

Probability and Statistics (MA6.101)

Monsoon 2021, IIIT Hyderabad
Assignment 3

Abhinav S Menon

Problem 1

We have a random variable X whose MGF is given by

$$M_X(t) = \frac{1}{10}e^{-20t} + \frac{1}{5}e^{-3t} + \frac{3}{10}e^{4t} + \frac{2}{5}e^{5t}.$$

We wish to find $P(|X| \leq 2)$.

We know that for any random variable Y , the MGF of Y is given by

$$M_Y(t) = E[e^{tY}] = \sum_S e^{tx} P(Y = t).$$

Therefore, comparing with M_X , we see that the PMF of X must be

$$P(X = k) = \begin{cases} \frac{1}{10} & k = -20 \\ \frac{1}{5} & k = -3 \\ \frac{3}{10} & k = 4 \\ \frac{2}{5} & k = 5. \end{cases}$$

Now, for all $x \in S_X$, $|x| > 2$; therefore $P(|X| \leq 2) = 0$.

Problem 2

We have a random variable X with MGF $M(t)$, *i.e.*,

$$M(t) = \sum_S e^{tx} P(x).$$

Part 1

We need to find the MGF $M_1(t)$ of the random variable given by kX . Let $Y = kX$.

We have

$$\begin{aligned}
M_1(t) &= \sum_S e^{ty} P(Y = y) \\
&= \sum_S e^{tkx} P(Y = kx) \\
&= \sum_S e^{(tk)x} P(X = x) \\
&= M(tk).
\end{aligned}$$

Thus, we have $M_1(t) = M(tk)$ for all t .

Part 2

We need to find the MGF $M_2(t)$ of the random variable given by X_k . Let $Z = X + k$.

We have

$$\begin{aligned}
M_2(t) &= \sum_S e^{tz} P(Z = z) \\
&= \sum_S e^{t(x+k)} P(Z = k + x) \\
&= \sum_S e^{tk} e^{tx} P(X = x) \\
&= e^{tk} M(t).
\end{aligned}$$

Thus, we have $M_2(t) = e^{tk} M(t)$ for all t .

Part 3

We know that the probability distribution of Y is the same as that of $X + k$, *i.e.*, $P(Y = s) = P(X = s + k)$. We need to find the MGF $M_3(t)$ of Y .

We have

$$\begin{aligned}
M_3(t) &= \sum_S e^{ty} P(Y = y) \\
&= \sum_S e^{ty} P(X = y + k) \\
&= \sum_S e^{t(x-k)} P(X = x) \\
&= \frac{1}{e^{tk}} \sum_S e^{tx} P(X = x) \\
&= \frac{1}{e^{tk}} M(t).
\end{aligned}$$

Thus, we have $M_3(t) = e^{-tk} M(t)$ for all t .

Part 4

We know that the probability distribution of Y is the same as that of $2X$, *i.e.*, $P(Y = s) = P(X = 2s)$. We need to find the MGF $M_4(t)$ of Y .

We have

$$\begin{aligned} M_4(t) &= \sum_S e^{ty} P(Y = y) \\ &= \sum_S e^{ty} P(X = 2y) \\ &= \sum_S e^{t \frac{x}{2}} P(X = x) \\ &= \sum_S e^{\frac{t}{2}x} P(X = x) \\ &= M\left(\frac{t}{2}\right). \end{aligned}$$

Thus, we have $M_4(t) = M\left(\frac{t}{2}\right)$ for all t .

Problem 3

n people have thrown their hats in a box and each one then picks a hat at random. The number of people who have got their own hat back is called X . We wish to find $E[X]$.

Suppose that $X = k$; then k is the number of fixed points of the permutation of hats.

Let us define n random variables F_k , where $F_k = 1$ iff the k^{th} person got their hat back. Then, clearly, $X = \sum_{k=1}^n F_k$ (the number of fixed points of the permutation is the same as the number of F_k whose value is 1).

Now, since expectation is linear, $E[X] = \sum_{k=1}^n E[F_k]$. To find $E[F_k]$, note that it is a Bernoulli variable with parameter $p =$ (the probability that the k^{th} person didn't get their hat back). Thus its expected value is $1 - p$, the probability that the k^{th} person got their hat back, which is

$$\frac{(n-1)!}{n!} = \frac{1}{n},$$

as there are $(n-1)!$ permutations assigning the k^{th} person's hat to them, and $n!$ total permutations.

