Chapter 8 Notes Binomial and Geometric Distribution

Often times we are interested in an event that has only two outcomes. For example, we may wish to know the outcome of a free throw shot (good or missed), the sex of a newborn (boy or girl), the result of a coin toss (heads or tails) or the outcome of a criminal trial (guilty or not). If these situations meet four conditions, we say it is a ***binomial setting***.

The Binomial Setting

1. Each observation falls into just one of two categories, which are generally referred to as success or failure.
2. There is a fixed number n of observations.
3. All n observations are independent. (Knowing the outcome tells us nothing about the other observations)
4. The probability of success, called p, is the same for each observation.

For example, if you draw a card from a deck, there are four possible suits (hearts, diamonds, clubs, or spades). If we want to know about drawing a heart, we can define a "success" as drawing a heart and a "failure" as drawing anything else.

If we have a binomial setting, we refer to the probability distribution of X as a binomial distribution, B(n, p) where n is the number of observations and p is the probability of success.

Binomial formulas:
The binomial coefficient is the number of ways of arranging k successes among n observations.

$$_nC_k = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Binomial Probability:
If X has a binomial distribution with n observations and probability p of success on each observation, the possible values of X are 0, 1,2,3,…,n. If k is any one of these values, then

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

The Probability Distribution Function (pdf) assigns a probability to each value of X. On your calculator, you can use binompdf (n, p, X).
Example:
Assume an inspector is examining switches. These switches are 10% faulty. If the inspector examines 10 of the switches, what is the probability that 2 switches are faulty? $P(X = 2)$.
1. Using the formula, n = 10, k = 2, p = .10, then

$$P(X = 2) = \binom{10}{2}(.10)^2(.90)^8 = .1937102445$$

Using binompdf (10, .1, 2) = .1937102445

What is the probability that at most 2 switches are faulty?
Using the formula,
$$P(X \le 2) = P(X = 0) + P(X = 1) + (P(X = 2)$$

$$P(X = 0) = \binom{10}{0}(.10)^0(.90)^{10} + P(X = 1) = \binom{10}{1}(.10)^1(.90)^9 + P(X = 2) = \binom{10}{2}(.10)^2(.90)^8$$

$$= .3487 + .3874 + .1937 = .9298$$

Using the calculator, binomcdf (n, p, X) is the cumulative density function of the binomial distribution. So, binomcdf $(10, .1, 2) = .9298091736$.

Binomial mean and standard deviation:
If a certain widget has a 75% chance of functioning correctly, then how many should work out of a lot of 100? Logically 75 should work. In fact, the formula for the **mean** of a binomial distribution is
$$\mu = np$$
The **standard deviation** is:
$$\sigma = \sqrt{np(1-p)}$$

So here: $\mu = 100(.75) = 75$

Standard deviation: $\sigma = \sqrt{100(.75)(.25)} = 4.33$

Geometric Distribution

In case of the binomial distribution, the number of trials was predetermined. Sometimes, however, we wish to know the number of trials needed before a certain outcome occurs. For example, we wish to play until we win, or until we lose; you roll dice until you get an 11; a mechanic waits for the first plane to arrive at the airport that needs repair; a basketball player shoots until he makes it. These situations fall under the ***geometric distribution.***

The Geometric Setting
1. Each event falls into just one of two categories, which are generally referred to as a "success" or "failure."
2. The probability of success, call it p, is the same for each observation.
3. The observations are all independent.
4. The variable of interest is the number of trials required to obtain the **first** success.

Geometric Formulas:
If X has a geometric distribution with probability p of success and (1- p) of failure on each observation, the possible values of X are 1, 2, 3, … If n is any one of these values, then the probability that the first success will occur on the nth trial is

$$P(X = n) = (1 - p)^{n-1} p$$

Example: On the leeward side of the island of Oahu in the small village of Nanakuli, about 80% of the residents are of Hawaiian ancestry (The Honolulu Advertiser). Suppose you fly to Hawaii and visit Nanakuli. What is the probability that the first villager you meet is Hawaiian? What is the probability that you do not meet a Hawaiian until the third villager?

Using the formula, $P(X = 1) = (1 - .8)^{1-1}(.80) = (.2)^0(.80) = .80$. That should be logical.

$P(X = 3) = (.2)^2(.80) = .032$

Using the calculator, geompdf(p,X)=geompdf(.8,1) = .8
Using the calculator, geompdf(.8, 3) = .032

What is the probability that you will meet at most three people to produce an Hawaiian person?
Using the formula, $P(X \leq 3) = P(X = 1) + PX = 2) + P(X = 3)$

$$= (.2)^0(.8) + (.2)^1(.8) + (.2)^2(.8)$$
$$= .8 + .16 + .032 = .992$$

Using your calculator, geomcdf (p,X) = geomcdf(.8,3) = .992

How many people should you expect to meet before you meet the first Hawaiian?

Geometric Distribution Mean:
If X is a geometric random variable with probability p of success on each trial, the

expected number of trials necessary to reach the first success is $\mu = \dfrac{1}{p}$.

Using the formula, $\mu = \dfrac{1}{.80} = 1.25$.

What is the probability that it takes more than three people before you meet an Hawaiian?

Furthermore, it can be shown that the probability that it takes more than n trials before we see the first success is $P(X > n) = (1 - p)^n$.

Using the formula, $P(X > 3) = (1 - .8)^3 = .008$

AP Statistics Chapter 8 Problems

1. The Los Angeles Times (Dec. 13, 1992) reported that 80% of airline passengers prefer to sleep on long flights rather than watch movies, read, etc. Consider randomly selecting 25 passengers from a particular long flight. Define a random variable X and answer the following questions.

a) What is the probability that exactly 12 of those selected passengers prefer sleeping?
b) What is the probability that all passengers selected prefer to sleep on the long flight?
c) What is the probability that at least 20 passengers refer sleeping on the long flight?
d) What is the average number of passengers that prefer sleeping on the long flight and calculate and interpret the standard deviation of X.

2. Sophie is a dog who loves to play catch. Unfortunately, she is not very good, and the probability that she catches a ball is 10%. Define X to be the number of tosses until Sophie catches the ball.

a) What is the probability that Sophie will catch the ball on her second try?
b) What is the probability that it will take more than three tosses for Sophie to catch her first ball?
c) What is the expected number of tosses before Sophie will catch her first ball?

3. You are to take a multiple choice exam consisting of 100 questions with five possible responses to each. Suppose you did not study and decided to guess randomly on each question. Let X = the number of correct answers.

a) What is your expected score on the exam?
b) Compute and interpret the standard deviation of X.
c) If 60% is passing, what is the probability you will pass the exam by guessing?

4. Suppose 15% of the cereal boxes contain a prize. You are determined to buy cereal boxes until you win a prize.

a) What is the probability you will have to buy at most 2 boxes?
b) What is the probability you will have to buy exactly 4 boxes?
c) What is the probability you will have to buy more than 4 boxes?
d) What is the average number of boxes you will need to buy before you get a prize?