MTL108: Solution to Problem Set-8

Prepared by TAs

IIT delhi

Problem 1. Suppose that X_1, X_2, X_3 are independent with the common probability mass function

$$P\{X_i = 0\} = 0.2, \quad P\{X_i = 1\} = 0.3, \quad P\{X_i = 3\} = 0.5, \qquad i = 1, 2, 3.$$

(a) Plot the probability mass function of $\bar{X}_2 = \dfrac{X_1 + X_2}{2}$.

(b) Determine $E[\bar{X}_2]$ and $\text{Var}(\bar{X}_2)$.

(c) Plot the probability mass function of $\bar{X}_3 = \dfrac{X_1 + X_2 + X_3}{3}$.

(d) Determine $E[\bar{X}_3]$ and $\text{Var}(\bar{X}_3)$.

Mean and Variance of each X_i

Each X_i has the same distribution, so:

$$\mu = E[X_i] = 0(0.2) + 1(0.3) + 3(0.5) = 0 + 0.3 + 1.5 = 1.8$$

$$E[X_i^2] = 0^2(0.2) + 1^2(0.3) + 3^2(0.5) = 0 + 0.3 + 4.5 = 4.8$$

$$\sigma^2 = \text{Var}(X_i) = E[X_i^2] - (E[X_i])^2 = 4.8 - (1.8)^2 = 4.8 - 3.24 = 1.56$$

Part (a): PMF of $\bar{X}_2 = \dfrac{X_1 + X_2}{2}$

We first find the PMF of $S_2 = X_1 + X_2$, then divide by 2.

Since each $X_i \in \{0, 1, 3\}$, the possible values of S_2 are $\{0, 1, 2, 3, 4, 6\}$.

$$P(S_2 = 0) = P(X_1 = 0)\,P(X_2 = 0) = (0.2)(0.2) = 0.04$$

$$P(S_2 = 1) = P(X_1 = 0)\,P(X_2 = 1) + P(X_1 = 1)\,P(X_2 = 0) = (0.2)(0.3) + (0.3)(0.2) = 0.12$$

$$P(S_2 = 2) = P(X_1 = 1)\,P(X_2 = 1) = (0.3)(0.3) = 0.09$$

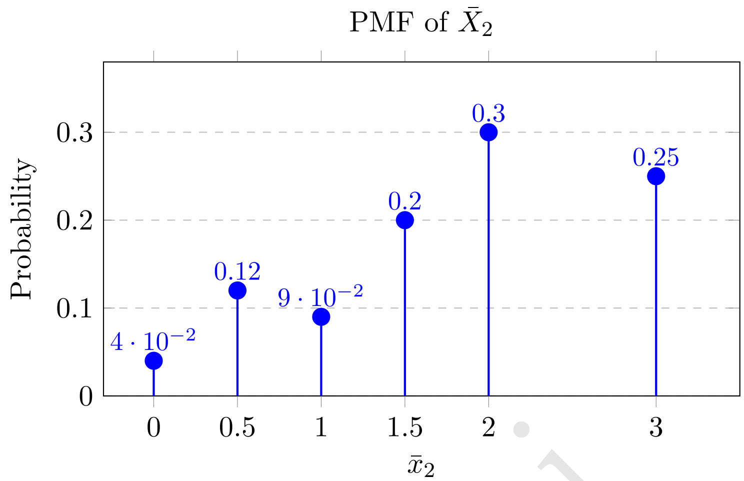
$$P(S_2 = 3) = P(X_1 = 0)\,P(X_2 = 3) + P(X_1 = 3)\,P(X_2 = 0) = (0.2)(0.5) + (0.5)(0.2) = 0.20$$

$$P(S_2 = 4) = P(X_1 = 1)\,P(X_2 = 3) + P(X_1 = 3)\,P(X_2 = 1) = (0.3)(0.5) + (0.5)(0.3) = 0.30$$

$$P(S_2 = 6) = P(X_1 = 3)\,P(X_2 = 3) = (0.5)(0.5) = 0.25$$

Since $\bar{X}_2 = S_2/2$, the PMF of $\bar{X}_2$ is:

$\bar{x}_2$	0	0.5	1	1.5	2	3
$P(\bar{X}_2 = \bar{x}_2)$	0.04	0.12	0.09	0.20	0.30	0.25



Part (b): $E[\bar{X}_2]$ and $\text{Var}(\bar{X}_2)$

Using the properties of the sample mean:

$$\boxed{E[\bar{X}_2] = E\left[\frac{X_1 + X_2}{2}\right] = \frac{E[X_1] + E[X_2]}{2} = \frac{\mu + \mu}{2} = \mu = 1.8}$$

Since X_1 and X_2 are independent:

$$\boxed{\text{Var}(\bar{X}_2) = \text{Var}\left(\frac{X_1 + X_2}{2}\right) = \frac{\text{Var}(X_1) + \text{Var}(X_2)}{4} = \frac{\sigma^2 + \sigma^2}{4} = \frac{2\sigma^2}{4} = \frac{\sigma^2}{2} = \frac{1.56}{2} = 0.78}$$

Verification by direct calculation:

$$\begin{aligned} E[\bar{X}_2] &= 0(0.04) + 0.5(0.12) + 1(0.09) + 1.5(0.20) + 2(0.30) + 3(0.25) \\ &= 0 + 0.06 + 0.09 + 0.30 + 0.60 + 0.75 = 1.80 \checkmark \end{aligned}$$

$$\begin{aligned} E[\bar{X}_2^2] &= 0^2(0.04) + 0.5^2(0.12) + 1^2(0.09) + 1.5^2(0.20) + 2^2(0.30) + 3^2(0.25) \\ &= 0 + 0.03 + 0.09 + 0.45 + 1.20 + 2.25 = 4.02 \end{aligned}$$

$$\text{Var}(\bar{X}_2) = 4.02 - (1.80)^2 = 4.02 - 3.24 = 0.78 \checkmark$$

Part (c): PMF of $\bar{X}_3 = \dfrac{X_1 + X_2 + X_3}{3}$

We find all possible values of $S_3 = X_1 + X_2 + X_3$ where each $X_i \in \{0, 1, 3\}$.

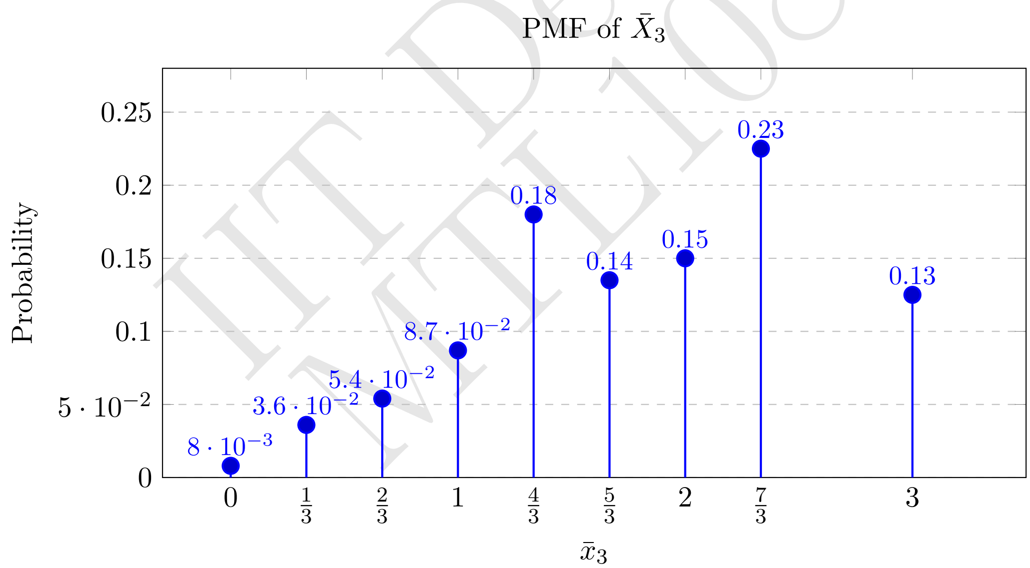
The possible values are $S_3 \in \{0, 1, 2, 3, 4, 5, 6, 7, 9\}$.

