Sampling Distributions (Page 331-343, Chapter 13)

TODAY YOU WILL BE ABLE TO...

- Describe a binomial experiment (the binomial setting)
- Describe binomial distributions and its parameters
- Calculate binomial probabilities
- Calculate and interpret the binomial mean and standard deviation
- Calculate the Normal approximation to binomial distributions

Like the Normal Distribution, the binomial distribution has a pattern that allows us to make inferences about populations that we know follow this distribution.

RECALL

- A **random variable** is a rule that assigns one (and only one) numerical value to each simple event of an experiment.

  Consider tossing a fair coin 3 times. What is the sample space (possible outcomes)?

  S={TTT, HTT, THT, TTH, THH, HTH, HHT, HHH}

  Define the random variable X to be the number of heads obtained.

X=no. of heads	0	1	2	3
Outcomes	TTT	HTT, THT, TTH	THH, HTH, HHT	HHH

- The **probability distribution** of a random variable gives the values of the random variable and their probabilities. Note that P(heads) = 0.5 and P(tails) = 0.5 so each outcome is equally likely...

X=no. of heads	0	1	2	3
Probability	1/8	3/8	3/8	1/8

- P(A or B) = P(A) + P(B) when events A and B are disjoint.

- P(A and B) = P(A) × P(B) when events A and B are independent.

THE BINOMIAL EXPERIMENT

A **binomial experiment** occurs when all of the following criteria are met. The experiment of tossing a coin three times meets the binomial criteria.

1. There are a **fixed number of observations** denoted by n.
   - How many observations were in the previous coin tossing experiment? $n=3$
2. The n observations are all **independent**, i.e., knowing the result of one observation does not change the probabilities we assign to other observations.
   - Is tossing a coin independent of other coin tosses? Yes
3. The various combinations of possible outcomes are disjoint.
   - Is one combination disjoint of another? Yes Ask: Can you obtain HHH and HTH for the same set of three coin tosses?
4. Each observation has ONLY **two possible outcomes**.
   - What are the possible outcomes of tossing a coin? How many? H, T 2 possible outcomes

   When counting a particular outcome, such as the number of heads, the outcome we are counting is called a "**success**" while the other outcome is called a "**failure.**"
5. The probability of **success**, call it ***p***, is the same for each observation. The probability of **failure**, for each observation, is ***1-p***.
   - Is the probability of tossing heads the same each time you toss the coin?

     Yes P(heads) can be regarded as the same each time you toss the same coin.

THE BINOMIAL DISTRIBUTION

The parameter **p** is called the **binomial (or population) proportion**. We say that the random variable X has a binomial distribution with parameters n and p.

In very many repetitions of binomial experiments…

- the average count of successes, i.e., the mean of a binomial random variable, is $\mu = np$, and
- the variance is $\sigma^2 = np(1-p)$ and standard deviation is $\sigma = \sqrt{np(1-p)}$.

BINOMIAL PROBABILITIES

Define the random variable X to be the number of heads obtained when tossing a coin three times. What is the probability of getting **two** heads?

With heads being the outcome of interest, the "success", define P(heads) to be 0.51.
$p = 0.51$
$1-p = 0.49$

Always toss starting with heads up to guarantee P(heads) is the same for each toss.

X = no. of heads	0	1	2	3
Probability			0.3822	

Using the probability formula for independent events…

$P(X = 2)$ is

P(**heads**) × P(**heads**) × P(tails) = 0.51 × 0.51 × 0.49 = 0.1274

OR

P(tails) × P(**heads**) × P(**heads**) = 0.49 × 0.51 × 0.51 = 0.1274

OR

P(**heads**) × P(tails) × P(**heads**) = 0.51 × 0.49 × 0.51 = 0.1274

NOTICE…

$P(X = 2) =$

$[\mathbf{(0.51)^2 \times (0.49)^1}] + [\mathbf{(0.51)^2 \times (0.49)^1}] + [\mathbf{(0.51)^2 \times (0.49)^1}] =$

0.1274 + 0.1274 + 0.1274

NOTICE ALSO…

$3 \times [\mathbf{(0.51)^2 \times (0.49)^1}] =$ 0.3822

********** "3" is the number of ways of arranging two heads! **********

What happens when n is large?

