Statistics – Ch 6.3 Notes **Name: ______________________**

Binomial and Geometric Random Variables

In the 2005 Conference USA basketball tournament, Memphis trailed Louisville by two points. At the buzzer, Memphis's Darius Washington attempted a 3-pointer; he missed but was fouled, and went to the line for three free throws. Each made free throw is worth 1 point. Was it smart to foul?

a) Darius Washington was a 72% free-throw shooter. Find the probability that Memphis will win, lose or go to overtime.

Win	Lose	Overtime
makes all 3 ✓✓✓ $(.72)(.72)(.72)$ $= \boxed{0.373248}$	misses all 3 or makes 1 XXX + XX✓ + X✓X + ✓XX $=(.28)^3+(.28)^2(.72)+(.28)^2(.72)+(.28)^2(.72)$ oR $=(.28)^3+3(.28)^2(.72)$ $= \boxed{0.191296}$	Makes 2 ✓✓X + ✓X✓ + X✓✓ $= 3(.72)^2(.28)$ $= \boxed{0.435456}$

b) Washington is a 40% 3-point shooter. Do you think Louisville was smart to foul? Why or why not?

Maybe → The chance of losing in regulation time goes down, but they may lose in overtime.

1. *Determine whether the conditions for using a binomial random variable are met.*

If we define **X = The number of free throws made in 3 shots**, then we may call this a binomial random variable.

A **binomial setting** arises when we perform several independent trials of the same chance process and record the number of times that a particular outcome occurs. The four conditions for a binomial setting are:

BINS

- Binary The possible outcomes of each trial can be classified as "success" or "failure."
- Independent Trials must be independent; that is, knowing the result of one trial must not tell us anything about the result of any other trial.
- Number The number of trials n of the chance process must be fixed in advance.
- Success There is the same probability p of success on each trial.

The count *X* of successes in a binomial setting is a **binomial random variable.** The probability distribution of *X* is a **binomial distribution** with parameters *n* and *p*, where *n* is the number of trials of the chance process and *p* is the probability of a success on any one trial. The possible values of *X* are the whole numbers from 0 to *n*.

2. *Compute and interpret probabilities involving binomial distributions.*

a) The probability of Memphis **winning** would be: $P(X = 3)$

$$P(X=3) = 1(.72)^3(.28)^0$$

b) The probability of Memphis **Going to Overtime** would be: $P(X = 2)$

$$P(X=2) = \binom{3}{2}(.72)^2(.28)^1 = 0.435456 \xrightarrow{\text{OR}} \text{binompdf}(3, .72, 2)$$

$$= 3(.72)^2(.28)^1$$

$= 0.435456$

c) The probability of Memphis **Losing** would be: $P(X = 1 \text{ or } X = 0)$

$$P(X \le 1) = P(X=1) + P(X=0) \xrightarrow{\text{OR}} \text{binomcdf}(3, .72, 1)$$

$$= \binom{3}{1}(.72)^1(.28)^2 + (.28)^3$$

$= 0.191296$

$$= 0.191296$$

Binomial Probabilities:

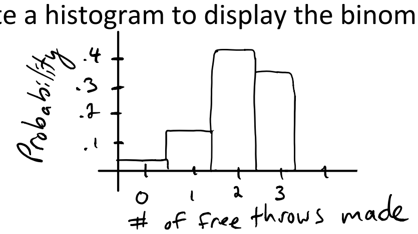
$$P(X=x) = \binom{n}{x} p^x \cdot (1-p)^{n-x}$$

\# of Successes   \# of trials   Probability of a success

Binomial Coefficient

d) Complete the probability model and create a histogram to display the binomial distribution.

X	0	1	2	3
P(X)	.022	.169	.436	.373



Independent Practice

1. Determine whether or not the given random variable has a binomial distribution. Justify your answer.

a) Genetics says that children receive genes from each of their parents independently. Each child of a particular set of parents has probability of 0.25 of having type O blood type. Suppose these parents have 5 children. Let **X = the number of children with type O blood.**

✓B - Success "Type O" and failure "Not type O"

✓I - yes, Knowing one child's blood type tells us nothing about another child's blood type

✓N - $n = 5$; 5 trials

✓S - Probability of a success is the same for each trial; $P = 0.25$

b) Shuffle a deck of cards. Turn the first 10 cards over, one at a time. Let **Y = the number of aces you observe.**

No b/c trials are not independent.

✓ ✗ ✓ ✗
B I N S

I - If the first card is an Ace, then the next card is less likely to be one.

