4. Discrete Random Variables and Probability Distributions:

(Reading Exercises: Montgomery and Runger Section 2.9, 3.1-3.2, 3.4-3.8 & Class notes; Yates and Goodman Chapter 3)

Learning outcomes:

You will be able to

- Define a discrete random variable.
- Define a discrete probability mass function and a discrete cumulative distribution function.
- Describe the important discrete random variables and use probability mass function to find probabilities for practical situations.
- Apply the material learned in this section to new problems.

Definition: Given a random experiment with sample space S, a random variable is a function that associates a unique real number to each outcome (or element) that belongs in the sample space S.

- A random variable is denoted by upper case letters, for example, X
- Lower case letters are used to denote the numerical value that a random variable takes, for example, $X = x$.

Definition: Let S be a sample space. A discrete random variable is one that takes on either

- A finite number of values $x_1, x_2 \dots, x_n$ or
- An infinite number of values $x_1, x_2 \dots$ from S.

Example: A coin is tossed 3 times and the sequence of Heads and Tails is noted. The sample space comprises outcomes, which are all possible combinations of heads and tails,

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

Denote an outcome by the symbol ξ. Let X denote the number of heads in the three tosses. The random variable, X, maps each outcome onto a real number. We see that X can take on four values as illustrated in the table below,

$$X: \xi \text{ } where \text{ } x\epsilon\{0, 1, 2, 3\}$$

ξ	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
$X(\xi)$	3	2	2	2	1	1	1	0

4.1 Probability Mass Function (PMF)

4.1.1 Characterization of Discrete Random Variables:

Random variables are characterized by their probability distributions.

Definition: Consider a discrete random variable X. The probability that X takes on a particular value, $X = x$, is denoted $p_X(x)$, that is,

$$p_X(x) = P[X = x]$$

Suppose that X can take on any value from a set $x \in \{x_1, x_2, \cdots\}$. The probability mass function (PMF) is the collection of the probabilities of all the values X can take, that is,

$$p_X(x) = P[X = x], \forall\, x \epsilon \{x_1, x_2, \dots\}$$

Properties of PMF: Since $p_X(x)$ is probability, it must satisfy the following conditions (axioms):

(i) $p_X(x) \geq 0, \forall\, x \in \{x_1, x_2, \cdots\}$

(ii) $\sum_x p_X(x) = 1$

(iii) $p_X(x) = 0$ if $x \notin \{x_1, x_2, \cdots\}$.

Example: A shipment of 20 RAM chips from the same company to a lab contains 3 that are defective. If 2 of these chips are randomly picked, find the PMF for the number of defective RAM chips.

Solution: Define random variables x and y, and events
A_1: {x defective chips are selected};A_2: {$y = 2 - x$ non-defective chips are selected}; $x = 0, 1, 2$.

For event A_1 (or Partition 1), the number of possible ways x chips can be selected from 3 defective chips is the combination $n_1 = \binom{3}{x}, x = 0, 1, 2,$
For event A_2 (or Partition 2), the number of possible ways y chips can be selected from 17 non-defective chips is the combination $n_2 = \binom{17}{y} = \binom{17}{2-x}$, $x = 0, 1, 2$.
The experiment involves picking 2 chips out of 20. The number of elements of the sample space is, therefore, the combination $n_S = \binom{20}{2}$.
The PMF of the random variable X, is

$$p(x) = P[X = x] = \frac{n_1 \times n_2}{n_S} = \frac{\binom{3}{x}\binom{17}{2-x}}{\binom{20}{2}}$$

$$p(0) = P[X = 0] = \frac{\binom{3}{0}\binom{17}{2}}{\binom{20}{2}} = \frac{68}{95} = \frac{136}{190}$$

$$p(1) = P[X = 1] = \frac{\binom{3}{1}\binom{17}{1}}{\binom{20}{2}} = \frac{51}{190}$$

$$p(2) = P[X = 2] = \frac{\binom{3}{2}\binom{17}{0}}{\binom{20}{2}} = \frac{3}{190}$$