The number of ways of arranging k successes among n observations is given by the **binomial coefficient**, denoted $\binom{n}{x}$ and pronounced "n choose x,"

$$\binom{n}{x} = \frac{n!}{x!\,(n - x!)}$$

for k = 0, 1, 2, 3, …, n.

$\binom{n}{x}$ is NOT a fraction!

$\binom{n}{x}$ can also be written ${}_nC_x$, and this is how the binomial coefficient is represented on your calculator!

Example 1

On average, a basketball player makes 80% of her free throw shots. Consider the player taking 5 free throw shots and assume that outcome of each shot is independent of the outcome of any of the other shots.

a. Examine the four conditions for a binomial experiment.
   - The experiment consists of n = 5 trials (free throw attempts).
   - The player will either make (success) or miss (failure) each shot.
   - The probability of making a shot on each attempt is approximately p = 0.8.
   - The outcome of each shot is independent of the outcome of any of the other shots.

b. Define the binomial random variable.

Let X be the number of free throw shots made out of 5 shots.

c. Specify (in words) the distribution of X.

X has a binomial distribution with parameters n=5 and p=0.8

d. Write the formula for the probability distribution function of X.

$p(x) = P(X=x) = {}_nC_x\,(0.8)^x\,(0.2)^{5-x}$ for $x = 0, 1, 2, 3, 4, 5$

e. Write the probability distribution in tabular form.

PDF of X	
x	p(x)
0	0.00032
1	0.00640
2	0.05120
3	0.20480
4	0.40960
5	0.32768
Total	1.0000

← $p(0) = P(X = 0) = 1 \cdot (0.80)^0 \cdot (0.10)^5$

← $p(1) = P(X = 1) = 5 \cdot (0.80)^1 \cdot (0.10)^4$

← $p(2) = P(X = 2) = 10 \cdot (0.80)^2 \cdot (0.10)^3$

← $p(3) = P(X = 3) = 10 \cdot (0.80)^3 \cdot (0.10)^2$

← $p(4) = P(X = 4) = 5 \cdot (0.80)^4 \cdot (0.10)^1$

← $p(5) = P(X = 5) = 1 \cdot (0.80)^5 \cdot (0.10)^0$

f. Find the mean and the standard deviation of X. Interpret these numbers.

$\mu = np = (5)(0.8) = 4 \qquad \sigma = \sqrt{np(1-p)} = \sqrt{(5)(0.8)(0.2)} = 0.894$

In repeated sets of 5 free throw shots, the player will make on average 4 out of the 5 shots. Typically, the number of shots made will vary from the mean of 4 by 0.894 shots.

Example 2

A player wins a certain game of chance about 32% of the time. Suppose the player plays the game 6 times. Find the probability that the player wins the game exactly twice. Begin by defining the random variable X and specifying the distribution of X.

1. Let X denote the number of wins in 6 plays of the game.
2. X has a binomial distribution with parameters $n=6$ and $p=0.32$.
3. $P(X=2) = {}_6C_2 \cdot (0.32)^2 (0.68)^4 = \boxed{0.3284}$

Example 3

A fair four-sided die is tossed 8 times. Find the probability of observing no more than 1 four. Begin by defining the random variable X and specifying the distribution of X.

1. Let X denote the number of 4s observed in 8 tosses.
2. X has a binomial distribution with parameters $n=8$ and $p=0.25$
3. $P(X \le 1) = P(X=0)$ OR $P(X=1)$

$= {}_8C_0 \cdot (0.25)^0 (0.75)^8 + {}_8C_1 \cdot (0.25)^1 (0.75)^7$

$= (1)(1)(0.75)^8 + (8)(0.25)(0.75)^7 = 0.3671$

20% chance of answering each question correctly

Example 4

A poorly prepared student is guessing at a 10-question multiple choice exam. Each question has 5 selections. Find the probability of observing at least 8 correct answers. Begin by defining the random variable X and specifying the distribution of X.

