Problem Set 7 Solution

Problem 1:

$$Y = X_1 + X_2 + \ldots + X_n,$$

where $n = 50$, $EX_i = \mu = 2$, and $\mathrm{Var}(X_i) = \sigma^2 = 1$. Thus, we can write

$$\begin{aligned} P(90 < Y \leq 110) &= P\left(\frac{90 - n\mu}{\sqrt{n}\sigma} < \frac{Y - n\mu}{\sqrt{n}\sigma} < \frac{110 - n\mu}{\sqrt{n}\sigma}\right) \\ &= P\left(\frac{90 - 100}{\sqrt{50}} < \frac{Y - n\mu}{\sqrt{n}\sigma} < \frac{110 - 100}{\sqrt{50}}\right) \\ &= P\left(-\sqrt{2} < \frac{Y - n\mu}{\sqrt{n}\sigma} < \sqrt{2}\right). \end{aligned}$$

By the CLT, $\frac{Y-n\mu}{\sqrt{n}\sigma}$ is approximately standard normal, so we can write

$$\begin{aligned} P(90 < Y \leq 110) &\approx \Phi(\sqrt{2}) - \Phi(-\sqrt{2}) \\ &= 0.8427 \end{aligned}$$

Problem 2:

Let us define X_i as the indicator random variable for the ith bit in the packet. That is, $X_i = 1$ if the ith bit is received in error, and $X_i = 0$ otherwise. Then the X_i's are i.i.d. and $X_i \sim Bernoulli(p = 0.1)$. If Y is the total number of bit errors in the packet, we have

$$Y = X_1 + X_2 + \ldots + X_n.$$

Since $X_i \sim Bernoulli(p = 0.1)$, we have

$$EX_i = \mu = p = 0.1, \qquad \mathrm{Var}(X_i) = \sigma^2 = p(1-p) = 0.09$$

Using the CLT, we have

$$\begin{aligned} P(Y > 120) &= P\left(\frac{Y - n\mu}{\sqrt{n}\sigma} > \frac{120 - n\mu}{\sqrt{n}\sigma}\right) \\ &= P\left(\frac{Y - n\mu}{\sqrt{n}\sigma} > \frac{120 - 100}{\sqrt{90}}\right) \\ &\approx 1 - \Phi\left(\frac{20}{\sqrt{90}}\right) \\ &= 0.0175 \end{aligned}$$

Problem 3:

If W is the total weight, then $W = X_1 + X_2 + \cdots + X_n$, where $n = 100$. We have

$$\begin{aligned} EW &= n\mu \\ &= (100)(170) \\ &= 17000, \\ \mathrm{Var}(W) &= 100\mathrm{Var}(X_i) \\ &= (100)(30)^2 \\ &= 90000. \end{aligned}$$

Thus, $\sigma_W = 300$. We have

$$\begin{aligned} P(W > 18000) &= P\left(\frac{W - 17000}{300} > \frac{18000 - 17000}{300}\right) \\ &= P\left(\frac{W - 17000}{300} > \frac{10}{3}\right) \\ &= 1 - \Phi\left(\frac{10}{3}\right) \quad \text{(by CLT)} \\ &\approx 4.3 \times 10^{-4}. \end{aligned}$$

Problem 4:

We have

$$\begin{aligned} EX_i &= (0.6)(1) + (0.4)(-1) \\ &= \frac{1}{5}, \\ EX_i^2 &= 0.6 + 0.4 \\ &= 1. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathrm{Var}(X_i) &= 1 - \frac{1}{25} \\ &= \frac{24}{25}; \\ \text{thus,} \quad \sigma_{X_i} &= \frac{2\sqrt{6}}{5}. \end{aligned}$$

Therefore,

$$\begin{aligned} EY &= 25 \times \frac{1}{5} \\ &= 5, \\ \mathrm{Var}(Y) &= 25 \times \frac{24}{25} \\ &= 24; \\ \text{thus,} \quad \sigma_Y &= 2\sqrt{6}. \end{aligned}$$

$$\begin{aligned} P(4 \leq Y \leq 6) &= P(3.5 \leq Y \leq 6.5) \quad \text{(continuity correction)} \\ &= P\left(\frac{3.5-5}{2\sqrt{6}} \leq \frac{Y-5}{2\sqrt{6}} \leq \frac{6.5-5}{2\sqrt{6}}\right) \\ &= P\left(-0.3062 \leq \frac{Y-5}{2\sqrt{6}} \leq +0.3062\right) \\ &\approx \Phi(0.3062) - \Phi(-0.3062) \quad \text{(by the CLT)} \\ &= 2\Phi(0.3062) - 1 \\ &\approx 0.2405 \end{aligned}$$

