Chapter_5 Probability Distributions (part B)

Ram C. Kafle, Ph.D.
Assistant Professor of Statistics
Sam Houston State University

Successes and Failures- Bernoulli Trials

Features of a binomial experiment

1. There is a *fixed number of trials*. We denote this number by the letter n.
2. The n trials are *independent* and repeated under identical conditions.
3. Each trial has only *two outcomes:* success, denoted by S, and failure, denoted by F.
4. For each individual 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 $q = 1 - p$.
5. The central problem of a binomial experiment is to find the *probability of r successes out of n trials*.

If elements are sampled from a dichotomous population at random and with replacement, the conditions for Bernoulli trials are satisfied.

When the sampling is made without replacement, the condition of the independence of trials is violated. However, if the population is large and only a small fraction of it (less than 10%, as a rule of thumb) is sampled, the effect of this violation is negligible and the model of the Bernoulli trials can be taken as a good approximation.

The Binomial Experiment

A binomial experiment must satisfy the following four conditions.

1. There are *n* identical trials.
2. Each trail has only two possible outcomes.
3. The probabilities of the two outcomes remain constant.
4. The trials are independent.

Example:

(a) Five percent of all DVD players manufactured by a large electronics company are defective. Three DVD players are randomly selected from the production line of this company. The selected DVD players are inspected to determine whether each of them is defective or good. Is this experiment a binomial experiment?

(b) A box contains 20 cell phones, and two of them are defective. Three cell phones are randomly selected from this box and inspected to determine whether each of them is good or defective. Is this experiment a binomial experiment?

Example

Five percent of all DVD players manufactured by a large electronics company are defective. Three DVD players are randomly selected from the production line of this company. The selected DVD players are inspected to determine whether each of them is defective or good. Is this experiment a binomial experiment?

(a) We check whether all four conditions of the binomial probability distribution are satisfied.
1. This example consists of three identical trials.
2. Each trial has two outcomes: a DVD player is defective or a DVD player is good.
3. The probability p that a DVD player is defective is .05. The probability q that a DVD player is good is .95.
4. Each trial (DVD player) is independent.

Because all four conditions of a binomial experiment are satisfied, this is an example of a binomial experiment.

Example:

(b)

1. This example consists of three identical trials.
2. Each trial has two outcomes: good or defective.
3. The probability p is that a cell phone is good. The probability q is that a cell phone is defective. These two probabilities do not remain constant for each selection. They depend on what happened in the previous selection.
4. Because p and q do not remain constant for each selection, trials are not independent.

Given that the third and fourth conditions of a binomial experiment are not satisfied, this is not an example of a binomial experiment.

Example

Consider a Bernoulli trials with success probability p=0.3.

1. Find the probability that four trials result in all failure.

2.Given that the first four trials results in all failures, what is the conditional probability that the next four trials are all success.

3.Find the probability that the first success occurs in the fourth trials.

We have $P(S) = p = 0.3, P(F) = q = 1 - 0.3 = 0.7$.

(a) $P(FFFF) = 0.7^4 = 0.2401$

(b) Because the trials are independent, the required conditional probability is the same as the (unconditional) probability of 4 trials resulting in all successes, which is $P(SSSS) = (0.3)^4 = 0.0081$

(c) $P(FFFS) = (0.7)^3(0.3) = 0.1029$

Probability model

A **probability model** is an assumed form of the probability distribution that describes the chance behavior for a random variable X.

Probabilities are expressed in terms of relevant population quantities, called the **parameters.**

The Binomial Distribution

The Binomial Distribution

Denote

$$n = \text{a fixed number of Bernoulli trials}$$
$$p = \text{the probability of success in each trial}$$
$$X = \text{the (random) number of successes in } n \text{ trials}$$

The random variable X is called a **binomial random variable.** Its distribution is called a **binomial distribution.**

The **binomial distribution** with n trials and success probability p is described by the function

$$f(x) = P[X = x] = \binom{n}{x} p^x (1 - p)^{n-x}$$

for the possible values $x = 0, 1, \ldots, n$.

Combination Formula

$$C(n,r) = {}_{n}C_{r} = \binom{n}{r} = \frac{n!}{(n-r)!\,r!}$$

A factorial is a counting method that uses consecutive whole numbers as factors.

The factorial symbol is !

Examples

7! = 7x6x5x4x3x2x1 = **5040**

Note:

1! = 1

0!=1

C(n,n) = 1 **Ex: C(5,5)= ……..**

C(n,1) = n **Ex: C(5,1) =………**

C(n,0)= 1 **Ex: C(5,0)=…………**

Example

In each case, find the probability of x successes in n Bernoulli trials with success probability p for each trial.

(a) $x = 2 \quad n = 3 \quad p = .35$

SOLUTION

With $n = 3$, $p = 0.35$, $q = 0.65$ we obtain

$$P[X = 2] = \binom{3}{2}(0.35)^2(0.65) = 0.2389$$

(b) $x = 3 \quad n = 6 \quad p = .25$

ANSWER

SOLUTION

$n = 6$, $p = 0.25$, $q = 0.75$

$$P[X = 3] = \binom{6}{3}(0.25)^3(0.75)^3 = 0.132$$

Example

The Binomial Distribution and Genetics

According to the Mendelian theory of inherited characteristics, a cross fertilization of related species of red- and white-flowered plants produces a generation whose offspring contain 25% red-flowered plants. Suppose that a horticulturist wishes to cross 5 pairs of the cross-fertilized species. Of the resulting 5 offspring, what is the probability that:

(a) There will be no red-flowered plants?

(b) There will be 4 or more red-flowered plants?

