6. Point Estimation

Chapter 6: Point Estimation

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

6.1 Some General Concepts of Point Estimation

- In order to get some population characteristics, statistical inference needs obtain sample data from the population under study, and achieve the conclusions can then be based on the computed values of various sample quantities (statistics).

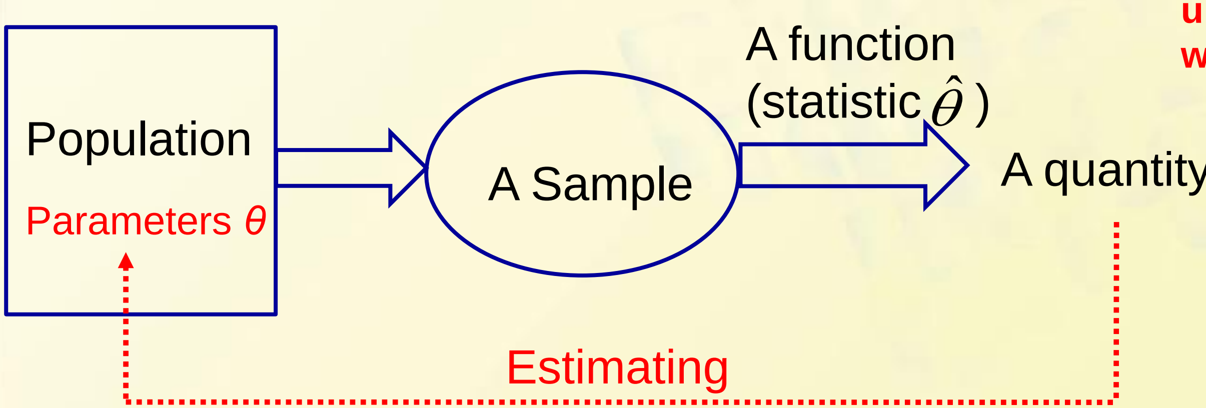
- Typically, we will use the Greek letter θ for the parameter of interest. The objective of point estimation is to select a single number, based on sample data (statistic $\hat{\theta}$), that represents a sensible value for θ.

6.1 Some General Concepts of Point Estimation

- Point Estimation

A point estimate of a parameter θ is a *single* number that can be regarded as a sensible value for θ.

A point estimate is obtained by selecting a suitable statistic and computing its value from the given sample data. The selected statistic is called the **point estimator** of θ.



Here, the type of population under study is usually known, while the parameters are unkown.

Q #1: How to get the candidate estimators based on the population?

Q #2: How to measure the candidate estimators?

6.1 Some General Concepts of Point Estimation

- Example 6.1

  The manufacturer has used this bumper in a sequence of 25 controlled crashes against a wall, each at 10 mph, using one of its compact car models. Let X = the number of crashes that result in no visible damage to the automobile. What is a sensible estimate of the parameter p = the proportion of all such crashes that result in no damage

  If X is observed to be $x = 15$, the most reasonable estimator and estimate are

$$\text{estimator} \quad \hat{p} = \frac{X}{n} \qquad\qquad \text{estimate} = \frac{x}{n} = \frac{15}{25} = 0.60$$

Example 6.2

Reconsider the accompanying 20 observations on dielectric breakdown voltage for pieces of epoxy resin first introduced in Example 4.29 (pp. 193)

24.46	25.61	26.25	26.42	26.66	27.15	27.31	27.54	27.74	27.94
27.98	28.04	28.28	28.49	28.50	28.87	29.11	29.13	29.50	30.88

The pattern in the normal probability plot given there is quite straight, so we now assume that the distribution of breakdown voltage is **normal** with mean value μ. Because normal distribution are symmetric, μ is also the median lifetime of the distribution. The given observation are then assumed to be the result of a random sample $X_1, X_2, \ldots, X_{20}$ from this normal distribution.

