BNAD 276
Lecture 5
Discrete Probability Distributions

Phuong Ho

May 22, 2017

Outline

Outline

Example 1

- Let's begin with an example.
- Suppose we conduct the experiment of tossing a coin.
- Possible outcomes in this experiment are either H or T.
- We want to work with numbers instead of alphabet letter, H or T, when we describe an outcome from this experiment.
- We assign a number for each outcome as follows.

$$X = \begin{cases} 0 & \text{if the outcome is } H \\ 1 & \text{if the outcome is } T \end{cases}$$

- X in the above is an example of random variables.

Random Variables

The "formal" definition of random variable is as follows.

Random Variables

A **random variable** is a **numerical description** of the outcome of an experiment.

- Recall that a specific outcome is randomly realized when we conduct an experiment.
- Note that a value of a random variable is necessarily associated with an outcome in an experiment.
- So, the value of the random variable is determined based on the realized outcome of the experiment.

Notations

- We use a capital letter to denote a random variable, (e.g. X).
- We use a lowercase letter to denote values that a random variable can take, (e.g. x).

e.g. Previously we have a random variable X which associates H to 1 and T to 0.

- Note that, in this example, our random variable is denoted by X.
- Possible values that X can take are denoted by $x = 0, 1$.

Example 1 cont'd

- An outcome, H or T, is randomly realized.
- If an outcome H is realized, we decided to record $X = 1$ instead of H.
- Because the value that X takes depends on the realized outcome and an outcome is randomly realized, the value of X is also randomly realized.
- Hence, X is called a random variable.

Example 2

- Sometimes, an outcome is a number by itself.
- Consider the experiment of rolling a die.
- The sample space of this experiment is

$$S = \{1, 2, 3, 4, 5, 6\}$$

- We can define our random variable X as follows.

$$X = \text{An outcome in the experiment}$$

- Since the outcome is already a number, X is of course a numerical description of the outcome, which fits the definition we had.

Example 3 (1/3)

- In the previous examples, we have a straightforward way to obtain a random variable. However, it is not always straightforward.
- And, we may be interested in only some aspect of the outcome in the experiment.

Suppose we conduct an experiment of tossing a coin three times.

- The sample space will be
  $S = \{(HHH), (HHT), (HTH), (HTT), (THH), (THT), (TTH), (TTT)\}$
- Now, define our random variable X as follows.

$$X = \text{The number of Heads in an outcome.}$$

- Note that X is a numerical description of an outcome since X will be either 0,1,2 or 3.

Example 3 cont'd (2/3)

We can summarize X as follows:

$$X = \begin{cases} 0 & \text{The outcome is } (TTT) \\ 1 & \text{The outcome is either } (HTT), (THT), \text{ or } (TTH) \\ 2 & \text{The outcome is either } (HHT), (THH) \text{ or } (HTH) \\ 3 & \text{The outcome is } (HHH) \end{cases}$$

Example 3 cont'd (3/4)

We can define a different random variable Y:

$$Y = \text{ The number of tails in an outcome}$$

$$Y = \begin{cases} 3 & \text{The outcome is } (TTT) \\ 2 & \text{The outcome is either } (HTT), (THT), \text{ or } (TTH) \\ 1 & \text{The outcome is either } (HHT), (THH) \text{ or } (HTH) \\ 0 & \text{The outcome is } (HHH) \end{cases}$$

Example 3 cont'd (4/4)

We can define another different random variable Z:

$Z =$ The number of heads in the last two tossing times of an outcome

$$Z = \begin{cases} 0 & \text{The outcome is } (TTT) \text{ or } (HTT) \\ 1 & \text{The outcome is either } (HHT), (HTH), (THT), \text{ or } (TTH) \\ 2 & \text{The outcome is either } (THH) \text{ or } (HHH) \end{cases}$$

- Hence, an experiment generates one sample space but there may be multiple random variables depending on what we are interested in.

Discrete Random Variable

- The examples we considered so far has a common characteristics.
- In every example, the values that X can take are *discrete*.

e.g. 1 X only can take values such as 0,1,2, or 3 in the previous example.

