CHAPTER PROBLEM

Did Mendel's results from plant hybridization experiments contradict his theory?

Gregor Mendel conducted original experiments to study the genetic traits of pea plants. In 1865 he wrote "Experiments in Plant Hybridization," which was published in *Proceedings of the Natural History Society*. Mendel presented a theory that when there are two inheritable traits, one of them will be dominant and the other will be recessive. Each parent contributes one gene to an offspring and, depending on the combination of genes, that offspring could inherit the dominant trait of the recessive trait. Mendel conducted an experiment using pea plants. The pods of pea plants can be green or yellow. When one pea carrying a dominant green gene and a recessive yellow gene is crossed with another pea carrying the same green/yellow genes, the offspring can inherit any one of four combinations of genes, as shown in the table below. Because green is dominant and yellow is recessive, the offspring pod will be green if either of the two inherited genes is green. The offspring can have a yellow pod only if it inherits the yellow gene from each of the two parents. We can see from the table that when crossing two parents with the green/yellow pair of genes, we expect that $\frac{3}{4}$ of the offspring peas should have green pods. That is, P(green pod) = $\frac{3}{4}$.

When Mendel conducted his famous hybridization experiments using parent pea plants with the green/yellow combination of genes, he obtained 580 offspring. According to Mendel's theory, $\frac{3}{4}$ of the offspring should have green pods, but the actual number of plants with green pods was 428. So the proportion of offspring with green pods to the total number of offspring is 428/580 = 0.738. Mendel expected a proportion of $\frac{3}{4}$=0.75, but his actual result is a proportion of 0.738. In this chapter we will consider the issue of whether the experimental results contradict the theoretical results, and, in so doing, we will lay a foundation for hypothesis testing, which is introduced in Chapter 8.

Gene from Parent 1	Gene from Parent 2	Offspring Genes	Color of Offspring Pod
green	green	green/green	green
green	yellow	green/yellow	green
yellow	green	yellow/green	green
yellow	yellow	yellow/yellow	yellow

MATH 103 CHAPTER 1 HOMEWORK
5.2 1, 3, 4, 5-16, 19, 20, 21, 23, 26, 27, 30
5.3 1, 2, 4, 5-12, 14, 15-20, 21, 24 25-28, 30, 31, 36, 39, 42
5.4 1, 2, 5-8, 10, 11, 14, 15, 16, 19

5.1 REVIEW AND PREVIEW

In this chapter we combine the methods of ___________________________ ______________________________ described in Chapters 2 and 3 and those of ______________________________ described in Chapter 4 to describe and analyze ________________________________ distributions. Probability distributions describe what will ___________________________ happen instead of what ____________________________ did happen, and they are often given in the form of a _____________________________________, ______________________________, or ____________________________.

Recall that in Chapter 2 we used ___________________________ sample ______________________ to construct ______________________________ distributions. In this chapter we use the possible _______________________ of a _________________________ (determined using methods from Chapter 4) along with the ______________________________ relative frequencies to construct ___________________________ _________________________,

which serve as models of ________________________________ perfect

frequency distributions. With this knowledge of population outcomes, we are

able to find important characteristics, such as the ___________________

and ____________________________ _________________________, and

to compare ___________________________ probabilities to ___________

results in order to determine whether outcomes are __________________.

In order to fully understand probability distributions, we must first

understand the concept of a ___________________________ variable, and

be able to distinguish between ________________________ and

__________________________ random variables. In this chapter we focus

on ____________________________ probability distributions. In

particular, we discuss _____________________ and

____________________________ probability distributions.

5.2 RANDOM VARIABLES

Key Concept…

In this section, we consider the concept of ________________________

________________________________ and how they relate to

________________________________ ________________________________.

We also discuss how to ____________________________ between

________________________ random variables and

________________________ random variables. In addition, we develop

formulas for finding the _____________________, _________________,

and ____________________________ ______________________________

for a ____________________________ ______________________________.

Most importantly, we focus on determining whether the outcomes are likely

to occur by _______________________ or they are ___________________ (in the sense that they are not likely to occur by chance).

DEFINITION

A **<u>random variable</u>** is a ___________________________ (typically represented by _____) that has a ______________________ ________________________ value, determined by _______________________, for each ________________________ of a ______________________________.

DEFINITION

A **probability distribution** is a ____________________________ that gives the

_____________________________ for each value of the ____________________

______________________________. It is often expressed in the format of a

__________________________, ____________________________, or

___________________________.

NOTE

If a probability value is very small, such as 0.000000123, we can represent it as 0+ in a table, where 0+ indicates that the probability value is a very small positive number. Why not represent this as 0?

