EE2012 Tutorial 3

Dr Feng LIN

Instructor

Feng LIN (Frank)

Email: feng_lin@nus.edu.sg

Location:

- E1A-04-01
- Level 6, #06-02, T-Lab Building

Summary of Chapter 3

- **Random Variables (RVs)**
  - A random variable maps outcomes of an experiment to real numbers.
  - Types of RVs:
    - Discrete (countable values, e.g., dice, coin flips)
    - Continuous (uncountable values, e.g., measurements)
    - Mixed-type
- **Why Use Random Variables?**
  - Simplify representation of events (e.g., number of heads in coin tosses).
  - Events can be described in terms of the RV (e.g., $\{X > 1\} = \text{Tom wins}$).
  - Easier to compute probabilities.

Summary of Chapter 3

- **Probability Mass Function (PMF)**

  Defines $p_X(x_i) = P(X = x_i)$ for all x_i in support S_X.
  - $p_X(x) \geq 0$
  - $\sum_{x \in S_X} p_X(x) = 1$
  - Event probabilities are sums of PMFs
- **Expectation (Mean) of RVs**

  $E[X] = \sum_{x \in S_X} x\, p_X(x) \rightarrow$ probabilistic weighted average.
  - **Median**: middle value of support
  - **Mode**: most likely value
- **Variance and Standard Deviation**

  Measures how spread out values are from the mean.
  - Variance: $\mathrm{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$
  - Standard deviation $= \sqrt{\mathrm{Var}(X)}$

Summary of Chapter 3

Important Discrete Random Variables

RV	PMF	S_X	m_X and σ_X^2
Discrete Uniform $U(a,b)$	$p_X(k) = \begin{cases} \dfrac{1}{b-a+1}, & k \in S_X \\ 0, & k \notin S_X \end{cases}$	$S_X = \{a, a+1, \cdots, b\}$	$m_X = (a+b)/2$ $\sigma_X^2 = \dfrac{(b-a+2)(b-a)}{12}$
Bernoulli (p)	$p_X(k) = \begin{cases} p, & k = 1 \\ 1-p, & k = 0 \end{cases}$	$S_X = \{0, 1\}$	$m_X = p$ $\sigma_X^2 = p(1-p)$
Binomial (n,p)	$p_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k}, & k \in S_X \\ 0, & k \notin S_X \end{cases}$	$S_X = \{0, 1, \cdots, n\}$	$m_X = np$ $\sigma_X^2 = np(1-p)$
Geometric (p)	$p_X(k) = (1-p)^{k-1} p$	$S_X = \{1, 2\ \cdots\}$	$m_X = p^{-1}$ $\sigma_X^2 = (1-p)/p^2$
Poisson (λ)	$p_X(k) = \dfrac{\lambda^k}{k!} e^{-\lambda}$	$S_X = \{0,1,2\ \cdots\}$	$m_X = \lambda$ $\sigma_X^2 = \lambda$

Summary of Chapter 3

Cumulative Distribution Function (CDF)

- Defined as $F_X(x) = P(X \leq x)$
- Properties: staircase-shaped (for discrete RVs), non-decreasing, right-continuous.
- Probabilities of ranges can be derived from differences of CDF values.

Q1

The random variable Y has the following probability mass function (PMF):

$$p_Y(y) = \begin{cases} \alpha y^2, & y = 1,2,3,4 \\ 0, & \text{otherwise} \end{cases}$$

(a) What is the value of the constant α?

(b) What is $P(Y \in \{v^2 | v = 1, 2, 3, \dots\})$?

(c) Find the probability that Y is even.

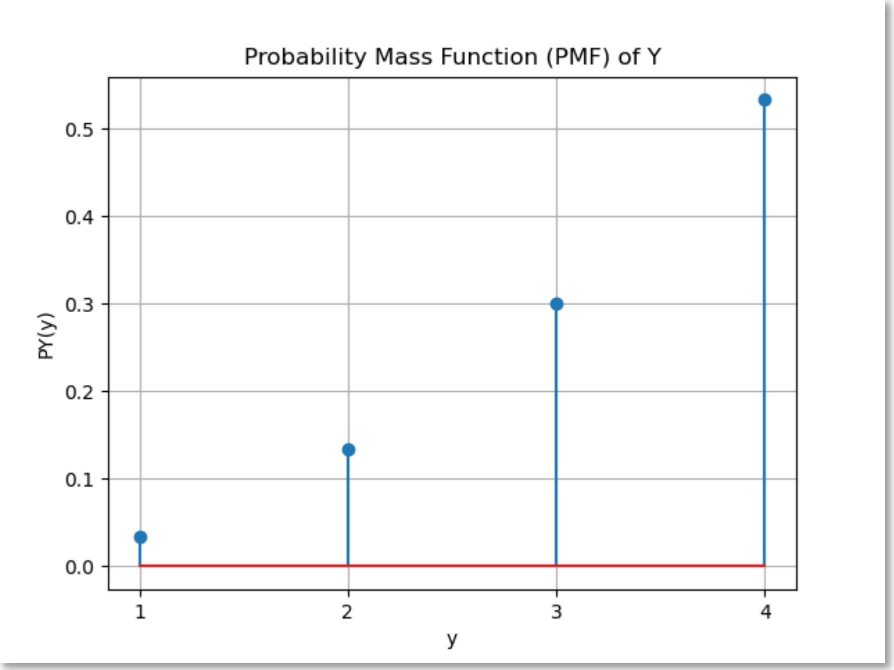
(d) Find $P(Y \geq 2)$.