Problem 5:

Let X_i be the number of sandwiches that the ith person needs, and let

$$Y = X_1 + X_2 + \cdots + X_{64}.$$

The goal is to find y such that

$$P(Y \leq y) \geq 0.95$$

First note that

$$\begin{aligned} EX_i &= \frac{1}{4}(0) + \frac{1}{2}(1) + \frac{1}{4}(2) \\ &= 1, \end{aligned}$$

$$\begin{aligned} EX_i^2 &= \frac{1}{4}(0^2) + \frac{1}{2}(1^2) + \frac{1}{4}(2^2) \\ &= \frac{3}{2}. \end{aligned}$$

Thus,

$$\begin{aligned} \text{Var}(X_i) &= EX_i^2 - (EX_i)^2 \\ &= \frac{3}{2} - 1 \\ &= \frac{1}{2} \quad \rightarrow \quad \sigma_{X_i} = \frac{1}{\sqrt{2}}. \end{aligned}$$

Thus,

$$\begin{aligned} EY &= 64 \times 1 \\ &= 64, \end{aligned}$$

$$\begin{aligned} \text{Var}(Y) &= 64 \times \frac{1}{2} \\ &= 32 \rightarrow \sigma_Y = 4\sqrt{2}. \end{aligned}$$

Now, we can use the CLT to find y

$$\begin{aligned} P(Y \leq y) &= P\left(\frac{Y-64}{4\sqrt{2}} \leq \frac{y-64}{4\sqrt{2}}\right) \\ &= \Phi\left(\frac{y-64}{4\sqrt{2}}\right) \quad \text{(by CLT).} \end{aligned}$$

Now, we can use the CLT to find y

$$\begin{aligned} P(Y \leq y) &= P\left(\frac{Y-64}{4\sqrt{2}} \leq \frac{y-64}{4\sqrt{2}}\right) \\ &= \Phi\left(\frac{y-64}{4\sqrt{2}}\right) \quad \text{(by CLT).} \end{aligned}$$

We can write

$$\Phi\left(\frac{y-64}{4\sqrt{2}}\right) = 0.95$$

Therefore,

$$\begin{aligned} \frac{y-64}{4\sqrt{2}} &= \Phi^{-1}(0.95) \\ &\approx 1.6449 \end{aligned}$$

Thus, $y = 73.3$.
Therefore, if you make 74 sandwiches, you are 95% sure that there is no shortage. Note that you can find the numerical value of $\Phi^{-1}(0.95)$ by running the norminv(0.95) command in MATLAB.

Problem 6:

Let $Y = X_1 + X_2 + \cdots + X_n$, so $\overline{X} = \frac{Y}{n}$. Since $X_i \sim Exponential(1)$, we have

$$E(X_i) = \frac{1}{\lambda} = 1, \qquad \text{Var}(X_i) = \frac{1}{\lambda^2} = 1.$$

Therefore,

$$E(Y) = nEX_i = n, \qquad \text{Var}(Y) = n\text{Var}(X_i) = n,$$

$$\begin{aligned} P(0.9 \leq \overline{X} \leq 1.1) &= P\left(0.9 \leq \frac{Y}{n} \leq 1.1\right) \\ &= P(0.9n \leq Y \leq 1.1n) \\ &= P\left(\frac{0.9n - n}{\sqrt{n}} \leq \frac{Y-n}{\sqrt{n}} \leq \frac{1.1n - n}{\sqrt{n}}\right) \\ &= P\left(-0.1\sqrt{n} \leq \frac{Y-n}{\sqrt{n}} \leq 0.1\sqrt{n}\right). \end{aligned}$$

By the CLT $\frac{Y-n}{\sqrt{n}}$ is approximately $N(0,1)$, so

$$\begin{aligned} P(0.9 \leq \overline{X} \leq 1.1) &\approx \Phi(0.1\sqrt{n}) - \Phi(-0.1\sqrt{n}) \\ &= 2\Phi(0.1\sqrt{n}) - 1 \quad (\text{since} \quad \Phi(-x) = 1 - \Phi(x)). \end{aligned}$$

We need to have

$$2\Phi(0.1\sqrt{n}) - 1 \geq 0.95, \qquad \text{so} \quad \Phi(0.1\sqrt{n}) \geq 0.975.$$

Thus,

$$\begin{aligned} 0.1\sqrt{n} &\geq \Phi^{-1}(0.975) = 1.96 \\ \sqrt{n} &\geq 19.6 \\ n &\geq 384.16 \end{aligned}$$

Since n is an integer, we conclude $n \geq 385$.