Discrete Random Variable

Discrete random variable is a random variable that takes a finite number of values, or an infinite sequence of integer values such as $0, 1, 2, \ldots$

Examples of discrete RVs

Experiment	Random Variable (X)	Possible Values for X
Contact five customers	Number of customers who place an order	0,1,2,3,4,5
Inspect 50 radios	Number of defective radios	0,1,2,3,....,50
Operate a restaurant one day	Number of customers	0,1,2,3,......
Sell an automobile to someone	Gender of the customer	0 if male, 1 if felmale

Continuous Random Variables

- Some experimental outcomes have non-discrete values.

e.g. An experimental outcome such as time, weight, distance, and temperature can take any value in an interval.

Experiment	Weigh a parcel in a USPS store
Random Variable	Temperature
Possible Values	$0 < x$

- Note that X (the weight) can take any positive values, for example 1lb, 2.3lbs
- When a random variable exhibits this property, we call it a **Continuous Random Variable.**

Summary

1. A random variable is a numerical description of the outcome in an experiment.
2. A random variable is called
   - Discrete Random Variable when it only takes clearly separated values such as 0,1,2,3,...
   - Continuous Random Variable when it can takes any values in an interval.
3. In this lecture note, we are going to learn probability theory of a discrete random variable. In the next lecture note (LN6), we will learn the continuous one.

Outline

- We only restrict our attention on discrete random variables in this section.
- We will talk about **Discrete Probability Distribution** and **Discrete Probability Function**.

Example 1

- Recall the experiment of tossing a coin 3 times.
- The sample space for this experiment is

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

- Following the classical method of assigning probabilities, we have

$$P(HHH) = P(HHT) = \cdots = P(TTT) = \frac{1}{8}$$

- Now, define discrete random variable X as the number of heads. i.e.
  X=the number of H in the outcome

Example 1 cont'd

- Then, X is summarized as follows before

$$X = \begin{cases} 0 & \text{The outcome is } (TTT) \\ 1 & \text{The outcome is either } (HTT), (THT), \text{ or } (TTH) \\ 2 & \text{The outcome is either } (HHT), (THH) \text{ or } (HTH) \\ 3 & \text{The outcome is } (HHH) \end{cases}$$

- Now, ask the probability that $X = 0$,
- It is the same question as "What is the probability that (TTT) happens?"
- The answer is simply $1/8$ based on the probabilities that we assigned.

Example 1 cont'd

- What is the probability that $X = 1$?
- It is the same question as "What is the probability that either $(HTT), (THT)$ or (TTH) happens?"
- Note that if (HTT) happens, neither of (THT) or (TTH) happens.
- They are mutually exclusive. Thus, we merely sum up probability of $(HTT), (THT)$ and (TTH).
- Finally, the probability that $X = 1$ is $\frac{3}{8}$.

Example 1 cont'd

- We can do the same procedure for $X = 2, 3$.
- Now define $f(x)$ is the function that shows the probability that $X = x$, where $x = 0, 1, 2, 3$.
- Finally we have the following.

$$f(x) = \begin{cases} \frac{1}{8} & \text{if } x = 0 \\ \frac{3}{8} & \text{if } x = 1 \\ \frac{3}{8} & \text{if } x = 2 \\ \frac{1}{8} & \text{if } x = 3 \end{cases}$$

- The above function is called the probability function and probability distribution is determined by the probability function $f(x)$.

Discrete Probability Distributions

- The probability distribution for a random variable describes how probabilities are distributed over the values of the random variables.

Discrete probability function

A probability function, $f(x)$, for a discrete random variable X is defined by

$$f(x) = Prob(X = x), \quad x \text{ is the possible values of } X$$

- The probability function $f(x)$ for a **discrete** random variable X is also called a **"pmf"**, i.e. **probability mass function.**

Remarks

1. Example of tossing a coin 3 times illustrates that probability function is closely related to the fundamental experiments and the probabilities of outcomes in an experiments.
2. Probability function $f(x)$ of a discrete random variable is the function that gives us the probability that $X = x$.

e.g. If $1, 2, 3$ are the possible values of X, small xs are 1,2, or 3. And, $f(x = 1)$ is the probability that X takes 1 among the possible values, 1,2, or 3.

3. Probability distribution is the description of how probabilities are distributed over the possible values of a random variable. It is described by $f(x)$ for all points x.