Q1(a) What is the value of the constant α?

To find α, we use the fact that the sum of the probabilities for all possible values of Y must equal 1. That is:

$$\sum_{y=1}^{4} p_Y(y) = \alpha(1 + 2^2 + 3^2 + 4^2) = 1.$$

This implies $\alpha = 1/30$



Q1(b) What is $P(Y \in \{v^2 | v = 1, 2, 3, \ldots\})$?

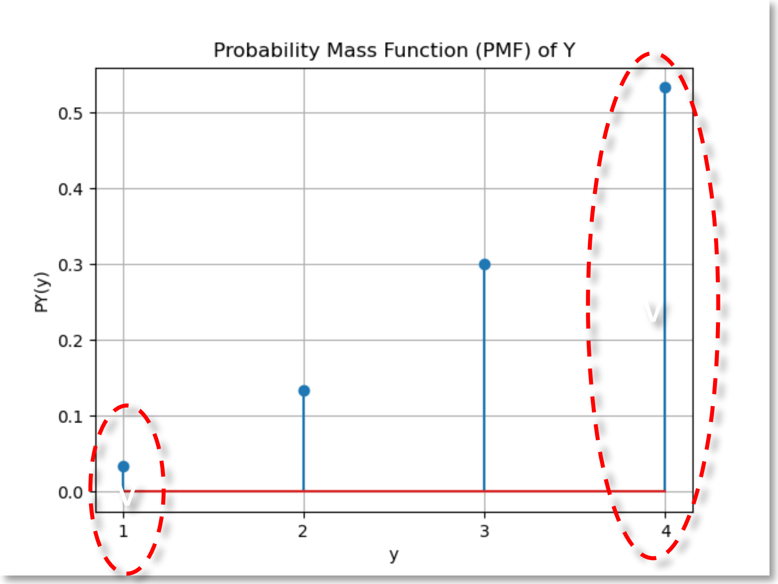
Y only has non-zero probabilities over it support $S = \{1, 2, 3, 4\}$. And $S \cap \{v^2 | v = 1, 2, 3, \ldots\} = \{1, 4\}$. This implies

$$P(Y \in \{v^2 | v = 1, 2, 3, \ldots\})$$

$$= P(Y = 1) + P(Y = 4)$$

$$= 1/30 + 4^2/30 = 17/30.$$

Note that Y only has nonzero probability values over its support. The set $\{v^2 | v = 1, 2, 3, \ldots\}$ defines a new set, which may not be the same as S. Also, it is the set of interested values y of Y, not the values of the probability $P(Y = y)$.



Q1(c) Find the probability that Y is even

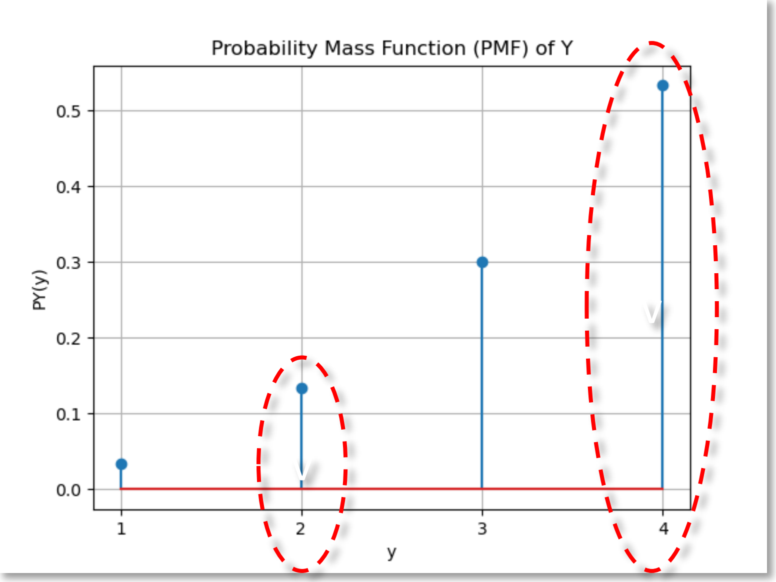
To find the probability that Y is even, we consider the even values of Y from the set $\{1,2,3,4\}$. These are $Y = 2$ and $Y = 4$.

Since $S \cap \{\text{even number}\} = \{2,4\}$, we have

$$P(Y \text{ is even}) = P(Y = 2) + P(Y = 4)$$

$$= \frac{2^2}{30} + \frac{4^2}{30}$$

$$= 2/3.$$



Q1(d) Find $P(Y \geq 2)$.

