$Y_1, \dots, Y_n$ IID $N(\cancel{\mu}, \sigma^2)$ model t_ν

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

goal: estimate μ, σ^2 using maximum likelihood estimation

$\mu = E(Y_i)$

$\sigma^2 = Var(Y_i)$

$\hat{\sigma}^2$ = sample variance $= \frac{1}{n}\sum(Y_t - \bar{Y})^2$, $\frac{1}{n-1}$

Set up likelihood function

independent:

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

$Y_2 \sim N(\mu, \sigma^2)$

$\vdots$

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

Because of independence, joint density of $Y_1, \dots, Y_n$ = product of marginal densities

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

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

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

$$\boxed{\begin{aligned} &\exp(a)\exp(b) = \exp(a+b) \\ &e^a \cdot e^b = e^{a+b} \end{aligned}}$$

Likelihood function → joint density viewed as a function of unknown parameters, and evaluated at the observed values of $Y_1, \dots, Y_n$.

Hard to maximize likelihood function directly! — by hand: calculus techniques become very difficult to use; by computer: overflow, underflow

Approach: maximize log-likelihood instead!

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

$\max_x f(x)$ transform $\max_x \log f(x)$

in computers → default is to minimize functions.

$\max_x f(x)$ $\quad \min_x -f(x)$

$\theta = (\mu, \sigma^2)'$

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

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

prop. of logs

$$" (2\pi\sigma^2)^{-1/2}$$

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

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

$$\left.\begin{aligned} \frac{\partial \log L(\mu,\sigma^2)}{\partial \mu} &= 0 \\ \frac{\partial \log L(\mu,\sigma^2)}{\partial \sigma^2} &= 0 \end{aligned}\right\}$$

Solve for optimal μ, σ^2.
↳ maximizes log-likelihood

MLE → general purpose estimation algorithm

In R:

MASS::fitdistr() univariate dist / fixed set that you can use

nlm() nonlinear minimization
↳ give function to be minimized

(book) optim() general-purpose optimization
↳ give function to be minimized

encode the log-likelihood!

optimx()

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

sample variance in R sd()

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