Example 7.4 Problem

In any interval of k seconds, the number N_k of packets passing through an Internet router is a Poisson random variable with expected value $E[N_k] = kr$ packets. Let $\hat{R}_k = N_k/k$ denote an estimate of r. Is each estimate $\hat{R}_k$ an unbiased estimate of r? What is the mean square error e_k of the estimate $\hat{R}_k$? Is the sequence of estimates $\hat{R}_1, \hat{R}_2, \ldots$ consistent?

$$E[\hat{R}_k] = E\left[\frac{N_k}{k}\right] = \frac{E[N_k]}{k} = \frac{kr}{k} = r.$$

$\therefore$ $\hat{R}_k$ is an unbiased estimator $\forall k$.

$$e_k = E\left[(\hat{R}_k - r)^2\right]$$
$$= E\left[(\hat{R}_k - E[\hat{R}_k])\right] = \mathrm{Var}[\hat{R}_k]$$

Theorem 7.4

Given:

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

Show that:

$$\lim_{n\to\infty} P\left[|R_n - r| > \epsilon \right] = 0$$

Use the Cheby ineq:

$$P\left[|R_n - r| > \epsilon \right] \le \frac{\mathrm{Var}[R_n]}{\epsilon^2}$$

$$\lim_{n\to\infty} P\left[|R_n - r| > \epsilon \right] \le 0$$

$\left[\because \text{ Given that } \lim_{n\to\infty} e_n = 0 \right]$

Theorem 7.5

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

$$M_n(X) = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

$$E[M_n(X)] = E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i]$$

$$X_i \sim X \quad \forall i$$

$$\therefore\ E[M_n(X)] = \frac{1}{n}\sum_{i=1}^{n} E[X]$$

$$= E[X]$$

Theorem 7.6

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

$$e_n = E\left[(M_n(X) - E[X])^2\right] = \mathrm{Var}[M_n(X)] = \frac{\mathrm{Var}[X]}{n}.$$

$$\mathrm{Var}[M_n(X)] = \mathrm{Var}\left[\frac{X_1 + X_2 + \cdots + X_n}{n}\right]$$

Given samples are independent:

$$\mathrm{Var}[M_n(X)] = \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}[X_i] = \frac{1}{n^2} n \mathrm{Var}[X)$$

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

This is also the variance of the sample mean estimator. Square-root of this is called the standard error.

Given the mean and variance of the estimator, what can we say about the sample mean that results from n IID trials? $\longrightarrow$ The sample mean is a consistent estimator

Example 7.5 Problem

How many independent trials n are needed to guarantee that $\hat{P}_n(A)$, the relative frequency estimate of $P[A]$, has standard error less than 0.1?

Let X_i be the Bern RV that models the outcome of the ith trial.

The rel freq estimate is $\dfrac{X_1 + X_2 + \cdots + X_n}{n} \triangleq M_n(X)$

where $X_i \sim \text{Bern}(P[A])$

$M_n(X)$ is an unbiased estimator. $\therefore$ MSE is $\text{Var}[M_n(X)] = \dfrac{\text{Var}(X)}{n}$

$$= \underbrace{\frac{P[A](1-P[A])}{n}}$$

We require:

$$\text{MSE} < 0.01 \Rightarrow \frac{P[A](1-P[A])}{n} < 0.01$$

$$\Rightarrow n > \frac{P[A](1-P[A])}{0.01};$$

We don't know $P[A]$. We want:

$$n > \max\left[\frac{P[A](1-P[A])}{0.01}\right] = 25.$$

Theorem 7.7

If X has finite variance, then the sample mean $M_n(X)$ is a sequence of consistent estimates of $\underbrace{E[X]}_{r = E[X]}$.

Consistent estimates:

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

Chebyshev's.

$$P\left[\,|M_n(X) - \underbrace{E[X]}_{= E[M_n(X)]}| \geq \epsilon\right] \leq \frac{\mathrm{Var}[M_n(X)]}{\epsilon^2} = \frac{\mathrm{Var}(X)}{n\epsilon^2}$$

Take the limit $n \to \infty$. and you have proved the Thm.

Theorem 7.8 Weak Law of Large Numbers

If X has finite variance, then for any constant $c > 0$,

$$P[|M_n(X) - \mu_X| \geq c]$$

(a) $\lim_{n\to\infty} P[|M_n(X) - \mu_X| \geq c] = 0$,

$$= 1 - P[\mu_X - c \leq M_n(X) \leq \mu_X + c]$$

(b) $\lim_{n\to\infty} P[|M_n(X) - \mu_X| < c] = 1$.

example seq of outcomes: (assuming $X \sim$ Bern)

n

seq 1 (sample fn 1): 1 1 0 1 0 0 0 1 1 - - -

2: 0 1 0 1 0 0 1 1 1 - -

⋮

Strong law vs. Weak law

↳ Not in syllabus.

The outcomes of your experiment are governed by the RV X.

Suppose you perform k trials of your experiment.