Two requirements for a discrete probability function

- Remind that $f(x)$ is the **probability** that X is equal to x.
- At the end of the day, $f(x)$ is a probability. Thus, it should satisfy the two requirements for probability.

Two Requirements for Discrete Probability Functions

$$0 \leq f(x) \leq 1, \quad \text{for all } x$$

$$\sum_x f(x) = 1$$

- The range of all values x at which $f(x) > 0$ is called the **support** of the random variable X (or of $f(x)$).

Example 2

Suppose we have a discrete random variable, X, and the probability function for X, $f(x)$, as follows.

$$f(x) = x/15, \quad x = 1, 2, 3, 4, 5$$

- We don't know what the underlying experiment and sample space for X but we still can check whether the above probability function is valid.
- Is $f(x)$ between 0 and 1?
- Is $\sum f(x)$ is equal to 1?

Example 2 cont'd

- Furthermore, we can answer to the questions such as
  - What is $Prob(X = 1)$?
  - Recall that $f(x) = Prob(X = x)$. The above question asks $Prob(X = 1)$. By definition, we know $Prob(X = 1) = f(1)$.
  - Thus, $Prob(X = 1) = f(1) = \frac{1}{15}$.
  - What is $Prob(X = 5)?$.
- Once we know the underlying experiment and what X means, we can give the appropriate interpretation on $Prob(X = 1)$.

Example 2 cont'd

- Once we know probability function, we can also calculate probabilities such as $Prob(X \leq x)$.
- What is $Prob(X \leq 3)$?
  - We know that the possible values that X can take is 1,2,3,4, or 5.
  - Thus, $Prob(X \leq 3) = Prob(X = 1$ or $X = 2$ or $X = 3)$.
  - It becomes $Prob(X = 1) + Prob(X = 2) + Prob(X = 3)$ since $X = 1$, $X = 2$ and $X = 3$ are mutually exclusive.
  - Finally, we have

$$\begin{aligned} Prob(X \leq 3) &= Prob(X = 1) + Prob(X = 2) + Prob(X = 3) \\ &= f(1) + f(2) + f(3) \\ &= \frac{1}{15} + \frac{2}{15} + \frac{3}{15} = \frac{6}{15} \end{aligned}$$

Cumulative Distribution Function

Discrete Cumulative Distribution Function

The **cumulative distribution function CDF**, or the cumulative probability distribution, $F(x_0)$, of a discrete random variable, X, represents the probability that X does not exceed a specific value x_0. That is,

$$F(x_0) = Prob(X \leq x_0) = \sum_{x \leq x_0} f(x).$$

e.g In the previous slide, we had $Prob(X \leq 3) = Prob(X = 1) + Prob(X = 2) + Prob(X = 3)$.

- It can be written as $F(3) = Prob(X \leq 3) = \sum_{x \leq 3} f(x) = f(1) + f(2) + f(3)$.

Example: Cumulative Distribution Function

- Once we have a probability distribution of X, it directly determines the cumulative probability distribution of X.
- We had $f(x) = x/15, \qquad x = 1, 2, 3, 4, 5$.

Random Variable	Probability Distribution	Cumulative Probability Distribution
X	$f(x)$	$F(x)$
$x = 1$	$f(1) = 1/15$	$F(1) = Prob(X \leq 1) = 1/15$
$x = 2$	$f(2) = 2/15$	$F(2) = Prob(X \leq 2) = 1/15 + 2/15 = 3/15$
$x = 3$	$f(3) = 3/15$	$F(3) = Prob(X \leq 3) = 1/15 + 2/15 + 3/15 = 6/15$
$x = 4$	$f(4) = 4/15$	$F(4) = Prob(X \leq 4) = 10/15$
$x = 5$	$f(5) = 5/15$	$F(5) = Prob(X \leq 5) = 15/15 = 1$

Exercise 1

Suppose we have the probability distribution for the random variable X as follows.

X	$f(x)$
20	.20
25	.15
30	.25
35	.40

- Is this probability distribution valid? Explain.
- What is the probability that $X = 30$.
- What is the probability that X is less than or equal to 25?
- What is the probability that X is greater than 30?