Check:

$$p(0) + p(1) + p(2) = 1$$

The PMF (probability distribution) can be represented in tabular form as

x	0	1	2
$p(x)$	$\frac{136}{190}$	$\frac{51}{190}$	$\frac{3}{190}$

4.2 Cumulative Distribution Function (CDF) for Discrete Random Variable:

Definition: The Cumulative Distribution Function (CDF), denoted $F_X(x)$, of a discrete random variable, X, with probability mass function $p_X(x)$, is a function defined as the probability that X does not exceed a value x, that is,

$$F_X(x) = P[X \leq x] = \sum_{y=-\infty}^{x} p_X(y), \quad -\infty < x < \infty$$

Properties of the CDF:

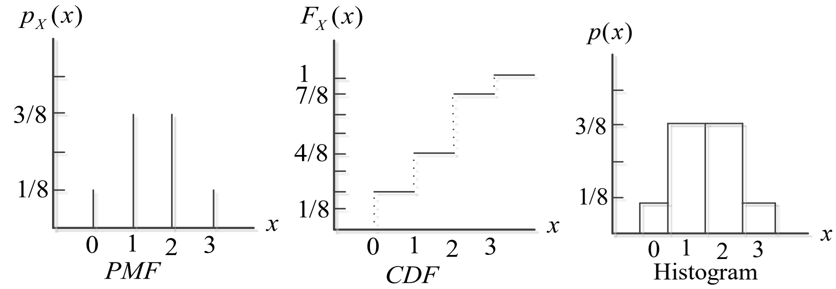
(1) $F(-\infty) = 0 \Rightarrow$ impossible and $F(\infty) = 1 \Rightarrow$ certain,

(2) $F(x)$ is monotonically non-decreasing in x.

Example: The table below represents the PMF and the CDF of random variable X, representing the number of heads in 3 tosses of a fair coin. There are 8 events in total.

Random variable	$x = 0$	$x = 1$	$x = 2$	$x = 3$
PMF	$p_X(0) = 1/8$	$p_X(1) = 3/8$	$p_X(2) = 3/8$	$p_X(3) = 1/8$
CDF	$F_X(0) = 1/8$	$F_X(1) = 4/8$	$F_X(2) = 7/8$	$F_X(3) = 8/8$

The PMF and CDF can be represented graphically as illustrated in the figures below.



Graphical methods of representing probability distributions

Example: Find the CDFs, $F_X(0)$and $P[-3 \le X \le 1]$, for the PMF of random variable X in the table below.

$X = x$	-8	-3	-1	0	1	2	6
$p_X(x)$	0.13	0.15	0.17	0.20	0.15	0.11	0.09

$$F_X(0) = P[X \le 0] = 0.13 + 0.15 + 0.17 + 0.20 = 0.65$$
$$P[-3 \le X \le 1] = 0.15 + 0.17 + 0.20 + 0.15 = 0.67$$

Example: The table below provides the PMF and CDF of 2 dice rolls when the random variable X represents the sum of the two numbers that face up.

$$p_X(x) = \frac{\text{\# of ways 2 dices can sum to } x}{\text{\# of all possible ways 2 dices can result in}}$$

$$F_X(x) = \sum_{i=2}^{x} p_X(i)$$

Probability Mass Function & CDF of two dice roll

x	$p_X(x)$	$F_X[x]$
2	1/36	1/36
3	2/36	3/36
4	3/36	6/36
5	4/36	10/36
6	5/36	15/36
7	6/36	21/36
8	5/36	26/36
9	4/36	30/36
10	3/36	33/36
11	2/36	35/36
12	1/36	36/36

4.3 Important Discrete Random Variables and their Distribution Functions:

Counting Processes:

There are three counting processes – (1) Binomial random variables, (2) Geometric random variable and (3) Poisson random variables. Fundamental to counting processes is the Bernoulli trial and Bernoulli random variable.

Bernoulli trial: (Daniel Bernoulli – 1700-1782, Swiss Mathematician and Physicist)

> **Definition:** A Bernoulli trial is a single random experiment in which there can be only two outcomes (success or failure). Probability of success is $0 < p < 1$ and of failure is $q = 1 - p$.

Examples of Bernoulli Trials: (1) Tossing of a fair coin (outcomes are H or T) and (2) classification of products as defective or non-defective (outcomes are D or N).

Bernoulli random variable:

> **Definition:** A Bernoulli random variable, X, is a random variable whose only possible values are 1 and 0, and whose probability mass density function is
>
> $$p_X(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1$$
>
> $$p_X(0) = 1 - p$$
>
> $$p_X(1) = p$$
>
> We denote this distribution by $Ber(p)$.

