Probability and Statistics (IT302) Class No. 19

29th September 2020 Tuesday 10:30AM - 11:00AM

Binomial Distribution

An experiment often consists of repeated trials, each with two possible outcomes that may be labeled **success** or **failure**.

The most obvious application deals with the testing of items as they come off an assembly line, where each trial **may indicate a defective or a nondefective item**. Choose to define either outcome as a success. The process is referred to as a **Bernoulli process**. Each trial is called a **Bernoulli trial**.

The Bernoulli Process

Strictly speaking, the Bernoulli process must possess the following properties:

1. The experiment consists of repeated trials.
2. Each trial results in an outcome that may be classified as a success or a failure.
3. The probability of success, denoted by p, remains constant from trial to trial.
4. The repeated trials are independent.

Consider the set of Bernoulli trials where three items are selected at random from a manufacturing process, inspected, and classified as defective or non-defective. **A defective item is designated a success**. The number of successes is a random variable X assuming integral values from 0 through 3. The eight possible outcomes and the corresponding values of X are

Outcome	*NNN*	*NDN*	*NND*	*DNN*	*NDD*	*DND*	*DDN*	*DDD*
x	0	1	1	1	2	2	2	3

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9^{th} Edition, Prentice Hall.

The Bernoulli Process Contd.

Since the items are selected independently and we **assume that the process produces 25% defectives**, then $P(NDN) = P(N)P(D)P(N) =(3/4)\ (1/4)\ (3/4) =9/64$

Similar calculations yield the probabilities for the other possible outcomes. The probability distribution of X is therefore

x	0	1	2	3
$f(x)$	27/64	27/64	9/64	1/64

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

What constitutes a Bernoulli Trial?

- To be considered a **Bernoulli trial**, an experiment must meet each of three criteria:

- There must be **only 2 possible outcomes**, such as: black or red, sweet or sour. One of these outcomes is called a **success**, and the other a **failure**. Successes and Failures are denoted as S and F, though the terms given do not mean one outcome is more desirable than the other.

- Each outcome has a **fixed probability** of occurring; a success has the probability of p, and a failure has the probability of $1 - p$.

- Each experiment and result are completely **independent** of all others.

Source : http://www.math.wichita.edu/history/topics/probability.html#bern-trials Accessed on 28th September 2020

Some examples of Bernoulli Trials

- Flipping a coin. In this context, obverse ("heads") denotes success and reverse ("tails") denotes failure. A fair coin has the probability of success 0.5 by definition.

- Rolling a die, where for example we designate a six as "success" and everything else as a "failure".

- In conducting a political opinion poll, choosing a voter at random to ascertain whether that voter will vote "yes" in an upcoming referendum.

Source: http://en.wikipedia.org/wiki/Bernoulli_trial Accessed on 28th September 2020

Binomial Distribution

A Bernoulli trial can result in a success with probability p and a failure with probability $q = 1 - p$. Then the probability distribution of the binomial random variable X, the number of successes in n independent trials, is

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

Note that when $n = 3$ and $p = 1/4$, the probability distribution of X, the number of defectives, may be written as

$$b\left(x; 3, \frac{1}{4}\right) = \binom{3}{x} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{3-x}, \quad x = 0, 1, 2, 3,$$

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Example 5.1

The probability that a certain kind of component will survive a shock test is 3/4. Find the probability that exactly 2 of the next 4 components tested survive.

Solution **:** Assuming that the tests are independent and p = 3/4 for each of the 4 tests, obtain

$$b\left(2;4,\frac{3}{4}\right)=\binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2=\left(\frac{4!}{2!\,2!}\right)\left(\frac{3^2}{4^4}\right)=\frac{27}{128}.$$

Example 5.2

The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that **(a) at least 10 survive**, **(b) from 3 to 8 survive**, and **(c) exactly 5 survive?**

Solution **:** Let X be the number of people who survive.

(a)
$$P(X \geq 10) = 1 - P(X < 10) = 1 - \sum_{x=0}^{9} b(x; 15, 0.4) = 1 - 0.9662 = 0.0338$$

(b)
$$P(3 \leq X \leq 8) = \sum_{x=3}^{8} b(x; 15, 0.4) = \sum_{x=0}^{8} b(x; 15, 0.4) - \sum_{x=0}^{2} b(x; 15, 0.4) = 0.9050 - 0.0271 = 0.8779$$

(c)
$$P(X = 5) = b(5; 15, 0.4) = \sum_{x=0}^{5} b(x; 15, 0.4) - \sum_{x=0}^{4} b(x; 15, 0.4) = 0.4032 - 0.2173 = 0.1859$$

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Example 5.3

A large chain retailer purchases a certain kind of electronic device from a manufacturer. The manufacturer indicates that the defective rate of the device is 3%.

a) The inspector randomly picks 20 items from a shipment. What is the probability that there will be at least one defective item among these 20?

a) Suppose that the retailer receives 10 shipments in a month and the inspector randomly tests 20 devices per shipment. What is the probability that there will be exactly 3 shipments each containing at least one defective device among the 20 that are selected and tested from the shipment?