Outline

- Recall that we measure the central location and variability of **data** by **sample** mean and **sample** variance.
- A random variable X possibly takes many different values, thus it also has central location and variability measures.
- A measure of central location of a **random variable** X is **EXPECTED VALUE** (or “TRUE” MEAN).
- A measure of variability of a **random variable** X is **VARIANCE** (or “TRUE” VARIANCE).

Expected Value (“True” Mean)

- Once we know the possible values x of a random variable X and the probability mass function of X, we can calculate the expected value, or **the mean, of a random variable X.**

Expected Value of a Discrete Random Variable

$$E(X) = \mu = \sum_{x} x f(x)$$

- Intuitively, it makes sense since each value x is weighted by the “relative frequency” that can become the probability of the event $X = x$.

Example 1.Calculate Expected Value/Mean

Suppose we have the following probability distribution for X.

X	$f(x)$
0	0.1
1	0.4
2	0.3
3	0.2

$$
\begin{aligned}
E(X) &= \mu = \sum_x x f(x) \\
&= 0 \cdot f(0) + 1 \cdot f(1) + 2 \cdot f(2) + 3 \cdot f(3) \\
&= 0 + 0.4 + 0.6 + 0.6 = 1.6
\end{aligned}
$$

Variance

- Since the random variable X can take many possible values x, we also have a measure of the variability of X. One of such measure is the variance:

Variance of a Discrete Random Variable

$$Var(X) = \sigma^2 = \sum_x (x - \mu)^2 f(x)$$

- Note that $x - \mu$ implies how far a particular value x of the random variable is from the expected value μ.

Example 1 cont'd. Calculate Variance

X	$f(x)$
0	0.1
1	0.4
2	0.3
3	0.2

- We know that $E(X) = \mu = 1.6$.
- At first we need to calculate $x - \mu$ for each x. Then, we square each number and multiply $f(x)$ to that number and sum up all numbers.

Example 1 cont'd. Calculate Variance

$$\begin{aligned}\sigma^2 = Var(X) &= \sum_x (x-\mu)^2 f(x) \\ &= (0-1.6)^2 \times 0.1 + (1-1.6)^2 \times 0.4 \\ &\quad + (2-1.6)^2 \times 0.3 + (3-1.6)^2 \times 0.2 \\ &= 0.84\end{aligned}$$

- The higher variance, the higher variability of X.
- With μ and σ^2, we can summarize the central location and variability of X.
- It is **IMPORTANT** to notice that μ is NOT the *sample* mean and σ^2 is NOT the *sample* variance.

Standard Deviation of a Random Variable

Standard Deviation of a Random Variable

$$\sigma = \sqrt{\sigma^2} = \sqrt{Var(X)}$$

- In words, standard deviation of a random variable is the square root of the variance of that random variable.
- Standard deviation has the same unit as the mean/expected value.
- Hence, standard deviation tells us how far the values of the random variable is from the mean/expected value in general.

Expected Value/Mean of a general function $g(X)$

Expected Value of a function $g(X)$

$$E[g(X)] = \sum_x g(x)f(x)$$

Example, Second Moment, or $E(X^2)$

$$E(X^2) = \sum_x x^2 f(x)$$

Useful Property 1: Linearity of the Expected Values

Expected value of a linear function of X

Let Y be a linear function of X. i.e. $Y = a + bX$, where a and b are constants. Then,

$$E(Y) = E(a + bX) = a + bE(X)$$

Proof of property 1

$$
\begin{aligned}
E(Y) &= E(a+bX) = \sum_x (a+bx)f(x) \\
&= \sum_x (af(x) + bxf(x)) \\
&= \sum_x af(x) + \sum_x bxf(x) \\
&= a\sum_x f(x) + b\sum_x xf(x) \\
&= a + bE(X)
\end{aligned}
$$

Useful Property 2: Variance of a linear function of X

Variance of a linear function of X

Let Y be a linear function of X. i.e. $Y = a + bX$ where a and b are constants. Then,

$$Var(Y) = Var(a + bX) = b^2 Var(X)$$

So, note that $Var(constant) = 0$

Proof of property 2

$$
\begin{aligned}
Var(Y) &= Var(a+bX) = E[\{a+bX-E(a+bX)\}^2] \\
&= E\{a+bX-a-bE(X)\}^2] \\
&= E[\{b(X-\mu\}^2] \\
&= E[b^2\{X-\mu\}^2] \\
&= b^2E[(X-\mu)^2] \\
&= b^2Var(X)
\end{aligned}
$$

