SQ Solutions

November 14, 2025

Question 1 (Memory test)

(i) Definitions.

Convergence in probability. A sequence of random variables X_n converges to a random variable X in probability, written

$$X_n \longrightarrow X,$$

if for every $\epsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$$

Convergence in the r-th mean (or in L^r). For a fixed $r > 0$, X_n converges to X in the r-th mean, written

$$X_n \xrightarrow{L^r} X$$

if

$$\lim_{n\to\infty} \mathbb{E}[|X_n - X|^r] = 0.$$

(ii) Strong consistency.

An estimator sequence $\hat{\theta}_n$ is said to be strongly consistent for the parameter θ if $\hat{\theta}_n$ converges to θ almost surely, i.e.

$$\hat{\theta}_n \longrightarrow \theta \quad a.s.$$

which means

$$\mathbb{P}(\{\omega : \lim_{n\to\infty} \hat{\theta}_n(\omega) = \theta\}) = 1$$

Equivalently, the set of sample outcomes for which $\hat{\theta}_n(\omega)$ fails to converge to θ has probability zero.

Example. If $X_1, X_2, \ldots$ are i.i.d. with finite mean $\theta = \mathbb{E}[X_1]$ then the sample mean $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ satisfies $\overline{X}_n \longrightarrow \theta$ a.s. by the Strong Law of Large Numbers; hence $\overline{X}_n$ is strongly consistent.

Question 2

Find the stationary distribution for the Markov chain with transition matrix (5 marks)

$$P = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.25 & 0.5 & 0.25 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

Solution

To find the stationary distribution $\boldsymbol{\pi} = [\pi_1, \pi_2, \pi_3]$, we must solve the system of equations given by $\boldsymbol{\pi} = \boldsymbol{\pi} P$, subject to the constraint that the elements sum to 1: $\pi_1 + \pi_2 + \pi_3 = 1$.

The matrix equation $\boldsymbol{\pi} = \boldsymbol{\pi} P$ is:

$$[\pi_1, \pi_2, \pi_3] = [\pi_1, \pi_2, \pi_3] \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.25 & 0.5 & 0.25 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

This gives the system of linear equations:

$$\pi_1 = 0.5\pi_1 + 0.25\pi_2 \tag{1}$$
$$\pi_2 = 0.5\pi_1 + 0.5\pi_2 + 0.5\pi_3 \tag{2}$$
$$\pi_3 = 0.25\pi_2 + 0.5\pi_3 \tag{3}$$

And the normalization constraint:

$$\pi_1 + \pi_2 + \pi_3 = 1 \tag{4}$$

From equation (1), we can simplify:

$$0.5\pi_1 = 0.25\pi_2 \Rightarrow \pi_2 = 2\pi_1$$

From equation (3), we can simplify:

$$0.5\pi_3 = 0.25\pi_2 \Rightarrow \pi_2 = 2\pi_3$$

This implies that $\pi_1 = \pi_3$.

Now, we substitute $\pi_2 = 2\pi_1$ and $\pi_3 = \pi_1$ into the constraint (4):

$$\pi_1 + (2\pi_1) + \pi_1 = 1$$

$$4\pi_1 = 1 \Rightarrow \pi_1 = 0.25$$

From this, we find the other components:

- $\pi_1 = 0.25$
- $\pi_2 = 2\pi_1 = 2(0.25) = 0.5$
- $\pi_3 = \pi_1 = 0.25$

The stationary distribution is $\boldsymbol{\pi} = [0.25, 0.5, 0.25]$.

Question 3

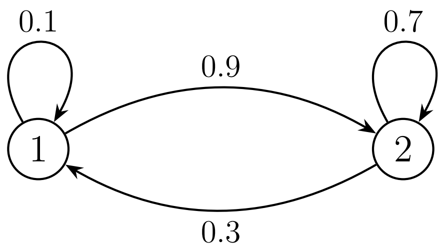
Give a method to simulate n successive states of a Markovian coin with initial distribution $\mu = [0.5, 0.5]$ and

$$P = \begin{bmatrix} 0.1 & 0.9 \\ 0.3 & 0.7 \end{bmatrix}$$

You are given access to only samples from a Uniform $[0, 1]$ random variable.

Solution

Interpret the two states as $\{1, 2\}$ (e.g. 1 = Heads, 2 = Tails). The matrix P means $P(1 \to 1) = 0.1$, $P(1 \to 2) = 0.9$ $P(2 \to 1) = 0.3$, $P(2 \to 2) = 0.7$.