This gives us

$$\begin{aligned} E[X] &= \sum_{k=0}^n E[F_k] \\ &= \sum_{k=0}^n \frac{1}{n} \\ &= 1. \end{aligned}$$

Problem 4

Let the lifetimes of bulbs A and B be L_A and L_B respectively. We have that $L_A \sim \text{Poisson}(0.25)$, and $L_B \sim \text{Poisson}(0.5)$, where the times are in years.

Part 1

We replace the bulbs alternately, *i.e.* A with B and vice versa. Thus, if we make 3 replacements, the lifetime can be modelled as a variable $L(3) = L_A + L_B + L_A + L_B$.

From this, we can see that $E[L(3)] = E[L_A] + E[L_B] + E[L_A] + E[L_B]$, by the linearity of expectation. We also know that the expected value of a Poisson variable is its parameter λ ; this gives us $E[L(3)] = 0.25 + 0.5 + 0.25 + 0.5 = 1.5$ years.

Part 2

Let $L(n)$ be a random variable denoting the total illumination time, given we do n replacements, starting with a bulb of type A . Therefore, $L(0) = L_A$. We wish to find $E[L(n)]$.

When the current bulb's life ends, we replace it with a bulb of type A with probability p , and one of type B with probability $1 - p$.

Let X_n be a variable denoting the n^{th} replacement; its value is 0 if we choose bulb A , and 1 if we choose B .

With this information, we can find the distribution of $L(n)$ in terms of $L(n-1)$, for any $n > 0$. When the time after n replacements is over, with probability p , we use a bulb of type A , and with probability $(1-p)$, we use one of type B . This means that

$$L(n) = \begin{cases} L(n-1) + 0.25 & \text{if } X_n = 0, \\ L(n-1) + 0.5 & \text{if } X_n = 1. \end{cases}$$

We can write this more concisely as $L(n) = L(n-1) + 0.25(1 - X_n) + 0.5(X_n)$.

With some rearrangement, we get

$$\begin{aligned} L_n &= L(n-1) + 0.25(1 + X_n) \\ &= L(0) + \sum_{i=1}^n 0.25(1 + X_n) \\ &= L_A + \sum_{i=1}^n 0.25(1 + X_n). \end{aligned}$$

Now, we can calculate the expected value of X_n from the fact that $X_n \sim \text{Bernoulli}(p)$, *i.e.*, $E[X_n] = 1 - p$. We can use this to find $E[L(n)]$ as follows.

$$\begin{aligned} E[L(n)] &= E \left[0.25 + \sum_{i=1}^n 0.25(1 + X_n) \right] \\ &= E[L_A] + \sum_{i=1}^n E[0.25(1 + X_n)] \\ &= 0.25 + \sum_{i=1}^n 0.25(1 + E[X_n]) \\ &= 0.25 + \sum_{i=1}^n 0.25(2 - p) \\ &= 0.25 + 0.5n - 0.25np. \end{aligned}$$

Problem 5

We know that the net gain G is a random variable with the PMF

$$P(G = g) = \begin{cases} \frac{1}{3} & g = -2 \\ \frac{1}{2} & g = 1 \\ \frac{1}{6} & g = 3 \\ 0 & \text{otherwise.} \end{cases}$$

Vinay and Mahesh play alternately, with Vinay starting.

Part 1

Let N be a variable denoting the number of rounds until Vinay and Mahesh lose in succession. Clearly, $N \geq 1$.

If $N = k$, it means that for $k-1$ rounds, either Vinay or Mahesh (or both) won their plays, and in the k^{th} round, they both lost. The probability that at least one of them won their play is

$$1 - \left(\frac{1}{3} \cdot \frac{1}{3} \right) = \frac{8}{9}.$$

Thus, we can say that $N \sim \text{Geometric}(\frac{1}{9})$, and it has the distribution

$$P(N = k) = \begin{cases} \left(\frac{1}{9}\right) \left(\frac{8}{9}\right)^{k-1} & k \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Part 2

We have said that Z is a variable denoting the time at which Mahesh has his third loss. Clearly, $Z \geq 6$. Suppose this happens at time $2k$ (since Mahesh only plays at even times).

