Binomial and Geometric Distributions - Terms and Formulas

Binomial Experiments - experiments having all four conditions:

1. Each observation falls into one of two categories – we call them "success" or "failure." (it is important to realize that success isn't necessarily a positive. If for instance, our experiment is concerned with the number of people who get colds after being doused with water on a winter day, success might be defined as getting a cold.

2. There is a fixed number of n observations.

3. The n observations are independent. Knowing the result of one observation tells you nothing about the other observations.

4. The probability of success p is the same for each observation.

Binomial Distributions - the distribution of the count X successes in the binomial experiment with parameters n and p. The possible values of X are the integers from 0 to n. We say that X is $B(n, p)$

Example 1) Tossing 20 coins and counting the number of heads.

1) Success is a heads, failure is a tails. 2. $n = 20$. 3. Independence is true – coins have no influence on each other. 4. $p = .5$. So X is $B(20, .5)$. The possible values of X are the integers from 0 to 20.

Example 2) Picking 5 cards from a standard deck and counting the number of hearts. We replace the card each time and reshuffle.

1) Success is a heart, failure is anything but a heart. 2. $n = 5$. 3. Independence is true. 4. $p = .25$. So X is $B(5, .25)$. The possible values of X are the integers from 0 to 25.

Example 3) Picking 5 cards from a standard deck and counting the number of hearts without reshuffling.

This is not binomial because of the independence issue.

Example 4) Choosing a card from a standard deck *until* you get a heart.

This is not binomial as there are not a fixed number of observations.

Example 5) It is estimated that 87% of computers users use Explorer as their default web browser. We choose 50 computer users and ask their default browser.

1) Success is Explorer, failure is anything else. 2. $n = 50$. 3. Independence seems logical. 4. $p = .87$. So X is $B(50, .87)$. The possible values of X are the integers from 0 to 50.

Example 6) A University of 10,000 students has 1,000 scholarship students. We choose 8 students and count the number of scholarship students.

1) Success is a scholarship student. 2. $n = 8$. 3. It could be argued we don't have independence, as choosing the first student as a scholarship student changes the probability of the second being a scholarship student. But the probabilities change so little that we still consider this an independent situation. 4. $p = .1$. So X is $B(8, .1)$. Possible values of X are integers from 0 to 8.

Calculating Binomial Probability

There are 2 ways to calculate binomial probabilities. One involves a formula and the second uses the calculator. However, since the formula will require use of the calculator to do the math, we will learn both.

If X has a binomial distribution with n observations with probability of p on each observation, the possible values of X are 0, 1, 2, … , n. If r is any one of these values, the probability that x equals r is given by

$P(X = r) = \binom{n}{r} p^r (1-p)^{n-r}$ where $\binom{n}{r} = \dfrac{n!}{r!(n-r)!}$ .We say that this is a combination of n things taken r at a time.

This formula looks hard but if you think about it, it makes sense. The formula is made up of 2 parts:

$p^r(1-p)^{n-r}$ is: $(\text{the probability of success})^{\text{number of desired successes}} \cdot (\text{the probability of failure})^{\text{number of desired failures}}$

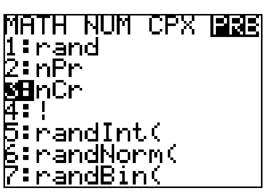
$\binom{n}{r}$ means how many combination of ways this can come about.

Example 1) You toss 5 coins. What is the probability that you get 3 heads?

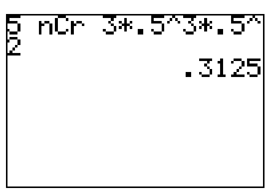
Solution: First, this is a binomial experiment $B(5, .5)$. The formula is:

$P(X = 3) = \binom{5}{3} .5^3 (1-.5)^{5-3} = \dfrac{5!}{3!(2)!} .5^3 (5)^{.2}$ … saying that we want 3 successes (heads) and 2 failures (tails).

Thankfully, we don't have to do all this computation. The calculator has a combination function. It is located in the MATH PRB menus.

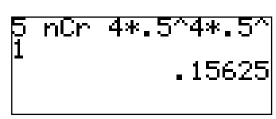


. Here is the full formula in use:



If we wish to have 4 heads instead of 3 heads, the formula changes: $P(X = 4) = \binom{5}{4} .5^4 (1-.5)^{5-4} = \dfrac{5!}{4!(1)!} .5^4 (5)^1$.

Again, this says that we want 4 successes (heads) and 1 failure (tails). The number of successes and failures must add up to the number of trials (n). The probability obviously goes down because it is harder to get 4 heads rather than 3.



There is an easier way to calculate these problems where the formula doesn't need to be known. On the AP exam, you will be asked to recognize the formula, but it is far easier to use the calculator's built-in ability to find binomial probability.

A **Binomial PDF** (Probability Density function) allows you to find the probability that X is any value in a binomial distribution. It is found in the Distribution Menu: 2nd | VARS | A : binompdf(. Its form is: **Binompdf**(n, p, X). (There are 3 important variables: n is the number of observations, p is the probability of success, and X is the number of successes you want. If you don't specify X, it will give you the probability for all values of X, from 0 to n).

So, in the preceding problem, if we want the probability of 3 heads from 5 tosses of a coin, here is the screen that gives it to us without the formula: We can easily adjust the formula to get the probability of 4 heads.

```
binompdf(5,.5,3
                .3125
binompdf(5,.5,4
                .1563
```

Example 2) Will Guess takes a true-false test of 6 questions and has absolutely no idea of any of the answers. So, true to his name, he guesses on all of them. If 4 questions correct is passing, what is the probability that he passes the exam?

Again, this is a binomial distribution with n = 6 and p = .5.

By formula:

$$P(X=4)=\binom{6}{4}.5^4(1-.5)^{6-4}=\frac{6!}{4!(2)!}.5^4(.5)^2=.234$$

$$P(X=5)=\binom{6}{5}.5^5(1-.5)^{6-5}=\frac{6!}{5!(1)!}.5^5(.5)^1=.094$$

$$P(X=6)=\binom{6}{6}.5^6(1-.5)^{6-6}=\frac{6!}{6!(0)!}.5^6(.5)^9=.016$$

The probability of passing is .234 + .094 + .016 = .344

Using the calculator is easier:

```
binompdf(6,.5,4)
+binompdf(6,.5,5
)+binompdf(6,.5,
6)
                .344
```

Example 3) Suppose the test above is now multiple choice with 4 answers per problem and again, Will Guesses. Find the probability that he passes the test and the expected number of passing students in a school of 1,500 if they all guessed.

