Dr. Saeed A. Dobbah Alghamdi

Department of Statistics
Faculty of Sciences
Building 90, Office 26F41
King Abdulaziz University
http://saalghamdy.kau.edu.sa

Dr. Saeed's Website

Statistics for Business & Economics

David R. Anderson
Dennis J. Sweeney
Thomas A. Williams

Discrete Probability Distributions

Chapter 5

Learning Objectives

LO1 Distinguish between *discrete* and *continuous random variable*.

LO2 Identify the characteristics of a *probability distribution.*

LO3 Compute the *mean*, *variance and standard deviation* of a probability distribution.

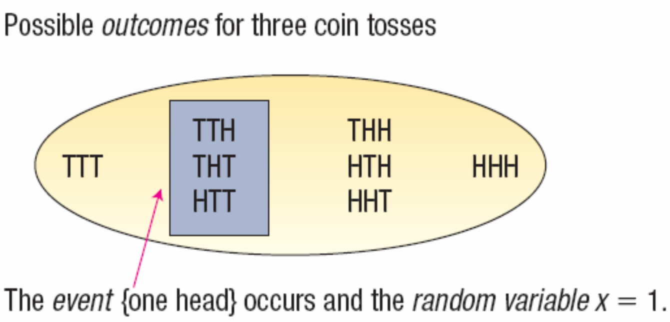
LO4 Describe and compute probabilities for a *binomial distribution*.

LO5 Describe and compute probabilities for a *Poisson distribution*.

LO6 Describe and compute probabilities for a *hypergeometric distribution*.

Random Variables

RANDOM VARIABLE is a numerical description of the outcome of an experiment



DISCRETE RANDOM VARIABLE is a random variable that can assume only certain clearly separated values. It is usually the result of <u>counting</u> something.

EXAMPLES

1. The number of students in a class.
2. The number of children in a family.
3. The number of cars entering a carwash daily.
4. Number of home mortgages approved by Coastal Federal Bank last week.

CONTINUOUS RANDOM VARIABLE is a random variable that can assume any numerical value in an interval or collection of intervals. It is usually the result of some type of measurement

EXAMPLES

1. The time between customer arrivals in minutes.
2. The weight of students in this class.
3. The Percentage of project complete after six months.
4. The amount of money earned by each of the players currently on a Major League.

What is a Probability Distribution?

PROBABILITY DISTRIBUTION is a listing of all the outcomes of an experiment and the probability associated with each outcome.

CHARACTERISTICS OF A PROBABILITY DISTRIBUTION

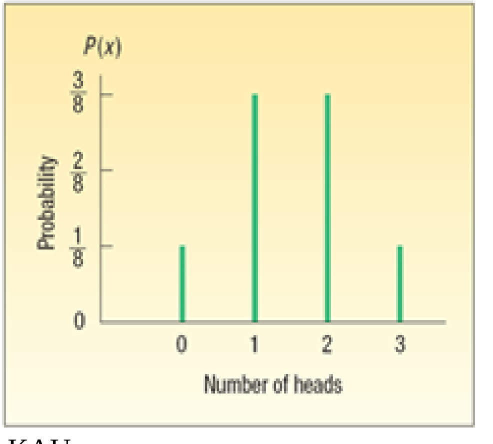
1. The probability of a particular outcome is between 0 and 1 inclusive, $0 \leq P(x) \leq 1$
2. The outcomes are mutually exclusive events.
3. The list is exhaustive. So the sum of the probabilities is equal to 1, $\sum P(x) = 1$.

Experiment

Toss a coin three times. Observe the number of heads. The possible results are:
Zero heads, One head, Two heads, and Three heads.
What is the probability distribution for the number of heads?

Possible Result	Coin Toss First	Coin Toss Second	Coin Toss Third	Number of Heads
1	T	T	T	0
2	T	T	H	1
3	T	H	T	1
4	T	H	H	2
5	H	T	T	1
6	H	T	H	2
7	H	H	T	2
8	H	H	H	3

Number of Heads, x	Probability of Outcome, $P(x)$
0	$\frac{1}{8} = .125$
1	$\frac{3}{8} = .375$
2	$\frac{3}{8} = .375$
3	$\frac{1}{8} = .125$
Total	$\frac{8}{8} = 1.000$



Discrete Probability Distribution

The Expected Value

The expected value or mean of a random variable is a measure of central location of a random variable.