S - $P = \frac{4}{52}$, but it is not the same for each trial

c) Shuffle a deck of cards. Turn over the top card. Put the car back in the deck, and shuffle again. Repeat this process until you get an ace. Let **W = the number of cards required.**

✓ ✓ ✗ ✓
B I N S

N - no set number of trials

2. Each child of a pair of parents has probability 0.25 of having type O blood. Genetics says that children receive genes from each of their parents independently. If these parents have 5 children, the count *X* of children with type O blood is a binomial random variable with *n* = 5 trials and probability *p* = 0.25 of a success on each trial. In this setting, a child with type O blood is a "success" (S) and a child with another blood type is a "failure" (F).

   a) Find $P(X = 2)$

   $P(X=2) = 10\,(0.25)^2(0.75)^3$

   $= 0.263$

   OR

   binompdf(5, .25, 2)

   $= 0.263$

   b) Find the probability that exactly 3 of the children have type O blood.

   $P(X=3) = \begin{bmatrix} 5 \\ 3 \end{bmatrix} (0.25)^3(.75)^2$

   $P(X=3) = 10\,(.25)^3(.75)^2$

   $= \boxed{0.08789}$

   OR

   binompdf(5, .25, 3)

   $= 0.08789$

   c) Should the parents be surprised if more than 3 of their children have type O blood? To answer this, we need to find *P*(X > 3).

   $P(X>3) = P(X=4) + P(X=5)$

   $= \begin{bmatrix} 5 \\ 4 \end{bmatrix} (0.25)^4(0.75)^1 + \begin{bmatrix} 5 \\ 5 \end{bmatrix} (0.25)^5(0.75)^0$

   $= 5\,(0.25)^4(0.75)^1 + 1\,(0.25)^5(0.75)^0$

   $= \boxed{0.01563}$

   OR

   1 – binomcdf(5, .25, 3)

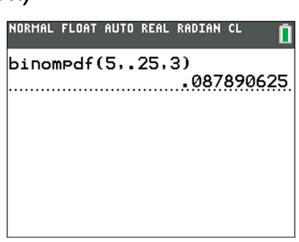
   $= 0.015625$

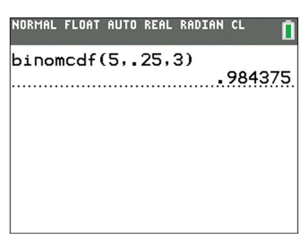
Binomial Probaility on a Calculator

- We can also use the calculator to find these probabilities.
- Binompdf(n,p,x) computes *P(X = x)*
- Binomcdf(n,p,x) computes *P(X ≤ x)*

Key Strokes

- 2^{nd} VARS
- Binompdf(
- Binomcdf(





What if we want to find *P(X > x)*?

Use the complement rule: 1 – *P(X ≤ x)*

3. Calculate the mean and standard deviation of a binomial random variable. Interpret these values in context.

4. When appropriate, use the Normal approximation to the binomial distribution to calculate probabilities.

In 2013 Skittles changed the flavor of green from lime to sour apple. Mr. McGee takes a small handful of 5 Skittles from a bin of 10,000 Skittles and gets ALL GREEN. If 20% of the Skittles are green, what is the probability of this happening?

1. Is this a binomial setting? Explain.

✓ B - success = green ; fail = not green
X I - No, sampling without replacement
✓ N - set number of trials is 5
X S - Probability of a success changes

2. Find the probability of getting 5 green Skittles if this is NOT a binomial setting. Think P(1st green AND 2nd green AND 3rd green…).

$$P(\text{all green}) = \frac{2{,}000}{10{,}000} \times \frac{1{,}999}{9{,}999} \times \frac{1{,}998}{9{,}998} \times \frac{1{,}997}{9{,}997} \times \frac{1{,}996}{9{,}996} = 0.000319$$

3. Find the probability of getting 5 green Skittles if this were a binomial setting.

$$P(X=5) = {}_5C_5\,(.20)^5(.80)^0 = 0.000320$$

4. How do your answers from #2 and #3 compare? Why do you think this is?

Almost the same. The probabilities change so little because the sample is small relative to the population.

10% Condition = Independence

When the population is much larger than the sample, a count of successes in an SRS of size n has approximately the binomial distribution with n equal to the sample size and p equal to the proportion of successes in the population.

