CS203B

Assignment 1

Priya Gole (200727)

---

1. A company asks you to set up a mechanism to send single bits (0 or 1) between two points A and B. The company has n identical communication channels available. However, these channels are not perfect. Each of these channels flips the bits send through it (0 becomes 1 or 1 becomes 0) with probability p. The company gives you two options for using these channels to connect A and B: either the channels can be connected in series between A and B or they can be connected in parallel between A and B. In the series mode, the output of the channel at the receiving end is taken as the received bit, while in the parallel mode, the bit which appears in the majority of the channel outputs is taken as the received bit. Answer the following questions:

   (a) If $p = \frac{1}{3}$ and $n = 5$, which mode of connecting the channels maximizes the probability that the correct bit is received at the receiving end?

   (b) Assume that $p = \frac{1}{3}$, $n = 3$ and the parallel mode is used for connecting A and B. A chooses the bit 0 with probability $\frac{2}{3}$ and the bit 1 with probability $\frac{1}{3}$ and sends this random bit to B. The majority of the bits received at B turns out to be equal to 1. Given this fact, what is the probability that the original bit sent by A was also equal to 1?

Sol. (a)

(i) Case 1 Series Connection:

We have 5 communication channels connected in series. Let's assume t number of channel flips the bit. Then for the output to be correct, the number of flips must be even, i.e., $t = 0, 2, 4$. Therefore, the probability that correct bit is received is

$$P_s = \sum_{t \in \{0,2,4\}} \binom{5}{t} p^t (1-p)^{5-t}$$

where, p is the probability that channel flips the bit.

(ii) Case 2 Parallel Connection:

We have 5 communication channels connected in parallel. Let's assume t number of channel flips the bit. Then for the output to be correct, the number of flipped bits must be less than the number of bits that do not flip, i.e., $t = 0, 1, 2$.

Therefore, the probability that correct bit is received is

$$P_p = \sum_{t \in \{0,1,2\}} \binom{5}{t} p^t (1-p)^{5-t}$$

where, p is the probability that channel flips the bit.

We are given $p = \frac{1}{3}$.

Subtracting P_s and P_p we get,

$$\begin{aligned}
P_s - P_p &= \sum_{t\in\{0,2,4\}} \binom{5}{t} p^t(1-p)^{5-t} - \sum_{t\in\{0,1,2\}} \binom{5}{t} p^t(1-p)^{5-t} \\
&= \binom{5}{4} \times \left(\frac{1}{3}\right)^4 \left(1-\frac{1}{3}\right)^1 - \binom{5}{1} \times \left(\frac{1}{3}\right)^1 \left(1-\frac{1}{3}\right)^4 \\
&= 5 \times \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^1 - 5 \times \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^4 \\
&= 5 \times \left(\frac{1}{3}\right) \times \left(\frac{2}{3}\right) \times \left(\left(\frac{1}{3}\right)^3 - \left(\frac{2}{3}\right)^3\right) \\
&= 5 \times \left(\frac{1}{3}\right) \times \left(\frac{2}{3}\right) \times \left(\frac{-7}{27}\right) \\
P_s - P_p &< 0
\end{aligned}$$

Therefore, $\boxed{P_s < P_p}$, i.e., the probability of getting the correct bit in parallel connection is more than that of series connection.

Sol. (b)

We are given 3 communication channels between A and B in parallel mode and it is given that A chooses 0 bit with probability $\frac{2}{3}$ and 1 bit with probability $\frac{1}{3}$.

We are given that majority of the output bits are 1, i.e., the outputs must be $\{1, 1, 1\}$ or $\{0, 1, 1\}$.

We are also given the probability with which the output flips $p = \frac{1}{3}$.

We have two cases here:

(i) When A gives input $= 0$ bit :

(1) When the output is $\{1, 1, 1\}$: bits from all three channels must be flipped. The probability of that happening is p^3.

(2) When the output is $\{0, 1, 1\}$: bits from two of the channels must be flipped. The probability of that happening is $\binom{3}{2}p^2(1-p) = 3p^2(1-p)$.

So, the probability when A gives input $= 0$ bit and majority of the output bits are 1 is $\boxed{P_0 = p^3 + 3p^2(1-p)}$

(i) When A gives input $= 1$ bit :

(1) When the output is $\{1, 1, 1\}$: bits from all three channels must pass unchanged. The probability of that happening is $(1-p)^3$.

(2) When the output is $\{0, 1, 1\}$: bits from one of the channels must be flipped. The probability of that happening is $\binom{3}{2}(1-p)^2p = 3p(1-p)^2$.

