3 : Basic Statistics

IT5506 – Mathematics for Computing II

Level III - Semester 5

Intended Learning Outcomes

At the end of this lesson, you will be able to;

- define Binomial random variables are and how they are used.
- compute the Cumulative Distribution Function of a Binomial random variable.
- compute the mean and the variance of a Binomial random variable.

List of sub topics

3.4 The Binomial probability distribution

3.4 The Binomial probability distribution

Binomial Experiment

- Suppose that we have a biased coin for which the probability of obtaining a head is 2/3.
- We toss the coin 100 times and count the number of heads obtained.
- This problem is typical of an entire class of problems that are characterized by the feature that there are exactly two possible outcomes (for each trial) of interest.
- These problems are called binomial experiments.

3.4 The Binomial probability distribution

Features of a Binomial experiment

1. There are a fixed number of trials. We denote this number by n.
2. The n trials are independent (result of one trial does not depend on any other trial), and are repeated under identical conditions.
3. Each trial has only two outcomes; success denoted by S, and failure denoted by F.
4. For each trial, the probability of success is the same. We denote the probability of success by p and that of failure by q. Since each trial results in either success or failure, $p + q = 1$ and $q = 1 - p$.

3.4 The Binomial probability distribution

- Suppose that we have a biased coin for which the probability of obtaining a head is 2/3. We toss the coin 100 times and count the number of heads obtained.
- In this experiment of tossing a biased coin, let us now see how it meets the above criteria of a binomial experiment. Consider the above features one at a time.
    1. The coin is tossed 100 times, so there are n = 100 trials (fixed) in this case.
    2. The trials can be considered as independent, as the outcome of one trial has no effect on the outcome of another trial.
    3. There are only two outcomes, head or tail. As we are interested in getting a head, it can be considered as a success, and getting a tail can be considered as a failure.
    4. On each trial, the probability p of success is 2/3 (same for all trials).
- In this type of binomial experiments, our interest is to find the probability of a certain number of successes (say r) out of n trials.

Here, if X is defined as the number of getting r successes, then we say X is distributed as Binomial with parameters n and p. That is denoted by:

3.4 The Binomial probability distribution

- In this type of binomial experiments, our interest is to find the probability of a certain number of successes (say r) out of n trials.
- Here, if X is defined as the number of getting r successes, then we say X is distributed as Binomial with parameters n and p. That is denoted by:

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

3.4 The Binomial probability distribution

Example:

- Suppose a student is taking a multiple-choice question paper, and he has only three more multiple-choice questions left to do. Each question has 4 suggested answers, and only one of the answers is correct. He has only few seconds left to do these three questions, so he decides to randomly select (guess) the answers. The interest here is to know the probability that he gets zero, one, two, or all three questions correct.

Solution:

- This is a binomial experiment. Each question can be considered as a trial, so the number of trials (n) is 3.
- There are only two outcomes for each trial – success (S) indicating a correct answer, and failure (F) indicating a wrong answer.
- The trials are independent – outcome (correct or incorrect) for any one question does not affect the outcome of the others.
- Sine he is guessing and there are 4 answers from which to select, the probability of a correct answer is $p=0.25$. The probability q of an incorrect answer is then $1– p = 1 - 0.25 = 0.75$.
- So, this is a binomial experiment with $n = 3$, $p = 0.25$. So $X \sim Bin(3, 0.25)$.

3.4 The Binomial probability distribution

- Now what are the possible outcomes in terms of success or failure for these three trials?
- Here we use the notation SFS to indicate a success on the first question, a failure on the second, and a success on the third. There are 8 possible combinations of S's and F's.
- They are:

SSS SSF SFS FSS SFF FSF FFS FFF

- The probability for each of the above combinations can be computed using the multiplication law as the trials are independent. To illustrate this, let us compute the probability of 'SFS' (success on the first question, failure on the second, and success on the third).
- $P(SFS) = P(S).P(F).P(S) = p.q.p = p2q = (0.25)2(0.75) \approx 0.047$

3.4 The Binomial probability distribution

In a similar way, probability of each of the above eight outcomes can be computed, and they are given in the table below.

Table 1 : Probabilities of outcomes for a Binomial experiment with n=3 & p=0.25

Outcome	No. of successes (r)	Probability
SSS	3	$P(SSS) = P(S)P(S)P(S) = p.p.p \approx 0.016$
SSF	2	$P(SSF) = P(S)P(S)P(F) = p.p.q \approx 0.047$
SFS	2	$P(SFS) = P(S)P(F)P(S) = p.q.p \approx 0.047$
FSS	2	$P(FSS) = P(F)P(S)P(S) = q.p.p \approx 0.047$
SFF	1	$P(SFF) = P(S)P(F)P(F) = p.q.q \approx 0.141$
FSF	1	$P(FSF) = P(F)P(S)P(F) = q.p.q \approx 0.141$
FFS	1	$P(FFS) = P(F)P(F)P(S) = q.q.p \approx 0.141$
FFF	0	$P(FFF) = P(F)P(F)P(F) = q.q.q \approx 0.422$

3.4 The Binomial probability distribution

- Let us now compute the probability that the student gets zero, one, two, or all three questions correct.

  P(X = 1) = P[SSF or FSF or FFS]

  = P(SFF) + P(FSF) + P(FFS) as SFF, FSF & FFS are mutually exclusive

  = 0.423

- In the same way, we can find that P(X =2), P(X = 3) and P(X = 0).
- Verify that P(X = 3) = 0.016, P(X =2) = 0.141, and P(X = 0) = 0.422. With these results you can see that there is very little chance (0.016) that the student gets all the questions correct.
- If X ~ Bin(n, p), then the general formula of computing the probability of getting r successes can be specified as:

$$P(X=r)=\frac{n!}{r!(n-r)!}p^r q^{n-r}, \quad \mathrm{r}=0, 1, 2, ..., \mathrm{n}$$

Where n = no. of trials, p = prob. of success
r = no. of successes

3.4 The Binomial probability distribution

- If X ~ Bin(n, p), then the general formula of computing the probability of getting r successes can be specified as:

$$P(X=r)=\frac{n!}{r!(n-r)!}p^r q^{n-r}, \quad r = 0, 1, 2, \ldots, n$$

Where n = no. of trials, p = prob. of success
r = no. of successes

- The $\frac{n!}{r!(n-r)!} = {}^nC_r$ is the binomial coefficient which represents the number of combinations with n trials having r successes.
- As an exercise, compute the probabilities in table 1 using the general formula of the binomial distribution.

3.4 The Binomial probability distribution

Example:

Suppose that a computer component has a probability of 0.7 of functioning more than 10,000 hours. If there are 6 such components, what is the probability that at least 4 computer components will function more than 10,000 hours?

Solution:

This is a binomial experiment with $n = 6$ and $p = 0.7$. Therefore, if we define X as the number of computer components functioning more than 10,000 hours, then

$$X \sim Bin(6, 0.7).$$

So, the required probability is:

$P(X \geq 4) = P(X=4 \text{ or } X=5 \text{ or } X=6)$

$= P(X=4) + P(X+5) + P(X=6)$

$= 0.324 + 0.303 + 0.118$

$= 0.745$

3.4 The Binomial probability distribution

Mean and Variance of Binomial distribution

Here we can use two formulas to compute the mean (μ) and variance (σ2) of the binomial distribution.

If X ~ Bin(*n, p*), then it can be shown that

Mean = μ = *n.p* is the expected number of successes and

Variance = σ^2 = *n.p.q* is the variance for the number of successes.