AP Statistics – Unit 3 Concepts (Chapter 5, 6, 7)

Baseline Topics: (must show mastery in order to receive a '3' or above

- **I can distinguish between a parameter and a statistic.**
- **I can use a probability distribution to answer questions about possible values of a random variable.**

Let W represent a random variable whose distribution if ***normal***, with a mean of 120 and a standard deviation of 18. What would be an ***equivalent*** statement to P(W < 156)?

- **I can calculate the mean of a discrete random variable.**

A nonprofit organization plans to hold a raffle to raise funds for its operations. A total of 2,000 raffle tickets will be sold for $5.00 each. After all the tickets are sold, one ticket will be selected at random and its owner will receive $700.00. The **expected value** for the net gain for each ticket is -$0.75. What is the meaning of the **expected value** in this context?

- **I can check whether the 10% and Normal conditions are met in a given setting.**

The distribution of an unknown situation is ***strongly skewed to the right*** with a mean of 458 and a standard deviation of 94. If all possible samples of size 49 are drawn from this population and the mean is calculated for each of these samples, describe the ***sampling distribution*** of the sample mean.

- **I can use simulation to model chance behavior.**
- **I can describe a probability model for a chance process.**
- **I can use basic probability rules, including the complement rule and the addition rule for mutually exclusive events.**
- **I can use a Venn diagram to model a chance process involving two events.**
- **I can use the general addition rule to calculate P(A $\cup$ B)**
- **I can use the general multiplication rule to solve probability questions.**
- **I can determine whether two events are independent.**
- **I can find the probability that an event occurs using a two-way table.**
- **I can, when appropriate, use the multiplication rule for independent events to compute probabilities.**
- **I can compute conditional probabilities.**
- **I can interpret the mean of a random variable in context.**
- **I can calculate the standard deviation of a discrete random variable.**
- **I can interpret the standard deviation of a random variable in context.**
- **I can describe the effects of transforming a random variable by adding or subtracting a constant and multiplying or dividing by a constant.**
- **I can find the mean and standard deviation of the sum or difference of independent random variables.**
- **I can determine whether two random variables are independent.**
- **I can find probabilities involving the sum or difference of independent Normal random variables.**
- **I can determine whether the conditions for a binomial random variable are met.**
- **I can compute and interpret probabilities involving binomial distributions.**
- **I can calculate the mean and standard deviation of a binomial random variable. Interpret these values in context.**
- **I can find probabilities involving geometric random variables.**
- **I can determine whether a statistic is an unbiased estimator of a population parameter.**
- **I can understand the relationship between sample size and the variability of an estimator.**
- **I can find the mean and standard deviation of the sampling distribution of a sample proportion $\hat{p}$ for an SRS of size *n* from a population having proportion *p* of successes.**
- **I can check whether the 10% and Normal conditions are met in a given setting.**
- **I can use Normal approximation to calculate probabilities involving $\hat{p}$.**
- **I can use the sampling distribution of $\hat{p}$ to evaluate a claim about a population proportion.**
- **I can find the mean and standard deviation of the sampling distribution of a sample mean $\bar{x}$ from an SRS of size *n*.**
- **I can calculate probabilities involving a sample mean $\bar{x}$ when the population distribution is Normal.**
- **I can explain how the shape of the sampling distribution of $\bar{x}$ is related to the shape of the population distribution.**
- **I can use the central limit theorem to help find probabilities involving a sample mean $\bar{x}$.**

AP Statistics – AP Exam Probability Review/CLIFF NOTES

Overview of Chapter 5, 6, 7

Some Key Vocabulary

- Random
- Probability
- independent trial
- Sample Space -- Understand what a sample space is AND how to find one.
- What is an event?
- Factorials – relating to Binomial Probability
- Simulation: *You need to be able to describe how you will perform a simulation in addition to actually doing it.*
    - ***Create a correspondence between random numbers and outcomes.***
    - ***Explain how you will obtain the random numbers (e.g., move across the rows of the random digits table, examining pairs of digits), and how you will know when to stop.***
    - ***Make sure you understand the purpose of the simulation -- counting the number of trials until you achieve "success" or counting the number of "successes" or some other criterion.***
    - ***Are you drawing numbers with or without replacement? Be sure to mention this in your description of the simulation and to perform the simulation accordingly.***
    - ***If you're not sure how to approach a probability problem on the AP Exam, see if you can design a simulation to get an approximate answer.***

Major Concepts to be mastered:

Basic Rules of Probability

#1. Know the FOUR basic rules of probability

***Rule #1:* The probability of any event is always between 0 and 1**

***Rule #2:* The probability of an entire Sample Space adds up to 1.**

***Rule #3:* The probability of an event NOT occurring is the 1 minus the probability that it WILL occur. (COMPLEMENT RULE)**

***Rule #4:* When two events are mutually exclusive/disjoint, P(A or B) = P(A) + P(B)**

#2. Understand what disjoint means (mutually exclusive).