So, the probability when A gives input $=$ 1 bit and majority of the output bits are 1 is $\boxed{P_1 = (1-p)^3 + 3p(1-p)^2}$

We have to find probability (P) that A gives input 1 given that majority of output bits are 1.

$$\begin{aligned} P &= \frac{\frac{1}{3}P_1}{\frac{2}{3}P_0 + \frac{1}{3}P_1} \\ &= \frac{\frac{1}{3}\left((1-p)^3 + 3p(1-p)^2\right)}{\frac{2}{3}\left(p^3 + 3p^2(1-p)\right) + \frac{1}{3}\left((1-p)^3 + 3p(1-p)^2\right)} \end{aligned}$$

On substituting the value of $p = \frac{1}{3}$ and simplifying P we get: $\boxed{P = \frac{10}{17}}$

---

2. Suppose U is a continuous random variable with the probability density function ($c \in \mathbb{R}$)

$$g(u) = \begin{cases} c - |u|, & \text{if } |u| < \frac{1}{2}. \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the constant c.

(b) The *cumulative distribution function* of a random variable X is the function $F_X(x) = P(X \leq x)$ for every $x \in \mathbb{R}$. Find the cumulative distribution function F_U of U.

(c) Evaluate the conditional probability $\Pr\left(\frac{1}{8} < U < \frac{2}{5} \mid \frac{1}{10} < U < \frac{1}{5}\right)$.

Sol. (a)

$g(u)$ is probability density function. Hence,

$$\int_{-\infty}^{\infty} g(u)\, du = 1$$

$$\begin{aligned} \int_{-\infty}^{-\frac{1}{2}} g(u)\, du + \int_{-\frac{1}{2}}^{0} g(u)\, du + \int_{0}^{\frac{1}{2}} g(u)\, du + \int_{\frac{1}{2}}^{\infty} g(u)\, du &= 1 \\ \int_{-\infty}^{-\frac{1}{2}} 0\, du + \int_{-\frac{1}{2}}^{0} (c+u)\, du + \int_{0}^{\frac{1}{2}} (c-u)\, du + \int_{\frac{1}{2}}^{\infty} 0\, du &= 1 \\ 0 + \left(cu + \frac{u^2}{2}\right)\Bigg|_{-\frac{1}{2}}^{0} + \left(cu - \frac{u^2}{2}\right)\Bigg|_{0}^{\frac{1}{2}} + 0 &= 1 \\ -c\left(-\frac{1}{2}\right) - \frac{1}{2}\left(\frac{1}{2}\right)^2 + c\left(\frac{1}{2}\right) - \frac{1}{2}\left(\frac{1}{2}\right)^2 &= 1 \\ c - \frac{1}{4} &= 1 \\ c &= 1 + \frac{1}{4} \\ \boxed{c = \frac{5}{4}} \end{aligned}$$

Sol. (b) The cumulative distribution function $F_U(u) = P(U \leq u)$:

Case 1: When $u \leq -\frac{1}{2}$:

$$\begin{aligned} F_U(u) &= \int_{-\infty}^{u} g(x)\, dx \\ &= \int_{-\infty}^{u} 0\, dx \\ &= 0 \end{aligned}$$

Case 2: When $0 \geq u > -\frac{1}{2}$:

$$\begin{aligned}
F_U(u) &= \int_{-\infty}^{u} g(x)\,dx \\
&= \int_{-\infty}^{-\frac{1}{2}} g(x)\,dx + \int_{-\frac{1}{2}}^{u} g(x)\,dx \\
&= \int_{-\infty}^{-\frac{1}{2}} 0\,dx + \int_{-\frac{1}{2}}^{u} (c+x)\,dx \\
&= 0 + \left(cx + \frac{x^2}{2}\right)\Big|_{-\frac{1}{2}}^{u} \\
&= cu + \frac{u^2}{2} - \left(-\frac{c}{2} + \frac{1}{2}\left(-\frac{1}{2}\right)^2\right) \\
&= cu + \frac{u^2}{2} + \frac{c}{2} - \frac{1}{8} \\
&= \frac{5u}{4} + \frac{u^2}{2} + \frac{5}{8} - \frac{1}{8} \\
&= \frac{u^2}{2} + \frac{5u}{4} + \frac{1}{2}
\end{aligned}$$