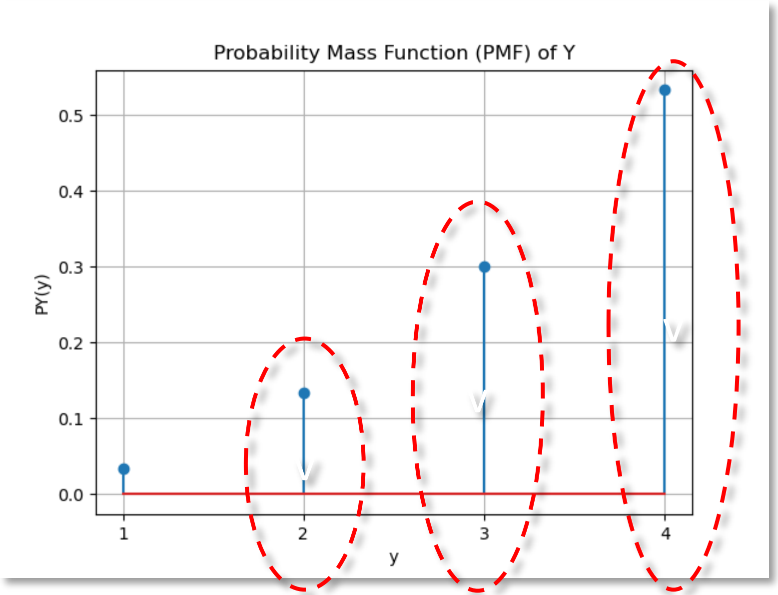
We want to find the probability that Y is greater than or equal to 2.

Let $A = S \cap \{Y \geq 2\} = \{2, 3, 4\}$, then we have

$$P(Y \geq 2) = \sum_{y \in A} P(Y = y)$$

$$= P(Y = 2) + P(Y = 3) + P(Y = 4)$$

$$= \frac{2^2}{30} + \frac{3^2}{30} + \frac{4^2}{30} = 29/30.$$



Q2

PMF & Conditional Probability

A smartphone attempts to send a message over a Wi-Fi link to an access point. Depending on the size of the message, it is transmitted as N packets where N has the PMF:

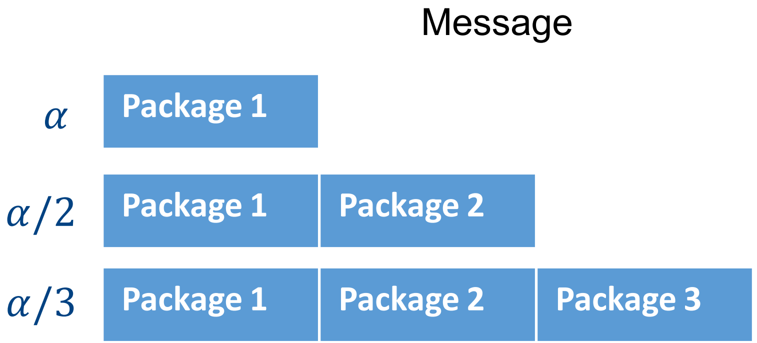
$$p_N(n) = \begin{cases} \alpha/n, & n = 1,2,3 \\ 0, & \text{otherwise} \end{cases}$$

(a) What is the value of the constant α?

(b) What is the probability that the packet size N is odd?

(c) Each packet is received correctly with probability p, and the message is received correctly if all packets are received correctly. What is the probability that the message is received correctly?

Q2 (a) What is the value of the constant α?

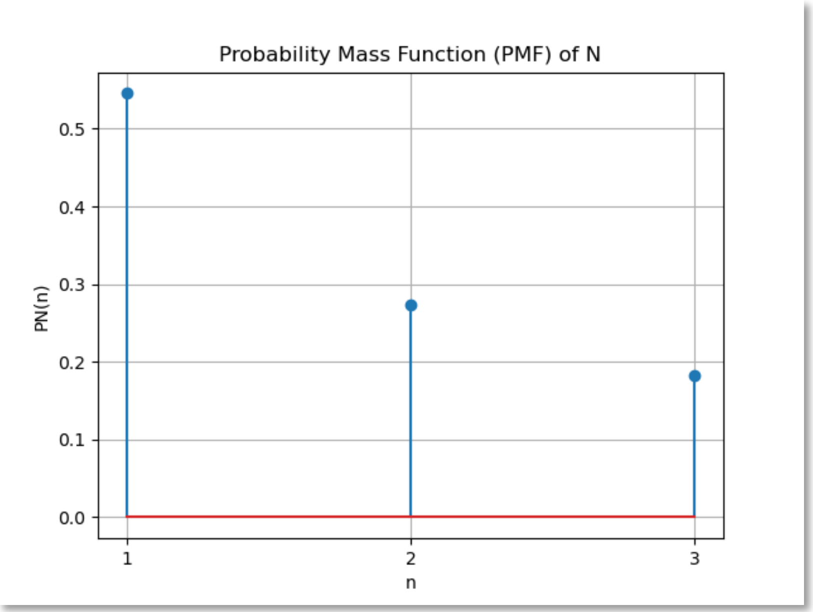


To find α, we use the fact that the sum of the probabilities over all possible values of N must equal 1. That is:

$$\sum_{y=1}^{3} p_N(n) = \alpha\left(1 + \frac{1}{2} + \frac{1}{3}\right) = 1$$

This implies

$$\alpha = \frac{6}{11}$$



Q2 (b) What is the probability that the packet size N is odd?