Since we are allowed only $U[0, 1]$ samples, we use the inverse CDF sampling method to sample discrete outcomes.

Method:

1. **Initial state:** draw $U_0 \sim U[0, 1]$. This produces $X_1 \sim \mu$. Set

$$X_1 = \begin{cases} 1, & \text{if } U_0 < 0.5, \\ 2, & \text{if } U_0 \geq 0.5. \end{cases}$$

2. **Iterate** for $t = 1, 2, \ldots, n-1$; draw $U_t \sim U[0, 1]$. Given the current state X_t, obtain X_{t+1} by thresholding according to the appropriate row of P:

$$X_{t+1} = \begin{cases} 1, & \text{if } X_t = 1 \text{ and } U_t < 0.1, \\ 2, & \text{if } X_t = 1 \text{ and } U_t \geq 0.1, \\ 1, & \text{if } X_t = 2 \text{ and } U_t < 0.3, \\ 2, & \text{if } X_t = 2 \text{ and } U_t \geq 0.3. \end{cases}$$

   (Equivalently: if $X_t = i$ compare U_t to the cumulative probabilities of row i of P.)

Remarks.

- If you need many independent trajectories, repeat the algorithm fresh for each trajectory.
- The same approach generalizes to any finite-state Markov chain: to move from state i, draw a uniform and compare against the cumulative sums $\sum_{j \leq k} P_{ij}$ to select the next state.

- If desired one can precompute the cumulative-transition matrix

$$C = \begin{bmatrix} 0.1 & 1.0 \\ 0.3 & 1.0 \end{bmatrix}$$

and choose the next state as the smallest j with $U < C_{ij}$.

- Accept-Reject method cannot be used for this question as the support sets for the uniform distribution and the target distribution are not same.

Section B

Mean Square Error (MSE) for Estimators

Final Answers

Estimator	Bias	Variance	MSE
$\hat{\Theta}_1 = 2\overline{X}_n - X_1$	0	σ^2	σ^2
$\hat{\Theta}_2 = \frac{1}{3}(X_1 + X_2)$	$-\frac{\theta}{3}$	$\frac{2\sigma^2}{9}$	$\frac{2\sigma^2+\theta^2}{9}$
$\hat{\Theta}_3 = c\overline{X}_n$	$(c-1)\theta$	$\frac{c^2\sigma^2}{n}$	$\frac{c^2\sigma^2}{n} + (c-1)^2\theta^2$

Useful note:

- The random variables X_i are mutually independent, so $Cov(X_i, X_j) = 0$ for $i \neq j$
- However, $\overline{X}_n$ is dependent on all X_i since it's constructed from them
- Therefore, $Cov(\overline{X}_n, X_i) \neq 0$ for any i

For any estimator of parameter θ:

$$MSE(\hat{\Theta}) = E[(\hat{\Theta} - \theta)^2] = Var(\hat{\Theta}) + [Bias(\hat{\Theta})]^2$$

where $Bias(\hat{\Theta}) = E[\hat{\Theta}] - \theta$

Solution 1: $\hat{\Theta}_1 = 2\overline{X}_n - X_1$

$$\begin{aligned} E[\hat{\Theta}_1] &= E[2\overline{X}_n - X_1] = 2E[\overline{X}_n] - E[X_1] = 2\theta - \theta = \theta \\ Bias(\hat{\Theta}_1) &= E[\hat{\Theta}_1] - \theta = \theta - \theta = 0 \end{aligned}$$

Since $\overline{X}_n$ and X_1 are not independent (as X_1 is part of $\overline{X}_n$):

$$\begin{aligned} Var(2\overline{X}_n - X_1) &= Var(2\overline{X}_n) + Var(X_1) - 2Cov(2\overline{X}_n, X_1) \\ &= 4Var(\overline{X}_n) + \sigma^2 - 4Cov(\overline{X}_n, X_1) \end{aligned}$$

Now, $Var(\overline{X}_n) = \frac{\sigma^2}{n}$. For the covariance:

$$
\begin{aligned}
Cov(\overline{X}_n, X_1) &= Cov\left(\frac{1}{n}\sum_{i=1}^{n} X_i, X_1\right) = \frac{1}{n}\sum_{i=1}^{n} Cov(X_i, X_1) \\
&= \frac{1}{n}Cov(X_1, X_1) = \frac{\sigma^2}{n} \quad \text{(since } Cov(X_i, X_1) = 0 \text{ for } i \neq 1\text{)}
\end{aligned}
$$