Case 3: When $\frac{1}{2} \geq u > 0$:

$$\begin{aligned}
F_U(u) &= \int_{-\infty}^{u} g(x)\,dx \\
&= \int_{-\infty}^{-\frac{1}{2}} g(x)\,dx + \int_{-\frac{1}{2}}^{0} g(x)\,dx + \int_{0}^{u} g(x)\,dx \\
&= \int_{-\infty}^{-\frac{1}{2}} 0\,dx + \int_{-\frac{1}{2}}^{0} (c+x)\,dx + \int_{0}^{u} (c-x)\,dx \\
&= 0 + \left(cx + \frac{x^2}{2}\right)\Big|_{-\frac{1}{2}}^{0} + \left(cx - \frac{x^2}{2}\right)\Big|_{0}^{u} \\
&= -\left(c\left(-\frac{1}{2}\right) + \frac{1}{2}\left(-\frac{1}{2}\right)^2\right) + cu - \frac{u^2}{2} \\
&= \frac{c}{2} - \frac{1}{8} + cu - \frac{u^2}{2} \\
&= \frac{5}{8} - \frac{1}{8} + \frac{5u}{4} - \frac{u^2}{2} \\
&= \frac{1}{2} + \frac{5u}{4} - \frac{u^2}{2}
\end{aligned}$$

Case 4: When $u > \frac{1}{2}$:

$$\begin{aligned}
F_U(u) &= \int_{-\infty}^{u} g(x)\,dx \\
&= \int_{-\infty}^{-\frac{1}{2}} g(x)\,dx + \int_{-\frac{1}{2}}^{0} g(x)\,dx + \int_{0}^{\frac{1}{2}} g(x)\,dx + \int_{\frac{1}{2}}^{u} g(x)\,dx \\
&= \int_{-\infty}^{-\frac{1}{2}} 0\,dx + \int_{-\frac{1}{2}}^{0} (c+x)\,dx + \int_{0}^{\frac{1}{2}} (c-x)\,dx + \int_{\frac{1}{2}}^{u} 0\,dx \\
&= 0 + \left(cx + \frac{x^2}{2}\right)\Bigg|_{-\frac{1}{2}}^{0} + \left(cx - \frac{x^2}{2}\right)\Bigg|_{0}^{\frac{1}{2}} + 0 \\
&= -\left(c\left(-\frac{1}{2}\right) + \frac{1}{2}\left(-\frac{1}{2}\right)^2\right) + \left(c\left(\frac{1}{2}\right) - \frac{1}{2}\left(\frac{1}{2}\right)^2\right) \\
&= \frac{c}{2} - \frac{1}{8} + \frac{c}{2} - \frac{1}{8} \\
&= c - \frac{1}{4} \\
&= \frac{5}{4} - \frac{1}{4} \\
&= 1
\end{aligned}$$

Hence, cumulative distribution function $F_U(u)$ is given by:

$$F_U(u) = \begin{cases} 0, & \text{if } u \leq -\frac{1}{2}. \\ \frac{1}{2} + \frac{5u}{4} + \frac{u^2}{2}, & \text{if } 0 \geq u > -\frac{1}{2}. \\ \frac{1}{2} + \frac{5u}{4} - \frac{u^2}{2}, & \text{if } \frac{1}{2} \geq u > 0. \\ 1, & \text{if } u > \frac{1}{2}. \end{cases}$$

Sol. (c) The conditional probability $\Pr\left(\frac{1}{8} < U < \frac{2}{5} \mid \frac{1}{10} < U < \frac{1}{5}\right)$:

$$\begin{aligned}
Pr\left(\frac{1}{8} < U < \frac{2}{5} \;\middle|\; \frac{1}{10} < U < \frac{1}{5}\right) &= \frac{Pr\left(\frac{1}{8} < U < \frac{2}{5} \bigcap \frac{1}{10} < U < \frac{1}{5}\right)}{Pr\left(\frac{1}{10} < U < \frac{1}{5}\right)} \\
&= \frac{Pr\left(\frac{1}{8} < U < \frac{1}{5}\right)}{Pr\left(\frac{1}{10} < U < \frac{1}{5}\right)} \\
&= \frac{\int_{\frac{1}{8}}^{\frac{1}{5}} g(x)\,dx}{\int_{\frac{1}{10}}^{\frac{1}{5}} g(x)\,dx} \\
&= \frac{\int_{\frac{1}{8}}^{\frac{1}{5}} (c-x)\,dx}{\int_{\frac{1}{10}}^{\frac{1}{5}} (c-x)\,dx}
\end{aligned}$$

