Ch10. Sample Mean

School of Electronics Engineering

Sample Mean

Definition 10.1

For *i.i.d.* random variables $X_1, \dots, X_n$ with PDF $f_X(x)$, the sample mean of X is the random variable

$$M_n(X) = \frac{X_1 + \cdots + X_n}{n}$$

Sample Mean

Theorem 10.1

The sample mean $M_n(X)$ has expected value and variance

$$E[M_n(X)] = E[X], \qquad Var[M_n(X)] = \frac{Var[X]}{n}$$

Inequalities

Theorem 10.2 (Markov inequality)

For a random variable X with $P[X < 0] = 0$ and a constant c

$$P[X \geq c^2] \leq \frac{E[X]}{c^2}$$

Inequalities

Example 10.1

Let X represent the height (ft) of a storm surge following a hurricane. If $E[X] = 5.5$, the Markov inequality provides an upper bound on $P[X \geq 11]$

Inequalities

Theorem 10.3 (Chebyshev inequality)

For an arbitrary random variable Y and constant $c > 0$,

$$P[|Y - \mu_Y| \geq c] \leq \frac{Var[Y]}{c^2}$$

Inequalities

Example 10.3

If the height X of a storm surge following a hurricane has $E[X] = 5.5$ ft and standard deviation $\sigma_X = 1$ ft, the Chebyshev inequality provides an upper bound on $P[X \geq 11]$

Inequalities

Theorem 10.4 (Chernoff bound)

For an arbitrary random variable X and a constant c,

$$P[X \geq c] \leq \min_{s \geq 0} e^{-sc} \phi_X(s)$$

Inequalities

Example 10.4

If the probability model of the height X (ft) of a storm surge following a hurricane is the Gaussian $(5.5, 1)$ random variable. The Chernoff bound provides an upper bound of $P[X > 11]$.

Laws of Large Numbers

Theorem 10.5 (Weak law of large numbers)

For any constant $c > 0$,

(a) $P[|M_n(X) - \mu_X| \geq c] \leq \frac{Var[X]}{nc^2}$

(b) $P[|M_n(X) - \mu_X| < c] \geq 1 - \frac{Var[X]}{nc^2}$

Laws of Large Numbers

Definition 10.2

The random sequence Y_n converges in probability to a constant y if for any $\epsilon > 0$,

$$\lim_{n\to\infty} P[|Y_n - y| \geq \epsilon] = 0$$

Laws of Large Numbers

Definition

$\hat{P}_A$ denotes the relative frequency of event A in n trials, which is defined as

$$\hat{P}_A = M_n(X_A) = \frac{X_{A1} + X_{A2} + \cdots + X_{An}}{n}$$

Laws of Large Numbers

Theorem 10.7

As $n \to \infty$, the relative frequency $\hat{P}_n(A)$ converges to $P[A]$; for any constant $c > 0$,

$$\lim_{n\to\infty} P\left[\left|\hat{P}_n(A) - P(A)\right| \geq c\right] = 0$$

Point Estimates of Model Parameters

Definition 10.3 (Consistent estimator)

The sequence of estimates $\hat{R}_1, \hat{R}_2, \ldots$ of parameter r is consistent if for any $\epsilon > 0$,

$$\lim_{n \to \infty} P\left[\left|\hat{R}_n - r\right| \geq \epsilon\right] = 0$$

Point Estimates of Model Parameters

Definition 10.4 (Unbiased estimator)

An estimate, $\hat{R}$, of parameter r is unbiased if $E[\hat{R}] = r$. Otherwise, $\hat{R}$ is biased

Point Estimates of Model Parameters

Definition 10.5 (Asymptotically unbiased estimator)

The sequence of estimator $\hat{R}_n$ of parameter r is asymptotically unbiased if

$$\lim_{n\to\infty} E\left[\hat{R}_n\right] = r$$

Point Estimates of Model Parameters

Definition 10.6 (Mean square error)

The mean square error of estimator $\hat{R}$ of parameter r is

$$e = E\left[\left(\hat{R} - r\right)^2\right]$$

Point Estimates of Model Parameters

Theorem 10.8

If a sequence of unbiased estimates $\hat{R}_1, \hat{R}_2, \ldots$ of parameter r has mean square error $e_n = Var[\hat{R}_n]$ satisfying $\lim_{n \to \infty} e_n = 0$, then the sequence $\hat{R}_n$ is consistent

Point Estimates of Model Parameters

Example 10.5

In any interval of k seconds, the number N_k of packets passing through an Internet router is a Poisson random variable with $E[N_k] = kr$ packets. Let $\hat{R}_k = N_k/k$ denote an estimate of the parameter r. Is $\hat{R}_k$ an unbiased? What is mean square error?

Point Estimates of the Expected Value

Theorem 10.9

The sample mean $M_n(X)$ is an unbiased estimate of $E[X]$

Point Estimates of the Expected Value

Theorem 10.10

The sample mean estimator $M_n(X)$ has mean square error

$$e_n = \frac{Var[X]}{n}$$

Point Estimates of the Expected Value

Theorem 10.11

If X has finite variance, then the sample mean $M_n(X)$ is a sequence of consistent estimates of $E[X]$

Point Estimates of the Variance

Definition 10.7

The sample variance of n independent observations of random variable X is

$$V_n(X) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - M_n(X)\right)^2$$

Point Estimates of the Variance

Theorem 10.12

$$E[V_n(X)] = \frac{n-1}{n} Var[X]$$

Point Estimates of the Variance

Theorem 10.13

The estimate

$$V_n'(X) = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - M_n(X)\right)^2$$

is an unbiased estimate of $Var[X]$