$5(10) = 500$

10 % Condition

When taking an SRS of size *n* from a population of size *N*, we can use a binomial distribution to model the count of successes in the sample if

$$n \le \frac{1}{10}N$$

n = Sample N = Population

* $n=5$ is less than 10% of 10,000. *

Statistics – Ch 6.3 Notes **Name: ______________________**

Binomial and Geometric Random Variables

Mean and Standard Deviation of Binomial Distributions

> If a count X has the binomial distribution with number of trials n and probability of success p, the **mean** and **standard deviation** of X are
>
> $$\mu_X = np$$
> $$\sigma_X = \sqrt{np(1-p)}$$

Mr. McGee now grabs a huge handful of 100 Skittles from the large bin of 10,000 Skittles. Let ***X* = number of green Skittles in a handful of 100 Skittles**. X can be considered a binomial random variable because the 10% condition is satisfied.

1. Calculate the mean and standard deviation of *X*.

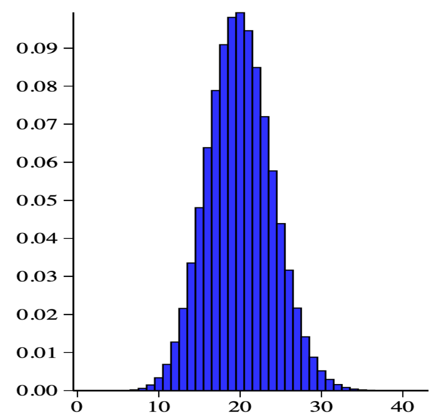
$\mu_x = 100(.20)$ $\quad\quad$ $\sigma_x = \sqrt{100(.2)(.8)}$

$\mu_x = 20$ $\quad\quad$ $\sigma_x = 4$

2. What is the probability of getting at most 11 green Skittles?

$P(x \leq 11) = \text{binomcdf}(100, .20, 11)$

$= 0.0126$



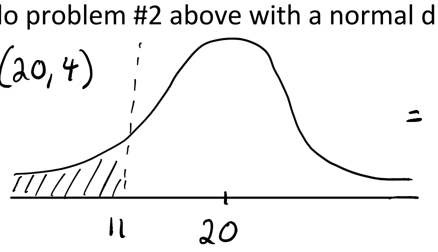
3. What do you think the shape of the distribution would be?

Normal b/c Normal (Large Counts) Condition

$100(.2) = 20$
$100(.8) = 80$ ← Both successes and failures are greater than ten.

4. Redo problem #2 above with a normal distribution and compare your answers.

$N(20, 4)$



$= \text{normalcdf}(-\infty, 11, 20, 4)$

$= 0.0122$

Normal Condition (Large Counts):

> As a rule of thumb, we will use the Normal approximation when n is so large that
>
> ***np* ≥ 10 and *n*(1 – *p*) ≥ 10.**
>
> That is, the expected number of successes and failures are both at least 10.

Statistics – Ch 6.3 Notes ***Name:*** _______________________

Binomial and Geometric Random Variables

Independent Practice:

In a survey of 500 U.S. teenagers aged 14 to 18, subjects were asked a variety of questions about personal finance. One of the questions asked whether teens had a debit card. Suppose that exactly 12% of teens aged 140 to 18 have debit cards. Let X = the number of teens in a random sample of size 500 who have a debit card.

a) Explain why X can be modeled by a binomial distribution even though the sample was selected without replacement?

$500(10) = 5{,}000$ → it is safe to assume the are more than 5,000 14 to 18 year olds in the population.

Even though trials aren't independent, 500 is less than 10% of the population.

b) Use a binomial distribution to estimate the probability that 50 or fewer teens in the sample have debit cards.

binomialcdf(500, .12, 50) = 0.093

c) Justify why X can be approximated by a Normal distribution.

$500(.12) = 60 \geq 10$

$500(.88) = 440 \geq 10$

Large Counts Condition was met.

d) Use a Normal distribution to estimate the probability that 50 or fewer teens in the sample have debit cards.

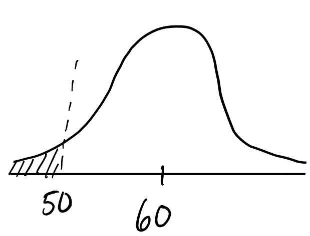
$\mu_x = 500(.12)$

$\mu_x = 60$

$\sigma_x = \sqrt{500(.12)(.88)}$

$\sigma_x = 7.27$

N(60, 7.27)



= normalcdf(-∞, 50, 60, 7.27)

= 0.08438