$$= \frac{\int_{\frac{1}{8}}^{\frac{1}{5}} (c - x)\, dx}{\int_{\frac{1}{10}}^{\frac{1}{5}} (c - x)\, dx}$$

$$= \frac{\left(cx - \frac{x^2}{2}\right)\Big|_{\frac{1}{8}}^{\frac{1}{5}}}{\left(cx - \frac{x^2}{2}\right)\Big|_{\frac{1}{10}}^{\frac{1}{5}}}$$

$$= \frac{\frac{5}{4}\left(\frac{1}{5} - \frac{1}{8}\right) - \frac{1}{2}\left(\left(\frac{1}{5}\right)^2 - \left(\frac{1}{8}\right)^2\right)}{\frac{5}{4}\left(\frac{1}{5} - \frac{1}{10}\right) - \frac{1}{2}\left(\left(\frac{1}{5}\right)^2 - \left(\frac{1}{10}\right)^2\right)}$$

$$= \frac{\frac{5}{4}\left(\frac{3}{40}\right) - \frac{1}{2}\left(\frac{39}{1600}\right)}{\frac{5}{4}\left(\frac{1}{10}\right) - \frac{1}{2}\left(\frac{3}{100}\right)}$$

$$= 0.741\, (approx.)$$

3. Alice has an unbiased 5-sided die and 5 different coins with her. The probabilities of obtaining a head on tosses of these coins are $\frac{1}{6}, \frac{2}{6}, \frac{3}{6}, \frac{4}{6}$ and $\frac{5}{6}$ respectively. She likes to observe patterns in subsequent tosses of these coins. Alice performs the following experiment. She rolls the 5-sided die and if the i^{th} side turns up, she chooses the i^{th} coin and starts tossing this coin repeatedly.

(a) What is the expected number of tosses required for obtaining 6 consecutive heads given that the side 1 turned up during the roll of the die?

(b) What is the expected number of tosses required for obtaining 6 consecutive heads while performing this random experiment?

(c) What is the probability that in the first n tosses, she obtains n consecutive heads?

(d) In the first n-tosses, she obtains n consecutive heads. Given this outcome, calculate the probability that the i^{th} side turned up during the roll of the die (the closed form expression for arbitrary i). How does these probabilities behave as $n \to \infty$?

Sol. (a) Side 1 turned up during the roll of the die, so we choose the coin whose probability of getting head is $p = \frac{1}{6}$.

Let x be the expected number of tosses required for obtaining 6 consecutive heads.

Start tossing the coin:

- If we get Tail (T), then expected number will be $x + 1$. The probability of getting (T) is $(1 - p)$.
- If we get Head then Tail (HT), then expected number will be $x + 2$. The probability of getting the sequence is $p(1 - p)$.
- If we get the sequence (HHT), then expected number will be $x + 3$. The probability of getting the sequence is $p^2(1 - p)$.
- If we get the sequence (HHHT), then expected number will be $x + 4$. The probability of getting the sequence is $p^3(1 - p)$.
- If we get the sequence (HHHHT), then expected number will be $x + 5$. The probability of getting the sequence is $p^4(1 - p)$.
- If we get the sequence (HHHHHT), then expected number will be $x + 6$. The probability of getting the sequence is $p^5(1 - p)$.
- If we get the sequence (HHHHHH), then expected number will be 6. The probability of getting the sequence is p^6.

Therefore,

$$x = (1-p)(1+x) + p(1-p)(2+x) + p^2(1-p)(3+x) + p^3(1-p)(4+x)+$$
$$p^4(1-p)(5+x) + p^5(1-p)(6+x) + p^6(6) \qquad \textbf{(1)}$$

$$
\begin{aligned}
x = &\left(\frac{5}{6}\right)(1+x) + \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)(2+x) + \left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)(3+x) + \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)(4+x) \\
&+ \left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)(5+x) + \left(\frac{1}{6}\right)^5\left(\frac{5}{6}\right)(6+x) + \left(\frac{1}{6}\right)^6(6)
\end{aligned}
$$

Simplifying this we get, $\boxed{x = 55986}$.