This means that Mahesh has played $(k-1)$ rounds, in which there are exactly two losses. Further, he has lost the k^{th} play. The chance of this is

$$\binom{k-1}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^{k-3} \cdot \left(\frac{1}{3}\right).$$

Thus, we can find the PMF of Z as

$$P(Z = 2k) = \begin{cases} \binom{k-1}{2} \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^{k-3} & k \geq 3 \\ 0 & \text{otherwise.} \end{cases}$$

Part 3

We have defined N as the number of rounds until each of Vinay and Mahesh has won at least once. If $N = k$, it means that one of Vinay and Mahesh has won their k^{th} play, and the other one has won one of the preceding plays.

If $N = k$, we distinguish two cases: Vinay won the k^{th} round, and Mahesh won it.

If Vinay won in the k^{th} round, Mahesh won one of his first $k-1$ plays, and they both lost all their other plays. Thus, the probability of this is

$$\left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^{k-1} \cdot (k-1) \left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^{k-2}.$$

If Mahesh won in the k^{th} round, Vinay won one of his first k plays, and they both lost all their other plays. Thus, the probability of this is

$$\left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^{k-1} \cdot k \left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^{k-1}.$$

The total probability of $N = k$ is, then, the sum of these, *i.e.*,

$$\left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^{k-1} \cdot \left(\frac{2}{3}\right) \left(\frac{1}{3}\right)^{k-1} (4k-3) = 4 \left(\frac{1}{9}\right)^k (4k-3).$$

Thus we can find the PMF of N as

$$P(N = k) = \begin{cases} 4 \left(\frac{1}{9}\right)^k (4k - 3) & k \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Now we need to find $E[N]$, for which we can proceed as follows.

$$\begin{aligned} E[N] &= \sum_{k=1}^{\infty} k P(N = k) = \sum_{k=1}^{\infty} 4k(4k - 3) \left(\frac{1}{9}\right)^k \\ &= 16 \sum_{k=1}^{\infty} k^2 \left(\frac{1}{9}\right)^k - 12 \sum_{k=1}^{\infty} k \left(\frac{1}{9}\right)^k \\ &= 16S_2 - 12S_1. \end{aligned}$$

To find S_1 , we will use the fact that the infinite sum of an AGP $a + (a + d)r + (a + 2d)r^2 + \dots$ is given by

$$\frac{a}{1 - r} + \frac{dr}{(1 - r)^2}$$

For S_1 , $a = 0, d = 1, r = \frac{1}{9}$, which gives us

$$S_1 = \frac{\frac{1}{9}}{\frac{8}{9}} = \frac{9}{64}.$$

To find S_2 ,

$$\begin{aligned} S_2 &= 1^2 \left(\frac{1}{9}\right) + 2^2 \left(\frac{1}{9}\right)^2 + 3^2 \left(\frac{1}{9}\right)^3 + \dots \\ \frac{1}{9}S_2 &= 1^2 \left(\frac{1}{9}\right)^2 + 2^2 \left(\frac{1}{9}\right)^3 + \dots \\ \frac{8}{9}S_2 &= \frac{1}{9} + 3 \left(\frac{1}{9}\right)^2 + 5 \left(\frac{1}{9}\right)^3 + \dots, \end{aligned}$$

which gives us

$$\frac{8}{9}S_2 = \frac{1}{9} \left[\sum_{i=0}^{\infty} (2i + 1) \left(\frac{1}{9}\right)^i \right],$$

which we can find using the above formula, keeping $a = 1, d = 2, r = \frac{1}{9}$:

$$\begin{aligned}
\frac{8}{9}S_2 &= \frac{1}{9} \left(\frac{1}{1 - \frac{1}{9}} + \frac{\frac{2}{9}}{(1 - \frac{1}{9})^2} \right) \\
&= \frac{1}{9} \left(\frac{9}{8} + \frac{18}{64} \right) \\
&= \frac{1}{8} + \frac{1}{32} \\
&= \frac{5}{32} \\
\Rightarrow S_2 &= \frac{45}{256}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
E[N] &= 16S_2 - 12S_1 \\
&= 16 \left(\frac{45}{256} \right) - 12 \left(\frac{9}{64} \right) \\
&= \frac{45 - 27}{16} = \frac{18}{16} \\
&= \frac{9}{8}.
\end{aligned}$$

Problem 6

We have assigned one of 225 colours to each of n balls, where each colour is equally likely.

Part 1

We wish to find the expected number of colours assigned to exactly k balls. Let N_k be a random variable denoting the number of colours assigned to k balls; then what we need is $E[N_k]$.