SOLUTION

Because the trials are conducted on different parent plants, it is natural to assume that they are independent. Let the random variable X denote the number of red-flowered plants among the 5 offspring. If we identify the occurrence of a red as a success S, the Mendelian theory specifies that $P(S) = p = \frac{1}{4}$, and hence X has a binomial distribution with $n = 5$ and $p = .25$. The required probabilities are therefore

(a) $P[X = 0] = f(0) = (.75)^5 = .237$

(b)
$$\begin{aligned} P[X \geq 4] &= f(4) + f(5) = \binom{5}{4}(.25)^4(.75)^1 \\ &\quad + \binom{5}{5}(.25)^5(.75)^0 = .015 + .001 = .016 \end{aligned}$$

Example

Suppose 15% of the trees in a forest have severe leaf damage from air pollution. If 5 trees are selected at random, find the probability of:

(a) Three of the selected trees have severe leaf damage.

(b) No more than two have severe leaf damage.

SOLUTION

Identify S : severe leaf damage.
X = Number of trees with severe leaf damage in a random sample of 5 trees.

X has a binomial distribution with $n = 5$, $p = 0.15$, $q = 0.85$.

$$f(x) = \binom{5}{x}(0.15)^x(0.85)^{5-x},\ x = 0, 1, \ldots, 5$$

(a) $P[X = 3] = f(3) = \binom{5}{3}(0.15)^3(0.85)^2 = 0.024$

(b) $P[X \leq 2] = f(0) + f(1) + f(2) = 0.974$.

Example (Acceptance Sampling)

A company purchases large shipments of computer chips and uses this acceptance sampling plan: randomly select and test one hundred computer chips, then accept the whole batch if there are fewer than three defective; that is, at most two chips do not work. If a particular shipment of computer chips actually has a 2% rate of defects, what is the probability that this whole shipment is accepted?

$$P(Accept) = P(\text{at most 2 defects}) = P(0) + P(1) + P(2)$$

$$= {}_{100}C_0 (0.02)^0 (0.98)^{100} + {}_{100}C_1 (0.02)^1 (0.98)^{99} + {}_{100}C_2 (0.02)^2 (0.98)^{98}$$

$$= (0.98)^{100} + 100(0.02)^1 (0.98)^{99} + \frac{100 \times 99}{2 \times 1} (0.02)^2 (0.98)^{98}$$

$$= 0.6767$$

Example 5

At the Express House Delivery Service, providing high-quality service to customers is the top priority of the management. The company guarantees a refund of all charges if a package it is delivering does not arrive at its destination by the specified time. It is known from past data that despite all efforts, 2% of the packages mailed through this company do not arrive at their destinations within the specified time. Suppose a corporation mails 10 packages through Express House Delivery Service on a certain day.

(a) Find the probability that exactly one of these 10 packages will not arrive at its destination within the specified time.

(b) Find the probability that at most one of these 10 packages will not arrive at its destination within the specified time.

Example 5: Solution

n = total number of packages mailed = 10
$p = P$(success) = .02
$q = P$(failure) = 1 – .02 = .98

x = number of successes = 1
$n – x$ = number of failures = 10 – 1 = 9

(a)

$$P(x=1) = {}_{10}C_1 (.02)^1 (.98)^9 = \frac{10!}{1!(10-1)!}(.02)^1(.98)^9$$

$$= (10)(.02)(.83374776) = .1667$$

Thus, there is a .1667 probability that exactly one of the 10 packages mailed will not arrive at its destination within the specified time.

Example 5-11: Solution

(b) At most one of the ten packages is given by the sum of the probabilities of $x = 0$ and $x = 1$

$$P(x \leq 1) = P(x = 0) + P(x = 1)$$

$$= {}_{10}C_0 (.02)^0 (.98)^{10} + {}_{10}C_1 (.02)^1 (.98)^9$$

$$= (1)(1)(.81707281) + (10)(.02)(.83374776)$$

$$= .8171 + .1667 = .9838$$

Thus, the probability that at most one of the 10 packages mailed will not arrive at its destination within the specified time is .9838.

Probability of Success and the Shape of the Binomial Distribution

1. The binomial probability distribution is symmetric if p =.50.

2. The binomial probability distribution is skewed to the right if p is less than .50.

3. The binomial probability distribution is skewed to the left if p is greater than .50.

Mean and Variance of Binomial Distribution

The binomial distribution with n trials and success probability p has

$$\begin{aligned}\text{Mean} &= np \\ \text{Variance} &= npq \quad (\text{Recall: } q = 1 - p) \\ \text{sd} &= \sqrt{npq}\end{aligned}$$

For the binomial distribution with $n = 3$ and $p = .5$, calculate the mean and the standard deviation.

SOLUTION

Employing the formulas, we obtain

$$\begin{aligned}\text{Mean} &= n\,p = 3 \times .5 = 1.5 \\ \text{sd} &= \sqrt{n\,p\,q} = \sqrt{3 \times .5 \times .5} = \sqrt{.75} = .866\end{aligned}$$

Example:

In a 2011 *Time* magazine poll, American adults were asked, “When children today in the U.S. grow up, do you think they will be better off or worse off than people are now?” Of these adults, 52% said *worse*. Assume that this result is true for the current population of U.S. adults. A sample of 50 adults is selected. Let x be the number of adults in this sample who hold the above-mentioned opinion. Find the mean and standard deviation of the probability distribution of x.

Example 5-14: Solution

n = 50, p = .52, and q = .48

Using the formulas for the mean and standard deviation of the binomial distribution,

$$\mu = np = 50(.52) = 26$$

$$\sigma = \sqrt{npq} = \sqrt{(50)(.52)(.48)} = 3.5327$$