Binomial Random Variable:

Consider the following:

1. An experiment is performed, in the same way, n times,

2. Each experiment is a Bernoulli trial,

3. The trials generate n independent Bernoulli random variables,

4. The probability of success, denoted p, is the same for each trial. The probability of failure is $q = 1 - p$.

Definition: A Binomial random variable, X, represents the number of successes (or failures) in n independent Bernoulli trials with p the probability of success and q the probability of failure at each trial. Let X_i denote the outcome at the i^{th} trial, then the number of successes is the summation,

$$X = X_1 + X_2 + \cdots + X_n = \sum_{i=1}^{n} X_i$$

$$X_i = 1 \rightarrow \text{success in the } \mathrm{i}^{th} \text{ experiment}$$
$$= 0 \rightarrow \text{success in the } \mathrm{i}^{th} \text{ experiment}$$

We note the following:

- The number of ways of arranging x successes in n trials is the combination $\binom{n}{x}$
- All arrangements of x successes and $(n-x)$ failures are equally likely with probability $p^x(1-p)^{n-x}$
- The PMF is, therefore, given by

$$p_X(x) = P[X = x] = \binom{n}{x} p^x(1-p)^{n-x}, x = 0, 1$$

- The above PMF is known as a Binomial distribution and is denoted $Bin(n, p)$ with parameters n and p, where $n = 1, 2, \ldots$ with $0 \leq p \leq 1$.
- The CDF of a Binomial random variable is given by,

$$F_X(x) = P[X \leq x] = \sum_{y=0}^{x} \binom{n}{y} p^y (1-p)^{n-y} \quad x = 0,1,\ldots,n$$

Example: A manufacturing process results in probability $p = 0.05$ of generating defective products. If we select 3 products from the manufacturing process and inspect them, what are the probabilities for the different numbers of defective products?

Solution: Each product selection is a Bernoulli trial because it has two outcomes: defective (D: failure) and non-defective (N: success). The number of failures (or successes) is a Bernoulli process because,

(1) There are 3 repeated trials,
(2) Each trial is a Bernoulli trial
(3) The probability of defective product $p = 0.05$ for each trial.
(4) The 3 trials are independent. The fact that the first product is D or N has no bearing on the second.

The following table summarizes all the possible probabilities:

Outcome	x	$P[X]$
NNN	0	$(1-0.05)^3$
NND	1	$(1-0.05)^2 \times 0.05$
NDN	1	$(1-0.05)^2 \times 0.05$
DNN	1	$(1-0.05)^2 \times 0.05$
DND	2	$(1-0.05) \times 0.05^2$
NDD	2	$(1-0.05) \times 0.05^2$
DDN	2	$(1-0.05) \times 0.05^2$
DDD	3	0.05^3

We note in the table that there are 1 combination of 3 successes, 3 combinations of 2 successes, and 2 combinations of 1 success.

Now, applying the formula $p_X(x) = P[X = x] = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1$ we obtain

$$p_X(0) = \binom{3}{0} p^0 (1-p)^{3-0} = (1-0.05)^3$$
$$p_X(1) = \binom{3}{1} p^1 (1-p)^{3-1} = 3(1-0.05)^2 \times 0.05$$
$$p_X(2) = \binom{3}{2} p^2 (1-p)^{3-2} = 3(1-0.05) \times 0.05^2$$
$$p_X(3) = \binom{3}{3} p^3 (1-p)^{3-3} = 0.05^3$$

These are the same results obtained in the table above.

Geometric Random Variable:

Definition: A geometric random variable X is the number of trials of any experiment (comprising n independent Bernoulli trials) until the first success. Therefore, the PMF of a geometric random variable X is given by,

$$p_X(x) = P(X = x) = (1-p)^{x-1} p, \quad x = 0,1,\ldots,n$$

This PMF, denoted by $Geo(p)$, is known as a geometric distribution with parameter $p, where\ 0 < p \leq 1$.