To find the probability that N is odd, we sum the probabilities for the odd values of N. From the given PMF, the possible odd values of N are $N = 1$ and $N = 3$.

$$P(N \text{ is odd}) = P(N = 1) + P(N = 3)$$

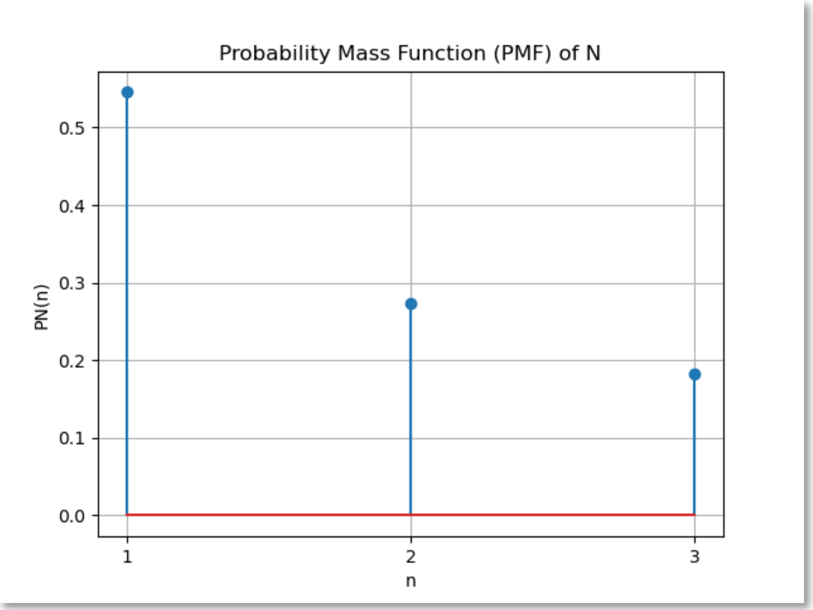
Substitute the values from the PMF:

$$P(N = 1) = \frac{\alpha}{1} = \frac{6}{11}$$

$$P(N = 3) = \frac{\alpha}{3} = \frac{6}{11} \cdot \frac{1}{3} = \frac{6}{33}$$

So the total probability is:

$$P(N \text{ is odd}) = \frac{6}{11} + \frac{6}{33} = \frac{8}{11}$$



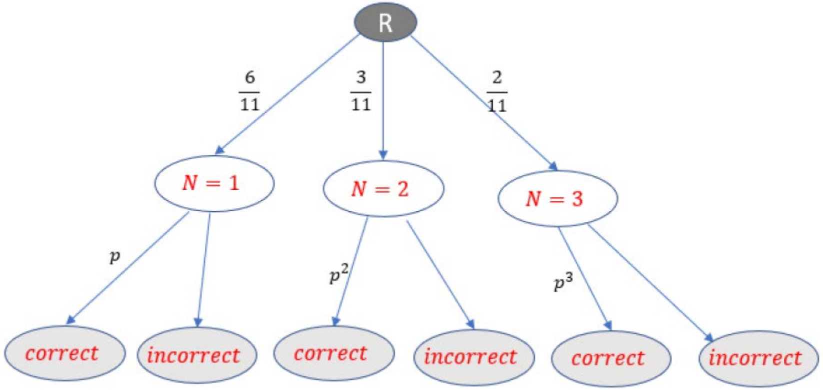
Q2 (c) Each packet is received correctly with probability p, and the message is received correctly if all packets are received correctly. What is the probability that the message is received correctly?

If a message includes i packages, it will be received correctly with probability

$$P(\text{msg received correctly} \mid i \text{ pkts in msg}) = p^i.$$

The number of packages is not fixed and follows the distribution of the $RV\ N$. The support of N is $S = \{1, 2, 3\}$ and i can be any value in S with probability $p_N(i)$. So the desired event probability is

$$\begin{aligned}
P(\text{msg received correctly}) &= \sum_{i \in S} P(i \text{ pkts in the msg and all received correctly}) \\
&= \sum_{i \in S} P(\text{msg received correctly} \cap i \text{ pkts in msg}) \\
&= \sum_{i \in S} P(\text{msg received correctly} | i \text{ pkts in msg}) \times P(i \text{ pkts in msg}) \\
&= \sum_{i \in S} p^i \times p_N(i) \\
&= p \times 6/11 + p^2 \times 3/11 + p^3 \times 2/11 = p(6 + 3p + 2p^2)/11.
\end{aligned}$$



The tree diagram

Q2 (c)