Recall the tree diagram we made for a couple having 3 children:

DEFINITION

A **discrete random variable** has either a ______________________________

number of __________________________ or a ____________________________

number of values, where ______________________________ refers to the fact that

there might be _______________________________ many values, but they can be

__________________________________ with a ______________________________
process, so that the number of values is 0 or 1 or 2 or 3, etc.

A **continuous random variable** has ___________________________ many values, and

those values can be associated with ________________________________ on a

______________________________ scale without ________________ or

________________________________.

Example 1: Give two examples of

a. Discrete random variables

b. Continuous random variables

GRAPHS

There are various ways to graph a ___________________________ distribution, but we will consider only the _______________________________ _______________________________. A probability histogram is similar to a relative frequency histogram, but the vertical scale shows ___________________________ instead of ___________________________ frequencies based on actual sample events.

REQUIREMENTS FOR A PROBABILITY DISTRIBUTION

1. $\sum P(x) = 1$ where x assumes all possible values. The sum of all probabilities must be ___________, but values such as 0.999 or 1.001 are acceptable because they result from ________________________ errors.
2. $0 \le P(x) \le 1$ for every individual value of x.

MEAN, VARIANCE, AND STANDARD DEVIATION

1. $\mu = \sum \left[x \cdot P(x) \right]$
2. $\sigma^2 = \sum \left[(x-\mu)^2 \cdot P(x) \right]$
3. $\sigma^2 = \sum \left[x^2 \cdot P(x) \right] - \mu^2$
4. $\sigma = \sqrt{\sum \left[x^2 \cdot P(x) \right] - \mu^2}$

ROUND-OFF RULE FOR μ, σ, and σ^2

Round results by carrying one more ______________________ place than the number of decimal places used for the ______________________ variable ________. If the values of _______ are ________________________, round to one decimal place.

IDENTIFYING UNUSUAL RESULTS WITH THE RANGE RULE OF THUMB

The range rule of thumb may be helpful in _________________________ the value of a ________________________ _______________________.

According to the ________________ __________________ of _________________, most values should lie within _________ standard deviations of the __________________; it is _______________________ for a value to differ from the mean by _____________ than ________ standard deviations.

Maximum usual value = _____ + ______

Minimum usual value = _____ - ______

IDENTIFYING UNUSUAL RESULTS WITH PROBABILITIES

x successes among *n* trials is an unusually high number of successes if the _______________________ of _____ or more ____________________ is unlikely with a probability of ____________ or ______________.

x successes among n trials is an unusually low number of successes if the

__________________________ of _____ or fewer ______________________ is

unlikely with a probability of _____________ or _______________.

RARE EVENT RULE FOR INFERENTIAL STATISTICS

If, under a given ___________________________, the probability of a particular

__________________________ event is extremely small, we conclude that the

___________________________ is probably not _________________________.

Example 2: Based on information from MRINetwork, some job applicants are required to have several interviews before a decision is made. The number of required interviews and the corresponding probabilities are: 1 (0.09); 2 (0.31); 3 (0.37); 4 (0.12); 5 (0.05); 6 (0.05).

a. Does the given information describe a probability distribution?

b. Assuming that a probability distribution is described, find its mean and standard deviation.

c. Use the range rule of thumb to identify the range of values for usual numbers of interviews.

d. Is it unusual to have a decision after just one interview. Explain.

EXPECTED VALUE

The ______________________ of a _______________________

______________________________ ____________________________ is the

theoretical mean outcome for __________________________ many trials. We

can think of that mean as the ________________________ ________________

in the sense that it is the _________________________ __________________

that we would expect to get if the trials could continue _____________________.

DEFINITION

The **expected value** of a ______________________ random variable is denoted by ________, and it represents the ____________________ ________________ of the ___________________. It is obtained by finding the value of $\Sigma\left[x\cdot P(x)\right]$.

$$E=\Sigma\left[x\cdot P(x)\right]$$

Example 3: There is a 0.9968 probability that a randomly selected 50-year old female lives through the year (based on data from the U.S. Department of Health and Human Services). A Fidelity life insurance company charges $226 for insuring that the female will live through the year. If she does not survive the year, the policy pays out $50,000 as a death benefit.

a. From the perspective of the 50-year-old female, what are the values corresponding to the two events of surviving the year and not surviving?

b. If a 50-year-old female purchases the policy, what is her expected value?

c. Can the insurance company expect to make a profit from many such policies? Why?