Example 6.2 (Cont')

Consider the following estimators and resulting estimates for μ

a. Estimator= $\overline{X}$, estimate= $\bar{x} = \sum x_i / n = 555.86/20 = \boxed{27.793}$

b. Estimator= $\tilde{X}$, estimate= $\tilde{x} = (27.94 + 27.98)/2 = \boxed{27.960}$

c. Estimator [min(X_i) + max(X_j)]/2 = the average of the two extreme lifetimes, estimate=[min(x_i)+max(x_i)]/2 = (24.46+30.88)/2 = $\boxed{27.670}$

d. Estimator = $\overline{X}_{tr(10)}$, the 10% trimmed mean (discard the smallest and largest 10% of the sample and then average)

$$\text{estimate} = \bar{x}_{tr(10)} = \frac{555.86 - 24.46 - 25.61 - 29.50 - 30.88}{16} = \boxed{27.838}$$

6.1 Some General Concepts of Point Estimation

- **Example 6.3**

In the near future there will be increasing interest in developing low-cost Mg-based alloys for various casting processes. It is therefore important to have practical ways of determining various mechanical properties of such alloys. Assume that the observations of a random sample X_1, X_2, …, X_8 from the population distribution of elastic modulus under such circumstances. We want to estimate the population variance σ^2

Method #1: sample variance

$$\hat{\sigma}^2 = S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1} = \frac{\sum X_i^2 - \left(\sum X_i\right)^2 / n}{n-1} \qquad \hat{\sigma}^2 = S^2 = \frac{\sum X_i^2 - \left(\sum X_i\right)^2 / 8}{7} \boxed{\approx 0.251}$$

Method #2: Divided by n rather than n-1

$$\hat{\sigma}^2 = S^2 = \frac{\sum (X_i - \bar{X})^2}{n} = \frac{\sum X_i^2 - \left(\sum X_i\right)^2 / n}{n} \qquad \hat{\sigma}^2 = S^2 = \frac{\sum X_i^2 - \left(\sum X_i\right)^2 / 8}{8} \boxed{\approx 0.220}$$

6.1 Some General Concepts of Point Estimation

- **Estimation Error Analysis**

Note that $\hat{\theta}$ is a function of the sample X_i's, so it is a random variable.

$$\hat{\theta} = \theta \text{ + error of estimation}$$

Therefore, an accurate estimator would be one resulting in small estimation errors, so that estimated values will be near the true value θ (unkown).

A good estimator should have the two properties:

1. unbiasedness (*i.e.* the average error should be zero)
2. minimum variance (*i.e.* the variance of error should be small)

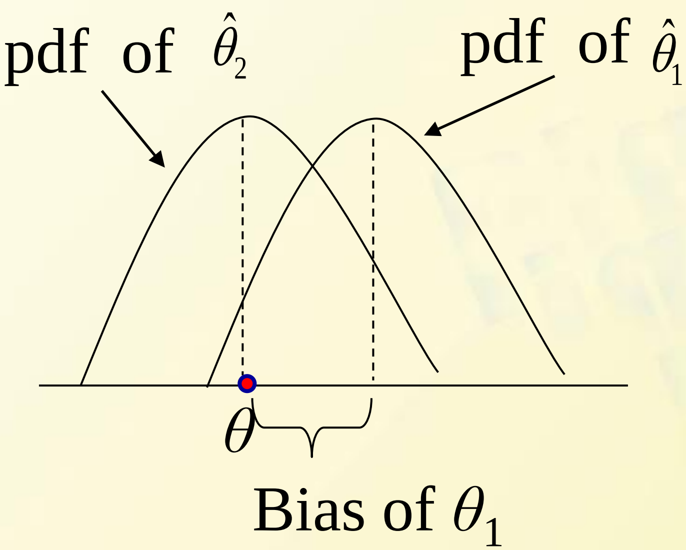
6.1 Some General Concepts of Point Estimation

- Unbiased Estimator (无偏估计)

A point estimator $\hat{\theta}$ is said to be an unbiased estimator of θ if

$E(\hat{\theta}) = \theta$ for every possible value of θ.

If $\hat{\theta}$ is not unbiased, the difference $E(\hat{\theta}) - \theta$ is called the bias of $\hat{\theta}$



Note: “centered” here means the expected vaule, not the median, of the distribution of $\hat{\theta}$ is equal to θ

6.1 Some General Concepts of Point Estimation

- Proposition

When X is a binomial rv with parameters n and p, the sample proportion $\hat{p}$ $=X/n$ is an unbiased estimator of p.

