Republic of the Philippines
POLYTECHNIC UNIVERSITY OF THE PHILIPPINES
Office of the Vice President for Academic Affairs
Department of Mathematics and Statistics

STAT 012 – Engineering Data Analysis
LAS 1.1: The Discrete Probability Distributions and its Types
2nd Semester

Learning Objectives:

1. Determine the types of Discrete Probability Distribution.
2. Recognize when and how to use Expected Value, Binomial, Poisson, and Hypergeometric Distribution.
3. Solve problems involving the Expected Value, Binomial, Poisson, and Hypergeometric Distribution.

Definition. Discrete Probability Distribution

Discrete probability distribution is a type of probability distribution that shows all possible values of a discrete random variable along with the associated probabilities. In other words, a discrete probability distribution gives the likelihood of occurrence of each possible value of a discrete random variable.

Note. The important property of a PMF is that the sum of all probabilities involved is 1. That is, if X is any random variable, then

$$\sum_{i=1}^{n} P(X = x_i) = 1$$

Example 1. The table below tells us the probability of selling a car unit in a given day.

x	0	1	2	3	4
$P(X = x)$	0.41	0.37	0.16	0.05	0.01

Example 2. Find the value of k. Find the probability of at most 3 and greater than 6.

x	0	1	2	3	4	5	6	7
$P(X = x)$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

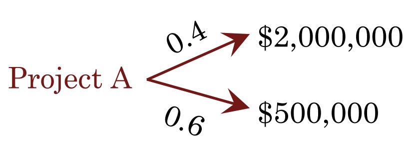
Definition. Expected Value

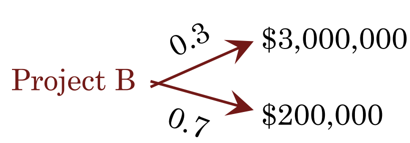
The expected value denoted as $E(X)$ is an anticipated value at some point in the future. Also called a "long-term" average or mean. It is calculated by multiplying each of the possible outcomes by the likelihood each outcome will occur and then summing all of those values.

$$E(X) = \sum_{i=1}^{n} x_i \cdot P(x_i)$$

Example 3. You are an engineer in a development company. Your manager just asked you to assess the viability of future development projects and select the most promising one. According to estimates, Project A, upon completion, shows a probability of 0.4 to achieve a value of $2 million and a probability of 0.6 to achieve a value of $500,000. Project B shows a probability of 0.3 to be valued at $3 million and a probability of 0.7 to be valued at $200,000 upon completion. Which project will you choose?

Solution.





$E(A) = (\$2{,}000{,}000)(0.4) + (0.6)(\$500{,}000)$

$E(A) = \$800{,}000 + \$300{,}000$

$E(A) = \$1{,}100{,}000$

$E(B) = (\$3{,}000{,}000)(0.3) + (0.7)(\$200{,}000)$

$E(B) = \$900{,}000 + \$140{,}000$

$E(B) = \$1{,}040{,}000$

Since Project A yields to higher expected value, I will choose the Project A.

Example 4. You decided to join a raffle and purchased one ticket. It costs \$10 each. The raffle manager is selling 7,000 tickets. One ticket will be randomly drawn and the winner will receive \$20,000. Assume that there were only 5,500 tickets sold. What is the expected value?

Definition. The Binomial Distribution

A binomial distribution can be thought of as simply the probability of a SUCCESS or FAILURE outcome in an experiment or survey that is repeated multiple times with replacement.

$$B(x; n,\ p) = C(n,\ x) \cdot p^x \cdot q^{n-x}$$

where b is the binomial probability.
n is the number of times the experiment runs.
x is the total number of "successes".
p is the probability of the success on an individual trial.
q is the probability of failure on an individual trial.
C indicates combination.

Johann Bernoulli
1667 - 1748
Swiss Mathematician

How it works? Recall the basic example of flipping a coin 2 times. As X be a random variable pertaining to number of heads. The probability mass function is

$$p(x) = \begin{cases} \dfrac{1}{2} & \text{if } x = 1 \\ \dfrac{1}{4} & \text{if } x = 2,\ 0 \end{cases}$$

At $x = 0$, (no head)

$$B(x; n,\ p) = C(n,\ x) \cdot p^x \cdot q^{n-x}$$
$$B(0; 2, ½) = C(2, 0) \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^{2-0}$$
$$= \frac{2!}{(2-0)! \cdot 0!} \cdot 1 \cdot \left(\frac{1}{4}\right)$$
$$= 1 \cdot 1 \cdot \frac{1}{4} = \frac{1}{4}$$

At $x = 1$, (1 head)

$$B(x; n,\ p) = C(n,\ x) \cdot p^x \cdot q^{n-x}$$
$$B(1; 2, ½) = C(2, 1) \cdot \left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^{2-1}$$
$$= \frac{2!}{(2-1)! \cdot 1!} \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right)$$
$$= 2 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$$

Note. We used the Bernoulli's Principle if the given experiment has binary outcomes. Binomial distributions play an important role in probability especially when a certain experiment performed a bunch of trials. That is, the value of n is way too large.