Therefore:

$$
Var(\hat{\Theta}_1) = \frac{4\sigma^2}{n} + \sigma^2 - 4 \cdot \frac{\sigma^2}{n} = \sigma^2
$$

Alternate method to compute variance: Decomposition into Independent Variables We can rewrite $\hat{\Theta}_1$ by expanding $\overline{X}_n$

$$
\begin{aligned}
\hat{\Theta}_1 &= 2\overline{X}_n - X_1 = 2\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - X_1 = \frac{2}{n}\sum_{i=1}^{n} X_i - X_1 \\
&= \frac{2X_1}{n} + \frac{2X_2}{n} + \cdots + \frac{2X_n}{n} - X_1 \\
&= \frac{2X_1 - nX_1}{n} + \frac{2X_2}{n} + \cdots + \frac{2X_n}{n} \\
&= \frac{(2-n)X_1}{n} + \frac{2X_2}{n} + \cdots + \frac{2X_n}{n} = -\frac{(n-2)X_1}{n} + \frac{2}{n}\sum_{i=2}^{n} X_i
\end{aligned}
$$

Since all X_i are independent:

$$
\begin{aligned}
Var(\hat{\Theta}_1) &= Var\left(-\frac{(n-2)X_1}{n}\right) + Var\left(\frac{2}{n}\sum_{i=2}^{n} X_i\right) \\
&= \left(-\frac{n-2}{n}\right)^2 Var(X_1) + \left(\frac{2}{n}\right)^2 \sum_{i=2}^{n} Var(X_i) \\
&= \frac{(n-2)^2}{n^2}\sigma^2 + \frac{4}{n^2}(n-1)\sigma^2 \\
&= \frac{(n-2)^2 + 4(n-1)}{n^2}\sigma^2 \\
&= \frac{n^2 - 4n + 4 + 4n - 4}{n^2}\sigma^2 = \frac{n^2}{n^2}\sigma^2 = \sigma^2
\end{aligned}
$$

Finally,

$$
MSE(\hat{\Theta}_1) = Var(\hat{\Theta}_1) + [Bias(\hat{\Theta}_1)]^2 = \sigma^2 + 0 = \sigma^2
$$

Solution 2: $\hat{\Theta}_2 = \frac{1}{3}(X_1 + X_2)$

$$
\begin{aligned}
E[\hat{\Theta}_2] &= \frac{1}{3}E[X_1 + X_2] = \frac{1}{3}(\theta + \theta) = \frac{2\theta}{3} \\
Bias(\hat{\Theta}_2) &= E[\hat{\Theta}_2] - \theta = \frac{2\theta}{3} - \theta = -\frac{\theta}{3}
\end{aligned}
$$

Since X_1 and X_2 are independent:

$$Var(\hat{\Theta}_2) = Var\left(\frac{X_1 + X_2}{3}\right) = \frac{1}{9}[Var(X_1) + Var(X_2)] = \frac{1}{9}(2\sigma^2) = \frac{2\sigma^2}{9}$$

$$MSE(\hat{\Theta}_2) = Var(\hat{\Theta}_2) + [Bias(\hat{\Theta}_2)]^2 = \frac{2\sigma^2}{9} + \left(-\frac{\theta}{3}\right)^2$$

$$MSE(\hat{\Theta}_2) = \frac{2\sigma^2}{9} + \frac{\theta^2}{9} = \frac{2\sigma^2 + \theta^2}{9}$$

Solution 3: $\hat{\Theta}_3 = c\overline{X}_n$

$$E[\hat{\Theta}_3] = cE[\overline{X}_n] = c\theta$$

$$Bias(\hat{\Theta}_3) = c\theta - \theta = (c-1)\theta$$

$$Var(\hat{\Theta}_3) = Var(c\overline{X}_n) = c^2 Var(\overline{X}_n) = c^2 \cdot \frac{\sigma^2}{n} = \frac{c^2\sigma^2}{n}$$

$$MSE(\hat{\Theta}_3) = Var(\hat{\Theta}_3) + [Bias(\hat{\Theta}_3)]^2 = \frac{c^2\sigma^2}{n} + (c-1)^2\theta^2$$