Again, this is a binomial distribution with n = 6 and p = .25.

$$P(X=4)=\binom{6}{4}.25^4(1-.25)^{6-4}=\frac{6!}{4!(2)!}.25^4(.75)^2=.033$$

$$P(X=5)=\binom{6}{5}.25^5(1-.25)^{6-5}=\frac{6!}{5!(1)!}.25^5(.75)^1=.004$$

$$P(X=6)=\binom{6}{6}.25^6(1-.25)^{6-6}=\frac{6!}{6!(0)!}.25^6(.75)^9=.000$$

The probability of passing is .033 + .004 + .000 = .037

```
binompdf(6,.25,4
)+binompdf(6,.25
,5)+binompdf(6,.
25,6)
                .038
```

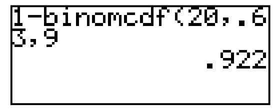
The discrepancy between the answers is round-off errors. Using the formula that Expected Value (mean number of passing students) = np, we get that 3.8% of 1,500 students or about 57 of them would pass the test by sheer guessing.

Example 4) In a particular city, 63% of the adults own their home and 37% rent. A sample of 20 adults is taken. Find the probability that the sample will have at least half home-owners.

This is binomial with $n = 20$ and $p = .63$. We want the value of $X = 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20$. That is a lot of work, even with the Binompdf formula.

To solve it, we turn to the **<u>Binomcdf</u>** formula found in the same menu. This gives the **cumulative probabilities** starting at $X = 0$. For instance, Binomcdf(20,.63,3) would give $P(X=0)+P(X=1)+P(X=2)+P(X=3)$.

In our case we can find the sum of the probabilities that $X = 0, 1, 2, 3, 4, 5, 6, 7, 8,$ and 9 and subtract that from 1. That will give us the probability that $X = 10, 11, 12, 13, 14, 15, 16, 17, 18, 19,$ or 20.



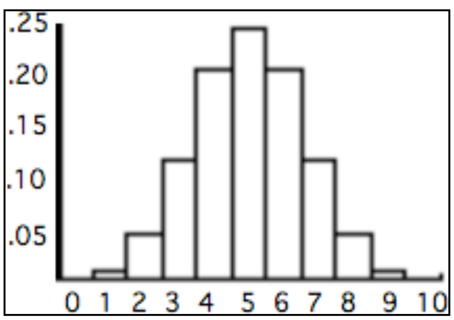
Mean and Standard Deviation of a Binomial Random Variable

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

Example 5) A basketball player is traditionally a 72% foul shooter. In a season, he takes 427 foul shots. Find the mean and standard deviation of the distribution.

Example 6) Suppose a coin is tossed 10 times. Create a probability distribution histogram.

X	0	1	2	3	4	5	6	7	8	9	10
P(X)	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.010	0.001



It should strike you that this distribution has the appearance of a normal distribution. It isn't. It is binomial. Normal distributions are for continuous data (like heights) where there are an infinite number of outcomes. Binomial distributions are for discrete data where there is only a finite number of outcomes. However, as n gets larger, a binomial distribution starts to appear more and more normal and each one is a good approximation for the other.

Geometric Experiments - experiments having all four conditions:

1. Each observation falls into one of two categories – we call them "success" or "failure".

2. The probability of success p is the same for each observation.

3. The n observations are independent. Knowing the result of one observation tells you nothing about the other observations.

4. The variable of interest in the number of trials required to obtain the *first* success.

Calculating Geometric Probabilities

If X has a geometric distribution with probability p of success and $(1-p)$ of failure, the possible values of X are 1, 2, 3, … The probability that the first success occurs on the nth trial is $P(X=n)=(1-p)^{n-1}p$.

The formula makes sense. If we want the first success on the nth trial, we need n -1 failures before our success. Since the trials are independent, we multiply the probabilities.

Example 1) You roll two dice and add them. Find the probability that we roll a 7 on the first trial, the second, the third, the 4^{th}, and the 5^{th}. Complete the chart.

Trial	1	2	3	4	5
Probability					

While not as important (because the formula is easy), the calculator has the ability to compute these probabilities. [2nd] [VARS] [E : geometpdf(] calculates geometric probabilities. The format is **geometpdf(**p, X) which will give the probability of success on the X^{th} trial.

```
geometpdf(1/6,1
             .167
geometpdf(1/6,2
             .139
geometpdf(1/6,3
             .116
```

Example 2) It is estimated that 45% of people in Fast-Food restaurants order a diet drink with their lunch. Find the probability that the fourth person orders a diet drink. Also find the probability that the first diet drinker of the day occurs before the 5^{th} person.

This last problem can also be done using the **geometcdf** function which will calculate the probability of success on or before the X^{th} trial.

```
geometcdf(.45,4
             .908
```

Finally, a couple more formulas:

The mean of a geometric random variable:

If X is a random variable with probability p on each trial, the mean (or expected value) is $\mu=\frac{1}{p}$. That means that the expected number of trials required for the first success is $\frac{1}{p}$.

The probability that its takes *more* than n trials to see the first success is $(1-p)^n$.

Example 3) In New York City at rush hour, the chance that a taxicab passes someone and is available is 15%. a) How many cabs can you expect to pass you for you to find one that is free and b) what is the probability that more than 10 cabs pass you before you find one that is free.

Example 4) A well-traveled highway has its traffic lights green for 82% of the time. If a person traveling the road goes through 8 traffic intersections, complete the chart to find a) the probability that the first red light occur on the *n*th traffic light and b) the cumulative probability that the person will hit the red light on or before the *n*th traffic light.

n	1	2	3	4	5	6	7	8
$P(x = n)$								
Cumulative Probability								

Also, find the expected number of traffic lights to hit a red light and the probability that you will go through more than 8 lights before seeing your first red light.

Finally, make a probability density histogram with the data above and a cumulative probability histogram.

Homework Problems

In each of the following situations, X is a count. Does X have a binomial distribution? Explain.

1. You observe the gender of the next 40 children born in a hospital. X is the number of boys born.

2. You decide that you will have children until you have a boy or have a maximum of 5 children. X is the number of boys born.

3. I roll 10 dice. X is the number of 6's.

4. I roll 2 dice and add them. I continue to roll until I get a 7. X is the number of 7's I get.

5. It is estimated that 43% of people sleep regularly with a nightlight in their room. I take a sample of 35 people. X is the number of people who sleep regularly with a nightlight.

6. In a classroom of 15 students, 10 of them wear glasses or contacts. I choose 6 students. X is the number of them wearing glasses or contacts.

7. In an office building of 1,500 workers, 1,000 of them wear glasses or contacts. I choose 6 workers. X is the number of them wearing glasses or contacts.