Problem 7:

Let X_i be the i^{th} measurement of a digital sample.

We can model

$X_i \sim \text{Uniform}(v - 0.5, v + 0.5), i = 1, \dots, 8$

where v is the exact value of the waveform sample.

The CD player produces the output

$U = \frac{W_8}{8}$ where $W_8 = \sum_{i=1}^{8} X_i$

we are interested in

$$P[|U-v|>0.1] = P\left[\left|W_8/8 - v\right| > 0.1\right]$$

$$= P\left[W_8/8 - v < -0.1\right] + P\left[W_8/8 - v > 0.1\right]$$

$$= P\left[W_8 < 8(v-0.1)\right] + P\left[W_8 > 8(v+0.1)\right]$$

(Assuming measurements are independent)
By the Central Limit Theorem,

$$W_8 \approx \mathcal{N}\left(8\mathbb{E}[X_i],\ 8\,\mathrm{Var}[X_i]\right)$$

$$\mathbb{E}[X_i] = \frac{v+0.5+v-0.5}{2} = v$$

$$\mathrm{Var}[X_i] = \frac{\left(v+0.5-(v-0.5)\right)^2}{12} = \frac{1}{12}$$

so $W_8 \approx \mathcal{N}\left(8v,\ 8/12\right)$

so

$$P[|U-v|>0.1] \approx 1 - Q\left(\frac{8v-0.8-8v}{\sqrt{8/12}}\right) + Q\left(\frac{8v+0.8-8v}{\sqrt{8/12}}\right)$$

$$= 2Q\left(\frac{0.8}{\sqrt{8/12}}\right) = 2Q\left(\frac{2\sqrt{6}}{5}\right)$$

$$\approx 0.3272$$

Problem 8:

Let X_i denote the i^{th} bit

We can model $X_i \sim \text{Bernoulli}(1/2)$, $i = 1, \ldots, 10^6$

(a) Let $W = \sum_{i=1}^{10^6} X_i$

$P[\text{at least } 502000 \text{ ones}]$

$= P[W \geq 502000]$

$\overset{CLT}{\approx} Q\left(\frac{502000 - 500000}{\sqrt{250000}}\right) \approx 3.17 \times 10^{-5}$

(where we used the CLT to approximate

$W \approx N(10^6 \underbrace{E[X_i]}_{=1/2}, 10^6 \underbrace{Var[X_i]}_{=1/4})$

$= N(500000, 250{,}000)$)

(b) We are interested in

$P[499{,}000 \leq W \leq 501{,}000]$

$= P[W \leq 501{,}000] - P[W \leq 499000]$

$\overset{CLT}{\approx} 1 - Q\left(\frac{501000 - 500000}{\sqrt{250000}}\right) - \left(1 - Q\left(\frac{499000 - 500000}{\sqrt{250000}}\right)\right)$

$= Q\left(\frac{-1000}{\sqrt{250000}}\right) - Q\left(\frac{1000}{\sqrt{25000}}\right)$

≈ 0.9544

Problem 9:

$$P\left(\sum_{1}^{100} R_i > 1030\right)$$

$$\sum_{1}^{100} R_i \sim N(1000, 200)$$ by central limit theorem

$$P\left(\sum_{1}^{100} R_i > 1030\right) \approx Q\left(\frac{1030-1000}{\sqrt{200}}\right)$$

$$= Q\left(30/\sqrt{200}\right) = Q\left(30/10\sqrt{2}\right)$$

$$= Q\left(\frac{3}{\sqrt{2}}\right)$$

$$= 0.0169$$

Problem 10:

$$S_N = \sum_{i=1}^{100} V_i \sim N(0, 100 \cdot 0.1)$$

$$= N(0, 10)$$ This is exact since sum of IID Gaussian random variables is Gaussian.

