Lecture 27

Last time we discussed point estimators and how to evaluate and compare them. We have not yet discussed how to obtain point estimators. Today we discuss two of the multitude of methods that might be used.

Definition Moments are of the form X^k with $k=1,2,3,\ldots$

$E[X^k]$ — theoretical moment

$\dfrac{\sum_{i=1}^{n} X_i^k}{n}$ — sample moment.

Method of Moments

Suppose we have a density function $f(x; \theta_1, \theta_2, \ldots, \theta_n)$ with the θ_i being the unknown parameters then by equating the first n sample moments with the first n theoretical moments and then solving the resulting system of equations will give us estimates of $\hat{\theta}_1, \ldots, \hat{\theta}_n$

Example 1

Consider a random sample from a distribution with two unknown parameters mean μ and variance σ^2. What are the MME (method of moments estimators) of μ and σ^2?

recall that $E[X] = \mu$ and $\sigma^2 = Var(X) = E[X^2] - [E[X]]^2$

$\Rightarrow E[X^2] = \sigma^2 + \mu^2$

$$\Rightarrow (1) \quad \frac{\sum X_i}{n} = \hat{\mu}$$

$$(2) \quad \frac{\sum X_i^2}{n} = \hat{\mu}^2 + \hat{\sigma}^2$$

From (1) $\mu^2 = \bar{X}^2$

$$\Rightarrow \frac{\sum X_i^2}{n} - \bar{X}^2 = \sigma^2$$

$$\Rightarrow \frac{\sum X_i^2 - n\bar{X}^2}{n} = \hat{\sigma}^2 \Rightarrow \frac{\sum (X_i - \bar{X})^2}{n} = \hat{\sigma}^2$$

Example 2

Consider exponential data ie

$$f(x;\lambda) = \lambda e^{-\lambda x} \quad x > 0$$

$$E[X] = \frac{1}{\lambda}$$

$$\Rightarrow \frac{\sum X_i}{n} = \frac{1}{\hat{\lambda}}$$

$$\Rightarrow \hat{\lambda} = \frac{1}{\bar{X}}$$

Example 3

Suppose $f(x;\theta) = \theta x^{\theta-1} \quad 0 < x < 1 \quad \theta > 0$

$= 0 \quad$ otherwise

$$E[X] = \int_0^1 \theta x^{\theta} \, dx$$

$$= \frac{\theta}{\theta+1} x^{\theta+1} \Big|_0^1 = \frac{\theta}{\theta+1}$$

$$\frac{\sum X_i}{n} = \frac{\hat{\theta}}{\hat{\theta}+1}$$

$$\Rightarrow \bar{x}(\hat{\theta}+1) = \hat{\theta}$$

$$\Rightarrow \bar{x}\hat{\theta} + \bar{x} - \hat{\theta} = 0$$

$$\Rightarrow \hat{\theta}(\bar{x}-1) = -\bar{x}$$

$$\Rightarrow \hat{\theta} = \frac{\bar{x}}{1-\bar{x}}$$

Example 4

Suppose $f(x;\theta) = \theta^2 x e^{-\theta x}$; $0 < x$ $\quad \theta > 0$

$= 0$ otherwise

$$E[X] = \int_0^\infty (\theta x)^2 e^{-\theta x}\,dx$$

$$= -\frac{(2+2\theta x+x^2\theta^2)}{\theta} e^{-\theta x}\Big|_0^\infty$$

$$= \frac{2}{\theta} \Rightarrow \bar{x} = \frac{2}{\hat{\theta}} \Rightarrow \hat{\theta} = \frac{2}{\bar{x}}$$

Example 5

Suppose we have gamma distribution data.

Recall that $E[X] = \alpha\beta$ and $Var(X) = \alpha\beta^2$

so $E[X^2] = \alpha\beta^2 + \alpha^2\beta^2 = \alpha\beta^2(1+\alpha)$

So MME are solutions of

(1) $\bar{X} = \hat{\alpha}\hat{\beta}$

(2) $\dfrac{\sum X_i^2}{n} = \hat{\alpha}\hat{\beta}^2(1+\hat{\alpha})$

from equation $1 \Rightarrow \hat{\alpha}^2\hat{\beta}^2 = (\bar{X})^2$ also from (1) $\alpha = \dfrac{\bar{X}}{\beta}$

so

$$\frac{\sum X_i^2}{n} = \hat{\alpha}\hat{\beta}^2 + \bar{X}^2$$

$$\Rightarrow \frac{\sum X_i^2}{n} = \bar{X}\hat{\beta} + \bar{X}^2$$

$$\Rightarrow \frac{\frac{\sum X_i^2}{n} - \bar{X}^2}{\bar{X}} = \hat{\beta}$$

and $\hat{\alpha} = \dfrac{\bar{X}^2}{\frac{\sum X_i^2}{n} - \bar{X}^2}$

Maximum Likelihood

The joint density of n r.v $X_1, \ldots X_n$ evaluated at $x_1, \ldots x_n$ say $f(x_1, \ldots x_n; \theta_1, \ldots, \theta_p)$ is refered to as the likelihood function $L(\underset{\sim}{\theta})$.

The $\hat{\theta}_1, \ldots \hat{\theta}_p$ which maximize this function are called the <u>Maximum Likelihood Estimators</u> (MLE).

<u>Note</u> if $X_1, \ldots, X_n$ are independent sample from some dist $f(x; \theta)$ then

$$L(\underset{\sim}{\theta}) = f(x_1; \underset{\sim}{\theta}) \cdots f(x_n; \underset{\sim}{\theta}) = \prod_{i=1}^{n} f(x_i; \underset{\sim}{\theta})$$

How do we find $\hat{\theta}_1, \ldots, \hat{\theta}_p$? differentiate $L(\underset{\sim}{\theta})$ wrt to each of $\theta_1, \ldots \theta_p$

system of equations

$$\frac{\partial}{\partial \theta_j} L(\underset{\sim}{\theta}) = 0$$

solve for $\hat{\theta}_1, \ldots, \hat{\theta}_p$.

Note It is almost always easier to deal with the log likelihood $l(\theta) = \ln L(\theta)$.

Note
MLE have an invariance property ie
if $\hat{\underset{\sim}{\theta}}$ is the MLE of $\underset{\sim}{\theta}$ then the mle of any function $h(\underset{\sim}{\theta})$ is $h(\hat{\underset{\sim}{\theta}})$

Note MLE have some asymptotic behaviors that make them desirable. Specifically as $n \to \infty$ the MLE $\hat{\theta}$ of θ is approximately unbiased and has minimum variance. ie mle $\hat{\theta}$ is asymptotically MVUE for θ.

Note because of these asymptotic results MLE is one of the most widely used techniques in statistics. However, it should be noted that it is not always possible to find the Mle using calculus.

Example 1

consider X_i from $N(\mu, \sigma^2)$. what are MLE of μ, σ^2?

recall that

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

so

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

and

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

$$= -\frac{1}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{\sum(x_i-\mu)^2}{2\sigma^2}$$

$$\frac{\partial l(\mu, \sigma^2)}{\partial \mu} = \frac{2\sum(x_i-\mu)}{2\sigma^2} = 0 \quad (1)$$

$$\frac{\partial l(\mu,\sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum (x_i-\mu)^2}{2\sigma^4} = 0 \quad (2)$$

from (1) $\quad \hat{\mu} = \bar{x}$

substituting into (2) gives

$$\hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

Example 2

Consider x_i from the gamma distribution

$$f(x;\alpha,\beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right)$$

Likelihood

$$L(\alpha,\beta) = \prod_{i=1}^{n} f(x_i;\alpha,\beta)$$

$$= \frac{1}{\beta^{n\alpha}[\Gamma(\alpha)]^n} \left[\prod x_i\right]^{\alpha-1} \exp\left(-\frac{\sum x_i}{\beta}\right)$$

So log-likelihood

$$\ell(\alpha,\beta) = -n\alpha\log(\beta) - n\log\Gamma(\alpha) + (\alpha-1)\log\prod x_i - \frac{\sum x_i}{\beta}$$

$$(1)\quad \frac{\partial \ell(\alpha,\beta)}{\partial \alpha} = -n\log(\beta) - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \log\prod x_i = 0$$

$$(2)\quad \frac{\partial \ell(\alpha,\beta)}{\partial \beta} = -\frac{n\alpha}{\beta} + \frac{\sum x_i}{\beta^2} = 0$$

from (2) $\quad \hat{\beta} = \dfrac{\bar{x}}{\hat{\alpha}}$

from (1) $\quad \log(\hat{\alpha}) - \dfrac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} - \log\left(\dfrac{\bar{x}}{(\prod x_i)^{1/n}}\right) = 0$

this is called the digamma function $\Psi(\hat{\alpha})$

No closed solution to these equations.

Instead we would have to solve the second equation numerically.

Note the the MLE differed from the MME.