1. Let X denote the number of correct answers on the exam.
2. X has a binomial distribution with parameters $n=10$ and $p=0.2$.
3. $P(X \geq 8) = P(X=8) + P(X=9) + P(X=10)$

$$= {}_{10}C_8 (0.2)^8 (0.8)^2 + {}_{10}C_9 (0.2)^9 (0.8)^1 + {}_{10}C_{10} (0.2)^{10} (0.8)^0$$

$$\boxed{= 0.0000778}$$

Example 5

Approximately 10% of the world's population is left-handed. Consider 25 randomly selected people. Define the random variable X to be the number of lefties in the sample.

n is finite, independent, two outcomes, p(leftie) is same for all selections

a. Specify (in words) the distribution of X.

X has a binomial distribution with parameters $n = 25$ and $p = 0.10$

b. Find the probability of observing no more than 4 lefties in the sample.

$P(X \leq 4) = P(X=0)$ or $P(X=1)$ or $P(X=2)$ or $P(X=3)$ or $P(X=4)$

$$= {}_{25}C_0 (0.1)^0 (0.9)^{25} + {}_{25}C_1 (0.1)^1 (0.9)^{24} + {}_{25}C_2 (0.1)^2 (0.9)^{23} + {}_{25}C_3 (0.1)^3 (0.9)^{22} + {}_{25}C_4 (0.1)^4 (0.9)^{21}$$

b. $\boxed{= 0.902}$

c. Find the probability of observing at least 5 lefties in the sample.

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.902 = \boxed{0.098}$$

d. Find the probability of observing between 2 and 5 lefties in the sample.

$P(2 \leq X \leq 5) = P(X=2)$ or $P(X=3)$ or $P(X=4)$ or $P(X=5)$

$$= 0.2659 + 0.2265 + 0.1384 + 0.0646$$

$$= \boxed{0.6954}$$

e. Find the mean and standard deviation of X.

$\mu = np = 25 \cdot (0.1) = 2.5$

$\sigma = \sqrt{25 \cdot 0.1 \cdot 0.9}$ $\boxed{= 1.5}$

$P(X=0) = 0.0718$

$P(X=1) = 0.1994$

$P(X=2) = 0.2659$

$P(X=3) = 0.2265$

$P(X=4) = 0.1384$

$P(X=5) = {}_{25}C_5 (0.1)^5 (0.9)^{20} = 0.0646$

NORMAL APPROXIMATION TO BINOMIAL DISTRIBUTIONS

Suppose that a count X has the binomial distribution with n observations and success probability p.

When n is large, the distribution of X is approximately Normal with mean $\mu = np$ and standard deviation $\sigma = \sqrt{np(1-p)}$. That is X has the distribution $N(np, \sqrt{np(1-p)})$.

This is nice, since we really do not want to explicitly calculate binomial probabilities when n > 100!

How large should n be? For binomial distributions, use the Normal approximation when n is so large that ...

$$np \geq 10$$

and

$$n(1-p) \geq 10$$

Example 6

If 2% of people have red hair, what is the probability that more than 15 in a random sample of 500 people have red hair?

✓ $np = 10$ ✓ $n(1-p) = 490$ Use Normal approximation

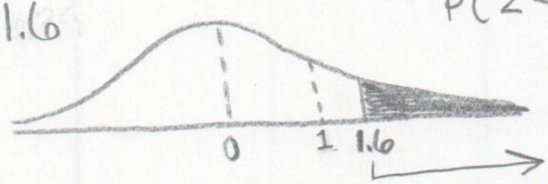
$\mu = np = 10$ $\sigma = \sqrt{np(1-p)} = 3.13$

$z = \frac{15-10}{3.13} = 1.6$

$P(z < 1.6) = 0.9452$

$P(z > 1.6) = 1 - 0.9452 = 0.0548$

$P(X > 15) \doteq \boxed{0.0548}$



Example 7

If 34% of professors at a university are women, what is the probability that fewer than 60 in a random sample of 250 professors are women?

✓ $np = 85$ ✓ $n(1-p) = 165$ Use Normal approximation

$\mu = np = 85$ $\sigma = \sqrt{np(1-p)} = 7.49$

$z = \frac{60-85}{7.49} = -3.34$

$P(z < -3.34) = 0.0004$

$P(X < 60) \doteq \boxed{0.0004}$