Consider a sequential experiment in which we repeat independent Bernoulli trials until the occurrence of the 1st success. Let random variable X denote the number of trials until the 1st success. Let $p \triangleq P['1^{st}\text{success after } x \text{ trials}]$. Then, the PMF is

$$\begin{aligned} p_X(x) &\triangleq P[1^{st}\text{success after } x \text{ trials}] \\ &= P[\text{failure in } (x-1) \text{ previous trials and success at the } x^{th} \text{ trial}] \\ &= P[\text{fail at } 1^{st}]P\left[\text{fail at } 2^{nd}\right] \cdot \cdots P[\text{fail at } (x-1)^{th}]P[\text{fail at } x^{th}] \\ &= (1-p)(1-p)\cdot\cdots(1-p)p \\ &= (1-p)^{x-1}p \equiv \text{Geometric Probability Law} \end{aligned}$$

Example: Let X denote the number of times a paging message needs to be transmitted until success.

(a) Find the PMF of X.
(b) Find the probability that X is an even number.

Solution: (a) The set is $S = \{1,2,3,\cdots\}$. Let $F \equiv$ "Failure" and $S \equiv$ "Success". The PMF is

$$p_X(m) = P[X = m] = P[F \cap F \cap \cdots \cap F \cap S] = P[F]P[F]\cdots P[F]P[S] = (1-p)^{m-1}p.$$

Solution: (b)

$$\begin{aligned} P[X \text{ is even}] = P[X=2] + P[X=4] + \cdots &= \sum_{m=1}^{\infty} p(2m) = p\sum_{m=1}^{\infty}(1-p)^{2m-1} \\ &= p\sum_{m=1}^{\infty}(1-p)^{2m-2+1} = p(1-p)\sum_{m=1}^{\infty}((1-p)^2)^{m-1} \\ &= p\frac{1-p}{1-(1-p)^2} = \frac{1-p}{2-p} \end{aligned}$$

Memoryless Property:

Theorem: A geometric distribution has the following memoryless property for all non-negative integers m and n,

$$P(X \geq m+n | X \geq m) = P[X \geq n]$$

Proof:

$$P[X \geq m+n | X \geq m] = \frac{P[\{X \geq m+n\} \cap \{X \geq m\}]}{P[X \geq m]}$$

$$= \frac{P[X \geq m+n]}{P[X \geq m]} = \frac{(1-p)^{m+n}}{(1-p)^m}$$

$$= (1-p)^n$$

$$= P[X \geq n]$$

Interpretation:
Suppose you are told that there have been m failures initially. Then the chance of at least n more failures before the first success; is the same as if you started the experiment for the first time and the information of initial m failures is irrelevant.

Example: What is the probability that more than K trials are required before 1st success?

Solution:

$$P[\{i > K\}] = P[\{K+1 \text{ or } K+2 \text{ or } \cdots\}]$$
$$= P[K+1] + P[K+2] + \cdots = \sum_{m=1}^{\infty} P[K+m]$$
$$= \sum_{m=1}^{\infty} (1-p)^{K+m-1} p = p(1-p)^K \sum_{m=1}^{\infty} (1-p)^{m-1}$$
$$= \frac{p(1-p)^K}{1-(1-p)} = (1-p)^K$$

Poisson Random Variables
Consider events occurring independently of each other, with an average number of events λ, in some fixed interval Δt.

Definition: A random variable$X$, is said to be a Poisson RV if it counts the number of events that occur, "completely at random", with a mean value λ in a specified time Δt interval.

Examples: A Poisson random variable can be used to model the following events:

- The number of calls arriving at a phone exchange centre,
- The number of data packets arriving at a router/server,
- The number of multipath arriving at a wireless radio receiver
- The number of photons arriving at a CCD pixel in some exposure time (astronomy observation)
- Number of customers arriving at a checkout counter
- The number of accidents occurring at an intersection.

PMF of Poisson RV

- Let X represent the number of events,
- α the average number of arrivals per unit time (rate of arrival)
- Δt the time interval of interest
- $\lambda = \alpha \times \Delta t$ is the mean or average number of events in an interval of interest.

Definition: A random variable X has a Poisson distribution, with parameter $\lambda > 0$, if its PMF is given by

$$P[X = x] = \frac{\lambda^x}{x!} e^{-\lambda}, x = 0, 1, 2, \ldots$$

Example: The number of queries arriving in a Δt seconds interval, at a call center is a Poisson RV with an average of 4 queries per minute.

(a) Find the probability of more than 4 queries in 10 seconds.

(b) Find the Probability of less than 5 queries in 2 minutes.

(c) Find the Probability of exactly 4 queries in 2 minutes.