When 2 or more events CANNOT occur at the same time.
If A and B are disjoint, then P(A and B) = 0
If A and B are disjoint, then A and B are not independent.

#3. Understand why disjoint events CANNOT be independent.

- **Independent events are not the same as mutually exclusive (disjoint) events.**
- **Two events, A and B, are independent if the occurrence or non-occurrence of one of the events has no effect on the probability that the other event occurs.**
- **Events A and B are mutually exclusive if they cannot happen simultaneously.**

Example:

Roll two fair six-sided dice. Let A = the sum of the numbers showing is 7,
B = the second die shows a 6, and C = the sum of the numbers showing is 3.

- By making a table of the 36 possible outcomes of rolling two six-sided dice, you will find that P(A) = 1/6, P(B) = 1/6, and P(C) = 2/36.
- Events A and B are independent. Suppose you are told that the sum of the numbers showing is 7. Then the only possible outcomes are {(1,6), (2,5), (3,4), (4,3), (5,2), and (6,1)}. The probability that event B occurs (second die shows a 6) is now 1/6. This new piece of information did not change the likelihood that event B would happen. Let's reverse the situation.
- Suppose you were told that the second die showed a 6. There are only six possible outcomes: {(1,6), (2,6), (3,6), (4,6), (5,6), and (6,6)}.
- The probability that the sum is 7 remains 1/6. Knowing that event B occurred did not affect the probability that event A occurs.
- Events A and B are not disjoint. Both can occur at the same time.
- Events B and C are mutually exclusive (disjoint). If the second die shows a 6, then the sum cannot be 3. Can you show that events B and C are not independent?

#4. Understand the <u>multiplication rule</u> for <u>independent</u> events.

P(A and B) = P(A) * P(B)

#5. What is a <u>complement</u> of an event? If two events are independent, then their complements are independent too.

Complement: The probability that something WON'T OCCUR

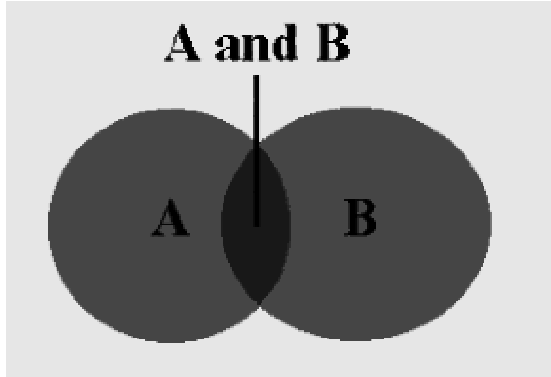
#6. Understand how to find the <u>union</u> of two events.

DISJOINT: P(A or B) = P(A) + P(B)

NON-DISJOINT: P(A or B) = P(A) + P(B) – P(A and B)

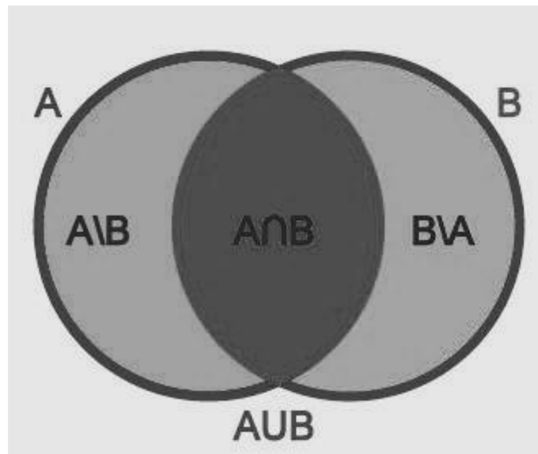
#7. What is the <u>intersection</u> of two or more events?

'OVERLAP'



#8. What is the *difference* between a <u>union</u> and an <u>intersection</u>?

UNION: 'or' ∪
INTERSECTION: 'and' ∩



#9. What is a <u>conditional probability</u>?

$$P(A|B)=\frac{P(A\cap B)}{P(B)}$$

#10. What is the <u>general multiplication rule</u>?

