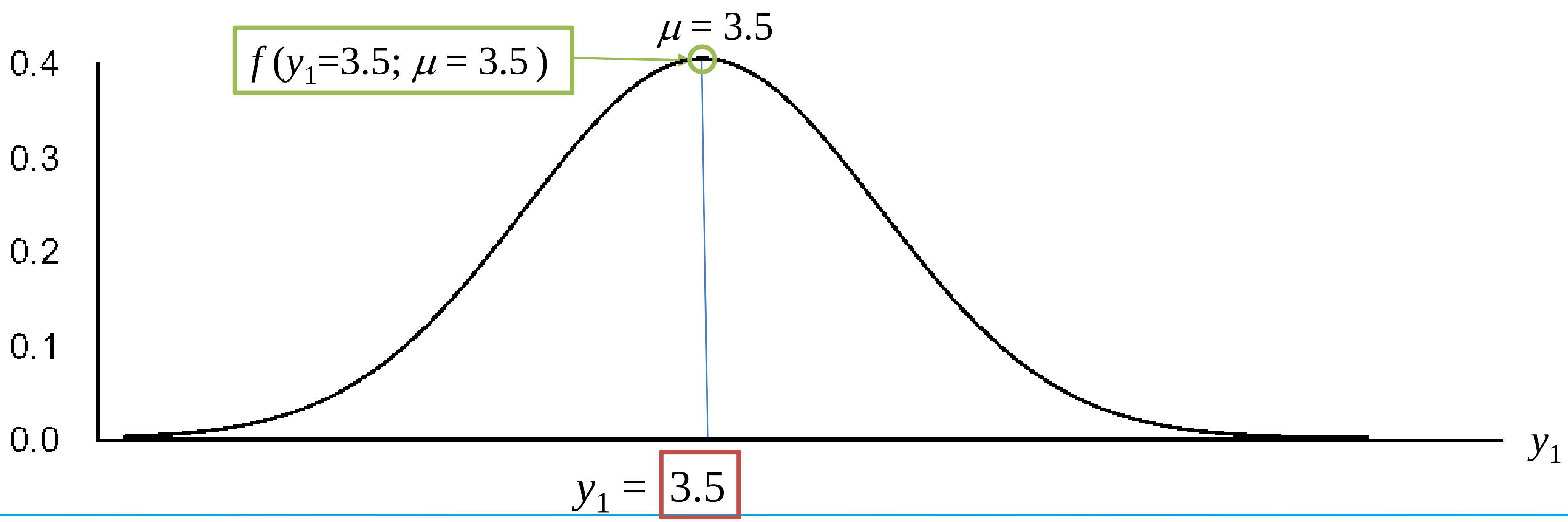
- The value of μ that makes the observed data most probable is $\mu = 3.5$



Maximum Likelihood Estimation

A Really Simple Example

- Maximum Likelihood Estimate of μ is 3.5



Maximum Likelihood Estimation

A Really Simple Example

Mathematical Derivations

- Model: $y_1 \sim N(\mu, \sigma^2 = 1)$
- The probability density function (pdf) of y_1 is

$$f(y_1 ; \mu) = (2\pi)^{-1/2} \exp[-(1/2)[(y_1 - \mu)^2]]$$

- If we consider $f(y_1 ; \mu)$ as a function of μ only, given the observed value of y_1, then the corresponding function $f(\mu)$ is known as the Likelihood Function
- The Maximum Likelihood Estimator of μ is found by

$$\text{Maximizing } f(\mu)$$

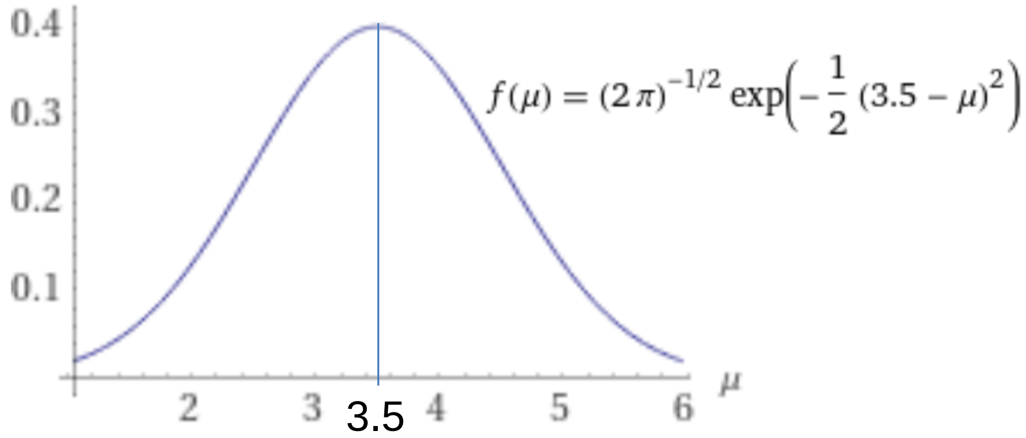
Maximum Likelihood Estimation

A Really Simple Example

Mathematical Derivations

- Example: $y_1 = 3.5$ is the observed data
- The likelihood function (a function of μ) is