Solution:

(a) $\Delta t = 10\text{s}; \quad \alpha = \frac{4 \text{ querries}}{\text{min}} = \frac{4}{60} \frac{\text{querries}}{s}; \quad \Rightarrow \lambda = \frac{4}{60} \times 10 = \frac{2}{3}$

$$P[X > 4] = 1 - P[X \leq 4]$$
$$= 1 - P[X = 0 \text{ or } X = 1 \text{ or } X = 2 \text{ or } X = 3 \text{ or } X = 4]$$
$$= 1 - (P[X = 0] + P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4])$$
$$= 1 - \left(\frac{\lambda^0}{0!} e^{-\lambda} + \frac{\lambda^1}{1!} e^{-\lambda} + \frac{\lambda^2}{2!} e^{-\lambda} + \frac{\lambda^3}{3!} e^{-\lambda} + \frac{\lambda^4}{4!} e^{-\lambda}\right)$$
$$= 1 - e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{6} + \frac{\lambda^4}{24}\right) = 6.33 \times 10^{-4}$$

(b) $\lambda = \alpha \times \Delta t = \left(\frac{4}{60}\right) \times (2 \times 60) = 8$

$$P[X < 5] = P[X \leq 4] = P[X = 0] + P[X = 1] + P[X = 2] + P[X = 3] + P[X = 4]$$
$$= e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{6} + \frac{\lambda^4}{24}\right) = 0.1$$

(c) $P[X = 4] = e^{-8} \frac{8^4}{24} = 0.06$

Example: Data packets arrive at a multiplexer at random and at an average rate of 1.2 per second.

(a) Find the probability of 5 messages arriving in a 2-seconds interval.

(b) For how long can the operation of the multiplexer be interrupted, if the probability of losing one or more packets is to be no more than 0.05?

Solution: Times of arrival form a Poisson process with rate $\alpha = 1.2/\sec$.

(a) Let X denote the number of messages arriving in a $\Delta t = 2\text{sec}$ interval. Then X is Poisson with mean number $\lambda = \alpha \times \Delta t = 1.2 \times 2 = 2.4$

$$P[X=5]=\frac{\lambda^5}{5!}e^{-\lambda}=\frac{2.4^5}{5!}e^{-2.4}=0.06$$

(b) Let the number of messages be denoted by random variable Y. The variable Y is Poisson with $\lambda=\alpha\times\Delta t=1.2\times\Delta t$. We need to find time interval Δt.

$$P[\{\text{at least one message}\}]=P[Y\geq 1]=1-P[Y=0]=1-\frac{2.4^0}{0!}e^{-1.2\times\Delta t}\leq 0.05$$
$$\Rightarrow e^{-1.2\times\Delta t}\geq 0.95\Rightarrow -1.2\times\Delta t\geq \ln(0.95)=-0.05129$$
$$\Rightarrow \Delta t=0.043\text{sec}$$

Discrete Uniform Distribution:

Definition: A discrete uniform random variable X, is one that is equally likely to take any integer value in a finite interval $x\in[k,l]$,

- X has a PMF defined as
$$p_X(x)=P[X=x]=\begin{cases}\frac{1}{l-k+1}, & x=k,k+1,k+2,\dots,l\\ 0, & \text{otherwise}\end{cases}$$
- A uniform random variable X, has a constant PMF over a finite range.

Useful Series in Probability:

$\sum_{k=1}^{n}q^{k-1}=\frac{1-q^n}{1-q},\quad (q>0)$	$\sum_{k=1}^{\infty}q^{k-1}=\frac{1}{1-q},\quad (0<q<1)$	$\sum_{k=1}^{\infty}kq^{k-1}=\frac{1}{(1-q)^2}, 0<q<1$
$\sum_{k=1}^{\infty}k^2q^{k-1}=\frac{1+q}{(1-q)^3},\quad (0<q<1)$	$\sum_{k=1}^{n}k=\frac{n(n+1)}{2}$	$\sum_{k=1}^{n}k^2=\frac{n(n+1)(2n+1)}{6}$
$\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{\lambda}$	$\sum_{k=0}^{n}\binom{n}{k}p^kq^{n-k}=(p+q)^n$ Binomial Theorem	$\sum_{x=k}^{l}\frac{x}{l-k+1}=\frac{l+k}{2}$

4.4 Mathematical expectation of discrete random variables:

Learning Outcomes:

You will be able to
- Define the expected value of a discrete random variable
- Define the expected value of a function of a discrete random variable.
- Apply the properties of mathematical expectations.
- Derive a formula for the mean of the special random variables.
- Define the variance and standard deviation of a discrete random variable.
- Apply a shortcut formula for the variance of a discrete random variable.
- Calculate the mean and variance of a linear function of a discrete random variable.
- Understand the steps involved in each of the proofs in the lesson.
- Apply the methods learned in this section to new problems.