- We also can view receiving a message correctly as a sequential experiment.
- Specifically, the number of packages N is selected according to its distribution first.
- Then these packets are transmitted and received (correctly or incorrectly).
- A tree diagram in Figure 1 is used to illustrate this sequential experiment. Let (C, k) denote the event that a message has k packets and they are all received correctly. From the tree diagram above, the probability that a message is received correctly is

$$\sum_{k=1}^{3} P(C,k) = P(C,1) + P(C,2) + P(C,3) = \frac{6p}{11} + \frac{3p^2}{11} + \frac{2p^3}{11} = \frac{p(6+3p+2p^2)}{11}.$$

Q3

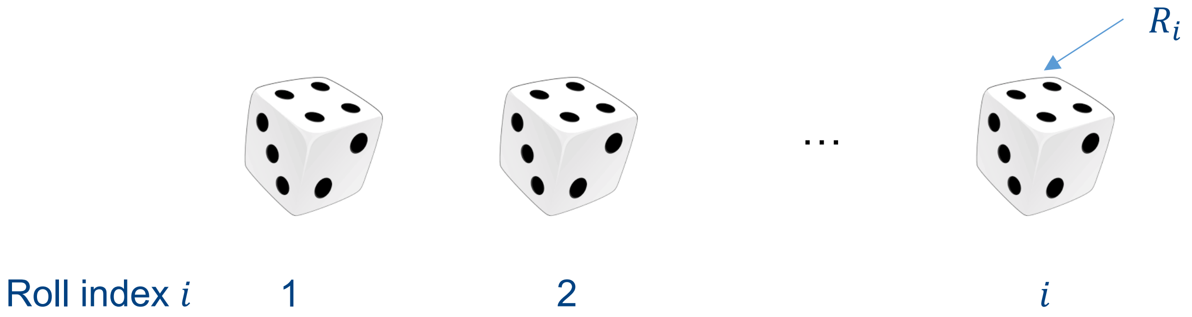
Sequential experiment

You roll a fair 6-sided dice repeatedly. Starting with roll $i = 1$, let R_i be the result of roll i. If $R_i > i$, then you roll again. Otherwise, you stop. Let N denote the number of rolls. What is $P(N > 3)$?

Q3

Problem Recap:

- You roll a fair 6-sided die repeatedly. The result of each roll is denoted by R_i for roll i. After each roll i, you stop rolling if $R_i \leq i$. If $R_i > i$, you roll again.
- We are asked to find $P(N > 3)$, which is the probability that the number of rolls N exceeds 3 (i.e., you roll more than 3 times).



Q3

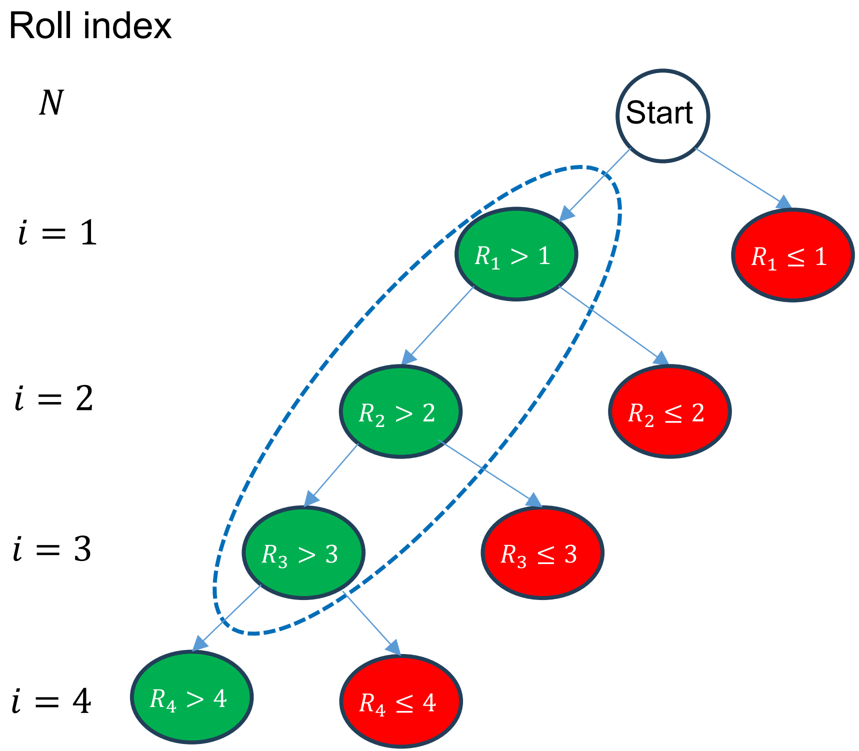
Step 1: Understand when you stop

The stopping condition is based on the comparison between R_i and i:

- You stop rolling if $R_i \leq i$.
- You continue rolling if $R_i > i$.

The possible values for R_i (since you are using a fair 6-sided die) are 1, 2, 3, 4, 5, and 6. For each roll, you compare the die result R_i to the roll index i.

That is there are more than N rolls only when $R_i > i$ for $1 \leq i \leq N$. So the event $\{N > 3\}$ is the same as $\{R_1 > 1, R_2 > 2, R_3 > 3\}$. Otherwise, it is impossible for $N > 3$.



