STAT 231

Roadmap

- The Theory of Estimation
- Likelihood functions and the MLE for discrete distributions.
- Continuous distributions.

$\{y_1, \ldots y_n\} \longrightarrow$ DATA SAMPLE

<u>Set-up</u>

MODEL: $Y_i \sim f(y_i; \theta) \ldots$ (1)

$i = 1, \ldots n.$

<u>Meaning</u>: Each data point y_i is an outcome of a. r.v. Y_i which has a distribution function f (DISCRETE).

<u>Objective</u>: To estimate θ = unknown parameter we are interested in.

Method: METHOD OF M. L.

Likelihood function:

$$L(\theta; y_1, \ldots y_n)$$

$$= P(Y_1 = y_1, Y_2 = y_2, \ldots Y_n = y_n).$$

= Probability of observing your sample

It is a function of θ

Choose the value of θ $(\hat{\theta})$.

that maximizes the likelihood function

$\hat{\theta}$ = Maximum Likelihood Estimate.

= MLE

$\hat{\theta}(y_1, \ldots, y_n)$. = function of the sample points.

Suppose all the data points are independently drawn.

$$L(\theta; y_1, \ldots y_n) = P(Y_1 = y_1, Y_2 = y_2, \ldots Y_n = y_n)$$

$$= \underline{P(Y_1 = y_1)} \cdot \underline{P(Y_2 = y_2)} \ldots P(Y_n = y_n).$$

(if they are independent).

$$\left(P(A \cap B) = P(A) \cdot P(B)\right)$$

$$= f(y_1) \cdot f(y_2) \ldots f(y_n).$$

$$\boxed{L(\theta; y_1, \ldots y_n) = \prod_{i=1}^{n} f(y_i; \theta)}$$

Given an independent sample, the likelihood function = Product of the distribution function, evaluated at the sample points

---

$\theta_1, \ldots, \theta_n$ $\longrightarrow$ Possible values of θ.

$\{y_1, \ldots, y_n\}$ $\longrightarrow$ Sample.

$\hat{\theta}$ = the value of θ that has the highest prob. of generating my sample.

$\{y_1, \ldots y_n\}$ $\quad Y_i \sim f(y_i ; \theta)$

$i = 1, \ldots n.$

The parameter θ in the model is

(a) An unknown constant

(b) A random variable

(c). a known #, given the sample.

The MLE $\hat{\theta}$ is

(a) a random variable

(b) a unknown constant

(c) a known # given the sample.

__Example__. Suppose a data set $\{\underline{y_1, \ldots, y_n}\}$ is drawn from a discrete distribution with pdf. (drawn independently).

$$f(y, \theta) = (1-\theta)^{y} \cdot \theta$$

$$y = 0, 1, 2, \ldots$$

Find the MLE for θ

Step 1: Construct L

$$L = \prod_{i=1}^{n} \underbrace{(1-\theta)^{y_i} \theta}$$

$$= (1-\theta)^{y_1} \cdot \theta \cdot (1-\theta)^{y_2} \cdot \theta \ldots (1-\theta)^{y_n} \cdot \theta.$$

$$L = \underline{(1-\theta)^{\sum y_i} \theta^n}$$

Step 2: Construct l = log likelihood function

$$l(\theta; y_1, .. y_n) = \sum y_i \cdot \ln(1-\theta) + n \ln \theta$$

Step 3 Maximize l

$dl/d\theta = 0$ and solve for θ.

$$-\frac{\sum y_i}{1-\theta} + \frac{n}{\theta} = 0$$

SOLVE $\Rightarrow$

2^{nd} order condition needs to be checked to guarantee maximum.

CONTINUOUS DISTRIBUTION

The definition in terms of probability does not quite carry over.

For a continuous distribution and an independent sample,

$$L(\theta; y_1, \dots y_n) = \prod_{i=1}^{n} f(y_i; \theta)$$

f = density function of Y

Example :

To estimate the average lifetime of an electric bulb produced by a company : μ

$\{y_1, \ldots y_n\}$ → SAMPLE

y_i = the lifetime of the i^{th} bulb in your sample

Given our data set, what is the MLE for μ ?

MODEL $Y_i \sim Exp(\mu)$

$i = 1, \ldots, n$

Y_i's independent

$$f(y;\mu) = \boxed{\frac{1}{\mu} e^{-y/\mu}} \quad y \geq 0$$

DENSITY FUNCTION.

Likelihood function

$$L(\mu; y_1, \ldots, y_n) = \frac{1}{\mu} e^{-y_1/\mu} \cdots \frac{1}{\mu} e^{-y_n/\mu}$$

$$= \frac{1}{\mu^n} e^{-\frac{1}{\mu}\sum_{i=1}^{n} y_i}$$

The log likelihood function

$$\ell(\mu) = -n \ln \mu - \frac{1}{\mu} \sum y_i$$

$$\frac{d\ell}{d\mu} = 0 \Rightarrow -\frac{n}{\mu} + \frac{1}{\mu^2} \sum y_i$$

$$= 0$$

$$\frac{n}{\cancel{\mu}} = \frac{1}{\mu^{\cancel{2}}} \sum y_i$$

$$\Rightarrow \hat{\mu} = \sum y_i / n$$

$$\boxed{\hat{\mu} = \bar{y}}$$

For the Exponential, the MLE $= \hat{\mu}$

$= \bar{y}$

Example 2

Objective: To estimate μ and σ

μ → Population average

σ → population s.d.

$\{y_1, \ldots y_n\}$ → SAMPLE of STAT 231 final grades

The "appropriate" is the GAUSSIAN model.

$N(\mu, \sigma^2)$ $\Uparrow$ Rest of the universe

$G(\mu, \sigma)$ $\downarrow$ UW

MODEL $\boxed{Y_i \sim G(\mu, \sigma)}$

$i = 1, \ldots, n$

independent.

$\hat{\mu}, \hat{\sigma} = ?$

DENSITY

$\downarrow$

$$\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(y_i-\mu)^2}$$

$$L(\mu,\sigma) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(y_i-\mu)^2}$$

$$\boxed{L = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{1}{2\sigma^2}\sum(y_i-\mu)^2}}$$

The log-likelihood

$$\ell = -\frac{n}{2}\ln(2\pi) - n\ln\sigma - \frac{1}{2\sigma^2}\sum(y_i-\mu)^2$$

To find the maximum

$$\left.\begin{aligned} \frac{\partial \ell}{\partial \mu} &= 0 \\ \frac{\partial \ell}{\partial \sigma} &= 0 \end{aligned}\right\} \text{ and SOLVE}$$

$\underset{\sim}{\theta} = (\mu, \sigma)$ = vector of unknown parameters

$$\boxed{\hat{\mu} = \bar{y}}$$

SAMPLE MEAN

$$\boxed{\hat{\sigma}^2 = \frac{1}{n}\sum (y_i - \hat{\mu})^2}$$

≠ SAMPLE VARIANCE