Definition: Let X be a discrete random variable with a possible set of values $S = \{x_1, x_2, \dots\}$ and PMF $p_X(x)$. The expected value or mean or average value of X, denoted $E[X]$ or μ_X is the weighted average defined as

$$\mu_X = E[X] = \sum_{x \in S} x p_X(x)$$

- The sum in the definition above, is known as the mathematical expectation of X.
- The expected value of X is also referred to as the first moment of$X$.
- μ_X is a measure of location of the PMF

Example: What is the average toss of a fair, six-sided die?

Example: Let random variable X denote the number of credit cards owned by an electrical engineering student. Using the data in the table below, find the expected number of credit cards a student will possess.

x	0	1	2	3	4	5	6
$p_X(x)$	0.08	0.28	0.38	0.16	0.06	0.03	0.01

Example: Find the mean of a Binomial random variable X, with parameters n and p.

Solution:

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

Example: Find the mean of a geometric RV, X, with $p = P[\text{Success}]$.

Solution: Use identity $\sum_{i=1}^{\infty} iq^i = \frac{q}{(1-q)^2}$; $q = 1-p$,

$$E[X] = \sum_{x=1}^{\infty} x p_X(x) = \sum_{x=1}^{\infty} x(1-p)^{x-1} p = \frac{p}{1-p} \sum_{x=1}^{\infty} x(1-p)^x = \frac{p}{1-p} \frac{1-p}{p^2} = \frac{1}{p}$$

Alternatively, use identity $\sum_{i=1}^{\infty} q^i = \frac{1}{1-q}$; $q = 1-p$

$$E[X] = \sum_{x=1}^{\infty} x(1-p)^{x-1} p = p \frac{d}{dp} \sum_{k=1}^{\infty} -(1-p)^x = -p \frac{d}{dp} \left(\frac{1}{1-(1-p)} \right) = \frac{1}{p}$$

Example: Find the mean of a discrete uniform random variable X. Using the series identity, we obtain

$$E(X) = \sum_{x=k}^{l} x \frac{1}{l-k+1} = \frac{k+l}{2}$$

Example: A quality control engineer is inspecting a batch of 7 electronic components. The batch consists of 4 good components and 3 defective components. The engineer takes a sample of 3 components. (a) Find the PMF of the good components and (b) Find the expected value of the number of good components in this sample.

Solution: Use partitions - good and defective. Let X represent the number of good components in the sample. There are 2 distinguishable sets with two combinations: x out of 4 good components and $3-x$ out of 3 defective components.

- For set 1, $n_1 = \binom{4}{x}$ possible number of combinations,
- For set 2 , $n_2 = \binom{3}{3-x}$ possible number of combinations.

- The total number of possible combinations is $n = \binom{7}{3}$. The probability distribution (PMF) of X is

$$p_X(x) = \frac{n_1 \times n_2}{n} = \frac{\binom{4}{x}\binom{3}{3-x}}{\binom{7}{3}}, x = 0, 1, 2, 3$$

The result is provided in the table below.

x	0	1	2	3
$p(x)$	$\frac{1}{35}$	$\frac{12}{35}$	$\frac{18}{35}$	$\frac{4}{35}$

- The expected value of X is,

$$\mu_X = 0 \times \left(\frac{1}{35}\right) + 1 \times \left(\frac{2}{35}\right) + 2 \times \left(\frac{18}{35}\right) + 3 \times \left(\frac{4}{35}\right) = \frac{17}{7} = 1.7$$

- This implies that if a sample size of 3 components is taken repeatedly, there will be on average, 1.7 good components.

Expected Value of a Function of a Discrete Random Variable:

Definition: consider a discrete random variable X, with probability mass density function $p_X(x)$. The expected value of a function $h(X)$, of random variable X is defined as,

$$\mu_{h(X)} = E[h(X)] = \sum_{x \in S} h(x) p_X(x)$$

Example: Suppose the number of cars passing through a car wash between 4:00 P.M. and 5:00 P.M. on any sunny Friday has the following probability distribution (PMF):

x	4	5	6	7	8	9
$p_X(x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Let $h(x) = 2X - 1$ represent the amount of money, in dollars, paid to the attendant by the manager. Find the attendant's expected (or average) earnings for this period.