Q3

Step 2: Analyze stopping at each roll

Roll 1 ($i = 1$):

- $R_1 \leq 1$ means you stop, i.e., $R_1 = 1$ (since the only possibility where $R_1 \leq 1$ is when $R_1 = 1$).
- $R_1 > 1$ means you continue to roll, which happens if R_1 is 2, 3, 4, 5, or 6.

The probability of continuing to the second roll is:

$$P(\text{Continue to Roll 2}) = P(R_1 > 1) = \frac{5}{6}$$

Roll 2 ($i = 2$):

- $R_2 \leq 2$ means you stop, i.e., $R_2 = 1$ or $R_2 = 2$.
- $R_2 > 2$ means you continue to roll, which happens if R_2 is 3, 4, 5, or 6.

The probability of continuing to the third roll (given that you reached the second roll) is:

$$P(\text{Continue to Roll 3} \mid \text{Continue to Roll 2}) = P(R_2 > 2) = \frac{4}{6}$$

Roll 3 ($i = 3$):

- $R_3 \leq 3$ means you stop, i.e., $R_3 = 1,\ 2,\ 3$.
- $R_3 > 3$ means you continue to roll, which happens if R_3 is 4, 5, or 6.

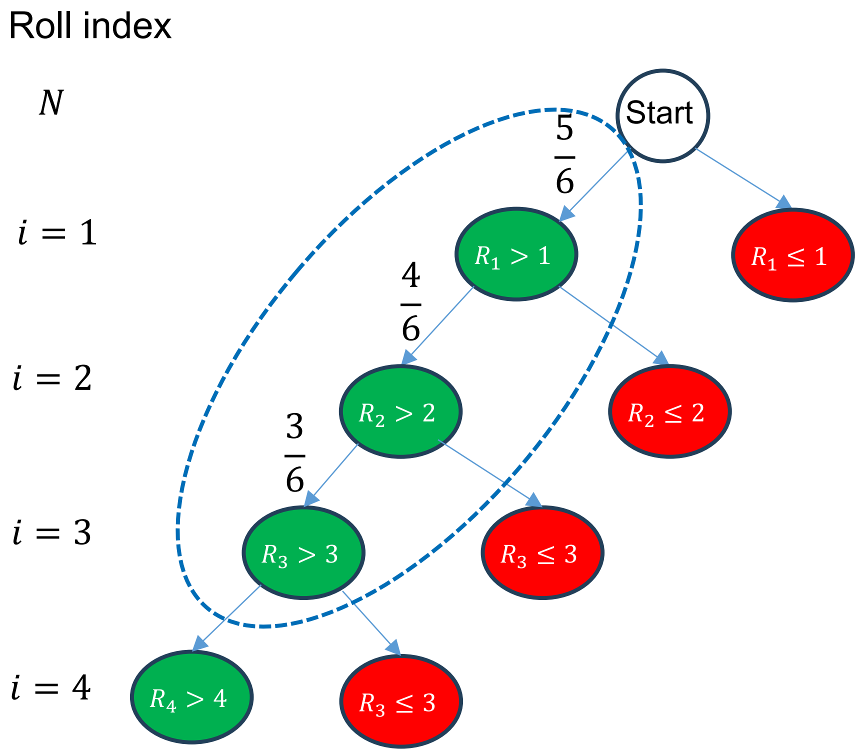
The probability of continuing to the fourth roll (given that you reached the third roll) is:

$$P(\text{Continue to Roll 4} \mid \text{Continue to Roll 3}) = P(R_3 > 3) = \frac{3}{6}$$

Q3

Step 3: Calculate $P(N > 3)$

- $P(N > 3)$ is the probability that you roll more than 3 times. This occurs if you do not stop after the first three rolls.
- The probability of this happening is the product of the probabilities of continuing past each of the first three rolls:



$$P(N > 3) = P(R_1 > 1, R_2 > 2, R_3 > 3) = P(R_1 > 1)P(R_2 > 2)P(R_3 > 3) = \frac{5}{6} \times \frac{4}{6} \times \frac{3}{6} = \frac{5}{18}$$

Q4

The number of buses that arrive at a bus stop in T minutes is a Poisson random variable B with the expected value $T/5$.

(a) What is the PMF of B, the number of buses that arrive in T minutes?

(b) What is the probability that in a two-minute interval, three buses arrive?

(c) What is the probability of no buses arriving in a 10-minute interval?

(d) How much time should you allow so that with probability 0.99 at least one bus arrives?

Q4

Problem Recap

The number of buses arriving at a bus stop follows a Poisson distribution with the expected value $\lambda = \frac{T}{5}$, where T is the time in minutes.

The Poisson PMF (Probability Mass Function) is given by:

$$P(B = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

where k is the number of events (buses arriving), and λ is the expected number of events in a given interval.

Q4 (a) What is the PMF of B, the number of buses that arrive in T minutes?