We count the number of each $(0, 1, 3)$ combination using multinomial probabilities. Let n_0, n_1, n_3 be the number of X_i's equal to 0, 1, 3 respectively ($n_0 + n_1 + n_3 = 3$).

$$P(S_3 = s) = \sum_{\substack{n_0+n_1+n_3=3 \\ 0n_0+1n_1+3n_3=s}} \frac{3!}{n_0!\,n_1!\,n_3!}(0.2)^{n_0}(0.3)^{n_1}(0.5)^{n_3}$$

The complete table is:

S_3	$\bar{X}_3$	Composition(s) (n_0, n_1, n_3)	$P(S_3)$
0	0	$(3,0,0)$: $(0.2)^3$	0.008
1	$\frac{1}{3}$	$(2,1,0)$: $3(0.2)^2(0.3)$	0.036
2	$\frac{2}{3}$	$(1,2,0)$: $3(0.2)(0.3)^2$	0.054
3	1	$(0,3,0)$: $(0.3)^3$; $(2,0,1)$: $3(0.2)^2(0.5)$	$0.027 + 0.060 = 0.087$
4	$\frac{4}{3}$	$(1,1,1)$: $6(0.2)(0.3)(0.5)$	0.180
5	$\frac{5}{3}$	$(0,2,1)$: $3(0.3)^2(0.5)$	0.135
6	2	$(1,0,2)$: $3(0.2)(0.5)^2$	0.150
7	$\frac{7}{3}$	$(0,1,2)$: $3(0.3)(0.5)^2$	0.225
9	3	$(0,0,3)$: $(0.5)^3$	0.125



Part (d): $E[\bar{X}_3]$ and $\text{Var}(\bar{X}_3)$

By linearity of expectation:

$$\boxed{E[\bar{X}_3] = E\left[\frac{X_1 + X_2 + X_3}{3}\right] = \frac{3\mu}{3} = \mu = 1.8}$$

By independence and properties of variance:

$$\boxed{\text{Var}(\bar{X}_3) = \text{Var}\left(\frac{X_1 + X_2 + X_3}{3}\right) = \frac{3\sigma^2}{9} = \frac{\sigma^2}{3} = \frac{1.56}{3} = 0.52}$$

Question 2

If 10 fair dice are rolled, approximate the probability that the sum of the values obtained (which ranges from 10 to 60) is between 30 and 40 inclusive. ——

Solution 2

Let X_i be outcome of a fair die.

$$E[X_i] = \frac{7}{2}, \quad \mathrm{Var}(X_i) = \frac{35}{12}$$

Let:

$$S = \sum_{i=1}^{10} X_i$$

$$E[S] = 10 \cdot \frac{7}{2} = 35, \quad \mathrm{Var}(S) = 10 \cdot \frac{35}{12} = \frac{350}{12}$$

By CLT:

$$Z = \frac{S - 35}{\sqrt{350/12}}$$

Using continuity correction:

$$P(30 \le S \le 40) \approx P(29.5 \le S \le 40.5)$$

$$= P\left(\frac{29.5 - 35}{\sqrt{350/12}} \le Z \le \frac{40.5 - 35}{\sqrt{350/12}}\right)$$

$$= P(-1.02 \le Z \le 1.02)$$

$$= \Phi(1.02) - \Phi(-1.02) \approx 0.6922$$

—

Question 3

Approximate the probability that the sum of 16 independent uniform (0, 1) random variables exceeds 10.