Sol. **(b)** The expected number of tosses required for obtaining 6 consecutive heads while performing the random experiment:

- On rolling the die, suppose we get 1 (with probability $\frac{1}{5}$). The expected number of tosses required to get consecutive 6 heads is 55986.
- On rolling the die, suppose we get 2 (with probability $\frac{1}{5}$). The expected number of tosses required to get consecutive 6 heads can be obtained from equation (1) by substituting $p = \frac{2}{6}$. On simplifying that, we get the expected number of tosses $= 1092$.
- On rolling the die, suppose we get 3 (with probability $\frac{1}{5}$). The expected number of tosses required to get consecutive 6 heads can be obtained from equation (1) by substituting $p = \frac{3}{6}$. On simplifying that, we get the expected number of tosses $= 126$.
- On rolling the die, suppose we get 2 (with probability $\frac{1}{5}$). The expected number of tosses required to get consecutive 6 heads can be obtained from equation (1) by substituting $p = \frac{4}{6}$. On simplifying that, we get the expected number of tosses $= 31.17 = 31(\textit{approx.})$ rounded off to nearest integer.
- On rolling the die, suppose we get 2 (with probability $\frac{1}{5}$). The expected number of tosses required to get consecutive 6 heads can be obtained from equation (1) by substituting $p = \frac{5}{6}$. On simplifying that, we get the expected number of tosses $= 11.92 = 12(\textit{approx.})$ rounded off to nearest integer.

Hence, the expected number of tosses required for obtaining 6 consecutive heads while performing the random experiment is:

$$
\begin{aligned}
P &= \left(\frac{1}{5}\right)55986 + \left(\frac{1}{5}\right)1092 + \left(\frac{1}{5}\right)126 + \left(\frac{1}{5}\right)31 + \left(\frac{1}{5}\right)12 \\
&= \frac{55986 + 1092 + 126 + 31 + 12}{5} \\
&= \frac{57247}{5} \\
&= 11449.4 \\
&= 11449 \quad (\textit{approx.})
\end{aligned}
$$

Sol. (c) Probability that in the first n tosses, n consecutive heads are obtained:

- On rolling the die, suppose we get 1 (with probability $\frac{1}{5}$). The probability that in the first n tosses, n consecutive heads are obtained is $\left(\frac{1}{6}\right)^n$.
- On rolling the die, suppose we get 2 (with probability $\frac{1}{5}$). The probability that in the first n tosses, n consecutive heads are obtained is $\left(\frac{2}{6}\right)^n$.
- On rolling the die, suppose we get 3 (with probability $\frac{1}{5}$). The probability that in the first n tosses, n consecutive heads are obtained is $\left(\frac{3}{6}\right)^n$.
- On rolling the die, suppose we get 4 (with probability $\frac{1}{5}$). The probability that in the first n tosses, n consecutive heads are obtained is $\left(\frac{4}{6}\right)^n$.
- On rolling the die, suppose we get 5 (with probability $\frac{1}{5}$). The probability that in the first n tosses, n consecutive heads are obtained is $\left(\frac{5}{6}\right)^n$.

Hence, the probability that in the first n tosses, n consecutive heads are obtained is:

$$P = \frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)$$

Sol. (d) Given that n consecutive heads are obtained:

(1) Probability that side 1 shows up during the roll of die:

$$P_1 = \frac{\frac{1}{5}\left(\frac{1}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)}$$

When $n \to \infty$,

$$\begin{aligned} P_1 &= \lim_{n\to\infty} \frac{\frac{1}{5}\left(\frac{1}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)} \\ &= \lim_{n\to\infty} \frac{1}{(1 + 2^n + 3^n + 4^n + 5^n)} \\ &= 0 \end{aligned}$$

(2) Probability that side 2 shows up during the roll of die:

$$P_2 = \frac{\frac{1}{5}\left(\frac{2}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)}$$

When $n \to \infty$,

$$\begin{aligned} P_2 &= \lim_{n\to\infty} \frac{\frac{1}{5}\left(\frac{2}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)} \\ &= \lim_{n\to\infty} \frac{1}{\left(\left(\frac{1}{2}\right)^n + \left(\frac{2}{2}\right)^n + \left(\frac{3}{2}\right)^n + \left(\frac{4}{2}\right)^n + \left(\frac{5}{2}\right)^n\right)} \\ &= 0 \end{aligned}$$

(3) Probability that side 3 shows up during the roll of die:

$$P_3 = \frac{\frac{1}{5}\left(\frac{3}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)}$$

When $n \to \infty$,

$$\begin{aligned} P_3 &= \lim_{n\to\infty} \frac{\frac{1}{5}\left(\frac{3}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)} \\ &= \lim_{n\to\infty} \frac{1}{\left(\left(\frac{1}{3}\right)^n + \left(\frac{2}{3}\right)^n + \left(\frac{3}{3}\right)^n + \left(\frac{4}{3}\right)^n + \left(\frac{5}{3}\right)^n\right)} \\ &= 0 \end{aligned}$$