Refer to Example 6.1, the sample proportion X/n was used as an estimator of p, where X, the number of sample successes, had a binomial distribution with parameters n and p, thus

$$E(\hat{p}) = E(\frac{X}{n}) = \frac{1}{n}E(X) = \frac{1}{n}(np) = p$$

- **Example 6.4**

Suppose that X, the reaction time to a certain stimulus, has a uniform distribution on the interval from 0 to an unknown upper limit θ. It is desired to estimate θ on the basis of a random sample $X_1, X_2, \ldots, X_n$ of reaction times. Since θ is the largest possible time in the entire population of reaction times, consider as a first estimator the largest sample reaction time:

$$\hat{\theta}_1 = \max(X_1, X_2, \cdots, X_n)$$ biased estimator, why?

Since $E(\hat{\theta}_1) = \frac{n}{n+1} \cdot \theta < \theta$ (refer to Ex. 32 in pp. 279)

Another estimator $\hat{\theta}_2 = \frac{n+1}{n} \cdot \max(X_1, X_2, \cdots, X_n)$ unbiased estimator

$$E(\hat{\theta}_2) = \frac{n+1}{n} \cdot \frac{n}{n+1} \cdot \theta = \theta$$

- **Proposition**

Let X_1, X_2, …, X_n be a random sample from a distribution with mean μ and variance σ^2. Then the estimator

$$\hat{\sigma}^2 = S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$$

is an unbiased estimator of σ^2, namely $E(S^2) = \sigma^2$

Refer to pp. 245 for the proof.

However,

$$E\left(\frac{\sum (X_i - \bar{X})^2}{n}\right) = \frac{n-1}{n} E(S^2) = \frac{n-1}{n} \sigma^2 < \sigma^2$$

6.1 Some General Concepts of Point Estimation

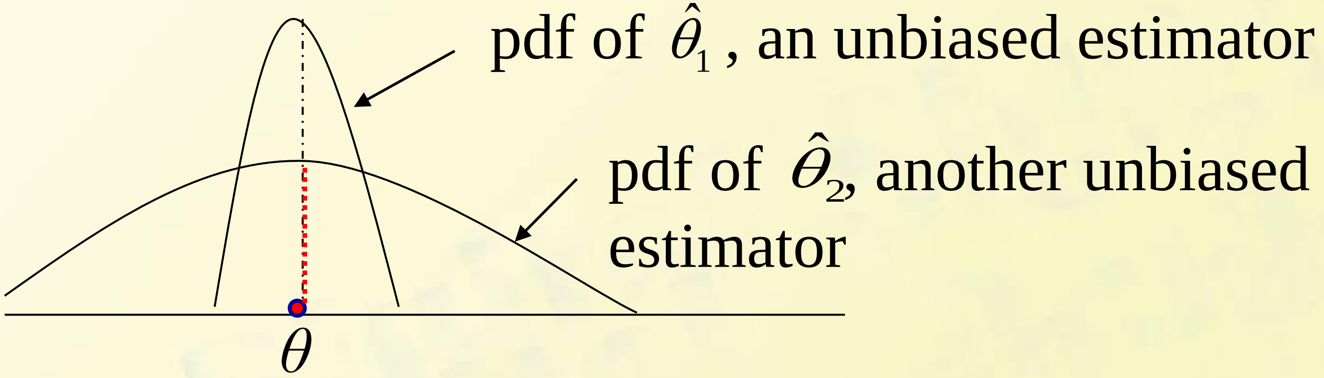
- **Proposition**

If X_1, X_2,…X_n is a random sample from a distribution with mean μ, then $\overline{X}$ is an unbiased estimator of μ. If in addition the distribution is **continuous and symmetric**, then $\tilde{X}$ and any trimmed mean are also **unbiased** estimator of μ