The formula for the expected value of a discrete random variable *x* follows

$$E(x) = \mu = \sum x\, P(x)$$

The VARIANCE and STANDARD DEVIATION

Even though the expected value provides the mean value for the random variable, we often need a measure of variability, or dispersion. **Variance** is used to summarize the variability in the values of a random variable.

$$Var(x) = \sigma^2 = \sum (x-\mu)^2\, P(x)$$

The **standard deviation**, σ, is the positive square root of the variance

$$\sigma = \sqrt{\sigma^2}$$

The standard deviation is measured in the same units as the random variable and therefore is used in describing the variability.

Discrete Probability Distribution - Example

John Ragsdale sells new cars for Pelican Ford. John usually sells the largest number of cars on Saturday. He has developed the following probability distribution for the number of cars he expects to sell on a particular Saturday.

Find the mean, variance and standard deviation.

Number of Cars Sold, x	Probability, $P(x)$
0	.10
1	.20
2	.30
3	.30
4	.10
Total	1.00

MEAN

$$\mu = \Sigma[xP(x)]$$
$$= 0(.10) + 1(.20) + 2(.30) + 3(.30) + 4(.10)$$
$$= 2.1$$

Number of Cars Sold, x	Probability, $P(x)$	$x \cdot P(x)$
0	.10	0.00
1	.20	0.20
2	.30	0.60
3	.30	0.90
4	.10	0.40
Total	1.00	$\mu = 2.10$

VARIANCE

$$\sigma^2 = \Sigma[(x-\mu)^2 P(x)]$$

Number of Cars Sold, x	Probability, $P(x)$	$(x-\mu)$	$(x-\mu)^2$	$(x-\mu)^2P(x)$
0	.10	0 − 2.1	4.41	0.441
1	.20	1 − 2.1	1.21	0.242
2	.30	2 − 2.1	0.01	0.003
3	.30	3 − 2.1	0.81	0.243
4	.10	4 − 2.1	3.61	0.361
				$\sigma^2 = 1.290$

STANDARD DEVIATION

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.290} = 1.136$$

Binomial Probability Distribution

- **A Widely occurring discrete probability distribution**
- **Characteristics of a Binomial Probability Distribution**

1. The experiment consists of a sequence of n identical trials.
2. There are only **two possible outcomes** on a particular trial of an experiment.
3. The **probability of success remains the same** from trial to another.
4. Each trial is **independent** of any other trial

BINOMIAL PROBABILITY FORMULA $P(x) = {}_nC_x \, \pi^x (1 - \pi)^{n-x}$

where:

C denotes a combination.
n is the number of trials.
x is the random variable defined as the number of successes.
π is the probability of a success on each trial.

MEAN OF A BINOMIAL DISTRIBUTION $\mu = n\pi$

VARIANCE OF A BINOMIAL DISTRIBUTION $\sigma^2 = n\pi(1 - \pi)$

Binomial Distribution - Examples

EXAMPLE

There are **five** flights daily from Pittsburgh via US Airways into the Bradford, Pennsylvania, Regional Airport. Suppose the probability that any flight arrives late is **.20**.

What is the probability that **none** of the flights are late today? What is the average number of late flights? What is the variance of the number of late flights?

X	P(X)	cumulative probability
0	0.32768	0.32768
1	0.40960	0.73728
2	0.20480	0.94208
3	0.05120	0.99328
4	0.00640	0.99968
5	0.00032	1.00000
	1.00000	
	1.000	expected value
	0.800	variance
	0.894	standard deviation

$$
\begin{aligned}
P(x=0) &= {}_nC_x \pi^x (1-\pi)^{n-x} \\
&= {}_5C_0 (.20)^0 (1-.20)^{5-0} \\
&= (1)(1)(.3277) \\
&= 0.3277
\end{aligned}
$$

$$
\begin{aligned}
\mu &= n\pi \\
&= (5)(0.20) = 1.0
\end{aligned}
$$

$$
\begin{aligned}
\sigma^2 &= n\pi(1-\pi) \\
&= (5)(0.20)(1-0.20) \\
&= (5)(0.20)(0.80) \\
&= 0.80
\end{aligned}
$$

EXAMPLE

Five percent of the worm gears produced by an automatic, high-speed Carter-Bell milling machine are defective.