4.5 Properties of Mathematical Expectations

Property 1: Consider a random variable X with PMF $p_X(x)$. Then the expectation of a function $g(X)$, of X is calculated by the following formula:

$$E[g(X)] = \sum_{x \in Range(X)} g(x) p_X(x)$$

Property 2:
Consider a random variable X with expectation $E[X]$. Consider a linear transformation $Y = aX + b$, of X, where a and b are two constants. Then the expectation of Y is given by

$$E[Y] = aE[X] + b$$

Examples:

4.6 Variance and Standard Deviation of Discrete Random Variables:

Definition: Consider a discrete random variable, X, with PMF $p_X(x)$ and expected value μ_X. The variance of X, denoted $Var[X]$ or σ_X^2 is defined as

$$\sigma_X^2 = E[(X - \mu_X)^2] = \sum_{x \in S} (x - \mu_X)^2 p_X(x)$$

Definition: The standard deviation of a random variable X is the square-root of the variance,

$$\sigma_X = \sqrt{E[(X - \mu_X)^2]} = \sqrt{\sigma_X^2}.$$

- The variance of a random variable is the second central moment of X
- The variance (or the standard deviation) of a random variable characterizes the variability or spread in the distribution of the random variable.
- It gives a description of the shape of the distribution.

Example:
The quiz scores for a particular student are 22, 25, 20, 18, 12, 20, 24, 20, 20, 25, 24, 25 and 18. Find the variance and standard deviation.

Value	12	18	20	22	24	25
Frequency	1	2	4	1	2	3
Probability	0.08	0.15	0.31	0.08	0.15	0.23

Shortcut formula for variance:

$$\sigma_X^2 = E[(X-\mu_X)^2] = E[X^2] - 2\mu_X E[X] + \mu_X^2 = E[X^2] - \mu_X^2$$

$E[X^2]$ is the Second Moment of X ; also called the mean-square value of X

Variance of a Function of a Discrete RV

Definition: Consider a discrete random variable X, with probability mass function $p_X(x)$. Let $Z = h(X)$ be a function of X with expected value $\mu_{h(X)}$. The variance of the discrete random variable $Z = h(X)$ is defined as

$$\sigma_{h(X)}^2 = E\left[\left(h(X) - \mu_{h(X)}\right)^2\right] = \sum_{x \in S} \left(h(X) - \mu_{h(X)}\right)^2 p_X(x)$$

Examples: Calculate the variance of $h(X) = 2X + 3$, where X is a random variable with probability distribution

x	4	5	6	7	8	9
$p(x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Solution:

$$\mu_{h(X)} = E[h(X)] = \sum_{x=0}^{3} (2x+3)\, p_X(x) = 6$$

$$\sigma_{h(X)}^2 = E\left[\left(h(X) - \mu_{h(X)}\right)^2\right] = \sum_{x=0}^{3} \left(h(X) - \mu_{h(X)}\right)^2 p_X(x) = \sum_{x=0}^{3} (2x+3-6)^2\, p_X(x) = 4$$

Example: A Square-Law Device (diode) has input noise voltage X, that is uniformly distributed with values in the set $S = \{-3, -1, +1, +3\}$. The output of a square-law device is $Z = X^2$. Find $E[Z]$.

Solution:
Since X is uniform we can assume the values are equally likely, therefore, the PMF is $p_X(x) = \frac{1}{4}$. Then

$$E[Z] = \sum_{x=-3}^{3} x^2\, p_X(x) = (-3)^3 \frac{1}{4} + (-1)^2 \frac{1}{4} + (1)^2 \frac{1}{4} + (3)^3 \frac{1}{4} = 5$$

Exercise: Suppose the output of the square-law device is now $Z = (2X + 10)^2$. Show that $E[Z] = 120$

Properties of Variance:

Consider a discrete random variable, X, with variance σ_X^2. Consider a linear transformation $Y = aX + b$, of X, where a and b are two constants. Then the variance of Y is given by

$$\sigma_Y^2 = a^2 \sigma_X^2$$

The variance of a constant $Y = b$, is $\sigma_Y^2 = 0$.

Examples:

Exercises: