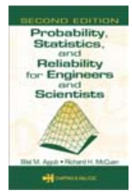
FUNDAMENTALS OF STATISTICAL ANALYSIS

• A. J. Clark School of Engineering •Department of Civil and Environmental Engineering

8b



Probability and Statistics for Civil Engineers

Department of Civil and Environmental Engineering
University of Maryland, College Park

CHAPMAN HALL/CRC

Estimation of Parameters

- Classification of Models
  - In terms of random variables
    - Univariate
    - Bivariate
    - Multivariate
  - In terms of parameters
    - One-parameter Model
    - Two-parameter Model
    - Three-parameter Model

Estimation of Parameters

- Examples:
  - Univariate with two parameters
    - Normal Distribution
  - Univariate with one parameter
    - Exponential Distribution
  - Bivariate with two parameters
    - Bivariate power model

$$\hat{Y} = aX^b$$

Estimation of Parameters

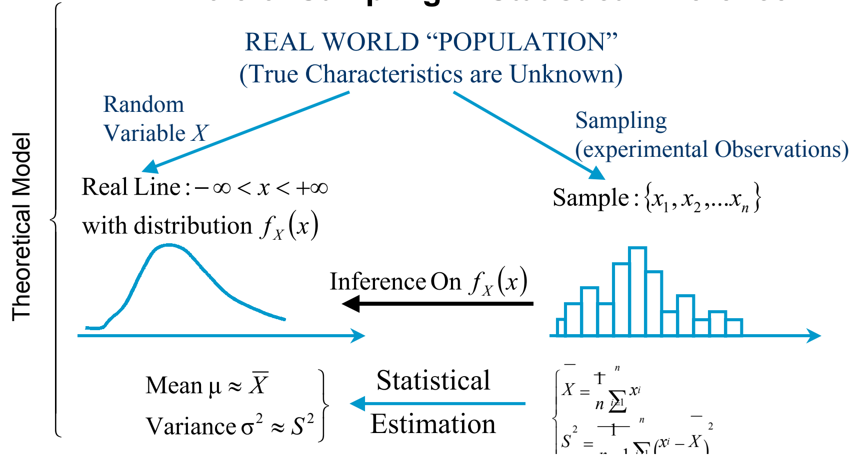
- Estimation of the Parameters
  - Samples are used to develop a model that can represent the population and to estimate the parameters of the population model
  - The parameters can be estimated in the form of
    - **Point Estimates** (single value)
    - **Interval Estimates** (range of values)

Estimation of Parameters

- Classical estimation of parameters are divided into:
  - Point Estimation
    - Point estimation is concerned with the calculation of a single number from a set of observed data to represent the parameter of underlying population.
  - Interval Estimation
    - Interval estimation goes further to establish a statement of confidence in the estimated quantity, resulting in the determination of an interval indicating the range where the population parameter may be located.

Estimation of Parameters

Role of Sampling in Statistical Inference



Estimation of Parameters

- Properties of Estimators
  - Bias
  - Precision
  - Accuracy
  - Consistency
  - Efficiency
  - Sufficiency

Estimation of Parameters

- Bias

  An estimate of a parameter θ made from sample statistic is said to be an unbiased estimate if the expected value of the sample quantity $\hat{\theta}$ is θ; that is

$$E(\hat{\theta}) = \theta$$

The bias is defined as

$$\left[E(\hat{\theta}) - \theta\right]$$

Estimation of Parameters

- Bias
  - Definition:

    Bias is a systematic deviation of values from the true value

Estimation of Parameters

- Bias
  - Consider four experiments where each experiment is repeated six times.
  - The following table shows the results of the four experiments:

	True (Population)	Exp. A	Exp. B	Exp. C	Exp D
Mean	15	15	24	7	14

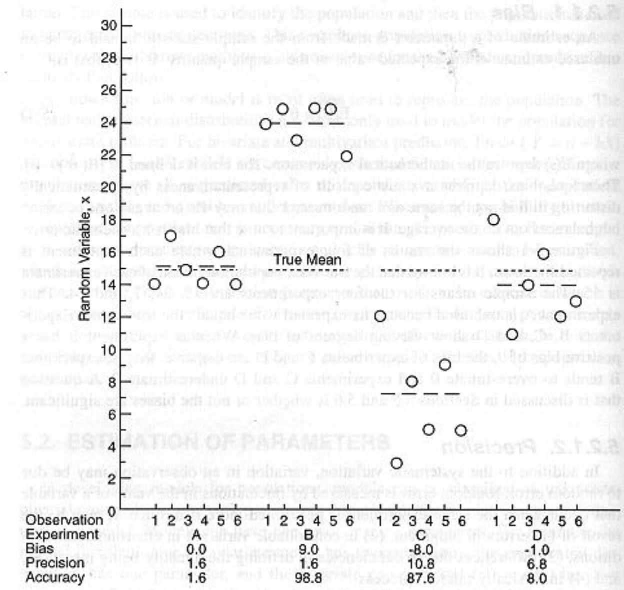


FIGURE 5-1 Bias, precision, and accuracy of experimental observations.

Estimation of Parameters

- Bias
  - Experiment A is unbiased because its expected value (mean) equals the true mean.
  - Experiments B, C, and D show varying degrees of bias.
  - Experiment B has a positive bias of 9, whereas the bias of C and D are negative.
  - Experiment B tends to overestimate θ, while C and D tend to underestimate θ.

Estimation of Parameters

- Precision
  - In addition to the systematic variation, variation in an observation may be due to ***random error***.
  - Random error is measured by fluctuations in the value of a variable that occur when the same experiment is performed more than once.

Estimation of Parameters

- Precision
  - Random error can be a result of:
    1. Errors in judgment.
    2. Uncontrollable variation in environmental conditions.
    3. Differences due to deficiencies in defining the quantity being measured, and
    4. Intrinsically random process

Estimation of Parameters

- Precision
  - Definition:

    Precision is defined as the ability of an estimator to give repeated estimates that are very close to each other.

Estimation of Parameters

- Precision
  - Precision can be expressed in terms of the variance of the estimator.

	Precision
Var(θ) ↑	Lack of precision
Var(θ) ↓	High precision
Var(θ) = 0	Absolute precision

Estimation of Parameters

- Precision
  - Consider four experiments where each experiment is repeated six times.
  - The following table shows the results of the four experiments:

	Exp. A	Exp. B	Exp. C	Exp D
Var	1.6	1.6	10.6	6.8

Note: Variance is about the sample mean for each experiment

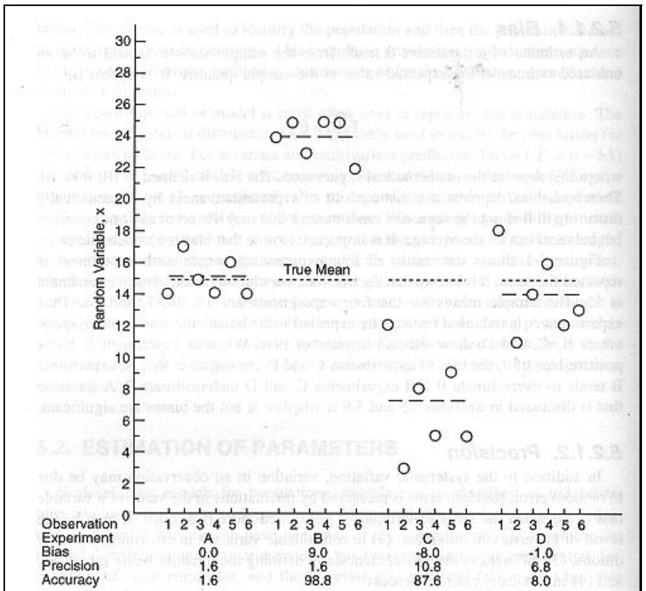


FIGURE 5-1 Bias, precision, and accuracy of experimental observations.

Estimation of Parameters

- Precision
  - Experiment A and B show considerably more precision (i.e., they have lower variances).
  - Experiment C has the largest variation, therefore, it is the least precise.
  - Experiments A and B have the same level of variation, however, A is unbiased, whereas B is highly biased.