$$P[S_N > 5] = Q\left(5/\sqrt{10}\right)$$

$$= 0.0569$$

Problem 11:

$$\begin{aligned} Y &= X_1 + X_2 + \cdots + X_{50} \\ X_i &\sim Bernoulli(\frac{1}{2}) \\ EX_i &= \frac{1}{2} \\ \text{Var}(X_i) &= \frac{1}{4} \\ EY &= 50\frac{1}{2} = 25 \\ \text{Var}Y &= \frac{50}{4} = 12.5 \end{aligned}$$

Problem 12:

Let $X \sim Exponential(1)$. For $x \leq 0$, we have

$$F_{X_n}(x) = F_X(x) = 0, \qquad \text{for } n = 2, 3, 4, \cdots.$$

For $x \geq 0$, we have

$$\begin{aligned} \lim_{n\to\infty} F_{X_n}(x) &= \lim_{n\to\infty} \left(1 - \left(1 - \frac{1}{n}\right)^{nx}\right) \\ &= 1 - \lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^{nx} \\ &= 1 - e^{-x} \\ &= F_X(x), \qquad \text{for all } x. \end{aligned}$$

Thus, we conclude that $X_n \xrightarrow{d} X$.

Problem 13:

We have

$$\begin{aligned} \lim_{n\to\infty} P(|X_n - 0| \geq \epsilon) &= \lim_{n\to\infty} P(X_n \geq \epsilon) && (\text{ since } X_n \geq 0\ ) \\ &= \lim_{n\to\infty} e^{-n\epsilon} && (\text{ since } X_n \sim Exponential(n)\ ) \\ &= 0, \qquad \text{for all } \epsilon > 0. \end{aligned}$$

Problem 14:

The PDF of X_n is given by

$$f_{X_n}(x) = \begin{cases} n & 0 \le x \le \frac{1}{n} \\ \\ 0 & \text{otherwise} \end{cases}$$

We have

$$\begin{aligned} E\left(|X_n - 0|^r\right) &= \int_0^{\frac{1}{n}} x^r n \quad dx \\ &= \frac{1}{(r+1)n^r} \to 0, \qquad \text{for all } r \ge 1. \end{aligned}$$

Problem 15:

a. To show $X_n \ \xrightarrow{p}\ 0$, we can write, for any $\epsilon > 0$

$$\begin{aligned} \lim_{n\to\infty} P(|X_n| \ge \epsilon) &= \lim_{n\to\infty} P(X_n = n^2) \\ &= \lim_{n\to\infty} \frac{1}{n} \\ &= 0. \end{aligned}$$

We conclude that $X_n \ \xrightarrow{p}\ 0$.

b. For any $r \ge 1$, we can write

$$\begin{aligned} \lim_{n\to\infty} E\left(|X_n|^r\right) &= \lim_{n\to\infty} \left(n^{2r} \cdot \frac{1}{n} + 0 \cdot \left(1 - \frac{1}{n}\right)\right) \\ &= \lim_{n\to\infty} n^{2r-1} \\ &= \infty \qquad (\text{since } r \ge 1). \end{aligned}$$

Therefore, X_n does not converge in the rth mean for any $r \ge 1$. In particular, it is interesting to note that, although $X_n \ \xrightarrow{p}\ 0$, the expected value of X_n does not converge to 0.

Problem 16:

By the Theorem above, it suffices to show that

$$\sum_{n=1}^{\infty} P\big(|X_n| > \epsilon\big) < \infty.$$

Note that $|X_n| = \frac{1}{n}$. Thus, $|X_n| > \epsilon$ if and only if $n < \frac{1}{\epsilon}$. Thus, we conclude

$$\begin{aligned} \sum_{n=1}^{\infty} P\big(|X_n| > \epsilon\big) &\le \sum_{n=1}^{\lfloor \frac{1}{\epsilon} \rfloor} P\big(|X_n| > \epsilon\big) \\ &= \lfloor \frac{1}{\epsilon} \rfloor < \infty. \end{aligned}$$

Problem 17:

a. $Y_n \xrightarrow{d} 0$: Note that

$$F_{X_n}(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

Also, note that $R_{Y_n} = [0, 1]$. For $0 \le y \le 1$, we can write

$$\begin{aligned} F_{Y_n}(y) &= P(Y_n \le y) \\ &= 1 - P(Y_n > y) \\ &= 1 - P(X_1 > y, X_2 > y, \cdots, X_n > y) \\ &= 1 - P(X_1 > y)P(X_2 > y) \cdots P(X_n > y) \qquad \text{(since } X_i\text{'s are independent)} \\ &= 1 - (1 - F_{X_1}(y))(1 - F_{X_2}(y)) \cdots (1 - F_{X_n}(y)) \\ &= 1 - (1 - y)^n. \end{aligned}$$