What is the probability that out of **six** gears selected at random none will be defective? Exactly one? Exactly two? Exactly three? Exactly four? Exactly five? Exactly six out of six?

X	P(X)	cumulative probability
0	0.73509	0.73509
1	0.23213	0.96723
2	0.03054	0.99777
3	0.00214	0.99991
4	0.00008	1.00000
5	0.00000	1.00000
6	0.00000	1.00000
	1.00000	
	0.300	expected value
	0.285	variance
	0.534	standard deviation

Binomial Distribution - Example

A study by the Illinois Department of Transportation concluded that 76.2% of front seat occupants used seat belts. A sample of 12 vehicles is selected.

What is the probability the front seat occupants in <u>exactly 7</u> of the 12 vehicles are wearing seat belts?

$$P(x = 7 | n = 12 \text{ and } \pi = .762)$$
$$= {}_{12}C_7(.762)^7(1 - .762)^{12-7}$$
$$= 792(.149171)(.000764) = .0902$$

What is the probability the front seat occupants in <u>at least 7</u> of the 12 vehicles are wearing seat belts?

$$P(x \geq 7 | n = 12 \text{ and } \pi = .762$$
$$= P(x = 7) + P(x = 8) + P(x = 9) + P(x = 10) + P(x = 11) + P(x = 12)$$
$$= .0902 + .1805 + .2569 + .2467 + .1436 + .0383$$
$$= .9562$$

Binomial distribution			
	12	n	
	0.762	p	
		cumulative	
X	P(X)	probability	
0	0.00000	0.00000	
1	0.00000	0.00000	
2	0.00002	0.00002	
3	0.00024	0.00026	
4	0.00172	0.00198	
5	0.00880	0.01078	
6	0.03287	0.04366	1-0.04366=0.95634
7	0.09022	0.13387	
8	0.18053	0.31441	
9	0.25689	0.57130	
10	0.24674	0.81804	
11	0.14364	0.96168	
12	0.03832	1.00000	
	1.00000		
	9.144	expected value	
	2.176	variance	
	1.475	standard deviation	

Poisson Probability Distribution

The **Poisson probability distribution** describes the number of times some event occurs during a specified interval. The interval may be time, distance, area, or volume.

Assumptions of the Poisson Distribution

(1) The probability is proportional to the length of the interval.

(2) The intervals are independent.

- The Poisson probability distribution is always positively skewed and the random variable has no specific upper limit.
- As μ becomes larger, the Poisson distribution becomes more symmetrical.

Examples include:

- The number of misspelled words per page in a newspaper.
- The number of calls per hour received a company.
- The number of vehicles sold per day at a car lot.
- The number of goals scored in a college soccer game.

POISSON DISTRIBUTION $$P(x) = \frac{\mu^x e^{-\mu}}{x!}$$

MEAN OF A POISSON DISTRIBUTION $$\mu = n\pi$$

where: μ (mu) is the mean number of occurrences (successes) in a particular interval.
e is the constant 2.71828 (base of the Napierian logarithmic system).
x is the number of occurrences (successes).
$P(x)$ is the probability for a specified value of x.

