## Maximum likelihood Estimation

Up to now, we know the $\bar{x}$ as sample mean is a point estimator of population mean $\mu$, the $s^2$ is an estimator of $\sigma^2$ and the estimator $x/_n$ for a binomial parameter p is an estimator but there are many situations in which it is not at all obvious what the proper estimator should be. **Maximum likelihood Estimation** is one of the most important approaches in all of statistical inference.

**Definition:** Given independent observations $x_1, x_2, \dots, x_n$ from a probability density function (continuous case) or probability mass function (discrete case) $f(x, \theta)$, the maximum likelihood estimator $\hat{\theta}$ is that which maximizes the likelihood function.

**Procedure**

1- Find $L(x,\theta) = f(x_1,\theta)f(x_2,\theta)\dots f(x_n,\theta)$
2- Calculate $L_n L(x,\theta)$
3- Obtain $dL_nL(x,\theta)/_{d\theta}$
4- Let $dL_nL(x,\theta)/_{d\theta} = 0$ and calculate $\hat{\theta}$

Remind: Properties of ln

$ln\, xy = lnx + lny$

$ln\frac{x}{y} = lnx - lny$

$lnx^y = ylnx$

$lne = 1$

**Example 1**: consider a poisson distribution with probability mass funtion

$$p(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \quad x = 0,1, \dots$$

Suppose that a random sample $x_1, x_2, \dots, x_n$ is taken from the distribution. What is the maximun likelihood estimete of $\lambda$?

1- $L(x,\lambda) = p(x_1,\lambda)p(x_2,\lambda)\dots p(x_n,\lambda) = \frac{e^{-\lambda}\lambda^{x_1}}{x_1!} \times \frac{e^{-\lambda}\lambda^{x_2}}{x_2!} \times \dots \times \frac{e^{-\lambda}\lambda^{x_n}}{x_n!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$

2- $L_nL(x,\lambda) = ln\frac{e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} = lne^{-n\lambda} + ln\lambda^{\sum_{i=1}^n x_i} - ln\prod_{i=1}^n x_i! = -n\lambda lne + \sum_{i=1}^n x_i \ln\lambda - \sum_{i=1}^n \ln x_i! = -n\lambda + \sum_{i=1}^n x_i \ln\lambda - \sum_{i=1}^n \ln x_i!$

3- $dL_nL(x,\lambda)/_{d\lambda} = -n + \frac{\sum_{i=1}^n x_i}{\lambda}$

4- $dL_nL(x,\lambda)/_{d\lambda} = 0 \to -n + \frac{\sum_{i=1}^n x_i}{\lambda} = 0 \to \frac{\sum_{i=1}^n x_i}{\lambda} = n \to \hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}$

**Example 2**: A sample of 3 observations, $(x_1 = 0.4,\ x_2 = 0.7, x_3 = 0.9)$ is collected from a continuous distribution with density $f(x) = \theta x^{\theta-1} \quad 0 < x < 1$. What is the maximum likelihood of $\theta$

1- $L(x,\theta) = f(x_1,\theta)f(x_2,\theta)f(x_3,\theta) = \theta {x_1}^{\theta-1} \times \theta {x_2}^{\theta-1} \times \theta {x_3}^{\theta-1} = \theta^3 \prod_{i=1}^{3} {x_i}^{\theta-1}$
2- $L_nL(x,\theta) = \ln\theta^3 + \ln\prod_{i=1}^{3} {x_i}^{\theta-1} = 3\ln\theta + (\theta-1)\sum_{i=1}^{3}\ln x_i$
3- $dL_nL(x,\theta)/_{d\theta} = \frac{3}{\theta} + \sum_{i=1}^{3}\ln x_i$
4- $dL_nL(x,\theta)/_{d\theta} = 0 \to \frac{3}{\theta} + \sum_{i=1}^{3}\ln x_i = 0 \to \hat{\theta} = \frac{-3}{\sum_{i=1}^{3}\ln x_i} = \frac{-3}{\ln 0.4 + \ln 0.7 + \ln 0.9} = 2.1766$

