MAXIMUM LIKELIHOOD METHOD

MAXIMUM LIKELIHOOD METHOD (MLM) IS A GENERAL METHOD FOR ESTIMATING PARAMETERS OF INTEREST.

STATEMENT OF THE PROBLEM

1. WE ARE GIVEN m ETTP MEASUREMENTS OF PARAMETER x. THESE ARE: $\{x_1, x_2, \ldots, x_m\}$

2. ASSUME THAT WE KNOW THE PROBABILITY DISTRIBUTION FUNCTION THAT DESCRIBES x: $f_X(x, \theta)$
3. X IS THE CORRESPONDING RANDOM VARIABLE. ALSO $x \in X$.
4. θ IS THE PARAMETER OF INTEREST, WHICH HAS TO BE ESTIMATED.

5. DATA POINTS ARE NOT NECESSARILY RANDOM.

6. THE MAXIMUM-LIKELIHOOD ESTIMATOR OF θ IS θ_{ML}. IT IS THAT θ WHICH MAXIMIZES IT.

MAXIMUM LIKELIHOOD ESTIMATION

1. MAXIMUM LIKELIHOOD ESTIMATOR OF θ, θ_{ML} IS THAT θ WHICH MAXIMIZES $f_{X_1, X_2, \ldots X_n}(x_1, x_2, \ldots, x_n ; \theta)$

WHERE $f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n)$ IS THE JOINT PROBABILITY DENSITY FUNCTION OF THE SAMPLE, WRITTEN AS A FUNCTION OF θ.

2. THIS JOINT DENSITY FUNCTION IS ALSO CALLED THE LIKELIHOOD FUNCTION, THAT IS

$$L(\theta ; x_1, x_2, \ldots, x_n) = f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n ; \theta)$$

3. θ_{ML} IS CALLED THE MAXIMUM LIKELIHOOD ESTIMATE OF θ, BECAUSE, IT IS THAT ESTIMATE OF θ THAT MAXIMIZES THE LIKELIHOOD OF DRAWING THE SAMPLE: $x_1, x_2, \ldots, x_n$.

4. θ_{ML} IS FOUND BY MAXIMIZING $L(\theta ; x_1, x_2, \ldots, x_n)$ WITH RESPECT TO θ.

5. IF THE SAMPLE IS RANDOM FROM $f_X(x, \theta)$ THEN

$$\begin{aligned} L(\theta ; x_1, x_2, \ldots, x_n) &= f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n ; \theta) \\ &= f_X(x_1, \theta)\, f_X(x_2, \theta) \cdots f_X(x_n, \theta) \\ &= \prod_{i=1}^{n} f_X(x_i, \theta) \end{aligned}$$

DENOTES PRODUCT OPERATOR (PI)

θ_{ML} IS THAT θ WHICH MAXIMIZES $\prod_{i=1}^{n} f_X(x_i ; \theta)$

6. THE θ THAT MAXIMIZES $\prod_{i=1}^{n} f_X(x_i, \theta)$ IS ALSO THE θ THAT MAXIMIZES

$$\ln \left\{ \prod_{i=1}^{n} f_X(x_i, \theta) \right\},$$

THIS IS GUARANTEED BECAUSE LOG FUNCTION IS MONOTONIC

DEFINITION A FUNCTION f DEFINED ON A SUBSET OF REAL NUMBERS WITH REAL VALUES IS CALLED MONOTONIC (ALSO MONOTONICALLY INCREASING; INCREASING; OR NON-DECREASING) IF FOR ALL x AND y SUCH THAT $x \leq y$ ONE HAS $f(x) \leq f(y)$. $\square$

NOTE THAT

$$\ln\left\{\prod_{i=1}^{n} f_X(x_i, \theta)\right\} = \sum_{i=1}^{n} \ln f_X(x_i, \theta)$$

EXAMPLE RV HAS BERNOULLI DISTRIBUTION WITH PARAMETER $p>0$

$$f_x(x,p) = \begin{cases} p & x=1 \\ 1-p & x=0 \end{cases}$$

GIVEN: $\{x_1, x_2, \ldots, x_n\}$; ESTIMATE p

SOLUTION $f_x(x,p)$ IS REWRITTEN AS

$$f_x(x,p) = p^x (1-p)^{1-x}$$

$$\ln L(p; x_1, x_2, \ldots, x_n) = \ln\left\{\prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i}\right\}$$

$$= \sum_{i=1}^{n} \ln\left\{p^{x_i}(1-p)^{1-x_i}\right\}$$

$$= \ln p\left\{\sum_{i=1}^{n} x_i\right\} + \ln(1-p)\left\{\sum_{i=1}^{n}(1-x_i)\right\}$$

LET $s \triangleq \sum_{i=1}^{n} x_i$; $\sum_{i=1}^{n}(1-x_i) = n-s$

$$\therefore \ln L(\cdot) = s\ln p + (n-s)\ln(1-p)$$

$$\frac{d\ln L}{dp} = \frac{s}{p} + \frac{(n-s)}{(1-p)}(-1) = 0 \Rightarrow (1-p)s = (n-s)p \Rightarrow s = np$$