5.3 BINOMIAL PROBABILITY DISTRIBUTIONS

Key Concept…

In this section we focus on one particular category of ________________

________________________ ________________________:

__________________________ probability distributions. Because binomial

probability distributions involve _________________________ used with

methods of _________________________ _________________________

discussed later in this book, it is important to understand

________________________________ properties of this particular class of probability distributions. We will present a basic definition of a ______________________ probability distribution along with _____________________, and methods for finding ____________________ values. ________________________ probability distributions allow us to deal with circumstances in which the __________________________ belong to _____________ relevant ____________________________, such as acceptable/defective or survived/___________.

DEFINITION

A **binomial probability distribution** results from a procedure that meets all of the following requirements:

1. The procedure has a _______________ _________________ of trials.
2. The trials must be ___________________________.
3. Each trial must have all _____________________ classified into _________ ____________________________ (commonly referred to as ___________________ and _______________________).
4. The probability of a ______________________ remains the _____________ in all trials.

NOTATION FOR BINOMIAL PROBABILITY DISTRIBUTIONS

S and F (success and failure) denote the two possible categories of outcomes

$P(S) = p$

$P(F) = 1 - p = q$

n

x

p

q

$P(x)$

Example 1: A psychology test consists of multiple-choice questions, each having four possible answers (a, b, c, and d), one of which is correct. Assume that you guess the answers to six questions.

a. Use the multiplication rule to find the probability that the first two guesses are wrong and the last four guesses are correct.

b. Beginning with WWCCCC, make a complete list of the different possible arrangements of 2 wrong answers and 4 correct answers, then find the probability for each entry in the list.

c. Based on the preceding results, what is the probability of getting exactly 4 correct answers when 6 guesses are made?

BINOMIAL PROBABILITY FORMULA

$$P(x)=\frac{n!}{(n-x)!x!}\cdot p^x \cdot q^{n-x} \qquad \text{for } x=0,1,2,\ldots,n$$

where

n

x

p

q

The factorial symbol _____, denotes the ______________________ of

decreasing _____________________________.

Example 2: Assuming the probability of a pea having a green pod is 0.75, use the binomial probability formula to find the probability of getting exactly 4 peas with green pods when 5 offspring peas are generated.

TABLE A-1 Binomial Probabilities

		p													
n	*x*	.01	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95	.99	*x*
2	0	.980	.902	.810	.640	.490	.360	.250	.160	.090	.040	.010	.002	0+	0
	1	.020	.095	.180	.320	.420	.480	.500	.480	.420	.320	.180	.095	.020	1
	2	0+	.002	.010	.040	.090	.160	.250	.360	.490	.640	.810	.902	.980	2
3	0	.970	.857	.729	.512	.343	.216	.125	.064	.027	.008	.001	0+	0+	0
	1	.029	.135	.243	.384	.441	.432	.375	.288	.189	.096	.027	.007	0+	1
	2	0+	.007	.027	.096	.189	.288	.375	.432	.441	.384	.243	.135	.029	2
	3	0+	0+	.001	.008	.027	.064	.125	.216	.343	.512	.729	.857	.970	3
4	0	.961	.815	.656	.410	.240	.130	.062	.026	.008	.002	0+	0+	0+	0
	1	.039	.171	.292	.410	.412	.346	.250	.154	.076	.026	.004	0+	0+	1
	2	.001	.014	.049	.154	.265	.346	.375	.346	.265	.154	.049	.014	.001	2
	3	0+	0+	.004	.026	.076	.154	.250	.346	.412	.410	.292	.171	.039	3
	4	0+	0+	0+	.002	.008	.026	.062	.130	.240	.410	.656	.815	.961	4
5	0	.951	.774	.590	.328	.168	.078	.031	.010	.002	0+	0+	0+	0+	0
	1	.048	.204	.328	.410	.360	.259	.156	.077	.028	.006	0+	0+	0+	1
	2	.001	.021	.073	.205	.309	.346	.312	.230	.132	.051	.008	.001	0+	2
	3	0+	.001	.008	.051	.132	.230	.312	.346	.309	.205	.073	.021	.001	3
	4	0+	0+	0+	.006	.028	.077	.156	.259	.360	.410	.328	.204	.048	4
	5	0+	0+	0+	0+	.002	.010	.031	.078	.168	.328	.590	.774	.951	5
6	0	.941	.735	.531	.262	.118	.047	.016	.004	.001	0+	0+	0+	0+	0
	1	.057	.232	.354	.393	.303	.187	.094	.037	.010	.002	0+	0+	0+	1
	2	.001	.031	.098	.246	.324	.311	.234	.138	.060	.015	.001	0+	0+	2
	3	0+	.002	.015	.082	.185	.276	.312	.276	.185	.082	.015	.002	0+	3
	4	0+	0+	.001	.015	.060	.138	.234	.311	.324	.246	.098	.031	.001	4
	5	0+	0+	0+	.002	.010	.037	.094	.187	.303	.393	.354	.232	.057	5
	6	0+	0+	0+	0+	.001	.004	.016	.047	.118	.262	.531	.735	.941	6
7	0	.932	.698	.478	.210	.082	.028	.008	.002	0+	0+	0+	0+	0+	0
	1	.066	.257	.372	.367	.247	.131	.055	.017	.004	0+	0+	0+	0+	1
	2	.002	.041	.124	.275	.318	.261	.164	.077	.025	.004	0+	0+	0+	2
	3	0+	.004	.023	.115	.227	.290	.273	.194	.097	.029	.003	0+	0+	3
	4	0+	0+	.003	.029	.097	.194	.273	.290	.227	.115	.023	.004	0+	4
	5	0+	0+	0+	.004	.025	.077	.164	.261	.318	.275	.124	.041	.002	5
	6	0+	0+	0+	0+	.004	.017	.055	.131	.247	.367	.372	.257	.066	6
	7	0+	0+	0+	0+	0+	.002	.008	.028	.082	.210	.478	.698	.932	7
8	0	.923	.663	.430	.168	.058	.017	.004	.001	0+	0+	0+	0+	0+	0
	1	.075	.279	.383	.336	.198	.090	.031	.008	.001	0+	0+	0+	0+	1
	2	.003	.051	.149	.294	.296	.209	.109	.041	.010	.001	0+	0+	0+	2
	3	0+	.005	.033	.147	.254	.279	.219	.124	.047	.009	0+	0+	0+	3
	4	0+	0+	.005	.046	.136	.232	.273	.232	.136	.046	.005	0+	0+	4
	5	0+	0+	0+	.009	.047	.124	.219	.279	.254	.147	.033	.005	0+	5
	6	0+	0+	0+	.001	.010	.041	.109	.209	.296	.294	.149	.051	.003	6
	7	0+	0+	0+	0+	.001	.008	.031	.090	.198	.336	.383	.279	.075	7
	8	0+	0+	0+	0+	0+	.001	.004	.017	.058	.168	.430	.663	.923	8

