6.1 Point Estimation

Suppose that the distribution of a random variable X depends on an unknown parameter θ. In order to estimate θ we consider a random sample $X_1, X_2, \dots, X_n$ from this distribution and we need to find a statistic $\hat{\theta} = \Theta(X_1, X_2, \dots, X_n)$ such that its observed value $\Theta\ (x_1, x_2, \dots, x_n)$ will be our estimate of an unknown parameter θ. Lower case letters $x_1, x_2, \dots, x_n$ are the observed values of $X_1, X_2, \dots, X_n$ (that is *sample data*).

DEFINITION

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.

DEFINITION

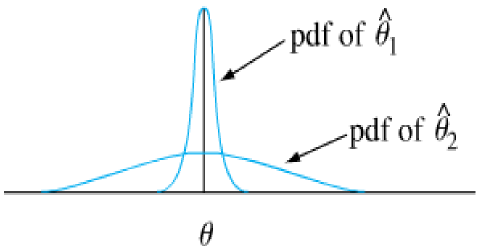
A point estimator $\hat{\theta}$ is said to be an **unbiased estimator** of θ if $E(\hat{\theta}) = \theta$ for every possible value of θ. If $\hat{\theta}$ is not unbiased, the difference $E(\hat{\theta}) - \theta$ is called the **bias** of $\hat{\theta}$.

Principle of Unbiased Estimation

When choosing among several different estimators of θ, select one that is unbiased.

Principle of Minimum Variance Unbiased Estimation

Among all estimators of θ that are unbiased, choose the one that has minimum variance. The resulting $\hat{\theta}$ is called the **minimum variance unbiased estimator (MVUE)** of θ.



Important Unbiased Estimators.

1. For the mean. The sample mean

$$\hat{\mu} = \bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

is an unbiased estimator of the mean μ. If the distribution is continuous and symmetric, then any trimmed mean is also unbiased estimators of the mean μ. If X has normal distribution, then $\bar{X}$ is MVUE for μ (so it is the best estimator).

2. For the proportion. Suppose that *p* the unknown probability of a certain event which we call *success* which we want to estimate. Consider a random variable which assumes only two values: 1 when a trial is *success* and 0 otherwise. Let $X_1, X_2, \ldots, X_n$ be a random sample from this distribution and let $X = X_1 + X_2 + \ldots + X_n$. Then X has a binomial distribution with parameters n, p. Take

$$\hat{p} = \frac{X}{n} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

Then $E(\hat{p}) = E\left(\frac{X}{n}\right) = \frac{np}{n} = p$, so $\hat{p}$ is an unbiased estimator of p

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.

3. For the variance.

Suppose that X is a random variable with unknown mean μ and unknown variance σ^2. We have already found that $\bar{X}$ is an unbiased estimator of the population mean μ.

PROPOSITION

Let $X_1, X_2, \ldots, X_n$ 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.

(proof is on page 245)

DEFINITION

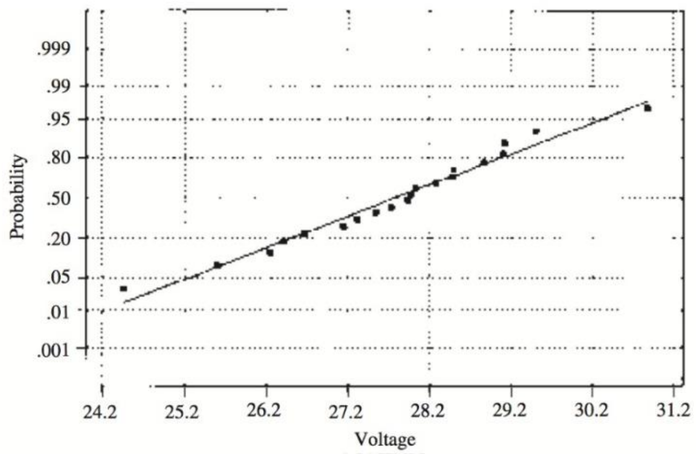
The **standard error** of an estimator $\hat{\theta}$ is its standard deviation $\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$. It is the magnitude of a typical or representative deviation between an estimate and the value of θ. If the standard error itself involves unknown parameters whose values can be estimated, substitution of these estimates into $\sigma_{\hat{\theta}}$ yields the **estimated standard error** (estimated standard deviation) of the estimator. The estimated standard error can be denoted either by $\hat{\sigma}_{\hat{\theta}}$ (the ^ over σ emphasizes that $\sigma_{\hat{\theta}}$ is being estimated) or by $s_{\hat{\theta}}$.

Example 1.

Reconsider the accompanying 20 observations on dielectric breakdown voltage for pieces of epoxy resin first introduced in Example 4.30 (Section 4.6).

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.



A computation shows that $\bar{x} = 27.793$ and $s = \sqrt{s^2} = 1.462$, where $s^2 = \frac{\sum_{k=1}^{n}(x_k - \bar{x})^2}{n-1}$

Assuming that breakdown voltage is normally distributed, $\hat{\mu} = \overline{X}$ is the best estimator of μ. If the value of σ is known to be 1.5, the standard error of $\overline{X}$ is $\sigma_{\overline{X}} = \sigma/\sqrt{n} = 1.5/\sqrt{20} = .335$. If, as is usually the case, the value of σ is unknown, the estimate $\hat{\sigma} = s = 1.462$ is substituted into $\sigma_{\overline{X}}$ to obtain the estimated standard error $\hat{\sigma}_{\overline{X}} = s_{\overline{X}} = s/\sqrt{n} = 1.462/\sqrt{20} = .327$. ■