$$f(\mu) = f(y_1{=}3.5;\, \mu) = (2\pi)^{-1/2} \exp\left[-(1/2)\left[(3.5-\mu)^2\right]\right]$$



Max $f(\mu) \Rightarrow$ MLE of μ is 3.5

Maximum Likelihood Estimation

A Second Example

- Observed Sample: $y_1, y_2, \ldots, y_T$

- Model: $y_t = \mu + \varepsilon_t$ with ε_t white noise $N(0, \sigma^2)$

  $\Leftrightarrow$ y_t i.i.d. $N(\mu, \sigma^2)$, $t = 1, \ldots, T$

- Objective: Estimate the unknown parameters μ and σ^2 by using the Maximum Likelihood Estimation method

Maximum Likelihood Estimation

A Second Example

- Observed Sample: $y_1 , y_2 , \ldots , y_T$
- Model: y_t i.i.d. N (μ, σ^2)
- Denote the probability density function (pdf) of the sample $(y_1 , y_2 , \ldots , y_T)$ given the parameters μ, σ^2 as

$$f(y_1 , y_2 , \ldots , y_T ; \mu, \sigma^2)$$

  - If we consider this $f(\cdots)$ as a function of μ, σ^2, given the observed sample values, it is called the Likelihood Function
- Maximum Likelihood Estimates of μ and σ^2

$$\text{Maximize } f(y_1 , y_2 , \ldots , y_T ; \mu, \sigma^2)$$

Maximum Likelihood Estimation

A Second Example

- If $y_1 , y_2 , \dots , y_T$ are independent random variables, their joint pdf is the product of their marginal pdfs

$$f(y_1 , y_2 , \dots , y_T) = f(y_1) \times f(y_2) \times \cdots \times f(y_T)$$

- In our model, $y_1 , y_2 , \dots , y_T$ are independent and so

$$f(y_1 , y_2 , \dots , y_T ; \mu , \sigma^2) = \Pi_{t=1,\dots,T} f(y_t ; \mu , \sigma^2)$$

where $f(y_t ; \mu , \sigma^2)$ is the pdf of each observation y_t

$\Rightarrow$ Maximum Likelihood Estimates of μ and σ^2

$$\text{Max } \Pi_{t=1,\dots,T} f(y_t ; \mu , \sigma^2)$$

Maximum Likelihood Estimation

A Second Example

- In many cases, as in this one, using the logs of the pdf/likelihood function, simplifies many calculations:

$$\log f\,(y_1 , y_2 , \ldots , y_T ; \mu , \sigma^2) = \Sigma_{t=1,\ldots,T} \log f\,(y_t ; \mu , \sigma^2)$$

- Note: the logarithm is a monotonic transformation. Therefore, the values of μ and σ^2 that Max $f\,(\cdots)$ are the same that Max $\log f\,(\cdots)$

$\Rightarrow$ Maximum Likelihood Estimates of μ and σ^2

$$\text{Max } \Sigma_{t=1,\ldots,T} \log f\,(y_t ; \mu , \sigma^2)$$

Maximum Likelihood Estimation

A Second Example

- In this model $y_t \sim N(\mu, \sigma^2)$, so that the pdf of y_t is

$$f(y_t ; \mu, \sigma^2) = (2\pi \sigma^2)^{-1/2} \exp [- (1/2) [(y_t - \mu)^2 / \sigma^2]]$$

- Taking the logarithm of the pdf, we get

$$\log f(y_t ; \mu, \sigma^2)$$

$$= -(1/2) \log(2\pi \sigma^2) -(1/2) [(y_t - \mu)^2 / \sigma^2]$$

$$= -(1/2) \log(2\pi) -(1/2) \log(\sigma^2) -(1/2) [(y_t - \mu)^2 / \sigma^2]$$

Maximum Likelihood Estimation

A Second Example

- Observed Sample: $y_1, y_2, \dots, y_T$
- Model: y_t i.i.d. N (μ, σ^2)
- $\log f(y_t; \mu, \sigma^2) =$

$$= -(1/2)\log(2\pi) - (1/2)\log(\sigma^2) - (1/2)\left[(y_t - \mu)^2/\sigma^2\right]$$