Estimation of Parameters

- Accuracy
  - Definition:

    Accuracy is defined as the closeness or nearness of the measurements to the true or actual value of the quantity being measured.

Estimation of Parameters

- Accuracy
  - Bias and Precision are considered elements of Accuracy.

$$\boxed{\text{Bias} + \text{Precision} \quad \Rightarrow \quad \text{Accuracy}}$$

  - Inaccuracy can result from either a *bias* or a lack of *precision*.

Estimation of Parameters

- Accuracy
  - Consider four experiments where each experiment is repeated six times.
  - The following table shows the results of the four experiments:

	Exp. A	Exp. B	Exp. C	Exp D
Var	1.6	98.8	87.6	8.0

Note: Variance is about the true mean of the population (i.e., 15)

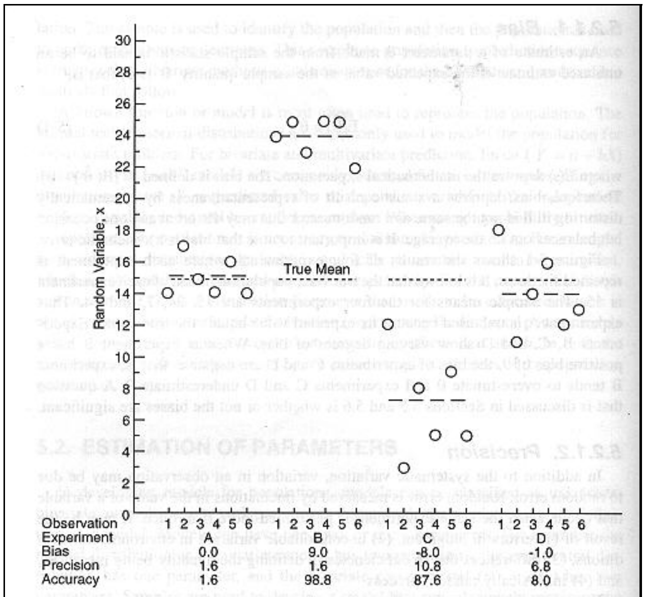


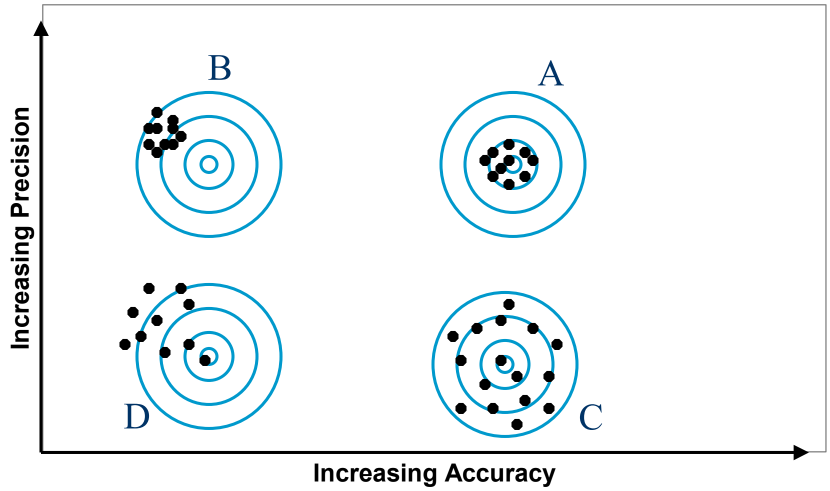
FIGURE 5-1 Bias, precision, and accuracy of experimental observations.

Estimation of Parameters

- Bias, Precision, and Accuracy
  - Consider the four dartboards of the following figure.
  - Assuming that these shooting at the targets were aiming at the center, the person shooting at target A was successful.

Estimation of Parameters

Bias, Precision, and Accuracy



Estimation of Parameters

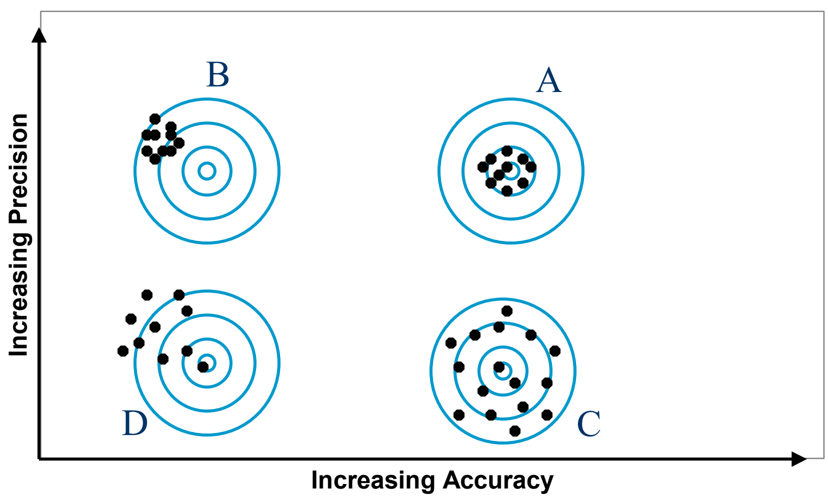
- Bias, Precision, and Accuracy
  - The holes in target B are similarly clustered as in target B, but they show large deviation from the center.
  - The holes in target C are very different in character from the holes in either target A or B.
  - They are not near the center, and they are not near each other.

Estimation of Parameters

- Bias, Precision, and Accuracy
  - The holes in target A and B show a measure of precision, therefore, the shooters were precise.
  - The shooters of targets C and D were imprecise since the holes show a lot of scatter.
  - The holes in targets B and D are consistently to the left, that is, there is a systematic distortion of the hole with respect to the center of the target.

Estimation of Parameters

Bias, Precision, and Accuracy



Estimation of Parameters

- Bias, Precision, and Accuracy
  - The holes in targets B and D show a systematic deviation to the left.
  - Targets A and C are considered to *unbiased* because there is no systematic deviation.
  - In the figure, accuracy increases as precision increases, therefore, the shooter of target A is the most accurate.

Estimation of Parameters

- Mean Square Error (MSE)
  - Accuracy reflects both
    - Systematic error
    - Random error
  - Both sources must be assessed when selecting one estimator from others, specially, a biasd estimator is preferable to an unbiased estimator if the precision of the biased estimator is significantly better than that of the unbiased estimator.

Estimation of Parameters

- Mean Square Error (MSE)
  - The MSE is defined as the expected value of the square of the deviation of the estimate from the true value.
  - It is equal to the variance of the estimate plus the square of the bias.
  - The MSE is considered a measure of accuracy.

Estimation of Parameters

- Consistency, Sufficiency, and Efficiency
  - **Consistency**
    - An estimator is said to be consistent if the probability that $\hat{\theta}$ will deviate from θ more than any fixed amount $\varepsilon > 0$, approaches zero as the sample size (n) becomes larger and larger.

$$P\left(\left|\hat{\theta} - \theta\right| \leq \varepsilon\right) \to 1 \quad \text{as } n \to \infty$$

Estimation of Parameters

– Consistency
  - Consistency is a “sample size” property of the estimator.
  - It can be shown that unbiased estimators are consistent estimators.
  - However, a consistent estimator is not necessarily unbiased.
  - For example, the sample variance is consistent but must be corrected for its biasedness by multiplying it by the factor $n/(n-1)$

Estimation of Parameters

- Consistency, Sufficiency, and Efficiency
  - **Sufficiency**
    - An estimator $\hat{\theta}$ is said to be sufficient if it exhausts all possible information on θ in a sample of any size.
    - Sufficiency implies that no other estimator computed from the same sample can provide additional information about θ

Estimation of Parameters

- Consistency, Sufficiency, and Efficiency
  - **Efficiency**
    - Efficiency is an important criterion for evaluating the quality of an estimator.
    - A consistent estimate $\hat{\theta}_1$ is said to be more efficient than another estimate $\hat{\theta}_2$ if

$$\mathrm{Var}\left(\hat{\theta}_1\right) < \mathrm{Var}\left(\hat{\theta}_2\right)$$

    - An estimate is considered efficient if it has the smallest variance of all available estimators.