**Example 3**: suppose that x is a descrete random variable with the following probability mass function. Where $0 \le \theta \le 1$ is a parameter. The following 10 independent observations.

| x | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| P(x) | $2\theta/_3$ | $\theta/_3$ | $2(1-\theta)/_3$ | $(1-\theta)/_3$ |

Where taken from such a distribution: (3, 0, 2, 1, 3, 2, 1, 0, 2, 1). What is the maximum likelihood estimate of $\theta$.

1- $L(x,\theta) = p(x_0,\theta)p(x_1,\theta)p(x_2,\theta)p(x_3,\theta) = (\frac{2\theta}{3})^2(\frac{\theta}{3})^3(2\frac{1-\theta}{3})^3(\frac{1-\theta}{3})^2$
2- $L_nL(x,\theta) = 2\ln\frac{2\theta}{3} + 3\ln\frac{\theta}{3} + 3\ln 2\frac{1-\theta}{3} + 2\ln\frac{1-\theta}{3} = 2\left(\ln\frac{2}{3} + \ln\theta\right) + 3\left(\ln\frac{1}{3} + \ln\theta\right) + 3\left(\ln\frac{2}{3} + \ln(1-\theta)\right) + 2(\ln\frac{1}{3} + \ln(1-\theta))$
3- $dL_nL(x,\theta)/_{d\theta} = \frac{2}{\theta} + \frac{3}{\theta} - \frac{3}{1-\theta} - \frac{2}{1-\theta} = \frac{5}{\theta} - \frac{5}{1-\theta}$
4- $dL_nL(x,\theta)/_{d\theta} = 0 \to \frac{5}{\theta} - \frac{5}{1-\theta} = 0 \to \theta = 1-\theta \to \hat{\theta} = \frac{1}{2}$

**Example 4**:
The Pareto distribution has a probability density function
$f(x,\alpha,\theta) = \theta\alpha^\theta x^{-\theta-1} \quad for\ x \ge \alpha, \theta > 1$
Where $\alpha$ and $\theta$ are positive parameters of the distribution. Assume that $\alpha$ is known and $x_1, x_2, \dots, x_n$ is a random sample of size n.
- Find the maximum likelihood estimator for $\theta$
- Assume that $\alpha$= 2, Estimate $\theta$ based on these data 3,5,2,3,4,1,4,3,3,3

1- $L(x,\theta) = f(x_1,\theta)f(x_2,\theta)\dots f(x_n,\theta) = (\theta\alpha^\theta {x_1}^{-\theta-1})(\theta\alpha^\theta {x_2}^{-\theta-1})\dots(\theta\alpha^\theta {x_n}^{-\theta-1}) = \theta^n\alpha^{n\theta}\prod_{i=1}^{n} {x_i}^{-\theta-1}$
2- $L_nL(x,\theta) = n\ln\theta + n\theta\ln\alpha + (-\theta-1)\sum_{i=1}^{n} ln\ x_i$
3- $dL_nL(x,\theta)/_{d\theta} = \frac{n}{\theta} + n\ln\alpha - \sum_{i=1}^{n} ln\ x_i$
4- $dL_nL(x,\theta)/_{d\theta} = 0 \to \frac{n}{\theta} + n\ln\alpha - \sum_{i=1}^{n} ln\ x_i = 0 \to \frac{n+\theta(n\ln\alpha - \sum_{i=1}^{n} ln\ x_i)}{\theta} \to$

$\hat{\theta} = \frac{-n}{n\ln\alpha - \sum_{i=1}^{n} ln\ x_i} = \frac{-10}{(10ln2)-(ln3+ln5+ln2+ln3+ln4+ln1+ln4+ln3+ln3+ln3)} = 2.73$