First, define 225 random variables C_1, \dots, C_{225} , where C_i represents the number of balls to which the i^{th} colour has been assigned. The PMF of each C_i is then

$$P(C_i = c) = \begin{cases} \binom{n}{c} \left(\frac{1}{225} \right)^c \left(\frac{224}{225} \right)^{n-c} & 0 \leq c \leq n \\ 0 & \text{otherwise.} \end{cases}$$

Let T_i^k be a random variable whose value is 1 iff $C_i = k$, and 0 otherwise. Then we know that

$$P(T_i^k = 1) = \binom{n}{k} \left(\frac{1}{225} \right)^k \left(\frac{224}{225} \right)^{n-k},$$

and that

$$E[T_i^k] = \binom{n}{k} \left(\frac{1}{225}\right)^k \left(\frac{224}{225}\right)^{n-k}.$$

Therefore $N_k = \sum_{i=0}^{225} T_i^k$. This means that

$$E[N_k] = \sum_{i=0}^{225} E[T_i^k],$$

since expected values can be summed even in the case of non-independent variables (which the T_i^k are), and so

$$E[N_k] = 225 \binom{n}{k} \left(\frac{1}{225}\right)^k \left(\frac{224}{225}\right)^{n-k}.$$

As an additional verification, note that $\sum_{k=0}^n N_k = 225$, which holds up as

$$\begin{aligned} \sum_{k=0}^n t_k &= \sum_{k=0}^n \binom{n}{k} \left(\frac{1}{225}\right)^k \left(\frac{224}{225}\right)^{n-k} \\ &= \left(\frac{1}{225} + \frac{224}{225}\right)^n \\ &= 1, \end{aligned}$$

as expected.

Therefore,

$$E[N_k] = 225 \binom{n}{k} \left(\frac{1}{225}\right)^k \left(\frac{224}{225}\right)^{n-k}.$$

Part 2

We want to find the expected number of colours assigned to more than one ball. Extending the notation of Part 1, we denote the number of colours assigned to more than one ball as $N_{>1}$. We want to find $E[N_{>1}]$.

We noted in Part 1 that $\sum_{k=0}^n N_k = 225$, which implies that

$$N_{>1} + N_{\leq 1} = 225;$$

but $N_{\leq 1} = N_0 + N_1$. This gives us

$$N_{>1} = 225 - N_0 - N_1,$$

which in turn implies that

$$E[N_{>1}] = 225 - (E[N_0] + E[N_1]).$$

We can calculate the required values by substitution:

$$\begin{aligned}
E[N_0] &= 225 \binom{n}{k} \left(\frac{1}{225}\right)^k \left(\frac{224}{225}\right)^{n-k} \Big|_{k=0} \\
&= 225 \binom{n}{0} \left(\frac{1}{225}\right)^0 \left(\frac{224}{225}\right)^n \\
&= 225 \cdot \left(\frac{224}{225}\right)^n. \\
E[N_1] &= 225 \binom{n}{k} \left(\frac{1}{225}\right)^k \left(\frac{224}{225}\right)^{n-k} \Big|_{k=1} \\
&= 225 \binom{n}{1} \left(\frac{1}{225}\right)^1 \left(\frac{224}{225}\right)^{n-1} \\
&= 225 \cdot n \cdot \left(\frac{1}{225}\right) \left(\frac{224}{225}\right)^{n-1}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
E[N_0] + E[N_1] &= 225 \cdot \left(\frac{224}{225}\right)^n + 225 \cdot n \cdot \left(\frac{1}{225}\right) \left(\frac{224}{225}\right)^{n-1} \\
&= 225 \left(\frac{224}{225}\right)^{n-1} \left[\frac{224}{225} + \frac{n}{225} \right] \\
&= (224 + n) \left(\frac{224}{225}\right)^{n-1},
\end{aligned}$$

and

$$E[N_{>1}] = 225 - (224 + n) \left(\frac{224}{225}\right)^{n-1}.$$

Problem 7

We are given that a 500-page book contains 500 misprints on average. We can assume that the probability of an error occurring on a page is independent of any other page and any other error; thus the expected number of errors per page is 1.