LET $\boxed{\hat{p} = \frac{s}{n}}$ $\Rightarrow$ ESTIMATE OF p

$$\frac{d^2\ln L}{dp^2} = -\frac{s}{p^2} - \frac{(n-s)}{(1-p)^2} \; ; \quad \left.\frac{d^2\ln L}{dp^2}\right|_{p=\hat{p}} = -\frac{s}{\hat{p}^2} - \frac{(n-s)}{(1-\hat{p})^2}$$

$$= -\left[\frac{sn^2}{s^2} + \frac{(n-s)n^2}{(n-s)^2}\right]$$

$$= -n^2\left[\frac{1}{s} + \frac{1}{n-s}\right]$$

$$= -\frac{n^3}{s(n-s)} < 0 \quad \text{IF } s<n$$

$\Rightarrow \hat{p}$ MAXIMIZES $\ln L(\,)$ $\boxed{\hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i}$

□

EXAMPLE RV X HAS POISSON DISTRIBUTION WITH PARAMETER $\lambda > 0$

$$b_X(x, \lambda) = e^{-\lambda} \frac{\lambda^x}{x!} \;;\quad x = 0, 1, 2, \ldots$$

GIVEN: $\{x_1, x_2, \ldots, x_n\}$; ESTIMATE λ

SOLUTION: $L(\lambda; x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$

$$\ln L(\lambda; x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \{-\lambda + x_i \ln \lambda - \ln x_i!\}$$

$$\frac{d \ln L}{d\lambda} = -n + \frac{1}{\lambda} \sum_{i=1}^{n} x_i = 0 \Rightarrow \lambda = \frac{1}{n} \sum_{i=1}^{n} x_i \triangleq \hat{\lambda}$$

$$\frac{d^2 \ln L}{d\lambda^2} = -\frac{1}{\lambda^2} \sum_{i=1}^{n} x_i$$

$$\left.\frac{d^2 \ln L}{d\lambda^2}\right|_{\lambda = \hat{\lambda}} = -\frac{1}{\hat{\lambda}^2} n \hat{\lambda} = -\frac{n}{\hat{\lambda}} < 0$$

$\Rightarrow$ $\hat{\lambda}$ MAXIMIZES $\ln L(\cdot)$. THUS $\boxed{\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i}$ $\square$

EXAMPLE RV X HAS A NORMAL DISTRIBUTION WITH MEAN $\mu \in \mathbb{R}$ AND VARIANCE $\sigma^2 > 0$.

$$f_X(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \quad ; \quad x \in \mathbb{R}$$

GIVEN $\{x_1, x_2, \ldots, x_n\}$; FIND ESTIMATES OF μ AND σ^2

SOLUTION: $L(\mu,\sigma^2; x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2\right]$

$$\ln L(\mu,\sigma^2; x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n}\left[\ln\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right) - \frac{1}{2\sigma^2}\left(\frac{x_i-\mu}{\cancel{\sigma}}\right)^2\right]$$

$$= n \ln\left(\frac{1}{\sqrt{2\pi\,\sigma^2}}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial \ln L}{\partial \mu} = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} 2(x_i-\mu)(-1) = 0 \quad — ①$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} - \frac{1}{2\sigma^4}(-1)\sum_{i=1}^{n}(x_i-\mu)^2 = 0 \quad — ②$$

① YIELDS $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$

② YIELDS $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$

$$\boxed{\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\hat{\mu})^2}$$

DO $\hat{\mu}$ AND $\hat{\sigma}^2$ MAXIMIZE $\ln L(\mu,\sigma^2; x_1, x_2, \ldots, x_n)$?

WE TAKE SECOND DERIVATIVES.

$$\frac{\partial \ln L}{\partial \mu} = +\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu) \; ; \; \frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial^2 \ln L}{\partial \mu^2} = -\frac{n}{\sigma^2} \; ; \; \frac{\partial^2 \ln L}{\partial \mu\,\partial \sigma^2} = -\frac{1}{\sigma^4}\sum_{i=1}^{n}(x_i-\mu)$$

$$\frac{\partial^2 \ln L}{\partial(\sigma^2)^2} = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^{n}(x_i-\mu)^2$$

THE SCOND DERIVATIVES ATL THE ESTIMATED VALUES OF μ AND σ^2 ARE:

$$\frac{\partial^2 \ln L}{\partial \mu^2} = -\frac{n}{\hat{\sigma}^2} \; ; \; \frac{\partial^2 \ln L}{\partial\mu\,\partial\sigma^2} = -\frac{1}{\hat{\sigma}^4}\left[n\hat{\mu} - n\hat{\mu}\right] = 0$$

$$\frac{\partial^2 \ln L}{\partial(\sigma^2)^2} = \frac{n}{2\hat{\sigma}^4} - \frac{n}{\hat{\sigma}^4} = -\frac{n}{2\hat{\sigma}^4}$$

HESSIAN $= H = \begin{bmatrix} -\frac{n}{\hat{\sigma}^2} & 0 \\ 0 & -\frac{n}{2\hat{\sigma}^4} \end{bmatrix}$ IS NEGATIVE DEFINITE

$\Rightarrow$ $\hat{\mu}$ AND $\hat{\sigma}^2$ MAXIMIZE $\ln L(\cdot\,;\mu;\sigma^2;x_1,x_2,\ldots,x_n)$

$\square$