Solution 3

Let $X_i \sim U(0,1)$.

$$E[X_i] = \frac{1}{2}, \quad \mathrm{Var}(X_i) = \frac{1}{12}$$

Let:

$$S = \sum_{i=1}^{16} X_i$$

$$E[S] = 8, \quad \mathrm{Var}(S) = \frac{16}{12} = \frac{4}{3}$$

Using CLT:

$$Z = \frac{S-8}{\sqrt{4/3}}$$

$$P(S > 10) = P\left(Z > \frac{10-8}{\sqrt{4/3}}\right)$$

$$= P\left(Z > \frac{2}{\sqrt{4/3}}\right) = P(Z > \sqrt{3})$$

$$\approx 1 - \Phi(\sqrt{3}) \approx 0.042$$

—

Question 22

Fifty-two percent of the residents of a certain city are in favor of teaching evolution in high school. Find or approximate the probability that at least 50 percent of a random sample of size n is in favor of teaching evolution, when- 1)n=10 2)n=100 3)n=1000 4)n=10000

Solution

Let $\hat{p} \sim N\left(0.52, \frac{0.52 \cdot 0.48}{n}\right)$.
We compute:

$$P(\hat{p} \geq 0.50)$$

$$Z = \frac{0.50 - 0.52}{\sqrt{0.52 \cdot 0.48/n}}$$

Compute for each n:

- $n = 10$: $P \approx 0.6711$
- $n = 100$: $P \approx 0.6918$
- $n = 1000$: $P \approx 0.9027$
- $n = 10000$: $P \approx 0.99997$

—

Question 23

The following table gives the percentages of individuals of a given city, categorized by gender, that follow certain negative health practices. Suppose a random sample of 300 men is chosen. Approximate the probability that:

(a) at least 150 of them rarely eat breakfast;

(b) fewer than 100 of them smoke.

	Sleeps 6 Hours or Less per Night	Smoker	Rarely Eats Breakfast	20% or More Overweight
Men	22.7	28.4	45.4	29.6
Women	21.4	22.8	42.0	25.6

Source: U.S. National Center for Health Statistics, Health Promotion and Disease Prevention.

Solution

Men sample size: $n = 300$

(a) Rarely eats breakfast ($p = 0.454$)

$$\mu = 300(0.454) = 136.2, \quad \sigma^2 = 300(0.454)(0.546) = 74.3652$$

$$\sigma = \sqrt{74.3652}$$

$$P(X \geq 150) \approx P\left(Z \geq \frac{149.5 - 136.2}{\sigma}\right)$$

$$\approx 0.0617$$

(b) Smoker ($p = 0.284$)

$$\mu = 85.2, \quad \sigma^2 = 61.0$$

$$P(X < 100) \approx P\left(Z < \frac{99.5 - 85.2}{\sigma}\right) \approx 0.9735$$

—

Question 24

(Use the table from Problem 23.) Suppose a random sample of 300 women is chosen. Approximate the probability that- (a) at least 60 of them are overweight by 20 percent or more; (b) fewer than 50 of them sleep 6 hours or less nightly

Solution

Women sample: $n = 300$

(a) Overweight ($p = 0.256$)

$$\mu = 76.8, \quad \sigma^2 = 57.14$$

$$P(X \geq 60) \approx 0.9904$$

(b) Sleep ≤ 6 hrs ($p = 0.214$)

$$\mu = 64.2, \quad \sigma^2 = 50.46$$

$$P(X < 50) \approx 0.0170$$

—

Question 25

(Use the table from Problem 23.) Suppose random samples of 300 women and of 300 men are chosen. Approximate the probability that more women than men rarely eat breakfast.

Solution

Let:

$$X \sim N(126, 73.08), \quad Y \sim N(136.2, 74.3652)$$

$$X - Y \sim N(-10.2, 147.4452)$$

$$P(X > Y) = P(X - Y > 0)$$

$$= P\left(Z > \frac{0 + 10.2}{\sqrt{147.4452}}\right)$$

$$= P(Z > 0.84) = \Phi(-0.84) \approx 0.2005$$

—