Useful Property 3: Variance

Another formula to calculate Variance

$$Var(X) = \sigma^2 = E(X^2) - \mu^2$$

In words, Variance (X) = second moment − mean squared

Proof of property 3

$$
\begin{aligned}
Var(X) &= \sum_x (x-\mu)^2 f(x) = \sum_x (x^2 - 2\mu x + \mu^2) f(x) \\
&= \sum_x (x^2 f(x) - 2\mu x f(x) + \mu^2 f(x)) \\
&= \sum_x x^2 f(x) - \sum_x 2\mu x f(x) + \sum_x \mu^2 f(x) \\
&= \sum_x x^2 f(x) - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x) \\
&= E(X^2) - 2\mu E(X) + \mu^2 \times 1 \\
&= E(X^2) - \mu^2 \\
\text{or} &= E(X^2) - [E(X)]^2
\end{aligned}
$$

Outline

Some Popular Discrete Probability Distributions

- Depending on how we define a random variable X and what kinds of experiment we have, we have infinitely many discrete probability distributions.
- However, there are several discrete probability distributions that are commonly used.
- Here, we will consider Binomial Probability Distribution and Poisson Probability Distribution.

Binomial Experiment

A Binomial Experiment is a multiple-step experiment such that

1. The experiment consists of a sequence of n identical trials.
2. Two outcomes are possible on each trial. So, we will refer to one outcome as a *success* and the other as a *failure*.
3. The probability of a success, denoted by p, does not change from trial to trial. (Hence, the probability of a failure is $1 - p$ for each trial.)
4. The trials are independent.

- Each trial is called a Bernoulli trial or binomial trial. A Bernoulli trial always has two possible outcomes, S or F. (e.g. toss a coin).
- In a binomial experiment, our interest is the number of successes occurring in the n trials.

Example 1

The experiment of tossing a coin 5 times. Our interests is the total number of H.

1. This experiment consists of a sequence of 5 identical trials: Each trial involves the tossing of one coin.
2. Two outcomes are possible for each trial: H or T. Define H as *Success* and F as *Failure*.
3. The probability of a head/success is the same 0.5 for each trial.
4. The trials or tosses are independent because the outcome on any one trial is not affected by what happens on other trials.

Thus, this experiment is a binomial experiments.

Example 2

Consider an insurance salesperson who visit 10 families to sell the insurance. If he can sell an insurance policy to a family successfully, we consider the outcome as $Success$. From the past experience, he knows that probability that a randomly selected family will purchase an insurance policy is 0.40. Our interest is the total number of insurance policies that he sold.

1. This experiment consists of 10 identical trials (Visiting a family)
2. There are only two outcomes, $Success$ (the family purchase the insurance) or $Failure$ (the family doesn't purchase the policy).
3. The probability of a purchase is $p = 0.4$ and are assumed to be the same for each sales call.
4. The trials are independent because the families are randomly selected.

Exercise

a) Consider the following experiment.

We toss a coin, roll a die, and toss a coin again. Is this experiment a binomial experiment? Explain.

b) Consider the following experiment.

We have a special coin that has the following property. Once a coin landed with H(or T), it is more likely that we have H(or T) in the next time. Using this coin we conduct the experiment of tossing this coin 3 times. Is this experiment a binomial experiment? Explain.

Binomial Distribution

We say our random variable X follows binomial distribution if

1. we have a binomial experiment, AND
2. X is defined by the number of $Success$ in that experiment.

If X follows a binomial distribution,

we can calculate $Prob(X = x)$ with $f(x)$, where $f(x)$ is the binomial probability function. Then, what is $f(x)$? How $f(x)$ looks like?

The Binomial Probability Function

The Binomial Probability Function

$$f(x) = \binom{n}{x} p^x \, (1-p)^{n-x},$$

where n is the total number of trials in the experiment, and x is the number of $Success$ in the experiment.

