6. Point Estimation

Chapter 6: Point Estimation

- **6.1. Some General Concepts of Point Estimation**
- **6.2. Methods of Point Estimation**

6.2 Methods of Point Estimation

- Two “constructive” methods for obtaining point estimators

> **Method of Moments**

> **Maximum Likehood Estimation**

6.2 Methods of Point Estimation

The Method of Moment

- The basic idea of this method is **equate** certain sample characteristics, such as the mean, **to the corresponding population** expected values. Then solving theses equations for unknown parameter values yields the estimators.

- **Moments**

Let $X_1, X_2, \ldots, X_n$ be a random sample from a pmf or pdf $f(x)$. For $k = 1, 2, 3, \ldots,$ the kth population moment, or kth moment of the distribution $f(x)$, is $E(X^k)$. The kth sample moment is .

$$(1/n)\sum_{i=1}^{n} X_i^k$$

6.2 Methods of Point Estimation

- **Moment Estimator**

Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with pmf or pdf $f(x;\theta_1,\ldots,\theta_m)$, where $\theta_1,\ldots,\theta_m$ are parameters whose values are unknown. Then the moment estimators $\hat{\theta}_1,\ldots,\hat{\theta}_m$ are obtained by equating the first m sample moments to the corresponding first m population moments and solving for $\theta_1,\ldots,\theta_m$.

With the given sample x_i

$$(1/n)\sum_{i=1}^{n} X_i^k \xrightarrow{\text{n is large}} E(X^k)$$

With unkonwn θ_i

K-th sample moment

K-th population moment

6.2 Methods of Point Estimation

General Algorithm :

$$\begin{cases} \mu_1 = \mu_1(\theta_1, \theta_2, \ldots, \theta_m) \\ \mu_2 = \mu_2(\theta_1, \theta_2, \ldots, \theta_m) \\ \ldots \\ \mu_m = \mu_m(\theta_1, \theta_2, \ldots, \theta_m) \end{cases} \Rightarrow \begin{cases} \theta_1 = \theta_1(\mu_1, \mu_2, \ldots, \mu_m) \\ \theta_2 = \theta_2(\mu_1, \mu_2, \ldots, \mu_m) \\ \ldots \\ \theta_m = \theta_m(\mu_1, \mu_2, \ldots, \mu_m) \end{cases}$$

The first m population moments

The solution of equations

Use the first m sample moment

$$A_l = \frac{1}{n}\sum_{i=1}^{n} X_i^l, l = 1, 2 \ldots m$$

to represent the population moments μ_l

$$\theta_i = \theta_i(A_1, A_2, \ldots, A_m), i = 1, \ldots, m$$

6.2 Methods of Point Estimation

- **Example 6.12**

Let $X_1, X_2, \ldots, X_n$ represent a random sample of service times of n customers at a certain facility, where the underlying distribution is assumed exponential with parameter λ. How to estimate λ by using the method of moments?

Step #1: The 1st population moment $E(X) = 1/\lambda$

then we have $\lambda = 1/ E(X)$

Step #2: Use the 1st sample moment $\overline{X}$ to represent 1st population moment $E(X)$, and get the estimator

$$\hat{\lambda} = 1/\,\overline{X}$$

6.2 Methods of Point Estimation

- **Example 6.13**

  Let $X_1, \ldots, X_n$ be a random sample from a gamma distribution with parameters α and β. Its pdf is

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}} & x\geq 0\\ 0 & \text{otherwise}\end{cases}$$

  There are two parameter need to be estimated, thus, consider the first two moments

- **Example 6.13 (Cont')**

Step #1: $E(X) = \mu = \alpha\beta$

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

$$\Rightarrow \quad \alpha = \frac{E(X)^2}{E(X^2) - E(X)^2}, \beta = \frac{E(x^2) - E(X)^2}{E(X)}$$

Step #2: $\bar{X} \rightarrow E(X), \frac{1}{n}\sum X_i^2 \rightarrow E(X^2)$

$$\hat{\alpha} = \frac{\bar{X}^2}{\frac{1}{n}\sum X_i^2 - \bar{X}^2} \qquad \hat{\beta} = \frac{\frac{1}{n}\sum X_i^2 - \bar{X}^2}{\bar{X}}$$

- **Example 6.14**

Let $X_1, \ldots, X_n$ be a random sample from a generalized negative binomial distribution with parameters r and p. Its pmf is

$$nb(x;r,p)=\binom{x+r-1}{r-1}p^r(1-p)^x,\quad x=0,1,2,\ldots.$$

Determine the moment estimators of parameters r and p.