P(A and B) = P(A) * P(B | A)

#11. What is the difference between the <u>general multiplication rule</u> and the <u>multiplication rule for independent events</u>?

The general multiplication rule uses a conditional probability

#12. How do you determine if two events are independent?

Test for independence:
P(A) = P(A | B) or P(B) = P(B | A)

#13. Understand the idea of "*replacement*" versus "*non-replacement*."

Events are only independent when 'replacement' is used.

Chapter 6

#1. What is the difference between a discrete random variable and continuous random variable?

Discrete: Countable

Continuous: Measurable (think: interval of numbers)

#2. How is a *normal distribution* related to a *probability distribution*?

It's a continuous random variable

#3. How do you find an expected value of a random variable X? Know the differences between the two symbols for means. Understand the rules for means AND rules for variances.

Example: Let X = the number of heads obtained when five fair coins are tossed.

Value of x	0	1	2	3	4	5
Probability	1/32 = 0.03125	5/32 = 0.15625	10/32 = 0.3125	10/32 = 0.3125	5/32 = 0.15625	1/32 = 0.03125

$E(X) = \mu_x$

$= 0(.03125) + 1(.15625) + 2(.3125) + 3(.3125) + 4(.15625) + 5(.03125) = 2.5.$

$Var(X) = \sigma_x^2$

$= 03125(0-2.5)^2 + .15625(1-2.5)^2 + .3125(2-2.5)^2 + .3125(3-2.5)^2$
$+ .15625(4-2.5)^2 + .03125(5-2.5)^2 = 1.25.$

So $\sigma_x = \sqrt{1.25} = 1.118.$

You need to be able to work with transformations and combinations of random variables.

- For any random variables X and Y:

  $\mu_{a+bX} = a + b\mu_X$ and $\sigma^2_{a+bX} = b^2\sigma^2_X$

  $\mu_{X+Y} = \mu_X + \mu_Y$

  $\mu_{X-Y} = \mu_X - \mu_Y$

- For independent random variables X and Y:

  $\sigma^2_{X+Y} = \sigma^2_x + \sigma^2_y$

Expected value ($E(X)$ or μ_X) does not have to be a whole number.

#4. Explain the Law of Large Numbers vs. Central Limit Theorem.

LAW: As 'n' gets larger, the sample mean ($\overline{x}$) gets closer to the population mean (μ)

THEOREM: As 'n' gets larger, a distribution gets closer and closer being 'normal.'

#5. What are the *four conditions* for a binomial setting?

The four requirements for a chance phenomenon to be a binomial situation are:

- ***There are a fixed number of trials.***
- ***On each trial, there are two possible outcomes that can be labeled "success" and "failure."***
- ***The probability of a "success" on each trial is constant.***
- ***The trials are independent.***

#6. What do the parameters 'n' and 'p' represent in a binomial distribution?

n = sample size; p = probability of success; q or (1 – p) = probability of failure.

Recognize a *binomial situation* when it arises.
Example: Consider rolling a fair die 10 times. There are 10 trials. Rolling a 6 constitutes a "success," while rolling any other number represents a "failure." The probability of obtaining a 6 on any roll is 1/6, and the outcomes of successive trials are independent.

Using the TI-83, the probability of getting exactly three sixes is $({}_{10}C_3)(1/6)^3(5/6)^7$ or binompdf(10,1/6,3) = 0.155045, or about 15.5 percent.

The probability of getting less than four sixes is binomcdf(10,1/6,3) = 0.93027, or about 93 percent. Hence, the probability of getting four or more sixes in 10 rolls of a single die is about 7 percent.

If X is the number of 6's obtained when ten dice are rolled, then

E(X) = μ_x = 10(1/6) = 1.6667, and $\sigma_x = \sqrt{10(1/6)(5/6)} = 1.1785$.

Did you notice that the coin-tossing example is also a binomial situation?

Realize that a binomial distribution can be approximated well by a normal distribution if the number of trials is sufficiently large. If n is the number of trials in a binomial setting, and if p represents the probability of "success" on each trial, then a good rule of thumb states that a normal distribution can be used to approximate the binomial distribution if np is at least 10 and n(1-p) is at least 10.

#7. Know the difference between a probability distribution function (pdf) and a cumulative distribution function (cdf). Be able to use the calculator commands correctly:

BINOMCDF BINOMPDF, GEOMETCDF, & GEOMETPDF.