- Maximum Likelihood estimator of μ and σ^2

$$\text{Max } \Sigma_{t=1,\dots,T} \log f(y_t; \mu, \sigma^2)$$

- After a few derivations, we obtain the ML estimates:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1,\dots,T} y_t \quad \text{and} \quad \widehat{\sigma}^2 = \frac{1}{T}\sum_{t=1,\dots,T}(y_t - \hat{\mu})^2$$

Maximum Likelihood Estimation

The AR(1) Model

- Model: $y_t = c + \rho\ y_{t\text{-}1} + \varepsilon_t$, ε_t white noise N(0 , σ^2)
- Three parameters to estimate: c , ρ , σ^2
- In this model $y_1 , y_2 , \dots , y_T$ are not independent

$$\Rightarrow f(y_1 , y_2 , \dots , y_T) \neq f(y_1) \times f(y_2) \times \cdots \times f(y_T)$$

- However, using the conditional pdfs, we have that

$$\begin{aligned} f(y_1 , y_2 , \dots , y_T) = f(y_1) &\times f(y_2 \mid y_1) \\ &\times f(y_3 \mid y_2 , y_1) \\ &\times f(y_4 \mid y_3 , y_2 , y_1) \\ &\times \cdots \\ &\times f(y_T \mid y_{T\text{-}1} , \dots , y_3 , y_2 , y_1) \end{aligned}$$

Maximum Likelihood Estimation

The AR(1) Model

- The conditional maximum likelihood estimator considers the joint pdf of $(y_2, \ldots, y_T)$ conditional on the initial value y_1

$$f(y_2, \ldots, y_T) = \Pi_{t=2,\ldots,T} f(y_t \mid I_{t-1})$$

where $f(y_t \mid I_{t-1})$ is the conditional pdf of y_t given the past values $I_{t-1} = \{y_{t-1}, \ldots, y_2, y_1\}$

- Using logarithms:

$$\log f(y_2, \ldots, y_T) = \Sigma_{t=2,\ldots,T} \log f(y_t \mid I_{t-1})$$

Maximum Likelihood Estimation

The AR(1) Model

- Model: $y_t = c + \rho\ y_{t-1} + \varepsilon_t$, ε_t white noise N(0 , σ^2)
- $\log f(y_t \mid I_{t-1}; c, \rho, \sigma^2) =$
  $$= -(1/2)\log(2\pi)$$
  $$-(1/2)\log(\sigma^2) - (1/2)\,[\,[y_t - (c + \rho\ y_{t-1})]^2/\sigma^2\,]$$
- Conditional Maximum Likelihood Estimator of c, ρ, σ^2

$$\text{Max}\ \Sigma_{t=2,\ldots,T} \log f(y_t \mid I_{t-1}; c, \rho, \sigma^2)$$

- After a few derivations, it can be shown that the estimates of c and ρ can be obtained by OLS, and the estimate of σ^2 is the sum of squared residuals divided by $T-1$

Maximum Likelihood Estimation

The AR(1) Model

- Note: To calculate the full (unconditional) maximum likelihood estimator, it is necessary to use the unconditional pdf of y_t for $t = 1$

$$f(y_t; c, \rho, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi(1-\rho^2)}} \exp\left\{ -\frac{(y_t - c/(1-\rho))^2}{2(\sigma^2/(1-\rho^2))} \right\} \qquad t = 1$$

$$f(y_t \mid I_{t-1}; c, \rho, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{(y_t - c - \rho y_{t-1})^2}{2(\sigma^2)} \right\} \qquad t > 1$$

Maximum Likelihood Estimation

The GARCH(1,1) Model

- Model: $y_t = c + u_t$, $u_t = (h_t)^{1/2}\, \varepsilon_t$, ε_t white noise N(0,1)

  or $$y_t \mid I_{t\text{-}1} \sim \mathrm{N}(c, h_t)$$

  with $$h_t = \alpha_0 + \alpha_1 u_{t\text{-}1}^2 + \beta_1 h_{t\text{-}1}$$

- Four parameters to estimate: c , α_0 , α_1 , β_1
- $\log f(y_t \mid I_{t\text{-}1}; c, \alpha_0, \alpha_1, \beta_1) =$
  $= -(1/2)\log(2\pi) - (1/2)\log(h_t) - (1/2)\left[(y_t - c)^2 / h_t\right]$
- Maximum Likelihood Estimator:

$$\text{Max}\ \Sigma_{t=1,\ldots,T} \log f(y_t \mid I_{t\text{-}1}; c, \alpha_0, \alpha_1, \beta_1)$$

(Note: the initial value of h_t has to be estimated)