***Solution* :** (a) Denote by X the number of defective devices among the 20. Then X follows a $b(x; 20, 0.03)$ distribution. Hence,

$$P(X \geq 1) = 1 - P(X = 0) = 1 - b(0; 20, 0.03)$$
$$= 1 - (0.03)^0(1 - 0.03)^{20-0} = 0.4562.$$

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Example 5.3 Contd.

Solution (b)

In this case, each shipment can either contain at least one defective item or not. Hence, testing of each shipment can be viewed as a Bernoulli trial with $p = 0.4562$ from part (a). Assuming independence from shipment to shipment and denoting by Y the number of shipments containing at least one defective item, Y follows another binomial distribution $b(y; 10, 0.4562)$. Therefore,

$$P(Y = 3) = \binom{10}{3} 0.4562^3 (1 - 0.4562)^7 = 0.1602.$$

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Example 5.4

It is conjectured that an impurity exists in 30% of all drinking wells in a certain rural community. In order to gain some insight into the true extent of the problem, it is determined that some testing is necessary. It is too expensive to test all of the wells in the area, so 10 are randomly selected for testing.

a) Using the binomial distribution, what is the probability that exactly 3 wells have the impurity, assuming that the conjecture is correct?
b) What is the probability that more than 3 wells are impure?

Solution: (a) We require

$$b(3; 10, 0.3) = \sum_{x=0}^{3} b(x; 10, 0.3) - \sum_{x=0}^{2} b(x; 10, 0.3) = 0.6496 - 0.3828 = 0.2668.$$

(b) In this case, $P(X > 3) = 1 - 0.6496 = 0.3504$.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Example 5.5

Find the mean and variance of the binomial random variable of Example 5.2, and then use Chebyshev's theorem (on page 137) to interpret the interval $\mu \pm 2\sigma$.

Solution: Since Example 5.2 was a binomial experiment with $n = 15$ and $p = 0.4$, by Theorem 5.1, we have

$$\mu = (15)(0.4) = 6 \text{ and } \sigma^2 = (15)(0.4)(0.6) = 3.6.$$

Taking the square root of 3.6, we find that $\sigma = 1.897$. Hence, the required interval is $6 \pm (2)(1.897)$, or from 2.206 to 9.794. Chebyshev's theorem states that the number of recoveries among 15 patients who contracted the disease has a probability of at least 3/4 of falling between 2.206 and 9.794 or, because the data are discrete, between 2 and 10 inclusive.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Exercise 5.2

Twelve people are given two identical speakers, which they are asked to listen to for differences, if any. Suppose that these people answer simply by guessing. Find the probability that three people claim to have heard a difference between the two speakers.

Solution : Binomial distribution with $n = 12$ and $p = 0.5$.

Hence, $P(X = 3) = P(X \leq 3) - P(X \leq 2) = 0.0730 - 0.0193 = 0.0537$.

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Exercise 5.6

According to a survey by the Administrative Management Society, one-half of U.S. companies give employees 4 weeks of vacation after they have been with the company for 15 years. Find the probability that among 6 companies surveyed at random, the number that give employees 4 weeks of vacation after 15 years of employment is

a) anywhere from 2 to 5;
b) fewer than 3.

Solution : For $n = 6$ and $p = 1/2$.

(a) $P(2 \leq X \leq 5) = P(X \leq 5) - P(X \leq 1) = 0.9844 - 0.1094.$

(b) $P(X < 3) = P(X \leq 2) = 0.3438..$

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.

Exercise 5.8

According to a study published by a group of University of Massachusetts sociologists, approximately 60% of the Valium users in the state of Massachusetts first took Valium for psychological problems. Find the probability that among the next 8 users from this state who are interviewed,

(a) exactly 3 began taking Valium for psychological problems;

(b) at least 5 began taking Valium for problems that were not psychological.

Solution : For $n = 8$ and $p = 0.6$, we have

a) $P(X = 3) = b(3; 8, 0.6) = P(X \leq 3) - P(X \leq 2) = 0.1737 - 0.0498 = 0.1239.$

b) $P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.4059 = 0.5941.$

Source : Probability & Statistics for Engineers & Scientists, by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, 9th Edition, Prentice Hall.