Refer to the estimators in Example 6.2

6.1 Some General Concepts of Point Estimation

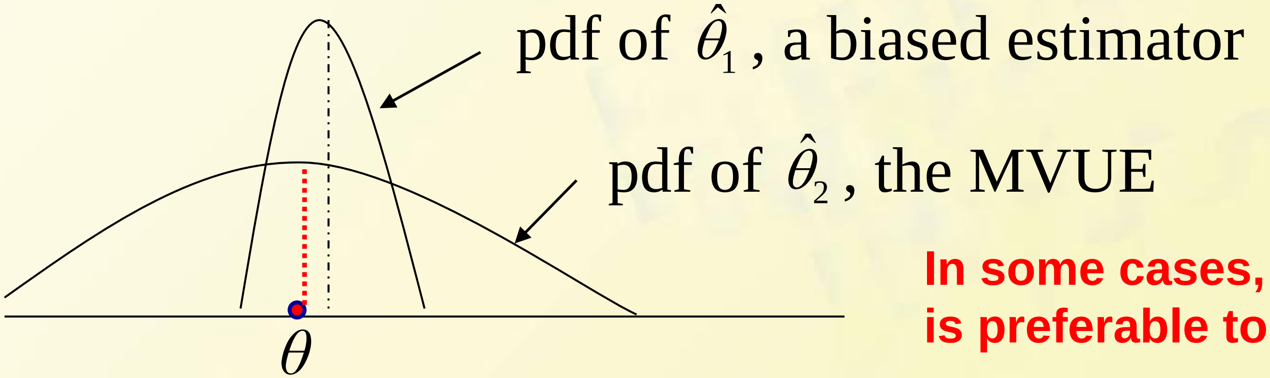
- **Estimators with Minimum Variance**



Obviously, the estimator $\hat{\theta}_1$ is better than the $\hat{\theta}_2$ in this example

6.1 Some General Concepts of Point Estimation

- **Estimator Selection**
  - When choosing among several different estimators of θ, select one that is unbiased.
  - Among all estimators of θ that are unbiased, choose the one that has **minimum variance**. The resulting $\hat{\theta}_1$ is called the minimum variance unbiased estimator (MVUE) of θ.



In some cases, a biased estimator is preferable to the MVUE

- **Example 6.6 (Ex. 6.4 Cont')**

When X_1, X_2, … X_n is a random sample from a uniform distribution on [0, θ], the estimator

$$\hat{\theta} = \frac{n+1}{n} \cdot \max(X_1, X_2, \cdots, X_n)$$

is unbiased for θ

It is also shown that $\hat{\theta} = \frac{n+1}{n} \cdot \max(X_1, X_2, \cdots, X_n)$ is the MVUE of θ.

6.1 Some General Concepts of Point Estimation

- **Theorem**

Let $X_1, X_2, \ldots, X_n$ be a random sample from a **normal distribution** with parameters μ and σ. Then the estimator

$$\hat{\mu} = \overline{X}$$

is the MVUE for μ.

How about those un-normal distributions?

6.1 Some General Concepts of Point Estimation

- **Example 6.7**

Suppose we wish to estimate the thermal conductivity μ of a certain material. We will obtain a random sample $X_1, X_2, \ldots, X_n$ of n thermal conductivity measurements. Let's assume that the population distribution is a member of one of the following three families:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)} \qquad -\infty < x < \infty$$

Gaussian Distribution

$$f(x) = \frac{1}{\pi[1+(x-\mu)^2]} \qquad -\infty < x < \infty$$

Cauchy Distribution

$$f(x) = \begin{cases} \dfrac{1}{2c} & -c \le x-\mu \le c \\ 0 & \text{otherwise} \end{cases}$$

Uniform Distribution

6.1 Some General Concepts of Point Estimation

1. If the random sample comes from a normal distribution, then $\overline{X}$ is the best of the four estimators, since it is the MVUE.

2. If the random sample comes from a Cauchy distribution, then $\overline{X}$ and $\overline{X}_e$ (the average of the two extreme observations) are terrible estimators for μ, whereas $\tilde{X}$ is quite good; $\overline{X}$ is bad because it is very sensitive to outlying observations, and the heavy tails of the Cauchy distribution make a few such observation likely to appear in any sample.

3. If the underlying distribution is uniform, the best estimator is $\overline{X}_e$; this estimator is greatly influenced by outlying observations, but the lack of tails makes such observations impossible.

4. The trimmed mean is best in none of these three situations, but works reasonably well in all three. That is, $\overline{X}_{tr(10)}$ does not suffer too much in any of the three situations.

A Robust estimator

6.1 Some General Concepts of Point Estimation

- The Standard Error

  The standard error of an estimator $\hat{\theta}$ is its standard deviation $\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$.