Example 5. A coin is tossed 10 times. What is the probability of getting exactly 6 heads?

Example 6. In a 20-item examination, each item contains 4-letter multiple choice labeled as A, B, C, and D. Find the probability of getting a perfect score.

Example 7. A couple has a plan to give birth of 5 babies. The doctor told them that the probability of getting a baby boy is about 38.7%. If a random variable Y corresponds to number of male babies, find the probability that that they will achieve 5 female babies in the future.

Siméon-Denis Poisson
1781 - 1840
French Mathematician

Definition. The Poisson Distribution

The probability distribution of the Poisson random variable X, representing the number of outcomes occurring in a given time interval or specified region denoted by t, is

$$p(x;\,\lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!},\ x = 0,\ 1,\ 2,\ \ldots$$

where λ is the average number of outcomes per unit time, distance, area, or volume.
t is the given time or space.
x is the number of outcomes to occur .
e is the "Euler number" $2.71828\ldots$.

Note. We used Poisson Distribution if there are any given average value and the experiments yielding numerical values of a random variable X, the number of outcomes occurring during a given time interval or in a specified region. The given time interval may be of any length, such as a minute, a day, a week, a month or even a year . The specified region could be a line segment, an area, a volume, or perhaps a piece of material.

Example 8. The secretary of ABC Company, a law firm, finds that there are on the average of 5 calls from clients per day. Find the probability that the firm will receive exactly 1 call and at least 1 call on a particular day.

Solution. Let X be the number of calls.

$\lambda = 5$ calls $t = 1$ day $x = 1$ call

By Poisson Distribution,

$$p(x;\,\lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!} \Rightarrow p(1;5) = \frac{e^{-5}(5)^1}{1!} = \frac{(0.006737946999085)(5)}{1} \approx 0.0337$$

$\therefore P(X = 1; 5) = 0.0337$ or 3.37%.

How about at least 1 call? That is, $P(X \geq 1; 5)$?

STAT 012 – Engineering Data Analysis
LAS 1.1: The Discrete Probability Distributions and its Types
2nd Semester

Example 9. The average number of class suspensions due to typhoons announced by Mayor Honey within a semester in Manila is 3.8. What is the probability that for this entire school year, classes in the City of Manila will be suspended due to typhoons for at most 4 days?

Example 10. You are a journalists, and based on your statistics performance declared by your department head, you are averaging 1 typographical errors per page in a CANDY magazine. In order to be a good journalist, you must not commit errors anymore. Based on this data, you asked yourself. What is the probability that in your 3-page content, there are 0 errors will occur in the said magazine?

Example 11. The mean number of bacteria per milliliter of a liquid is known to be 6. Find the probability that in 1 ml of the liquid, there will be (a) 0 bacteria and (b) at most 3 bacteria.

Definition. The Hypergeometric Distribution

The hypergeometric distribution is a distribution in which selections are made from two groups without replacing members of the groups. It can be expressed as

$$H(x;\ N, n, k) = \frac{\binom{k}{x}(\binom{N-k}{n-x}}{\binom{N}{n}}$$

where N is the total objects in the experiment
n is the number of trial
x is the possible value
k is the total number of success
$N-k$ is the total number of failure

John Wallis
1616 - 1703
English Mathematician

Note. The relationship of Binomial Distribution and Hypergeometric Distribution is that in a binomial distribution, sampling is done with replacement, meaning the probability of success remains constant for each trial, while in a hypergeometric distribution, sampling is done without replacement, causing the probability of success to change with each trial as the population composition alters with each selection.

Example 12. Lots of 40 components each are deemed unacceptable if they contain 3 or more defectives. The procedure for sampling a lot is to select 5 components at random and to reject the lot if a defective is found. What is the probability that exactly 1 defective is found in the sample if there are 3 defectives in the entire lot?

STAT 012 – Engineering Data Analysis
LAS 1.1: The Discrete Probability Distributions and its Types
2nd Semester

Example 13. A group of 10 individuals is used for a biological case study. The group contains 3 people with blood type O, 4 with blood type A, and 3 with blood type B. What is the probability that a random sample of 5 will contain 1 person with blood type O, 2 people with blood type A, and 2 people with blood type B?

Example 14. How many sample points are there if you bet a 6/49 Mega Lotto? What is the probability that you will the jackpot prize in a single bet? Construct the PMF.