Note: There are two parameters needs to estimate, thus the first two moments are considered.

6.2 Methods of Point Estimation

- **Example 6.14 (Cont')**

Step #1: $E(X) = r(1-p)/p$

$$E(X^2) = r(1-p)(r - rp + 1)/p^2$$

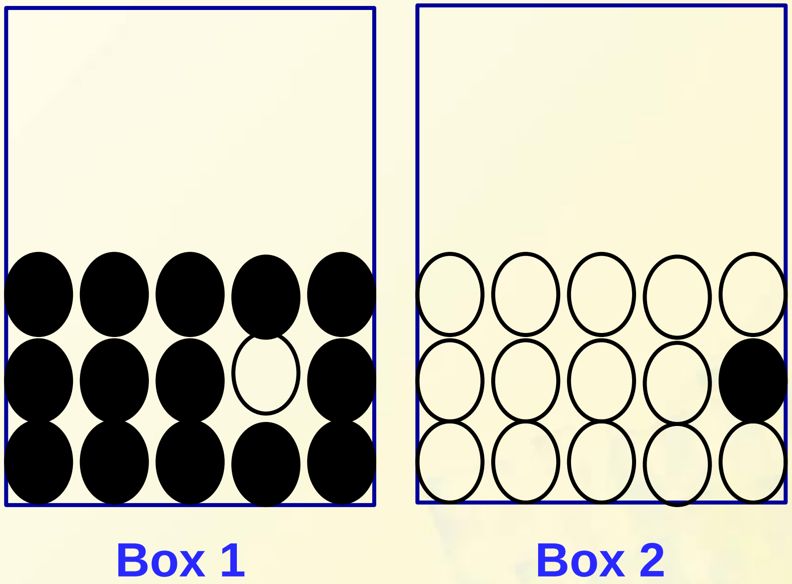
$$\Rightarrow \quad p = \frac{E(X)}{\mathrm{E(X^2)\text{-}E(X)^2}}, r = \frac{E(X)^2}{E(X^2) - E(X)^2 - E(X)}$$

Step #2: $\bar{X} \to E(X), \frac{1}{n}\sum X_i^2 \to E(X^2)$

$$\hat{p} = \frac{\bar{X}}{\frac{1}{n}\sum X_i^2 - \bar{X}^2} \qquad \hat{r} = \frac{\bar{X}^2}{\frac{1}{n}\sum X_i^2 - \bar{X}^2 - \bar{X}}$$

6.2 Methods of Point Estimation

- Maximum Likelihood Estimation (Basic Idea)



Experiment:
We firstly randomly choose a box,
And then randomly choose a ball.

Q: If we get a white ball, which box has the Maximum Likelihood being chosen?

$$P(W \mid Box1) = 1/15$$

$$P(W \mid Box2) = 14/15$$

Maximum Likelihood Estimation (Basic Idea)

The basis idea of the method of maximum likelihood is that we look at the sample values and then choose as our estimates of the unknown parameters the values for which the probability or probability density of getting the sample values is a maximum.

6.2 Methods of Point Estimation

- Maximum Likelihood Estimation (Basic Idea)

Q: What is the probability p of hitting the target?

$$f(\mathrm{p}) = p^3(1-p)^{5-3} = p^3(1-p)^2$$

$f(0.2) \approx 0.0051$ $f(0.4) \approx 0.0230$ $f(0.6) \approx 0.0346$ $f(0.8) \approx 0.0205$ ...

The best one among the four options

6.2 Methods of Point Estimation

- **Example 6.15**

A sample of ten new bike helmets manufactured by a certain company is obtained. Upon testing, it is found that the first, third, and tenth helmets are flawed, whereas the others are not. Let $p = P(\text{flawed helmet})$ and define $X_1, \ldots, X_{10}$ by $X_i = 1$ if the ith helmet is flawed and zero otherwise. Then the observed x_i's are 1,0,1,0,0,0,0,0,0,1.

The Joint pmf of the sample is

$$f(x_1, x_2, \ldots, x_{10}) = p(1-p)p\cdots p = p^3(1-p)^7$$

For what value of *p* is the observed sample most likely to have occurred? Or, equivalently, what value of the parameter *p* should be taken so that the joint pmf of the sample is maximized?

- **Example 6.15 (Cont')**

$$f(x_1, x_2, \ldots, x_{10}) = p(1-p)p \cdots p = p^3(1-p)^7$$

$$\ln[f(x_1, x_2, \ldots. x_{10}; p)] = 3\ln(p) + 7\ln(1-p)$$

Equating the derivative of the logarithm of the pmf to zero gives the maximizing value **(why?)**

$$\frac{d}{dp}\ln[f(x_1, x_2, \ldots. x_{10}; p)] = \frac{3}{p} - \frac{7}{1-p} = 0 \Rightarrow p = \frac{3}{10} = \frac{x}{n}$$

where x is the observed number of successes (flawed helmets). The estimate of p is now $\hat{p} = 3/10$. It is called the **maximum likelihood estimate** because for fixed $x_1, \ldots, x_{10}$, it is the parameter value that maximizes the likelihood of the observed sample.

6.2 Methods of Point Estimation

▪ Maximum Likelihood Estimation

Let $X_1, X_2, \ldots, X_n$ have joint pmf or pdf

$$f(x_1, x_2, \ldots, x_n; \theta_1, \ldots, \theta_m)$$

where the parameters $\theta_1, \ldots, \theta_m$ have unknown values. When $x_1, \ldots, x_n$ are the observed sample values and f is regarded as a function of $\theta_1, \ldots, \theta_m$, it is called the likelihood function.

The maximum likelihood estimates(mle's) $\hat{\theta}_1, \ldots, \hat{\theta}_m$ are those values of the θ_i's that maximize the likelihood function, so that

$$f(x_1, \ldots, x_n; \hat{\theta}_1, \ldots, \hat{\theta}_m) \geq f(x_1, \ldots, x_n; \theta_1, \ldots, \theta_m) \quad \text{for all } \theta_1, \ldots, \theta_m$$

When the X_i's are substituted in place of the x_i's, the maximum likelihood estimators result.

6.2 Methods of Point Estimation

- **Three steps**

1. Write the joint pmf/pdf (i.e. Likelihood function)

$$f(x_1, x_2, \ldots, x_n; \theta_1, \ldots, \theta_m) = \prod_{i=1}^{n} f(x_i; \theta_1, \ldots, \theta_m)$$

2. Get the ln(likelihood) **(if necessary)**

$$\ln[f(x_1, x_2, \ldots, x_n; \theta_1, \ldots, \theta_m)] = \sum_{i=1}^{n} \ln(f(x_i; \theta_1, \ldots, \theta_m))$$

3. Take the partial derivative of ln(f) with respect to θ_i, equal them to 0, and solve the resulting m equations.

$$\frac{d}{d\theta_i} \ln[f(x_1, x_2, \ldots, x_n; \theta_1, \ldots, \theta_m)] = 0$$

6.2 Methods of Point Estimation

- **Example 6.16**

Suppose X_1, X_2, …, X_n is a random sample from an exponential distribution with the unknown parameter λ. Determine the maximum likelihood estimator of λ.

The joint pdf is (independence)

$$f(x_1,\dots,x_n;\lambda) = (\lambda e^{-\lambda x_1})\cdots(\lambda e^{-\lambda x_n}) = \lambda^n e^{-\lambda \sum x_i}$$

The ln(likelihood) is $\ln[f(x_1,\dots,x_n;\lambda)] = n\ln(\lambda) - \lambda\sum x_i$

Equating to zero the derivative w.r.t. λ:

The estimator

$$\frac{d\ln[f(x_1,\dots,x_n;\lambda)]}{d\lambda} = \frac{n}{\lambda} - \sum x_i = 0 \Rightarrow \lambda = \frac{n}{\sum x_i} = \frac{1}{\bar{x}} \Rightarrow \hat{\lambda} = 1/\bar{X}$$

6.2 Methods of Point Estimation

- **Example 6.17**

Let $X_1, X_2, \ldots, X_n$ is a random sample from a normal distribution $N(\mu, \sigma^2)$. Determine the maximum likelihood estimator of μ and σ^2.

The joint pdf is

$$f(x_1, \ldots, x_n; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_2-\mu)^2}{2\sigma^2}} \cdots \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_n-\mu)^2}{2\sigma^2}} = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} e^{-\frac{\sum (x_i-\mu)^2}{2\sigma^2}}$$

$$\ln[f(x_1, \ldots, x_n; \mu, \sigma^2)] = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum (x_i - \mu)^2$$

Equating to 0 the partial derivatives w.r.t. μ and σ^2, finally we have

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

Here the mle of σ^2 is not the unbiased estimator.