Objectives:

- Binomial Distribution and its moments: Section 4-3,4-4 of text.
- Poisson Distribution and its moments: Section 4-5 of text.

- In this section we study a discrete probability distribution known as the *Binomial* probability distribution.
- First, we need to define the concept of a *Bernoulli* trial.
- Consider a random variable, Y that takes on one of two possible values, say, 0 (failure) or 1 (success).
- And, Y has the following probability distribution

y	0	1
$P(y)$	1-p	p

- We say, Y is a Bernoulli random variable with success probability p and write

$$Y \sim Bernoulli(p).$$

- Note: $q = 1 - p$.

- Experiment: Flip a biased coin where Heads is expected to appear 60% of the time, once and observe the outcome.
- Random variable: Outcome of flipping this 'unfair' coin once.
- Possible Outcomes: Heads, Tails.
- Call observing Heads as a success (S). Call observing Tails a failure (F).
- $P(S) = 0.6 = p$, $P(F) = 0.4 = q$.
- $X \sim Bernoulli(0.6)$.

Definition

A **binomial probability distribution** results from a procedure that meets all of the following requirements:

1. The procedure has a *fixed number of trials*.
2. The trials must be *independent*. That is, the outcome of any individual trial doesn't affect the probabilities in the other trials.
3. Each trial must have all outcomes classified into *two categories*.i.e. each trial is a Bernoulli random variable.
4. The probabilities must remain *constant* for each trial.

Notation for Binomial Probability Distributions

S and **F** (success and failure) denote the two possible categories of all outcomes.

- $P(S) = p$; p is the probability of a success
- $P(F) = 1 - p = q$; q is the probability of a failure
- n: the number of fixed trials
- x: the observed number of successes in n trials; x can take values $0, 1, \ldots, n$

Definition of a Binomial Random Variable

A random variable that counts the number of *successes* in n *independent* Bernoulli trials with probability of success p on each trial is called a **binomial** random variable with *parameters n and p*. We write

$$X \sim Bin(n, p).$$

Example 2

Let X, be the random variable that counts the number of girls born to a couple planning to have four children. Suppose that the probability of observing a boy is **0.53**. Is X a binomial random variable?

We need to check if all the requirements of a binomial random variable are met.

1. X counts the number of successes (girls) out of four trials.
2. Only two possible values are possible for each outcome.
3. Trials are independent; the outcome of each birth is not affected by the outcome of any other birth.
4. $P(\text{“}success\text{”}) = 0.47$ for all trials.

$$X \sim Bin(4, 0.47).$$

Example 3

Genetics says that children receive genes from their parents independently. Each child has probability **0.25** of having blood type O. If a set of parents has **5** children, does X, the random variable which counts the number of children with blood type O, follow a binomial distribution?

We need to check if all the requirements of a binomial random variable are met.

1. X counts the number of successes (blood type O children) out of five trials.
2. Only two possible values are possible for each outcome if we view the outcomes as "type O" and "not type O".
3. Trials are independent; the outcome of each birth is not affected by the outcome of any other birth.
4. $P(\text{``}success\text{''}) = 0.25$ for all trials.

$$X \sim Bin(5, 0.25).$$

When sampling without replacement, the events can be treated as if they were independent if the sample size is small relative to the population size, i.e. $\frac{n}{N} < 0.05$.

Example 4

A company inspects a random sample of **10** empty plastic containers from a shipment of **10,000** and examines them for traces of benzene. Suppose **10%** of the containers have traces of benzene. Is X, the number of containers contaminated with benzene a binomial random variable?

We need to check if all the requirements of a binomial random variable are met.

- Technically, $P(success) = P(\text{a container with benzene})$ is not 10% for all 10 containers due to sampling without replacement.
- However, $10/10,000 < 0.05$.
- Therefore, by p. 114 of textbook, we may view the events as independent Bernoulli trials with $P(success) = 0.1$.
- i.e.
$$X \sim Bin(10, 0.1).$$

Let $X \sim Bin(n, p)$. We find $P(x)$, the probability that X takes the value x using the following.

The Binomial Probability Formula

$$P(x) = {}_nC_x\, p^x q^{n-x}, \text{for } \; x = 0, 1, \ldots, n$$

Recall: $q = 1 - p$.

Example 5

Let X, be the random variable that counts the number of girls born to a couple planning to have four children. Suppose that the probability of observing a boy is **0.5**. Verify the probability distribution of X is the following:

x	0	1	2	3	4
$P(x)$	1/16	1/4	3/8	1/4	1/16

Example 5

Let X, be the random variable that counts the number of girls born to a couple planning to have four children. Suppose that the probability of observing a boy is **0.5**. Verify the probability distribution of X is the following:

x	0	1	2	3	4
$P(x)$	1/16	1/4	3/8	1/4	1/16

$$X \sim Bin(4, 0.5).$$

Therefore, $P(2) = {}_4C_2\, 0.5^2(1-0.5)^{4-2} = \dfrac{3}{8}$.

Example 4 Revisited:

A company inspects a random sample of **10** empty plastic containers from a shipment of **10,000** and examines them for traces of benzene. Suppose **10%** of the containers have traces of benzene. What is the probability that the sample contains at most one contaminated container?

Example 4 Revisited:

$$\begin{aligned} P(X \leq 1) &= P(X = 1, \text{or} X = 0) \\ &= P(0) + P(1) \\ &= {}_{10}C_0\, 0.1^0 (1 - 0.1)^{10-0} + {}_{10}C_1\, 0.1^1 (1 - 0.1)^{10-1} \\ &= 0.7361. \end{aligned}$$

Example 5

When a survey calls residential telephone numbers at random, **80%** of calls fail to reach a live person. A random dialling machine makes **15** calls.

(a) What is the probability that exactly three calls reach a person?

(b) What is probability that at least one call reaches a person?

Example 5 Revisited:

- Put $X =$ the number of live calls out of 15.
- Then $X \sim Bin(15, 0.2)$.
- $P(3) = {}_{15}C_3\, 0.2^3 0.8^{12} = 0.250$.
- $P(X \geq 1) = P(1) + P(2) + \ldots + P(15)$. Too much work!
- $P(X \geq 1) = 1 - P(X < 1) = 1 - P(0) = 1 - 0.8^{15} = 0.965$.

Suppose $X \sim Bin(n, p)$.

Mean

$$\mu = E(X) = np$$

Variance

$$\sigma^2 = npq = np(1 - p)$$

Standard Deviation

$$\sigma = \sqrt{npq} = \sqrt{np(1 - p)}$$

These expressions are given on the final exam.

Example 4 Revisited:

A company inspects a random sample of **10** empty plastic containers from a shipment of **10,000** and examines them for traces of benzene. Suppose **10%** of the containers have traces of benzene. Find the mean and standard deviation of the number of containers out of a sample of ten with benzene.

Example 4 Revisited:

- $X \sim Bin(10, 0.1)$.
- Mean: $\mu = np = 10(0.1) = 1$
- Standard Deviation: $\sigma = \sqrt{npq} = \sqrt{10(0.1)(0.9)} = 0.9487$

Definition

The **Poisson distribution** is a discrete probability distribution that applies to occurrences of some event *over a specified interval*. The interval can be time, distance, area, volume, or some similar unit.

Poisson Probability Distribution Formula

$$P(x) = \frac{e^{-\mu} . \mu^{x}}{x!}, x = 0, 1, 2, \ldots$$

This expression is provided for the final exam.

Requirements of the Poisson distribution:

- The random variable X is the number of occurrences of an event *over some interval*.
- The occurrences must be *random*.
- The occurrences must be *independent* of each other.
- The occurrences must be *uniformly distributed* over the interval being used.

Mean

The mean is μ.

Standard Deviation

The standard deviation is $\sigma = \sqrt{\mu}$.

Example Poisson-1

2.3 patients arrive at a hospital's emergency room on Fridays between 10:00 p.m. and 11:00 p.m. Define X to be the random variable that counts the number of patients in the 10-11 p.m. slot.

(a) What is the probability that exactly four patients arrive between 10:00 p.m. and 11:00 p.m. on a typical Friday?

(b) Find the mean and standard deviation of the number of patients that arrive between 10:00 p.m. and 11:00 p.m. on a typical Friday?

Example Poisson-1 continued:

- $P(4) = e^{-2.3}\dfrac{2.3^4}{4!} = 0.12.$
- Mean: $\mu = 2.3$; Standard deviation: $\sigma = \sqrt{2.3} = 1.52$

Example Poisson-2

The U.S. Centres for Disease Control reports **7.7** cases of typhoid fever per week, on average from all over the United States. Define X to be the random variable that counts the number of typhoid cases per week.

(a) What is the probability of at least one typhoid case per week?

(b) Find the mean and standard deviation of the number of typhoid cases per week.

Example Poisson-2 continued:

- $P(X \geq 1) = 1 - P(X < 1) = 1 - P(0) = 1 - e^{-7.7} = 1 - 0.000453 = 0.9995.$
- Mean: $\mu = 7.7$; Standard deviation: $\sigma = \sqrt{7.7} = 2.77$

Example Poisson-3

For a recent period of **100** years, there were **93** major earthquakes in the world. Find the mean number of earthquakes per year. Suggest a suitable probability model for the number of earthquakes per year. What is the probability of no earthquakes in a randomly selected year?

Example Poisson-3 continued:

- Put $\mu = 93/100 = 0.93$ as the rate parameter.
- Then, $X \sim \textit{Poisson}(0.93)$.
- $P(0) = e^{-0.93} = 0.39$.

Example Poisson-4

Suppose that the average number of zooplankten per litre of lake water is **2**. What is the probability that **5** zooplankten will be found in a random sample of **3** litres of lake water?

Example Poisson-4 continued:

- Let X be the number of zooplankten per litre. Then $X \sim Poisson(2)$.
- Put Y to be the number of zooplankten per 3 litres. Then, $Y \sim Poisson(6)$. i.e. the mean of Y is 3×2.
- Now find $P(5) = e^{-6}\dfrac{6^5}{5!} = 0.161$.

Binomial vs. Poisson

- Binomial: counts the number of *successes* in *n* independent Bernoulli trials.
- Poisson: counts the number of *successes* in an *interval*.