- Thus, $f(x)$ gives us the probability that the number of $Success$ is equal to x.

e.g. Suppose that we have $n = 5$ and $p = 0.4$. The probability that the number of $Success$ is 3 is:

$$P(X = 3) = f(3) = \binom{5}{3}(0.4)^3(1-0.4)^{5-3} = \binom{5}{3}(0.4)^3 \times (0.6)^2$$

Example. Selling insurance to 3 families

Let's consider simpler version of example 2. Instead of visiting 10 families, let us suppose we visit 3 families.

We will define the outcome that a family buy the insurance as S and the outcome that a family does not buy it as F.

The probability that a family buys the insurance is 0.40, i.e. $p = 0.4$.

Let's define $X =$ the number of sales after visiting 3 families.

$$X = \begin{cases} 0 & \text{if } (FFF) \text{ occurs} \\ 1 & \text{if } (SFF), (FSF), \text{ or } (FFS) \text{ occurs} \\ 2 & \text{if } (SSF),(SFS), \text{ or } (FSS) \text{ occurs} \\ 3 & \text{if } (SSS) \text{ occurs} \end{cases}$$

Selling insurance to 3 families, cont'd

- Question: What is the probability that the salesperson can sell exactly one insurance policy successfully after visiting 3 families? ie. $P(X = 1)$?
- We see that the experiment is binomial experiment with $n = 3, p = 0.4$. Our interest is to calculate $P(X = 1) = f(1)$.
- Now let's use the formula.

$$\begin{aligned} P(X = 1) = f(1) &= \binom{3}{1}(0.4)^1(1 - 0.4)^3 \\ &= \binom{3}{1}(0.4)(0.6)^2 \end{aligned}$$

Intuition of the binomial distribution formula

- Question: What is the probability that the salesperson can sell exactly one insurance policy successfully after visiting 3 families? ie. $P(X = 1)$?
- $X = 1$ happens when either (SFF), (FSF) or (FFS) happens.
- Hence, $P(X = 1)$ is

$$\begin{aligned} P(X=1) &= P(\ (SFF) \text{ or } (FSF) \text{ or } (FFS)\ ) \\ &= P(SFF) + P(FSF) + P(FFS) \end{aligned}$$

To calculate $P(X = 1)$, we need to calculate $P(SFF), P(FSF)$ and $P(FFS)$.

Intuition of the binomial distribution formula

- First, look at $P(SFF)$.

$$\begin{aligned} P(SFF \text{ occurs}) &= P(trial_1 = S \text{ and } trial_2 = F \text{ and } trial_3 = F) \\ &= P(trial_1 = S) \cdot P(trial_2 = F) \cdot (trial_3 = F) \\ &= (0.4)(0.6)(0.6) \\ &= (0.4)(0.6)^2 \end{aligned}$$

- We can immediately notice that $P(SFF) = P(FSF) = P(FFS)$.
- Thus, we have

$$\begin{aligned} P(X = 1) &= Prob(SFF) + Prob(FSF) + Prob(FFS) \\ &= 3 \times (0.4)(0.6)^2 \end{aligned}$$

So, $(0.4)(0.6)^2$ is the probability of a particular sequence of trial outcomes with 1 success in 3 trials.

Exercise 6

Consider a binomial experiment with $n = 10$ and $p = 0.10$.

a. Compute $f(0)$

b. Compute $f(2)$

c. Compute $P(X \leq 2)$.

d. Compute $P(X \geq 1)$.

Expected Value and Variance for the Binomial Distribution

- Recall that $E(X) = \sum xf(x)$, where x is a possible value of X.
- Also Recall that $Var(X) = \sum (x-\mu)^2 f(x)$, where x is a possible value of X.
- We can use this to calculate the True Mean and Variance of X when X follows a binomial distribution with (n, p).

