Stats 2MB3, Tutorial 6

Feb 27th, 2015

Maximum Likelihood Estimation

- The likelihood function is the joint density of all the observations, $X_1,\dots,X_n$,

  $f(x_1,\dots,x_n;\theta_1,\dots,\theta_m)$ where $\theta_1,\dots,\theta_m$ are parameters.
- We need to find $\hat{\theta}_1,\dots,\hat{\theta}_m$ such that

  $$f(x_1,\dots,x_n;\hat{\theta}_1,\dots,\hat{\theta}_m) = \max_{\theta_1,\dots,\theta_m} f(x_1,\dots,x_n;\theta_1,\dots,\theta_m)$$

  for all $\theta_1,\dots,\theta_m$. These estimators are called maximum likelihood estimators (MLE).

Ex 15, page 255

- Let X_1 , X_2 ,…, X_n represent a random sample from a Rayleigh distribution with pdf

$$f(\mathrm{x};\theta) = \frac{x}{\theta} e^{-x^2/(2\theta)}, x > 0$$

- a) It can be shown that $E(X^2)=2\theta$. Use this fact to construct an unbiased estimator of θ based on $\sum X_i^2$ (and use rules of expected value to show that it is unbiased).

- b) Estimate θ from the following n=10 observations on vibratory stress of a turbine blade under specified conditions:

  16.88, 10.23, 4.59, 6.66, 13.68, 14.23, 19.87, 9.40, 6.51, 10.95.

Solution:

a) Since $E(X^2) = 2\theta$ implies $\sum X_i^2 / n = 2\hat{\theta}$, then

$\hat{\theta} = \frac{\sum X_i^2}{2n}$ is the unbiased estimator of θ.

- b) $\sum X_i^2 = 1490.1058$, then $\hat{\theta} = \frac{1490.1058}{20} = 74.505$.

Ex 28, page 265

- Let X_1 , X_2 ,…, X_n represent a random sample from a Rayleigh distribution with density function given in Ex 15. Determine

  a) The maximum likelihood estimator of θ, and then calculate the estimate for the vibratory stress data. Is this estimator the same as the unbiased estimator suggested in Ex 15?

- b) The MLE of the median of the vibratory stress distribution.

- Solution:

  a) From the likelihood function

$$L(\theta; \mathrm{x}) = \prod_{i=1}^{n} \frac{x_i}{\theta} \exp(-\frac{x_i^2}{2\theta}) = \frac{\prod_{i=1}^{n} x_i}{\theta^n} \exp(-\frac{\sum x_i^2}{2\theta})$$

then take logarithm for both sides

$$l(\theta; \mathrm{x}) = \log L(\theta; \mathrm{x}) = \sum_{i=1}^{n} \log x_i - n \log \theta - \frac{\sum x_i^2}{2\theta}$$

- Take the derivative with respect to θ and obtain

$$\frac{\partial l}{\partial \theta} = -\frac{n}{\theta} + \frac{\sum x_i^2}{2\theta^2} = 0 \Rightarrow \hat{\theta} = \frac{\sum x_i^2}{2n}$$

which is the same as the expression of unbiased estimator in Ex 15.

- b)

  If we set m is the median, then P(X<m)=1/2.

$$\int_0^m \frac{x}{\theta}\exp\left(-\frac{x^2}{2\theta}\right)\mathrm{dx} = 1-\exp\left(-\frac{m^2}{2\theta}\right) = \frac{1}{2}$$

$$\Rightarrow m = \sqrt{(2\log 2)\theta}$$

By Invariance Principle, the MLE of the median

$$\widehat{m} = \sqrt{(2\log 2)\hat{\theta}} = \sqrt{(2\log 2)\frac{\sum x_i^2}{2n}} = \sqrt{\log 2\frac{\sum x_i^2}{n}}$$

Plug in the actual data and we get $\widehat{m} = 6.698$.

Exercise 3

- Consider a random sample of random variables $X_1, X_2, \ldots, X_n$ from a negative binomial population with a known parameter r and unknown parameter p.

a) Derive the maximum likelihood estimator of p.

- b) A numerical sample of size 6 from this population was gathered, resulting in the values.

  7, 2, 10, 3, 5, 14

  Calculate the maximum likelihood estimate of p arising from this particular numerical sample.

- a)

  The likelihood function is

$$L(p; x, r) = \prod_{i=1}^{n} \binom{x_i + r - 1}{r - 1} p^r (1 - p)^{x_i} = \prod_{i=1}^{n} \binom{x_i + r - 1}{r - 1} p^{nr} (1 - p)^{\sum x_i}$$

  and take the logarithm

$$l(p; x, r) = \log L(p; x, r) = \sum_{i=1}^{n} \log \binom{x_i + r - 1}{r - 1} + nr \log(p) + \sum x_i \log(1 - p)$$

  take the derivative with respect to p

$$\frac{\partial l}{\partial p} = \frac{nr}{p} - \frac{\sum x_i}{1 - p} = 0 \Rightarrow \hat{p} = \frac{nr}{nr + \sum x_i} \quad .$$

- b)

  Plug in the real numbers,

  $$\hat{p} = \frac{nr}{nr + \sum x_i} = \frac{6r}{6r + 41}$$. (r is known)