Poisson Probability Distribution - Example

Assume baggage is rarely lost by Northwest Airlines. Suppose a random sample of 1,000 flights shows a total of 300 bags were lost. Thus, the arithmetic mean number of lost bags per flight is 0.3 (300/1,000). If the number of lost bags per flight follows a Poisson distribution with ***u* = 0.3**,

find the probability of not losing any bags,

$$P(0) = \frac{\mu^{x} e^{-u}}{x!} = \frac{0.3^{0} e^{-.3}}{0!} = .7408$$

what is the probability **exactly one bag** will be lost on a particular flight?

$$P(1) = \frac{\mu^{x} e^{-\mu}}{x!} = \frac{0.3^{1} e^{-0.3}}{1!} = 0.2222$$

Poisson distribution			
0.3	mean rate of occurrence		
		cumulative	
X	P(X)	probability	
0	0.74082	0.74082	
1	0.22225	0.96306	
2	0.03334	0.99640	
3	0.00333	0.99973	
4	0.00025	0.99998	
5	0.00002	1.00000	
6	0.00000	1.00000	
7	0.00000	1.00000	
8	0.00000	1.00000	
…	….	…	

Hypergeometric Probability Distribution

Characteristics of a Hypergeometric Probability Distribution

1. An outcome on each trial of an experiment is classified into one of two mutually exclusive categories—a *success or a failure*.
2. The *probability of success and failure changes* from trial to trial.
3. The *trials are not independent*, meaning that the outcome of one trial affects the outcome of any other trial.

Note: *Use hypergeometric distribution if an experiment is binomial, but sampling is without replacement from a finite population where n/N is more than 0.05.*

$$P(x) = \frac{({}_SC_x)({}_{N-S}C_{n-x})}{{}_NC_n}$$

where:

N is the size of the population.
S is the number of successes in the population.
x is the number of successes in the sample. It may be 0, 1, 2, 3, . . .
n is the size of the sample or the number of trials.
C is the symbol for a combination.

Hypergeometric Probability Distribution

EXAMPLE

PlayTime Toys, Inc., employs 50 people in the Assembly Department. Forty of the employees belong to a union and ten do not. Five employees are selected at random to form a committee to meet with management regarding shift starting times. What is the probability that four of the five selected for the committee belong to a union?

$$P(x) = \frac{({}_SC_x)({}_{N-S}C_{n-x})}{{}_NC_n}$$

Here's what's given:

$N = 50$ (number of employees)

$S = 40$ (number of union employees)

$x = 4$ (number of union employees selected)

$n = 5$ (number of employees selected)

$$P(4) = \frac{({}_{40}C_4)({}_{50-40}C_{5-4})}{{}_{50}C_5} = \frac{\left(\frac{40!}{4!36!}\right)\left(\frac{10!}{1!9!}\right)}{\frac{50!}{5!45!}} = \frac{(91{,}390)(10)}{2{,}118{,}760} = .431$$

Hypergeometric distribution		
	50	N, population size
	40	S, number of possible occurrences
	5	n, sample size
		cumulative
X	P(X)	probability
0	0.00012	0.00012
1	0.00396	0.00408
2	0.04418	0.04826
3	0.20984	0.25810
4	0.43134	0.68944
5	0.31056	1.00000
	1.00000	
	4.000	expected value
	0.735	variance
	0.857	standard deviation

Summary

Random variable A numerical description of the outcome of an experiment.

Discrete random variable A random variable that may assume either a finite number of values or an infinite sequence of values.

Continuous random variable A random variable that may assume any numerical value in an interval or collection of intervals.

Probability distribution A description of how the probabilities are distributed over the values of the random variable.

Probability function A function, denoted by $P(x)$ or $f(x)$, that provides the probability that x assumes a particular value for a discrete random variable.

Discrete uniform probability distribution A probability distribution for which each possible value of the random variable has the same probability.

Expected value A measure of the central location of a random variable.

Variance A measure of the variability, or dispersion, of a random variable.

Standard deviation The positive square root of the variance.

Summary

Binomial probability distribution A probability distribution showing the probability of x successes in n trials of a binomial experiment.

Binomial probability function The function used to compute binomial probabilities.

Poisson probability distribution A probability distribution showing the probability of x occurrences of an event over a specified interval of time or space.

Poisson probability function The function used to compute Poisson probabilities.

Hypergeometric probability distribution A probability distribution showing the probability of x successes in n trials from a population with r successes and $N - r$ failures.

Hypergeometric probability function The function used to compute hypergeometric probabilities.