The expected value of B is $E[B] = T/5$. Since $\lambda = E[B]$ for Poisson RV, the PMF for B is

$$p_B(k) = \begin{cases} \dfrac{\left(\frac{T}{5}\right)^k}{k!} e^{-\frac{T}{5}}, & k = 0,1,2,\ldots \\ 0, & \text{otherwise} \end{cases}$$

Note that both the expectation and variance for this Poisson RV are λ.

Q4 (b) What is the probability that in a two-minute interval, three buses arrive?

For a two-minute interval, $T = 2$, so $\lambda = \frac{2}{5}$.

We are asked to find $p_B(3)$ for $k = 3$ and $\lambda = \frac{2}{5}$.

Using the PMF formula:

$$p_B(3) = \frac{\left(\frac{2}{5}\right)^3}{3!} e^{-2/5} \approx 0.0072$$

Q4 (c) What is the probability of no buses arriving in a 10-minute interval?

The observation time interval is 10 minutes.

Since $T = 10$, We have $\lambda = T/5 = 2$.

The probability that zero bus arrives in a 10-minute interval is

$$p_B(0) = \frac{2^0}{0!} e^{-2} \approx 0.1353$$

Q4 (d) How much time should you allow so that with probability 0.99 at least one bus arrives?

- At least one bus means it can be $1,2,3,\ldots$ buses. We can consider the complement event "no bus in this interval".
- Denote this interval length by T. Then $\lambda = T/5$ and the probability of this complement event is

$$P(B=0) = \frac{\left(\frac{T}{5}\right)^0}{0!} \times e^{-\frac{T}{5}} = \exp\left(-\frac{T}{5}\right)$$

- The event probability $P(B \geq 1) = 1 - P(B=0) = 1 - e^{\frac{-T}{5}}$ should be 0.99.
- Solve T from $1 - e^{\frac{-T}{5}} = 0.99$. We have

$$T = 5\ln(100) \approx 23 \text{ minutes.}$$

Q5

A random variable U has PMF

$$p_U(u) = \begin{cases} \frac{1}{5}, & u = 1,2,3,4,5 \\ 0 & \text{otherwise} \end{cases}$$

(a) What is $P(U = E[U])$?

(b) What is $P(U > E[U])$?

Q5 (a) What is $P(U = E[U])$?

- **Step 1**, we need to calculate $E[U]$, the expected value of U. The expected value for a discrete random variable is given by:

$$E[U] = \sum u \cdot p_U(u)$$

- **Step 2**: Using the given PMF, we compute $E[U]$:

$$E[U] = 1 \cdot \frac{1}{5} + 2 \cdot \frac{1}{5} + 3 \cdot \frac{1}{5} + 4 \cdot \frac{1}{5} + 5 \cdot \frac{1}{5} = 3$$

- **Step 3**: we are asked to find $P(U = E[U])$, which is:

$$P(U = 3)$$

From the PMF, we know that:

$$P(U = 3) = \frac{1}{5}$$

Q5 (b) What is $P(U > E[U])$?

We have $E[U] = 3$, so we need to find the probability that U is greater than 3:

$$P(U > 3) = P(U = 4) + P(U = 5)$$

From the PMF, we know:

$$P(U = 4) = \frac{1}{5}, \qquad P(U = 5) = \frac{1}{5}$$

Thus:

$$P(U > 3) = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

Q6

Let X be a binomial distributed random variable with parameters n and p.

Find $E[(1-p)^X]$.

Q6

We are given that X is a binomially distributed random variable with parameters n and p. This means that $X \sim \textbf{Binomial}(n, p)$, where the probability mass function (PMF) is given by:

$$p_X(i) = \binom{n}{i} p^i (1-p)^{n-i} \text{ for } 0 \le i \le n\,.$$

Step 1: Recall the formula for expected value

The expected value of a function $g(X)$ for a discrete random variable X is given by:

$$E[g(X)] = \sum_{i=0}^{n} g(X=i)\,P(X=i) = \sum_{i=0}^{n} g(i)\,p_X(i)$$

Recall page 21 Chapter 4

In our case, $g(X) = (1-p)^X$, so we need to compute:

$$E[(1-p)^X] = \sum_{i=0}^{n} (1-p)^i p_X(i) = \sum_{i=0}^{n} (1-p)^i \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^n$$

PMF

Q6

Step 2: Simplify the expression

Notice that $(1-p)^i$ is already part of the PMF, and it simplifies with the $(1-p)^{n-i}$ term from the PMF. So the expression simplifies to:

$$E[(1-p)^X] = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^n$$

We can factor out $(1-p)^n$, which does not depend on k:

$$E[(1-p)^X] = (1-p)^n \sum_{i=0}^{n} \binom{n}{i} p^i$$

Binomial expansion

$$(\underset{1}{\uparrow a} + \underset{p}{\uparrow b})^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i$$

Step 3: Recognize the binomial expansion

The remaining summation is a binomial expansion of the form:

$$\sum_{i=0}^{n} \binom{n}{i} p^i = (1+p)^n$$

Thus, the expression becomes:

$$E[(1-p)^X] = (1-p)^n \sum_{k=0}^{n} \binom{n}{i} p^i = (1-p)^n (1+p)^n = (1-p^2)^n$$

Q7

Suppose X has the binomial PMF

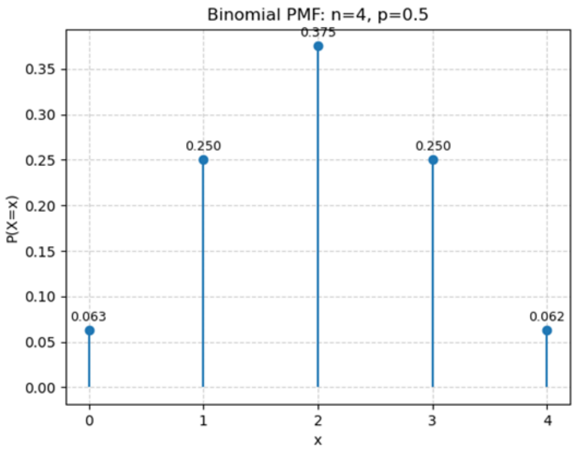
$$p_X(x) = \binom{4}{x}\left(\frac{1}{2}\right)^4$$

where $\binom{n}{x}$ is the binomial coefficient

(a) What is the standard deviation of X.

(b) What is $P(m_X - \sigma_X \leq X \leq m_X + \sigma_X)$? This is the probability that X is within one standard deviation away from the expected value m_X.

Q7(a) What is the standard deviation of X



Method I

From the expression of $p_X(x)$, we can see the support for X is $S = \{0, 1, 2, 3, 4\}$. To find the standard deviation, the square root of variance, we make use of the relation

$$\mathrm{Var}[X] = E[X^2] - (E[X])^2$$

Step 1: Compute $E[X]$

$$m_X = \mathrm{E}[X] = \sum_{x=0}^{4} x p_X(x) = 1 \times \binom{4}{1}\frac{1}{2^4} + 2 \times \binom{4}{2}\frac{1}{2^4} + 3 \times \binom{4}{3}\frac{1}{2^4} + 4 \times \binom{4}{4}\frac{1}{2^4} = 2.$$

Step 2: Compute the expected value of X^2

$$\mathrm{E}[X^2] = \sum_{x=0}^{4} x^2 p_X(x) = 1 \times \binom{4}{1}\frac{1}{2^4} + 4 \times \binom{4}{2}\frac{1}{2^4} + 9 \times \binom{4}{3}\frac{1}{2^4} + 16 \times \binom{4}{4}\frac{1}{2^4} = 5.$$

Step 3: Compute variance and standard deviation

The variance of X is thus $\mathrm{Var}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2 = 5 - 2^2 = 1$ and the standard deviation is $\sigma_X = \sqrt{\mathrm{Var}[X]} = 1$.

Q7(a) What is the standard deviation of X

Method II

From the expression of $p_X(x)$, we have

$$p_X(x) = \binom{4}{x}\left(\frac{1}{2}\right)^4 = \binom{4}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{4-x}, \quad \text{for } x \in S.$$

It is clear that $X \sim$ Binomial $(n = 4, p = 1/2)$. Its variance $\mathrm{Var}[X] = np(1-p) = 1$ as discussed in lecture. So the standard deviation is $\sigma_X = \sqrt{\mathrm{Var}[X]} = 1.$

A binomial random variable X:

- Expectation $m_X = np$
- Variance $\sigma_X^2 = np(1-p)$

Q7(b) What is $P(m_X - \sigma_X \le X \le m_X + \sigma_X)$?

The probability that X is within one standard deviation away of its expected value is

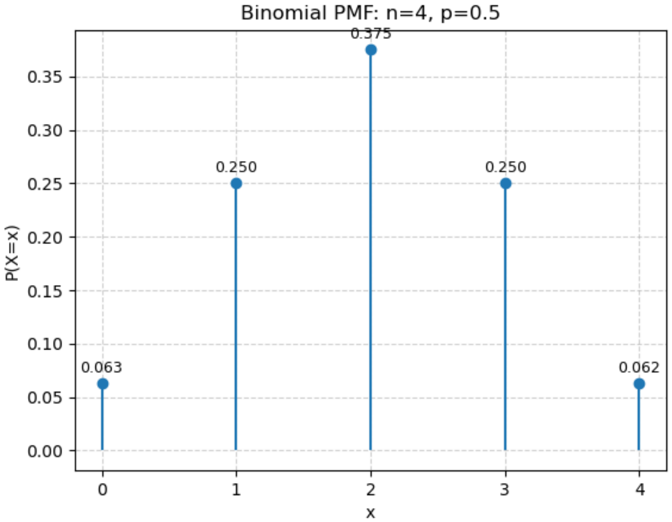
$$P(2-1 \le X \le 2+1) = P(1 \le X \le 3)$$

This can be easily computed by the PMF of X. We have

$$\{X \mid 1 \le X \le 3\} \cap S = \{1,2,3\}$$

And

$$P(1 \le X \le 3) = P(X=1) + P(X=2) + P(X=3) = 7/8.$$



THANK YOU