(4) Probability that side 4 shows up during the roll of die:

$$P_4 = \frac{\frac{1}{5}\left(\frac{4}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)}$$

When $n \to \infty$,

$$\begin{aligned} P_4 &= \lim_{n\to\infty} \frac{\frac{1}{5}\left(\frac{4}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)} \\ &= \lim_{n\to\infty} \frac{1}{\left(\left(\frac{1}{4}\right)^n + \left(\frac{2}{4}\right)^n + \left(\frac{3}{4}\right)^n + \left(\frac{4}{4}\right)^n + \left(\frac{5}{4}\right)^n\right)} \\ &= 0 \end{aligned}$$

(5) Probability that side 5 shows up during the roll of die:

$$P_5 = \frac{\frac{1}{5}\left(\frac{5}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)}$$

When $n \to \infty$,

$$\begin{aligned} P_5 &= \lim_{n\to\infty} \frac{\frac{1}{5}\left(\frac{5}{6}\right)^n}{\frac{1}{5}\left(\left(\frac{1}{6}\right)^n + \left(\frac{2}{6}\right)^n + \left(\frac{3}{6}\right)^n + \left(\frac{4}{6}\right)^n + \left(\frac{5}{6}\right)^n\right)} \\ &= \lim_{n\to\infty} \frac{1}{\left(\left(\frac{1}{5}\right)^n + \left(\frac{2}{5}\right)^n + \left(\frac{3}{5}\right)^n + \left(\frac{4}{5}\right)^n + \left(\frac{5}{5}\right)^n\right)} \\ &= 1 \end{aligned}$$

---

4. In this programming exercise, let us explore the behavior of the averages,

$$\frac{X_1 + X_2 + X_3 + \cdots + X_n}{n}$$

of independent and identically distributed random variables X_i as $n \to \infty$.

(a) Implement the following program using any language and graph plotting library of your choice (You are encouraged to use Python with Matplotlib for this assignment). Consider the random variable X which takes the values $0, 1, 2, 3, \cdots, m-1$ with the respective probabilities $p_0, p_1, p_2, p_3 \cdots p_{m-1}$ such that $p_0 + p_1 + p_2 + p_3 + \cdots + p_{m-1} = 1$. Your program should take as inputs, the value m, the probabilities $p_0, p_1, p_2, p_3 \cdots p_{m-1}$ and n which is the number of samples to be generated. Generate n samples according to the distribution of X and calculate the average value of the samples generated. Repeat this sampling and averaging process for a fairly large number of iterations and store the average values obtained. Round each of the average values to the nearest integer and generate a plot of the frequency of the rounded averages thus obtained against the range of possible values. The above programs should be included in the submitted .zip file with the name [Your roll number]_Q4 and the appropriate extension according to your choice of the programming language (for e.g. 20xxxx_Q4.py). You should also include a readme document explaining how this program can be executed.

(b) The following questions should be answered in your main answer script. Give a brief account of how you implemented random sampling according to the required distributions in your program. Also, use your program to answer the following questions:

(i) How does the frequency plot of the averages behave as $n \to \infty$?

(ii) Does the shape of the frequency plot change on varying m or the values of the probabilities? Can you interpret the shape of the plots for these distributions in terms of any of the concepts that were discussed in class?

Sol. (a) The program is implemented in 200727_Q4.zip

Implementation:

- Program takes the input m, n and probabilities from $p_0, p_1, p_2, p_3 \cdots p_{m-2}$ (probability p_{m-1} is calculated by this equation: $p_0 + p_1 + p_2 + p_3 + \cdots + p_{m-1} = 1$)
- For number of iterations itr, n samples are generated according the given probability distribution and those samples are stored in the list named *samples*.
- Average for each sample is calculated and rounded upto 2 decimal places. We store the average values in the list named *avg*.
- Then the frequency of average is calculated and the graph is plotted using matplotlib.

Sol. (b) The random sampling is implemented using *random.choice*, a built-in function which generates a random sample of elements. It takes parameter p which denotes the probability of the respective elements.

(i) Frequency plot of the averages as $n \to \infty$:

- The plot mimics the normal distribution curve.
- The peak is obtained near the expected value of X.

(ii)

- If we take a fairly large value of n, the value of m does not affect the value of shape of the curve, the shape still mimics bell curve.
- On changing the values of probabilities, we change the expected value of X hence shifting the peak of the curve. Therefore, probabilities affect the maxima of the curve.