$Y_1, \ldots, Y_n$ IID $N(\mu, \sigma^2)$ θ

ECONSTA: $\hat{\mu} = \bar{Y}$

$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$ $\quad \frac{1}{n-1}$

$\mu = \mathbb{E}(Y_i)$

$\sigma^2 = \mathrm{Var}(Y_i)$

Goal: Estimate μ, σ^2.

indpt

$Y_1 \sim N(\mu, \sigma^2)$ $\quad f(y_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(y_1-\mu)^2}{\sigma^2}\right)$ ← marginal density Y_1

$Y_2 \sim N(\mu, \sigma^2)$ $\quad f(y_2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(y_2-\mu)^2}{\sigma^2}\right)$ ← marginal density of Y_2

$\vdots$

$Y_n \sim N(\mu, \sigma^2)$ $\quad f(y_n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(y_n-\mu)^2}{\sigma^2}\right)$ ← marginal density of Y_n

Likelihood function = joint density of the observations.

→ product of marginal densities because of independence

joint density of $Y_1, \ldots, Y_n$

$$\hookrightarrow \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(y_1-\mu)^2}{\sigma^2}\right) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(y_2-\mu)^2}{\sigma^2}\right) \cdot \ldots \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(y_n-\mu)^2}{\sigma^2}\right)$$

algebra

$\exp(a)\exp(b) = \exp(a+b)$

$e^a e^b = e^{a+b}$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2}\frac{(y_1-\mu)^2}{\sigma^2} + \left(-\frac{1}{2}\right)\frac{(y_2-\mu)^2}{\sigma^2} + \ldots + \left(-\frac{1}{2}\right)\frac{(y_n-\mu)^2}{\sigma^2}\right)$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\left((y_1-\mu)^2 + (y_2-\mu)^2 + \ldots + (y_n-\mu)^2\right)\right]}_{\text{function of } n,\ \sigma^2,\ \mu,\ y_1, y_2, \ldots, y_n}$$

Likelihood function → treats joint density of observations as a function of the unknown parameters and evaluated at the observed data

Maximum likelihood estimator → maximize likelihood function! by choosing optimal μ, σ^2

$$L(\mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\left((y_1-\mu)^2 + (y_2-\mu)^2 + \ldots + (y_n-\mu)^2\right)\right]$$

$\frac{\partial L}{\partial \mu} = 0$, $\frac{\partial L}{\partial \sigma^2} = 0$ ⇒ solve for optimal μ, σ^2

very hard (easier)

computation log-likelihood

$\log(ab) = \log(a) + \log(b)$

$$\log L(\mu, \sigma^2) = \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n + \log \exp\left[-\frac{1}{2\sigma^2}\left((y_1-\mu)^2 + (y_2-\mu)^2 + \ldots + (y_n-\mu)^2\right)\right]$$

$$(2\pi\sigma^2)^{-\frac{1}{2}} \qquad = n \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{1}{2\sigma^2}\left((y_1-\mu)^2+(y_2-\mu)^2+\cdots+(y_n-\mu)^2\right)$$

$$= -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\left((y_1-\mu)^2+(y_2-\mu)^2+\cdots+(y_n-\mu)^2\right)$$

$$= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\left((y_1-\mu)^2+(y_2-\mu)^2+\cdots+(y_n-\mu)^2\right)$$

$$\left.\begin{aligned} \frac{\partial \log L(\mu,\sigma^2)}{\partial \mu} &= 0 \\ \frac{\partial \log L(\mu,\sigma^2)}{\partial \sigma^2} &= 0 \end{aligned}\right\}$$ Solve system of equations involving μ, σ^2

general purpose estimation algorithm → maximum likelihood estimation

$\max_{x} f(x)$ $\qquad$ $\max_{x} \log f(x)$

x^*

In R

(book) MASS::fitdistr( ) → univariate distributions (but limited to a fixed set of known distributions)

nlm( ) → nonlinear minimization (specify the function to be minimized)

(book) optim( ) → general purpose optimization (specify the function to be minimized)

→ encode log-likelihood in R!

in R, default is minimization of a function.

$\max_{x} f(x)$ $\qquad$ $\min_{x} -f(x)$

x^*