NOTE: 0+ represents a positive probability less than 0.0005.

(continued)

5.4 MEAN, VARIANCE, AND STANDARD DEVIATION FOR THE BINOMIAL DISTRIBUTION

Key Concept…

In this section, we consider important ______________________________ of

a _____________________________ distribution, including _______________,

_____________________________, and ______________________________.

That is, given a particular binomial probability distribution, we can find its

__________________________, ____________________________, and

______________________________ ______________________________. A

_____________________________ distribution is a ______________________

type of _______________________________ _________________________

_________________________________, so we could use the formulas from

5.2. However, it is easier to use the following formulas:

Any Discrete pdf	**Binomial Distributions**
1. $\mu = \Sigma\left[x \cdot P(x)\right]$	1. $\mu = np$
2. $\sigma^2 = \Sigma\left[(x-\mu)^2 \cdot P(x)\right]$	
3. $\sigma^2 = \Sigma\left[x^2 \cdot P(x)\right] - \mu^2$	2. $\sigma^2 = npq$
4. $\sigma = \sqrt{\Sigma\left[x^2 \cdot P(x)\right] - \mu^2}$	3. $\sigma = \sqrt{npq}$

RANGE RULE OF THUMB

Maximum usual value:

Minimum usual value:

Example 1: Mars, Inc. claims that 24% of its M&M plain candies are blue. A sample of 100 M&Ms is randomly selected.

a. Find the mean and standard deviation for the numbers of blue M&Ms in such groups of 100.

b. Data Set 18 in Appendix B consists of 100 M&Ms in which 27 are blue. Is this result unusual? Does it seem that the claimed rate of 24% is wrong?

Example 2: In a study of 420,095 cell phone users in Denmark, it was found that 135 developed cancer of the brain or nervous system. If we assume that the use of cell phones has no effect on developing such cancer, then the probability of a person having such a cancer is 0.000340.

a. Assuming that cell phones have no effect on developing cancer, find the mean and standard deviation for the numbers of people in groups of 420,095 that can be expected to have cancer of the brain or nervous system.

b. Based on the results from part (a), is it unusual to find that among 420,095 people, there are 135 cases of cancer of the brain or nervous system? Why or why not?

c. What do these results suggest about the publicized concern that cell phones are a health danger because they increase the risk of cancer of the brain or nervous system?