Expected Value for the Binomial Distribution

- We know $f(x)$ for the binomial distribution with (n, p).

$$f(x) = C_x^n \, p^x \, (1-p)^{n-x},$$

Thus, the True Mean can be calculated as follows.

$$\begin{aligned} E(X) &= \sum x \times C_x^n \, p^x \, (1-p)^{n-x} \\ &= 1 \times C_1^n \, p^1 \, (1-p)^{n-1} + 2 \times C_2^n \, p^2 \, (1-p)^{n-2} + \\ \cdots \; & +n \times C_n^n \, p^n \, (1-p)^{n-n} \end{aligned}$$

The above complicated expression can be simplified and we have

$$E(X) = np$$

Variance for the Binomial Distribution

The variance of a binomial random variable also can be calculated as follows.

$$\begin{aligned} Var(X) &= \sum (x-np)^2 \times C_x^n \, p^x \, (1-p)^{n-x} \\ &= (1-np)^2 \times C_1^n \, p^1 \, (1-p)^{n-1} + (2-np)^2 \times C_2^n \, p^2 \, (1-p)^{n-2} \\ &\quad + \cdots + (n-np)^2 \times C_n^n \, p^n \, (1-p)^{n-n} \end{aligned}$$

The above complicated expression can be also simplified and we have

$$Var(X) = np(1-p)$$

Expected Value and Variance for the Binomial Distribution

Expected Value and Variance for the Binomial Distribution

$$E(X) = np$$

$$Var(X) = np(1 - p)$$

- Every Binomial Distribution is characterized by two numbers, the total number of trial, n, and the probability of *Success* in each trial p.
- Given n and p, we can easily calculate the mean and variance of X that follows the binomial distribution.

Exercise 6 cont'd

Consider a binomial experiment with $n = 10$ and $p = 0.10$.

e. Compute $E(X)$

f. Compute $Var(X)$.

Poisson Distribution

- Sometimes, we are interested in the number of occurrences of an event within a time period or a space.
- That is, we may be interested in
  - e.g. the number of cars that arrived at a car wash place in one hour.
  - e.g. the number of leaks in 100 miles of pipeline.
  - e.g. the number of accidents in 10 miles of highway.
- The Poisson distribution is the most popular used distribution when one wants to study problems such as above.

Assumptions of the Poisson Distribution. Poisson Experiment

Poisson Experiment

1. The probability of an occurrence is the same for any two intervals of equal length.
2. The occurrence or nonoccurrence in any interval is independent of the occurrence or nonoccurrence in any other interval.

Example

- We are interested in the number of cars arriving at drive-through window of a restaurant during a 15-minute period. We somehow assume/agree that the two above requirements are satisfied in general.

The Poisson Probability Function

The Poisson Probability Function

$$f(x) = \frac{\mu^x e^{-\mu}}{x!},$$

where $f(x)$ is the probability that the number of occurrence is equal to x in an interval,
μ is the expected value (or mean) of the number of occurrences in an interval,
$e = 2.71828$.

- Sometime μ is referred to as the arrival rate.

The mean and variance of the Poisson distribution

If X follows the Poisson distribution,

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

Example

Consider the example of counting the number of cars that arrive at the drive-up teller window of a bank in 15-min. period. Suppose that an analysis of data from past shows that the average number of cars arriving in 15-min. is 10. (i.e. $\mu = 10$). Thus, the following probability function is applied for this case.

$$f(x) = \frac{10^x e^{-10}}{x!}$$

- What is $E(X)$?
- What is $Var(X)$?

Example cont'd

- What is the probability that the number of cars arriving in 15-min. period of time is equal to 5?
- The above question is the same as "what is $Prob(X = 5) = f(5)?$. Thus, the answer will be

$$Prob(X = 5) = f(5) = \frac{10^5 e^{-10}}{5!} = 0.0378$$

Example cont'd

- Sometime we may be interested in the number of cars arriving in a different time period instead of 15-min.
- We can still use the above probability function with a small change.
- Note that the average number of cars arriving in 15-min. is 10.
- Thus, the average number of cars arriving in 1-min. is $\frac{10}{15} = \frac{2}{3}$.
- Suppose we are interested in the number of cars arriving in 3-min.
- Then, the average number of cars arriving in 3-min. is $\frac{2}{3} \times 3 = 2$.

Example cont'd

- Thus, when we are interested in the number of cars arriving in 3 min. We will use 2 as our μ in the Poisson probability distribution function.

- Finally, define Y as the number of cars arriving in 3 min. Then, we have

$$f(y) = \frac{2^y e^{-2}}{y!} = 0.0378$$

- Then, the probability that the number of cars arriving in 3 min. is equal to 3 is

$$Prob(Y = 3) = f(3) = \frac{2^3 e^{-2}}{3!} = 0.541$$