In each trial you draw n outcomes

Trial 1 / Sample Fn 1: $X_1^{(1)}, X_2^{(1)}, X_3^{(1)}, \ldots, X_n^{(1)}$ $M_n^{(1)}(X)$

Trial 2 / Sample Fn 2: $X_1^{(2)}, X_2^{(2)}, X_3^{(2)}, \ldots, X_n^{(2)}$ $M_n^{(2)}(X)$

$\vdots$

Trial k / Sample Fn k: $X_1^{(k)}, X_2^{(k)}, X_3^{(k)}, \ldots, X_n^{(k)}$. $M_n^{(k)}(X)$

$\sim M_n(X)$

Clearly, by assumption, $X_j^{(q)}$ are iid $\forall\, j, q$.

$$\lim_{n\to\infty} P\left[\,|M_n(X) - \mu_X| \geq \epsilon\right] = 0$$ [Weak law. Convergence in prob]

$$P\left[\lim_{n\to\infty} \left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu_X\right| \geq \epsilon\right] = 0$$ [Strong Law]

almost sure convergence

Theorem 7.9

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

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

The weak law...

Definition 7.6 Convergence in Probability

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

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

Problem 7.3.1

When X is Gaussian, verify the claim of Equation (7.16) that the sample mean is within one standard error of the expected value with probability 0.68.

$$P\left[\left|M_n(X) - E[X]\right| \le \frac{\sigma_X}{\sqrt{n}}\right]$$

$M_n(X)$ is Gaussian. Note that the sum of independent Gaussians is Gaussian. Also any linear transf. of a Gaussian (in the above case scaling by $1/n$) is a Gaussian.

$$\mathrm{Var}[M_n(X)] = \frac{\sigma_X^2}{n}.$$

We want

$$P\left[-\frac{\sigma_X}{\sqrt{n}} \le M_n(X) - E[X] \le \frac{\sigma_X}{\sqrt{n}}\right]$$

$$= P[-1 \le Z \le 1] = 0.68$$

Estimating the Variance

- Mean is known → RV X, $E[X]$ is known

  $\sigma^2 = \mathrm{Var}[X]$

  $= E[(X - E[X])^2]$

  is estimated by the sample mean of

  $Y \triangleq (X - E[X])^2$.

- Mean Unknown

Definition 7.7 Sample Variance

The sample variance of a set of n *independent observations of random variable* X *is*

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

$$E[V_n(X)] = \frac{1}{n}\sum_{i=1}^{n} E[(X_i - M_n(X))^2] = \frac{1}{n}\sum_{i=1}^{n} E[X_i^2 + M_n^2(X) - 2X_i M_n(X)]$$

Note that $M_n(X) = \dfrac{X_1 + X_2 + \cdots + X_i + \cdots + X_n}{n}$

Also $E[X_i X_j] = E[X_i]E[X_j]$ $i \neq j$,

since X_i & X_j are independent.

Theorem 7.10

$$E\left[V_n(X)\right] = \frac{n-1}{n}\,\mathrm{Var}[X].$$

How to make it unbiased?

Multiply by $\left(\frac{n}{n-1}\right)$.

Theorem 7.11

The estimate

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

is an unbiased estimate of $\mathrm{Var}[X]$.

Problem 7.3.3

An experimental trial produces random variables X_1 and X_2 with correlation $r = E[X_1 X_2]$. To estimate r, we perform n independent trials and form the estimate

$$\hat{R}_n = \frac{1}{n} \sum_{i=1}^{n} X_1(i) X_2(i)$$

where $X_1(i)$ and $X_2(i)$ are samples of X_1 and X_2 on trial i. Show that if $\text{Var}[X_1 X_2]$ is finite, then $\hat{R}_1, \hat{R}_2, \ldots$ is an unbiased, consistent sequence of estimates of r.

Problem 7.3.4

An experiment produces random vector $\mathbf{X} = \begin{bmatrix} X_1 & \cdots & X_k \end{bmatrix}'$ with expected value $\boldsymbol{\mu}_{\mathbf{X}} = \begin{bmatrix} \mu_1 & \cdots & \mu_k \end{bmatrix}'$. The ith component of $\mathbf{X}$ has variance $\operatorname{Var}[X_i] = \sigma_i^2$. To estimate $\boldsymbol{\mu}_{\mathbf{X}}$, we perform n independent trials such that $\mathbf{X}(i)$ is the sample of $\mathbf{X}$ on trial i, and we form the vector mean

$$\mathbf{M}(n) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{X}(i).$$

(a) Show $\mathbf{M}(n)$ is unbiased by showing $E[\mathbf{M}(n)] = \boldsymbol{\mu}_{\mathbf{X}}$.

(b) Show that the sequence of estimates $\mathbf{M}_n$ is consistent by showing that for any constant $c > 0$,

$$\lim_{n\to\infty} P\left[\max_{j=1,\ldots,k} \left|M_j(n) - \mu_j\right| \geq c\right] = 0.$$

Hint: Let $A_i = \{|M_i(n) - \mu_i| \geq c\}$ and apply the union bound (see Problem 1.4.5) to upper bound $P[A_1 \cup A_2 \cup \cdots \cup A_k]$. Then apply the Chebyshev inequality.

Section 7.4

Confidence Intervals

Theorem 7.12

For any constant $c > 0$,

(a) $P[|M_n(X) - \mu_X| \geq c] \leq \dfrac{\text{Var}[X]}{nc^2} = \alpha,$

(b) $P\underbrace{[|M_n(X) - \mu_X| < c]}_{d} \geq 1 - \dfrac{\text{Var}[X]}{nc^2} = 1 - \alpha.$

Confidence *Coefficient* of your estimate being in an interval of length 2c

Event : $\{|M_n(X) - \mu_X| < c\}$