8. Only 12% of people like pineapple pizza. I choose 25 people and give pizza with pineapple. Participants are allowed to take the pineapple off the pizza if they wish. X is the number of people who like the pizza.

9. An ice hockey player scores on 5.5% of his shots. In a particular game he gets 8 shots on goal. X is the number of goals he scores in the game.

10. When a paperback book is published, the probability that it has defects is .03%. A sample of 100 books are examined. X is the number of defective books in the sample.

Problems on Binomial Distributions

For each problem, be sure that the situation fits the criteria for binomial distributions. If so, answer the questions (show the formula) and then find the mean and standard deviation of the distribution.

1) 80% of the graduates of Northeast High who apply to Penn State University are admitted. Last year, there were 6 graduates from Northeast who applied to Penn State. What is the probability that

a) 4 were admitted

b) more than 4 were admitted

Mean of distribution: ____________

Standard deviation of distribution: ____________

2) Tires from the Apex Tire Corp. are traditionally 5% defective. A truck carries 10 tires, 8 in use and 2 spares. If 10 tires are chosen from Alex, what is the probability that not more than two defective tires are chosen.

Mean of distribution: ____________

Standard deviation of distribution: ____________

3) Studies indicate that in 70% of the families of Blue Bell, both the husband and wife work. If 7 families are randomly selected from Blue Bell, what is the probability that

a) exactly 4 of them work.

b) more than 4 work

Mean of distribution: ____________

Standard deviation of distribution: ____________

4) According to the National Institute of Health, 32% of all women will suffer a hip fracture because of osteoporosis by the age of 90. If 10 women aged 90 are selected at random, find the probability that

a) 2 or more of them suffer/will suffer a hip fracture

b) none of them suffer/will suffer a hip fracture

Mean of distribution: ____________

Standard deviation of distribution: ____________

5) According to the Internal Revenue Service, the chances of your tax return being audited are 3 in 100 if your income is $60,000 or less and 8 in 100 if your income is more than $60,000. 8 tax payers are chosen

a. earning less than $60,000. Find the probability that none will be audited.

b. earning more than $60,000. Find the probability that 4 or more are audited.

Mean of distribution: ____________

Standard deviation of distribution: ____________

6) According to FBI statistics, only 52% of all rape cases are reported to the police. If 10 rape cases are randomly selected, what is the probability that at least one is reported to the police?

Mean of distribution: ____________

Standard deviation of distribution: ____________

7) In a school, typically only $\frac{1}{10}$ of the student body returns surveys. 20 students are chosen randomly to receive a survey. What is the probability that

a) they get no surveys back.

b) they get more 4 or more surveys back.

Mean of distribution: ____________

Standard deviation of distribution: ____________

8) The probability that a driver making a gas purchase will pay by credit card is $\frac{3}{5}$. If 50 cars pull up to the station to buy gas, what is the probability that at least half of the drivers will pay by credit card?

Mean of distribution: ____________

Standard deviation of distribution: ____________

9) Light bulbs work out of the box 99.6% of the time. A contractor buys 50 bulbs. What is the probability that no more than two fail?

Mean of distribution: ____________ Standard deviation of distribution: ____________

10) in Lansdale, 44% of all fire alarms are false alarms. On a certain day, there were 12 fire alarms. Find the probability that

a) none is a false alarm

b) there are at least 2 false alarms.

Mean of distribution: ____________ Standard deviation of distribution: ____________

11) An ice hockey player scores on 5.5% of his shots. In a particular game he gets 8 shots on goal. Find the probability he scores 3 or more goals.

Mean of distribution: ____________ Standard deviation of distribution: ____________

12) An insurance company sells flood insurance to 1,000 customers. Statistics show that the probability of a flood in these homes in the coming year is 2.6%. What is the probability that they will have to pay a flood claim on

a) 20 or more homes

b) 40 or more homes

Mean of distribution: Standard deviation of distribution:

Geometric Probability Questions

1) A basketball player hits 76% of his free throws. He shoots until he misses. Complete the chart to find a) the probability that the first miss will occur on the *n*th free throw and b) the cumulative probability that the player will miss on or before the *n*th free throw.

n	1	2	3	4	5	6	7	8	9	10
$P(x=n)$										
Cumulative Probability										

Also, find the expected number of free throws to miss the first free throw and the probability that the player will shoot more than 10 free throws before missing

Finally, make a probability density histogram with the data above and a cumulative probability histogram.

2) In Buffalo, NY in January, the probability that it will snow on any day is 15%. Complete the chart to find a) the probability that the first day of snow in January will be on the given day and b) the cumulative probability that the first day of snow will occur that day or before.

Day (n)	Jan 1	Jan 2	Jan 3	Jan 4	Jan 5	Jan 6	Jan 7
$P(x=n)$							
Cumulative Probability							

Also, find the expected day of the first day of snow. And the probability that Buffalo will go more than a week before seeing snow.

Finally, make a probability density histogram with the data above and a cumulative probability histogram.

AP Statistics - Random Variables, Binomial, Geometric - Practice Test

Problems 1 – 9 deal with this situation: In a popular ice cream shop, people order ice cream cones with different number of scoops. If X represents the number of scoops ordered for randomly selected customers, then the following table gives the probability distribution of X.

X	1	2	3	4	5
$P(X)$	.21	.38	.27	.11	.03

1. Find $P(X \le 2)$

2. Find $P(X < 2 \text{ or } X > 4)$ - what does this mean in words?

3. Find the mean μ for this distribution. Show how you got your answer.

4. Find the variance for this distribution. Show how you got your answer.

5. Find the standard deviation for this distribution. Show how you got your answer.

6. Suppose the number of scoops of 25 randomly chosen customers is recorded. Call this number Y. Find the mean of Y.

7. Find the variance of Y.

8. Find the standard deviation of Y.

9. Suppose the business owners wanted to keep track of its costs. Ice cream scoops cost the business 62 cents and ice cream cones cost the business 8 cents. That would involve multiplying the number of ice cream scoops by $0.62 and adding $.08. Let Z = .62X + .08.

   a. Find μ_Z

   b. Find σ_Z

Problems 10 – 13 deal with this situation: According to statistics, 22.8% of people in the U.S. are vegetarians. 18 people are chosen at random.

10. If X is a random variable for this situation, define X in words. What kind of distribution does X have?

11. Find the following probabilities: (calculators allowed)

    a. that less than 4 are vegetarians.

    b. that at least half are vegetarians.

12. Find

    a. μ_X

    b. σ_X

13. For 18 people, find the probability that the number of people who are vegetarians is within 1.5 standard deviations of its mean.

Problems 14 – 16 deal with this situation: I take a multiple choice test with 6 questions, each having choice A, B, C, D, E. Let *X* be the number of questions I get right if I guess at each answer.

14. Complete the chart (3 decimal places)

X	0	1	2	3	4	5	6
P(*X*)							

15. Find μ_X

16. Show that σ_X can be calculated in two ways.

Problem 17 – 21 deal with this situation: At a gas station, only 19% of customers purchase premium gasoline.

17. What is the probability that 4 consecutive non-premium purchases will go by before a premium customer comes?

18. How many customers does the station expect to have before it gets a premium gas purchase?

19. What is the probability that more than 10 purchases are made before there is a premium gas purchase?

20. Construct a probability distribution table (out to n = 6) for the number of cars that will come into the station for a purchase to be premium.

X	1	2	3	4	5	6
P(X)						

21. Construct a probability histogram for the table you just constructed.

Binomial and Geometric Distributions - Terms and Formulas

Binomial Experiments - experiments having all four conditions:

1. Each observation falls into one of two categories – we call them "success" or "failure." (it is important to realize that success isn't necessarily a positive. If for instance, our experiment is concerned with the number of people who get colds after being doused with water on a winter day, success might be defined as getting a cold.

2. There is a fixed number of n observations.

3. The n observations are independent. Knowing the result of one observation tells you nothing about the other observations.

4. The probability of success p is the same for each observation.

Binomial Distributions - the distribution of the count X successes in the binomial experiment with parameters n and p. The possible values of X are the integers from 0 to n. We say that X is $B(n, p)$

Example 1) Tossing 20 coins and counting the number of heads.

1) Success is a heads, failure is a tails. 2. $n = 20$. 3. Independence is true – coins have no influence on each other. 4. $p = .5$. So X is $B(20, .5)$. The possible values of X are the integers from 0 to 20.

Example 2) Picking 5 cards from a standard deck and counting the number of hearts. We replace the card each time and reshuffle.

1) Success is a heart, failure is anything but a heart. 2. $n = 5$. 3. Independence is true. 4. $p =.25$. So X is $B(5, .25)$. The possible values of X are the integers from 0 to 25.

Example 3) Picking 5 cards from a standard deck and counting the number of hearts without reshuffling.

This is not binomial because of the independence issue.

Example 4) Choosing a card from a standard deck *until* you get a heart.

This is not binomial as there are not a fixed number of observations.

Example 5) It is estimated that 87% of computers users use Explorer as their default web browser. We choose 50 computer users and ask their default browser.

1) Success is Explorer, failure is anything else. 2. $n =50$. 3. Independence seems logical.
4. $p = .87$. So X is $B(50, .87)$. The possible values of X are the integers from 0 to 50.

Example 6) A University of 10,000 students has 1,000 scholarship students. We choose 8 students and count the number of scholarship students.

1) Success is a scholarship student. 2. $n = 8$. 3. It could be argued we don't have independence, as choosing the first student as a scholarship student changes the probability of the second being a scholarship student. But the probabilities change so little that we still consider this an independent situation. 4. $p = .1$. So X is $B(8, .1)$. Possible values of X are integers from 0 to 8.

Calculating Binomial Probability

There are 2 ways to calculate binomial probabilities. One involves a formula and the second uses the calculator. However, since the formula will require use of the calculator to do the math, we will learn both.

If X has a binomial distribution with n observations with probability of p on each observation, the possible values of X are 0, 1, 2, … , n. If r is any one of these values, the probability that x equals r is given by

$P(X = r) = \binom{n}{r} p^r (1-p)^{n-r}$ where $\binom{n}{r} = \dfrac{n!}{r!(n-r)!}$ .We say that this is a combination of n things taken r at a time.

This formula looks hard but if you think about it, it makes sense. The formula is made up of 2 parts:

$p^r(1-p)^{n-r}$ is: $(\text{the probability of success})^{\text{number of desired successes}} \cdot (\text{the probability of failure})^{\text{number of desired failures}}$

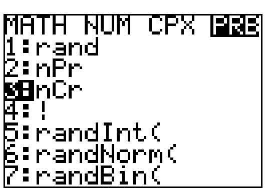
$\binom{n}{r}$ means how many combination of ways this can come about.

Example 1) You toss 5 coins. What is the probability that you get 3 heads?

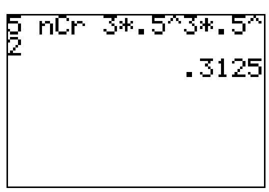
Solution: First, this is a binomial experiment $B(5, .5)$. The formula is:

$P(X = 3) = \binom{5}{3} .5^3 (1-.5)^{5-3} = \dfrac{5!}{3!(2)!} .5^3 (5)^{.2}$ … saying that we want 3 successes (heads) and 2 failures (tails).

Thankfully, we don't have to do all this computation. The calculator has a combination function. It is located in the MATH PRB menus.

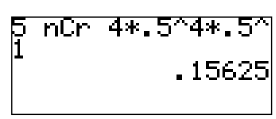


. Here is the full formula in use:



If we wish to have 4 heads instead of 3 heads, the formula changes: $P(X = 4) = \binom{5}{4} .5^4 (1-.5)^{5-4} = \dfrac{5!}{4!(1)!} .5^4 (5)^1$.

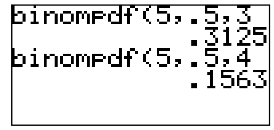
Again, this says that we want 4 successes (heads) and 1 failure (tails). The number of successes and failures must add up to the number of trials (n). The probability obviously goes down because it is harder to get 4 heads rather than 3.



There is an easier way to calculate these problems where the formula doesn't need to be known. On the AP exam, you will be asked to recognize the formula, but it is far easier to use the calculator's built-in ability to find binomial probability.

A **Binomial PDF** (Probability Density function) allows you to find the probability that X is any value in a binomial distribution. It is found in the Distribution Menu: [2nd] [VARS] [A : binompdf(] . Its form is: **Binompdf**(n, p, X). (There are 3 important variables: n is the number of observations, p is the probability of success, and X is the number of successes you want. If you don't specify X, it will give you the probability for all values of X, from 0 to n).

So, in the preceding problem, if we want the probability of 3 heads from 5 tosses of a coin, here is the screen that gives it to us without the formula: We can easily adjust the formula to get the probability of 4 heads.



Example 2) Will Guess takes a true-false test of 6 questions and has absolutely no idea of any of the answers. So, true to his name, he guesses on all of them. If 4 questions correct is passing, what is the probability that he passes the exam?

Again, this is a binomial distribution with $n = 6$ and $p = .5$.

By formula:

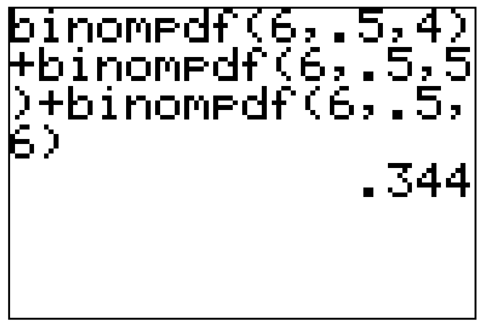
$$P(X=4)=\binom{6}{4}.5^4(1-.5)^{6-4}=\frac{6!}{4!(2)!}.5^4(.5)^2=.234$$

$$P(X=5)=\binom{6}{5}.5^5(1-.5)^{6-5}=\frac{6!}{5!(1)!}.5^5(.5)^1=.094$$

$$P(X=6)=\binom{6}{6}.5^6(1-.5)^{6-6}=\frac{6!}{6!(0)!}.5^6(.5)^9=.016$$

The probability of passing is $.234 + .094 + .016 = .344$

Using the calculator is easier:



Example 3) Suppose the test above is now multiple choice with 4 answers per problem and again, Will Guesses. Find the probability that he passes the test and the expected number of passing students in a school of 1,500 if they all guessed.

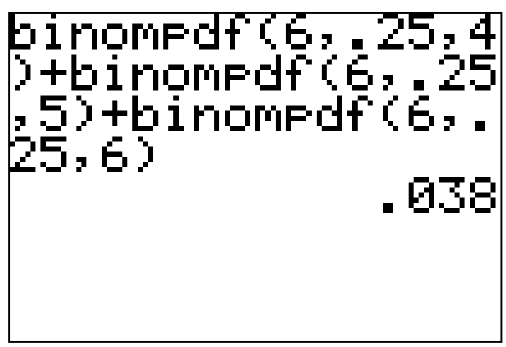
Again, this is a binomial distribution with $n = 6$ and $p = .25$.

$$P(X=4)=\binom{6}{4}.25^4(1-.25)^{6-4}=\frac{6!}{4!(2)!}.25^4(.75)^2=.033$$

$$P(X=5)=\binom{6}{5}.25^5(1-.25)^{6-5}=\frac{6!}{5!(1)!}.25^5(.75)^1=.004$$

$$P(X=6)=\binom{6}{6}.25^6(1-.25)^{6-6}=\frac{6!}{6!(0)!}.25^6(.75)^9=.000$$

The probability of passing is $.033 + .004 + .000 = .037$



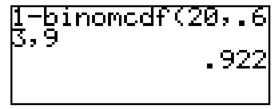
The discrepancy between the answers is round-off errors. Using the formula that Expected Value (mean number of passing students) = np, we get that 3.8% of 1,500 students or about 57 of them would pass the test by sheer guessing.

Example 4) In a particular city, 63% of the adults own their home and 37% rent. A sample of 20 adults is taken. Find the probability that the sample will have at least half home-owners.

This is binomial with $n = 20$ and $p = .63$. We want the value of $X = 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20$. That is a lot of work, even with the Binompdf formula.

To solve it, we turn to the **Binomcdf** formula found in the same menu. This gives the **cumulative probabilities** starting at $X = 0$. For instance, Binomcdf(20,.63,3) would give $P(X=0)+P(X=1)+P(X=2)+P(X=3)$.

In our case we can find the sum of the probabilities that $X = 0, 1, 2, 3, 4, 5, 6, 7, 8,$ and 9 and subtract that from 1. That will give us the probability that $X = 10, 11, 12, 13, 14, 15, 16, 17, 18, 19,$ or 20.



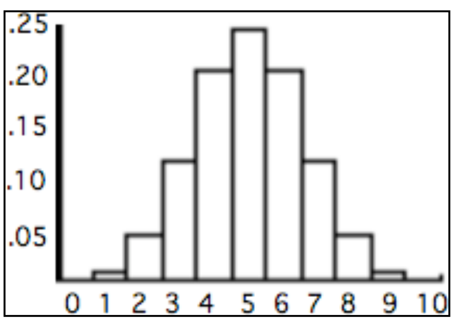
Mean and Standard Deviation of a Binomial Random Variable

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

Example 5) A basketball player is traditionally a 72% foul shooter. In a season, he takes 427 foul shots. Find the mean and standard deviation of the distribution.

Example 6) Suppose a coin is tossed 10 times. Create a probability distribution histogram.

X	0	1	2	3	4	5	6	7	8	9	10
P(X)	0.001	0.010	0.044	0.117	0.205	0.246	0.205	0.117	0.044	0.010	0.001



It should strike you that this distribution has the appearance of a normal distribution. It isn't. It is binomial. Normal distributions are for continuous data (like heights) where there are an infinite number of outcomes. Binomial distributions are for discrete data where there is only a finite number of outcomes. However, as n gets larger, a binomial distribution starts to appear more and more normal and each one is a good approximation for the other.

Geometric Experiments - experiments having all four conditions:

1. Each observation falls into one of two categories – we call them "success" or "failure".

2. The probability of success p is the same for each observation.

3. The n observations are independent. Knowing the result of one observation tells you nothing about the other observations.

4. The variable of interest in the number of trials required to obtain the *first* success.

Calculating Geometric Probabilities

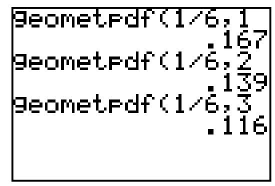
If X has a geometric distribution with probability p of success and $(1-p)$ of failure, the possible values of X are 1, 2, 3, … The probability that the first success occurs on the nth trial is $P(X=n)=(1-p)^{n-1}p$.

The formula makes sense. If we want the first success on the nth trial, we need n -1 failures before our success. Since the trials are independent, we multiply the probabilities.

Example 1) You roll two dice and add them. Find the probability that we roll a 7 on the first trial, the second, the third, the 4th, and the 5th. Complete the chart.

Trial	1	2	3	4	5
Probability	$\frac{1}{6}$	$\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)=\frac{5}{36}$	$\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)=\frac{25}{216}$	$\left(\frac{5}{6}\right)^3\left(\frac{1}{6}\right)=\frac{125}{1296}$	$\left(\frac{5}{6}\right)^4\left(\frac{1}{6}\right)=\frac{625}{7776}$

While not as important (because the formula is easy), the calculator has the ability to compute these probabilities. 2nd VARS E : geometpdf(calculates geometric probabilities. The format is **geometpdf(**p, X) which will give the probability of success on the X^{th} trial.



Example 2) It is estimated that 45% of people in Fast-Food restaurants order a diet drink with their lunch. Find the probability that the fourth person orders a diet drink. Also find the probability that the first diet drinker of the day occurs before the 5th person.

$$.45+(.55)(.45)+(.55)^2(.45)+(.55)^3(.45)=.908$$

This last problem can also be done using the **geometcdf** function which will calculate the probability of success on or before the X^{th} trial.

Finally, a couple more formulas:

The mean of a geometric random variable:

If X is a random variable with probability p on each trial, the mean (or expected value) is $\mu=\frac{1}{p}$. That means that the expected number of trials required for the first success is $\frac{1}{p}$.

The probability that its takes *more* than n trials to see the first success is $(1-p)^n$.

Example 3) In New York City at rush hour, the chance that a taxicab passes someone and is available is 15%. a) How many cabs can you expect to pass you for you to find one that is free and b) what is the probability that more than 10 cabs pass you before you find one that is free.

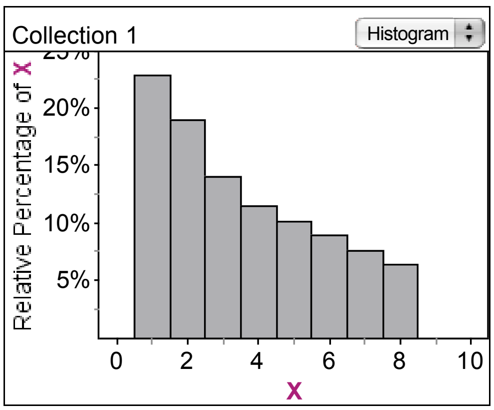
a. $\frac{1}{.15}=6.667$ b. $(1-.15)^{10}=.197$

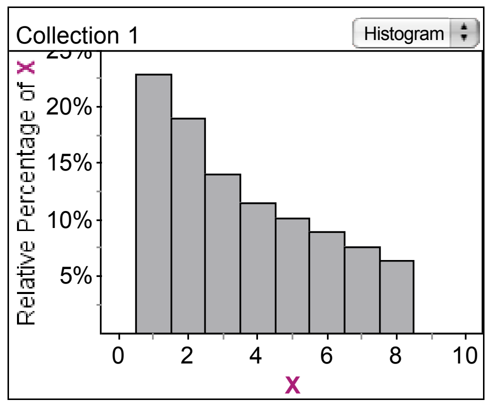
Example 4) A well-traveled highway has its traffic lights green for 82% of the time. If a person traveling the road goes through 8 traffic intersections, complete the chart to find a) the probability that the first red light occur on the *n*th traffic light and b) the cumulative probability that the person will hit the red light on or before the *n*th traffic light.

n	1	2	3	4	5	6	7	8
$P(x=n)$	.180	.148	.121	.089	.081	.067	.055	.045
Cumulative Probability	.180	.328	.449	.548	.629	.696	.751	.796

Also, find the expected number of traffic lights to hit a red light and the probability that you will go through more than 8 lights before seeing your first red light. a. $\frac{1}{.18} = 5.556$ b. $(1-.18)^8 = .204$

Finally, make a probability density histogram with the data above and a cumulative probability histogram.





Homework Problems

In each of the following situations, X is a count. Does X have a binomial distribution? Explain.

1. You observe the gender of the next 40 children born in a hospital. X is the number of boys born.

Yes. $n = 40, p = .5$. Children born are independent.

2. You decide that you will have children until you have a boy or have a maximum of 5 children. X is the number of boys born.

No. No value for n.

3. I roll 10 dice. X is the number of 6's.

No. There are more than two results (unless you consider a 6 is a success and not a 6 a failure).

4. I roll 2 dice and add them. I continue to roll until I get a 7. X is the number of 7's I get.

No. No value for n.

5. It is estimated that 43% of people sleep regularly with a nightlight in their room. I take a sample of 35 people. X is the number of people who sleep regularly with a nightlight.

Yes. $n = 35, p = .43$. Assume that the people chosen have independent sleeping habits.

6. In a classroom of 15 students, 10 of them wear glasses or contacts. I choose 6 students. X is the number of them wearing glasses or contacts.

No. Once you choose a person, the probability changes.

7. In an office building of 1,500 workers, 1,000 of them wear glasses or contacts. I choose 6 workers. X is the number of them wearing glasses or contacts.

Theoretically no. Once you choose a person, the probability changes, but so little that it acts like a binomial distribution.

8. Only 12% of people like pineapple pizza. I choose 25 people and give pizza with pineapple. Participants are allowed to take the pineapple off the pizza if they wish. X is the number of people who like the pizza.

No. The probability is not fixed.

9. An ice hockey player scores on 5.5% of his shots. In a particular game he gets 8 shots on goal. X is the number of goals he scores in the game.

Yes. $n = 8, p = .055$. Assume that his shooting any shot is independent.

10. When a paperback book is published, the probability that it has defects is .03%. A sample of 100 books are examined. X is the number of defective books in the sample.

Yes. $n = 100, p = .003$. Assume that defective books are independent.

Problems on Binomial Distributions

For each problem, be sure that the situation fits the criteria for binomial distributions. If so, answer the questions (show the formula) and then find the mean and standard deviation of the distribution.

1) 80% of the graduates of Northeast High who apply to Penn State University are admitted. Last year, there were 6 graduates from Northeast who applied to Penn State. What is the probability that

a) 4 were admitted

$$\binom{6}{4}(.8)^4(.2)^2 = \text{binompdf}(6,.8,4) = .246$$

b) more than 4 were admitted

$$\binom{6}{5}(.8)^5(.2)^1 + (.8)^6 = .655$$

Mean of distribution: $6(.8) = 4.8$

Standard deviation of distribution: $\sqrt{6(.8)(.2)} = .980$

2) Tires from the Apex Tire Corp. are traditionally 5% defective. A truck carries 10 tires, 8 in use and 2 spares. If 10 tires are chosen from Alex, what is the probability that not more than two defective tires are chosen.

$$(.05)^{10} + \binom{10}{1}(.05)^1(.95)^9 + \binom{10}{2}(.05)^2(.95)^8 = \text{binomcdf}(10,.05,2) = .988$$

Mean of distribution: $10(.05) = 0.5$

Standard deviation of distribution: $\sqrt{10(.05)(.95)} = .689$

3) Studies indicate that in 70% of the families of Blue Bell, both the husband and wife work. If 7 families are randomly selected from Blue Bell, what is the probability that

a) exactly 4 of them work.

$$\binom{7}{4}(.7)^4(.3)^3 = \text{binompdf}(7,.7,4) = .227$$

b) more than 4 work

$$\binom{7}{5}(.7)^5(.3)^2 + \binom{7}{6}(.7)^6(.3)^1 + (.7)^7 = .647$$

Mean of distribution: $7(.7) = 4.9$

Standard deviation of distribution: $\sqrt{7(.7)(.3)} = 1.212$

4) According to the National Institute of Health, 32% of all women will suffer a hip fracture because of osteoporosis by the age of 90. If 10 women aged 90 are selected at random, find the probability that

a) 2 or more of them suffer/will suffer a hip fracture

$$1 - (.68)^{10} - \binom{10}{1}(.32)^1(.68)^9 = .879$$

b) none of them suffer/will suffer a hip fracture

$$(.68)^{10} = .021$$

Mean of distribution: $10(.32) = 3.2$

Standard deviation of distribution: $\sqrt{10(.32)(.68)} = 1.475$

5) According to the Internal Revenue Service, the chances of your tax return being audited are 3 in 100 if your income is $60,000 or less and 8 in 100 if your income is more than $60,000. 8 tax payers are chosen

a. earning less than $60,000. Find the probability that none will be audited.

$\boxed{(.97)^8 = .784}$

b. earning more than $60,000. Find the probability that 4 or more are audited.

$\boxed{1 - \text{binomcdf}(8,3,.08) = .002}$

Mean of distribution: $8(.03) = 24 \quad 8(.08) = .64$

Standard deviation of distribution:
$\sqrt{8(.03)(.97)} = .482, \sqrt{8(.08)(.92)} = .767$

6) According to FBI statistics, only 52% of all rape cases are reported to the police. If 10 rape cases are randomly selected, what is the probability that at least one is reported to the police?

$$\boxed{1 - (.48)^{10} = .999}$$

Mean of distribution: $\boxed{10(.52) = 5.2}$

Standard deviation of distribution: $\boxed{\sqrt{10(.52)(.48)} = 1.580}$

7) In a school, typically only $\frac{1}{10}$ of the student body returns surveys. 20 students are chosen randomly to receive a survey. What is the probability that

a) they get no surveys back.

$\boxed{(.9)^{20} = .122}$

b) they get 4 or more surveys back.

$\boxed{1 - \text{binomialcdf}(20,.1,3) = .133}$

Mean of distribution: $\boxed{20(.1) = 2}$

Standard deviation of distribution: $\boxed{\sqrt{20(.1)(.9)} = 1.342}$

8) The probability that a driver making a gas purchase will pay by credit card is $\frac{3}{5}$. If 50 cars pull up to the station to buy gas, what is the probability that at least half of the drivers will pay by credit card?

$$\boxed{1 - \text{binomialcdf}(50,.6,24) = .943}$$

Mean of distribution: $\boxed{50(.6) = 3}$

Standard deviation of distribution: $\boxed{\sqrt{50(.6)(.4)} = 3.464}$

9) Light bulbs work out of the box 99.6% of the time. A contractor buys 50 bulbs. What is the probability that no more than two fail?

$$\boxed{(.996)^{50}+\binom{50}{1}(.004)^{1}(.996)^{49}+\binom{50}{2}(.004)^{2}(.996)^{48}=\text{binomcdf}(50,004,2)=.999}$$

Mean of distribution: $\boxed{50(.996)=49.}$ Standard deviation of distribution: $\boxed{\sqrt{50(.996)(.004)}=.446}$

10) In Lansdale, 44% of all fire alarms are false alarms. On a certain day, there were 12 fire alarms. Find the probability that

a) none is a false alarm

$\boxed{(.56)^{12}\approx 0}$

b) there are at least 2 false alarms.

$\boxed{1-\text{binomialcdf}(12,.44,1)=.990}$

Mean of distribution: $12(.44)=5.280$ Standard deviation of distribution: $\boxed{\sqrt{12(.44)(.56)}=1.720}$

11) An ice hockey player scores on 5.5% of his shots. In a particular game he gets 8 shots on goal. Find the probability he scores 3 or more goals.

$$\boxed{1-\text{binomialcdf}(8,.055,2)=.008}$$

Mean of distribution: $\boxed{8(.055)=.440}$ Standard deviation of distribution: $\boxed{\sqrt{8(.055)(.945)}=.645}$

12) An insurance company sells flood insurance to 1,000 customers. Statistics show that the probability of a flood in these homes in the coming year is 2.6%. What is the probability that they will have to pay a flood claim on

a) 20 or more homes

$\boxed{1-\text{binomialcdf}(1000,.026,19)=.906}$

b) 40 or more homes

$\boxed{1-\text{binomialcdf}(1000,.026,19)=.006}$

Mean of distribution: $\boxed{1000(.026)=26}$ Standard deviation of distribution: $\boxed{\sqrt{1000(.026)(.974)}=5.032}$

Geometric Probability Questions

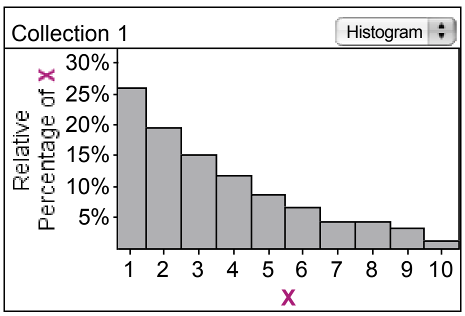
1) A basketball player hits 76% of his free throws. He shoots until he misses. Complete the chart to find a) the probability that the first miss will occur on the *n*th free throw and b) the cumulative probability that the player will miss on or before the *n*th free throw.

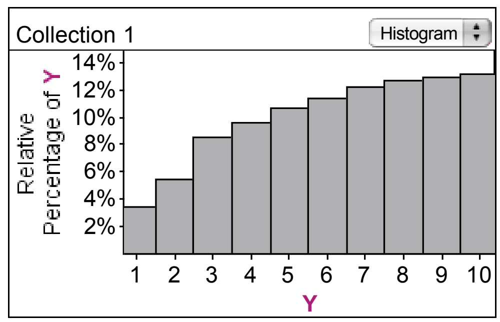
n	1	2	3	4	5	6	7	8	9	10
$P(x=n)$	.242	.182	.139	.105	.080	.061	.046	.035	.027	.020
Cumulative Probability	.240	.422	.561	.666	.746	.807	.854	.889	.915	.936

Also, find the expected number of free throws to miss the first free throw and the probability that the player will shoot more than 10 free throws before missing.

$$\frac{1}{.24} = 4.167 \qquad (.76)^{10} = .064$$

Finally, make a probability density histogram with the data above and a cumulative probability histogram.





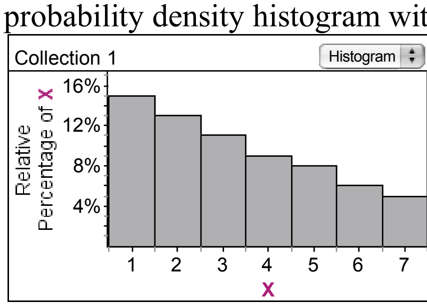
2) In Buffalo, NY in January, the probability that it will snow on any day is 15%. Complete the chart to find a) the probability that the first day of snow in January will be on the given day and b) the cumulative probability that the first day of snow will occur that day or before.

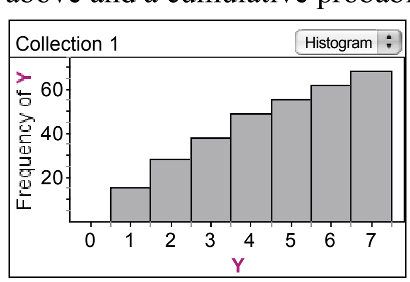
Day (n)	Jan 1	Jan 2	Jan 3	Jan 4	Jan 5	Jan 6	Jan 7
$P(x=n)$	.150	.128	.108	.092	.078	.067	.057
Cumulative Probability	.150	.278	.386	.478	.556	.623	.679

Also, find the expected day of the first day of snow. And the probability that Buffalo will go more than a week before seeing snow at the beginning of a year.

$$\frac{1}{.15} = 6.667 \ (\text{January } 7) \qquad (.85)^{7} = .321$$

Finally, make a probability density histogram with the data above and a cumulative probability histogram.





AP Statistics - Random Variables, Binomial, Geometric - Practice Test

Problems 1 – 9 deal with this situation: In a popular ice cream shop, people order ice cream cones with different number of scoops. If *X* represents the number of scoops ordered for randomly selected customers, then the following table gives the probability distribution of *X*.

X	1	2	3	4	5
P(*X*)	.21	.38	.27	.11	.03

1. Find $P(X \le 2)$

   $\boxed{.59}$

2. Find $P(X < 2 \text{ or } X > 4)$ - what does this mean in words?

   $\boxed{.24}$ - people either order very little or a lot of ice cream.

3. Find the mean μ for this distribution. Show how you got your answer.

   $\boxed{\mu = 1(.21) + 2(.38) + 3(.27) + 4(.11) + 5(.03) = 2.37}$

4. Find the variance for this distribution. Show how you got your answer.

   $\boxed{\sigma^2 = 1(.21 - 2.37)^2 + 2(.38 - 2.37)^2 + 3(.27 - 2.37)^2 + 4(.11 - 2.37)^2 + 5(.03 - 2.37)^2 = 1.053}$

5. Find the standard deviation for this distribution. Show how you got your answer.

   $\boxed{\sigma = \sqrt{1.053} = 1.026}$

6. Suppose the number of scoops of 25 randomly chosen customers is recorded. Call this number *Y*. Find the mean of *Y*.

   $\boxed{\mu_Y = \mu_{25X} = 25\mu_X = 59.25}$

7. Find the variance of *Y*.

   $\boxed{\sigma^2{}_{25X} = 25^2\sigma^2{}_X = 658.125}$

8. Find the standard deviation of *Y*.

   $\boxed{\sigma_{25X} = \sqrt{658.125} = 25.654}$

9. Suppose the business owners wanted to keep track of its costs. Ice cream scoops cost the business 62 cents and ice cream cones cost the business 8 cents. That would involve multiplying the number of ice cream scoops by $0.62 and adding $.08. Let Z = .62X + .08.

   a. Find μ_Z

   $\boxed{\mu_{.62X+.08} = .08 + .62\mu_X = 1.549}$

   b. Find σ_Z

   $\boxed{\sigma_{.62X+.08} = .62\sigma_X = .636}$

Problems 10 – 13 deal with this situation: According to statistics, 22.8% of people in the U.S. are vegetarians. 18 people are chosen at random.

10. If X is a random variable for this situation, define X in words. What kind of distribution does X have?

    X represents the number of people out of 18 (0 – 18) who are vegetarians. It is a binomial distribution.

11. Find the following probabilities: (calculators allowed)

    a. that less than 4 are vegetarians.

    $\boxed{P(\text{X} < 4) = .386}$

    Binomcdf(18,.228,3)

    b. that at least half are vegetarians.

    $\boxed{P(\text{X} \ge 9) = .011}$

    1 – Binomcdf(18,.228,8)

12. Find

    a. $\boxed{\mu_X = np = 18(.228) = 4.10}$

    b. $\boxed{\mu_X = \sqrt{np(1-p)} = \sqrt{18(.228)} = 1.78}$

13. For 18 people, find the probability that the number of people who are vegetarians is within 1.5 standard deviations of its mean.

You just found the standard deviation to be 1.78. 1.5(1.78) = 2.67. So you want the probability that X is between 4.10 - 2.67 and 4.10 + 2.67 meaning $P(1.43 < X < 6.77)$. Since X must be whole numbers, you want $P(X = 2 \text{ or } 3 \text{ or } 4 \text{ or } 5)$. These are binomial pdf's.

binompdf(18,.228,2) + binompdf(18,.228,3) + binompdf(18,.228,4) + binompdf(18,.228,5) = .730

Problems 14 – 16 deal with this situation: I take a multiple choice test with 6 questions, each having choice A, B, C, D, E. Let *X* be the number of questions I get right if I guess at each answer.

14. Complete the chart (3 decimal places)

X	0	1	2	3	4	5	6
P(*X*)	.262	.393	.246	.082	.015	.002	.000

15. Find μ_X

$\mu_X = np = 6(.2) = 1.2$

16. Show that σ_X can be calculated in two ways.

$\sigma_X = \sqrt{(0-1.2)^2(.262)+(1-1.2)^2(.393)+...}$ or $\sigma_X \sqrt{6(.25)(.75)}$

Problem 17 – 21 deal with this situation: At a gas station, only 19% of customers purchase premium gasoline.

17. What is the probability that 4 consecutive non-premium purchases will go by before a premium customer comes?

$(.81)^4(.19) = .082$

18. How many customers does the station expect to have before it gets a premium gas purchase?

$\frac{1}{.19} = 5.263$

19. What is the probability that it takes more than 10 purchases to see the first premium sale?

$.81^{10} = .122$

20. Construct a probability distribution table (out to *n* = 6) for the number of cars that will come into the station for a purchase to be premium.

X	1	2	3	4	5	6
P(X)	.19	.154	.125	.101	.082	.066

21. Construct a probability histogram for the table you just constructed.