Since we need to consider the errors as a Poisson process, if X represents the number of errors on a given page, we know that $X \sim \text{Poisson}(1)$. We need to find $P(X \geq 3)$, which by the law of total probability is $1 - P(X = 0) - P(X = 1) - P(X = 2)$.

$$\begin{aligned}
P(X = 0) &= \frac{1^k}{k!} e^{-1} \Big|_{k=0} \\
&= \frac{1}{1} e^{-1} \\
&\approx 0.3679 \\
P(X = 1) &= \frac{1^k}{k!} e^{-1} \Big|_{k=1} \\
&= \frac{1}{1} e^{-1} \\
&\approx 0.3679 \\
P(X = 2) &= \frac{1^k}{k!} e^{-1} \Big|_{k=2} \\
&= \frac{1}{2} e^{-1} \\
&\approx 0.1839
\end{aligned}$$

This gives us

$$\begin{aligned}
P(X \geq 3) &= 1 - [P(X = 0) + P(X = 1) + P(X = 2)] \\
&\approx 1 - (0.3679 + 0.3679 + 0.1839) \\
&= 1 - 0.9197 \\
&= 0.0803.
\end{aligned}$$

Problem 8

It is given that $G = (V, E)$ is an undirected graph and $a : V \rightarrow \{R, G, B\}$ is a 3-colour assignment to the vertices of G . The set of monochromatic edges $E(a) = \{(u, v) \in E \mid a(u) = a(v)\}$ is the set of edges with the same colours on its endpoints. a is randomly chosen.

Part 1

For any edge $e \in E$, X_e is a random variable such that $X_e = 1$ when $e \in E$ and $X_e = 0$ otherwise. We are given that all X_e are pairwise independent; we wish to prove that some X_i is not independent of all the others.

This can be proven by finding a case in which the value of X_e for some $e \in E$ is determined by the value of all the others. Let $Z = X_{e'}$ be some variable, where $e' \in E$. Consider the case where $X_e = 1$ for all $e \in E - e'$. This means that all the edges are monochromatic. **Assuming that the graph is connected**, this tells us that all vertices have the same colour – including the endpoints of e' . This means that $Z = 1$, which shows that Z is not independent of all variables in $\{X_e\} - Z$.

Part 2

Y is given to be the random variable denoting the number of non-monochromatic edges. We need to find $E[Y]$.

Let $|E| = n$, and let the edges be numbered e_1, \dots, e_n . Then, with the X_{e_i} defined as in Part 1, we can see that $Y =$ (the number of X_{e_i} whose value is 0), which means that

$$Y = n - \sum_{i=1}^n X_{e_i}.$$

As the expectation value is linear for non-independent variables also (like the X_{e_i}), it follows that

$$E[Y] = n - \sum_{i=1}^n E[X_{e_i}].$$

It can easily be seen that for each i ,

$$P(X_{e_i} = x) = \begin{cases} \frac{2}{3} & x = 0 \\ \frac{1}{3} & x = 1. \end{cases}$$

Therefore, $E[X_{e_i}] = \frac{1}{3}$ for all i , from which we get

$$E[Y] = \frac{2n}{3}.$$

Problem 9

The experiment is a series of independent trials, each of which consists of simultaneously flipping a set of Z fair coins.

Part 1

Given that the previous trial resulted in Z tails, we wish to find the probability of the next two trials also having this result.

We know that the trials are all independent; therefore the precondition can be ignored. The probability is then straightforwardly

$$\left(\frac{1}{2^Z}\right)^2.$$

Part 2

Now, the first trial uses M coins, and each time all coins land on the same side, a coin is removed (until only one coin is left). X is a random variable denoting the number of trials performed.

Suppose $X = n$. This means that after the n^{th} trial, one coin was removed from 2; thus the n^{th} trial landed both of the remaining coins on the same side. Furthermore, of the first $n - 1$ trials, $M - 2$ trials consisted of coins landing on the same side (each with a successively decreasing number of coins). The probability of this happening is

$$\binom{n-1}{M-2} \left(\frac{2}{2^M}\right) \cdot \left(\frac{2}{2^{M-1}}\right) \cdots \left(\frac{2}{2^3}\right) \frac{1}{2},$$

which is equivalent to

$$\binom{n-1}{M-2} \left(\frac{2^{M-2}}{2^{\frac{M(M+1)}{2}-2}}\right),$$

or

$$\binom{n-1}{M-2} \left(\frac{1}{2^{\frac{M(M-1)}{2}}}\right).$$