Therefore, we conclude

$$\lim_{n\to\infty} F_{Y_n}(y) = \begin{cases} 0 & y \le 0 \\ 1 & y > 0 \end{cases}$$

Therefore, $Y_n \xrightarrow{d} 0$.

b. $Y_n \xrightarrow{p} 0$: Note that as we found in part (a)

$$F_{Y_n}(y) = \begin{cases} 0 & y < 0 \\ 1 - (1 - y)^n & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}$$

In particular, note that Y_n is a continuous random variable. To show $Y_n \xrightarrow{p} 0$, we need to show that

$$\lim_{n\to\infty} P\big(|Y_n| \ge \epsilon\big) = 0, \qquad \text{for all } \epsilon > 0.$$

Since $Y_n \ge 0$, it suffices to show that

$$\lim_{n\to\infty} P\big(Y_n \ge \epsilon\big) = 0, \qquad \text{for all } \epsilon > 0.$$

For $\epsilon \in (0, 1)$, we have

$$\begin{aligned} P\big(Y_n \ge \epsilon\big) &= 1 - P(Y_n < \epsilon) \\ &= 1 - P(Y_n \le \epsilon) \qquad \text{(since } Y_n \text{ is a continuous random variable)} \\ &= 1 - F_{Y_n}(\epsilon) \\ &= (1 - \epsilon)^n. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{n\to\infty} P\big(|Y_n| \ge \epsilon\big) &= \lim_{n\to\infty} (1 - \epsilon)^n \\ &= 0, \qquad \text{for all } \epsilon \in (0, 1]. \end{aligned}$$

c. $Y_n \xrightarrow{L^r} 0$, for all $r \geq 1$: By differentiating $F_{Y_n}(y)$, we obtain

$$f_{Y_n}(y) = \begin{cases} n(1-y)^{n-1} & 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Thus, for $r \geq 1$, we can write

$$\begin{aligned} E|Y_n|^r &= \int_0^1 ny^r(1-y)^{n-1}dy \\ &\leq \int_0^1 ny(1-y)^{n-1}dy \qquad \text{(since } r \geq 1\text{)} \\ &= \Big[-y(1-y)^n\Big]_0^1 + \int_0^1 (1-y)^n dy \qquad \text{(integration by parts)} \\ &= \frac{1}{n+1}. \end{aligned}$$

Therefore

$$\lim_{n\to\infty} E\left(|Y_n|^r\right) = 0.$$

d. $Y_n \xrightarrow{a.s} 0$: We will prove

$$\sum_{n=1}^{\infty} P\Big(|Y_n| > \epsilon\Big) < \infty,$$

which implies $Y_n \xrightarrow{a.s} 0$. By our discussion in part (b),

$$\begin{aligned} \sum_{n=1}^{\infty} P\Big(|Y_n| > \epsilon\Big) &= \sum_{n=1}^{\infty} (1-\epsilon)^n \\ &= \frac{1-\epsilon}{\epsilon} < \infty \qquad \text{(geometric series).} \end{aligned}$$

Problem 18:

Solution: For $x > 1$, we have

$$\lim_{n \to \infty} F_{X_n}(x) = \lim_{n \to \infty} \frac{e^{n(x-1)}}{1 + e^{n(x-1)}}$$
$$= 1$$

For $0 \leq x < 1$,

$$\lim_{n \to \infty} F_{X_n}(x) = \lim_{n \to \infty} \frac{e^{n(x-1)}}{1 + e^{n(x-1)}}$$
$$= 0$$

For $x < 0$,

$$F_{X_n}(x) = 0$$

Therefore,

$$\lim_{n \to \infty} F_{X_n}(x) = \begin{cases} 1 & x > 1 \\ 0 & x < 1 \end{cases}$$

Thus,

$$X \xrightarrow{d} 1$$

Problem 19:

MATLAB CODE:

```
clear all
clc

N=1000;
sampleMeans=zeros(N,1);%this will hold the sample mean for each n up to
N
for n=1:N
    sampleMeans(n)=mean(randn(n,1)+1);%compute the mean of n RVs
distributed as N(1,1)
end

plot(1:N,sampleMeans)
```