PDF: = ***CDF: <, >, ≥, ≤***

#8. Know how to find the mean and standard deviation of a binomial random variable.

$$\mu_x = n \cdot p \qquad \sigma_x = \sqrt{n \cdot p \cdot q}$$, ***recall that q = 1 – p.***

#9. What are the four conditions for a geometric setting?

The four requirements for a chance phenomenon to be a geometric situation are:

- ***There IS NOT a fixed number of trials. You are looking for the first instance of 'success'.***
- ***On each trial, there are two possible outcomes that can be labeled "success" and "failure."***
- ***The probability of a "success" on each trial is constant.***
- ***The trials are independent.***

#10. Explain the key differences between a binomial and geometric setting.

The primary difference between a binomial random variable and a geometric random variable is what you are counting. A binomial random variable counts the number of "successes" in n trials. A geometric random variable counts the number of trials up to and including the first "success."

#11. Know how to calculate the expected value of a geometric random variable.

Mean/expected value of a geometric random variable:

$$\mu_x = \frac{1}{p}$$

Chapter 7: Major Concepts to be mastered:

- ***Understand the difference between a statistic and a parameter***
- ***What is sampling variability?***
- ***Be able to distinguish between the Law of Large Numbers and the Central Limit Theorem***

#1. You need to know the difference between a *population parameter*, a *sample statistic*, and the *sampling distribution of a statistic*

a. In sample proportions, what do we represent the population parameter as?

$$\mu_{\hat{p}} = p$$

b. Be familiar with the formulas to find the mean and standard deviations of a sampling distribution of (p-hat).

$\sigma_{\hat{p}} = \sqrt{\frac{p \cdot q}{n}}$**, where q = 1 – p**

#2. Be able to check the 10% & Normal conditions.

- ***10% condition/Independence condition: The population size must be at least 10 times larger than the sample size.***
- ***Normal Condition:***

<u>MEANS</u>: n should be at least 30.

<u>PROPORTIONS</u>: $\boldsymbol{n \cdot p \geq 10}$ ***and*** $\boldsymbol{n \cdot q \geq 10}$

#4. Be able to calculate probabilities with sampling distributions.

$$z = \frac{\bar{x} - \mu}{s_x / \sqrt{n}} \text{ OR } z = \frac{\hat{p} - p}{\sqrt{\frac{p \cdot q}{n}}}$$

AP Statistics – Practice Probability Multiple Choice Questions

#1. Dr. Stats plans to toss a fair coin 10,000 times in the hope that it will lead him to a deeper understanding of the law of probability. Which of the following statements is true?

a. It is unlikely that Dr. Stats will get more than 5000 heads.
b. Whenever Dr. Stats gets a string of 15 tails in a row, it becomes more likely that the next toss will be a head.
c. The fraction of tosses resulting in heads should be close to ½.
d. The chance that the 100th toss will be a head depends somewhat on the results of the first 99 tosses.
e. All of the above statements are true.

#2. China has 1.2 billion people. Marketers want to know which international brands they have heard of. A large study showed that 62% of all Chinese adults have heard of Coca-Cola. You want to simulate choosing a Chinese at random and asking if he or she has heard of Coca-Cola. One correct way to assign random digits to simulate the answer is:

a. One digit simulates one person's answer; odd means "Yes" and even means "No."
b. One digit simulates one person's answer; O to 6 mean "yes" and 7 to 9 mean "no."
c. One digit simulates the result; 0 to 9 tells how many in the sample said "yes."
d. Two digits simulate one person's answer; 00 to 61 mean "yes" and 62 to 99 mean "no."
e. Two digits simulate one person's answer; 00 to 62 mean "Yes" and 63 to 99 mean "no."

#3. Choose an American household at random and record the number of vehicles they own. Here is the probability model if we ignore the few households that own more than 5 cars:

Number of cars	**0**	**1**	**2**	**3**	**4**	**5**
Probability	**0.09**	**0.36**	**0.35**	**0.13**	**0.05**	**0.02**

A housing company builds homes with two-car garages. What percent of households have more cars than the garage can hold?

a. 7%
b. 13%
c. 20%
d. 45%
e. 55%

#4. Computer voice recognition software is getting better. Some companies claim that their software correctly recognizes 98% of all words spoken by a trained user. To simulate recognizing a single word when the probability of being correct is 0.98, let two digits simulate one word; 00 to 97 mean 'correct.' The program recognizes words (or not) independently. To simulate the program's performance on 10 words, use these random digits:

60970 70024 17868 29843 61790 90656
87964 18883

The number of words recognized correctly out of the 10 is

a. 10
b. 9
c. 8
d. 7
e. 6

Refer to the following to answer the next THREE questions:

One thousand students at a city high school were classified according to both GPA and whether or not they skipped classes. The two-way table below summarizes the data:

	GPA		
Skipped Classes	**<2.0**	**2.0-3.0**	**>3.0**
Many	80	25	5
Few	175	450	265

#5. What is the probability that a student has a GPA under 2.0?

a. 0.227
b. 0.255
c. 0.450
d. 0.475
e. 0.506

#6. What is the probability that a student has a GPA under 2.0 or has skipped many classes?

a. 0.080
b. 0.281
c. 0.285
d. 0.365
e. 0.727

#7. What is the probability that a student has a GPA under 2.0 given that he or she has skipped many classes?

a. 0.080
b. 0.281
c. 0.285
d. 0.314
e. 0.727

#8. For events A and B related to the same chance process, which of the following statements is true?

a. If A and B are mutually exclusive, then they must be independent.
b. If A and B are independent, then they must be mutually exclusive.
c. If A and B are not mutually exclusive, then they must be independent.
d. If A and B are not independent, then they must be mutually exclusive.
e. If A and B are independent, then they cannot be mutually exclusive.

#9. Choose an American adult at random. The probability that you choose a woman is 0.52. The probability that the person you choose has never married is 0.25. The probability that you choose a woman who has never married is 0.11. The probability that the person you choose is either a woman or has never been married (or both) is therefore about

a. 0.77
b. 0.66
c. 0.44
d. 0.38
e. 0.13

#10. A deck of playing cards has 52 cards, of which 12 are face cards. If you shuffle the deck well and turn over the top 3 cards, one after the other, what's the probability that all 3 are face cards?

a. 0.001
b. 0.005
c. 0.010
d. 0.012
e. 0.02

<u>Refer to the following to answer the next TWO questions:</u>

A psychologist studied the number of puzzles that subjects were able to solve in a five-minute period while listening to soothing music. Let X be the number of puzzles completed successfully by a subject. The psychologist found that X had the following probability distribution:

Value of X:	**1**	**2**	**3**	**4**
Probability:	**0.2**	**0.4**	**0.3**	**0.1**

#11. What is the probability that a randomly chosen subject completes at least 3 puzzles in the five minute period while listening to soothing music?

a. 0.3
b. 0.4
c. 0.6
d. 0.9
e. Cannot be determined

#12. Suppose that three randomly selected subjects solve puzzles for five minutes each. The expected value of the total number of puzzles solved by the three subjects is

a. 1.8
b. 2.3
c. 2.5
d. 6.9
e. 7.5

#13. Suppose a student is randomly selected from your school. Which of the following pairs or random variables are most likely independent?

a. X = student's height; Y = student's weight
b. X = student's IQ; Y = student's GPA
c. X = student's PSAT Math score; Y = student's PSAT Verbal score
d. X = average amount of homework the student does per night; Y = student's GPA
e. X = average amount of homework the student dos per night; Y = student's height

#14. A certain vending machine offers 20-ounce bottles of soda for $1.50. The number of bottles X brought from the machine on any day is a random variable with mean 50 and standard deviation 15. Let the random variable Y equal the total revenue from this machine on a given day. Assume that the machine works properly and that no sodas are stolen from the machine. What are the mean and standard deviation of Y?

a. $\mu_Y = \$1.50, \sigma_Y = \$22.50$
b. $\mu_Y = \$1.50, \sigma_Y = \$33.75$
c. $\mu_Y = \$75, \sigma_Y = \$18.37$
d. $\mu_Y = \$75, \sigma_Y = \$22.50$
e. $\mu_Y = \$75, \sigma_Y = \$33.75$

<u>Refer to the following to answer the next TWO questions:</u>
The weight of tomatoes chosen at random from a bin at the farmer's market is a random variable with mean $\mu = 10$ ounces and standard deviation $\sigma = 1$ ounce. Suppose we pick four tomatoes at random from the bin and find their total weight, T

#15. The random variable T has a mean of

a. 2.5 ounces
b. 4 ounces
c. 10 ounces
d. 40 ounces
e. 41 ounces

#16. The random variable T has a standard deviation (in ounces) of

a. 0.25
b. 0.50
c. 0.71
d. 2
e. 4

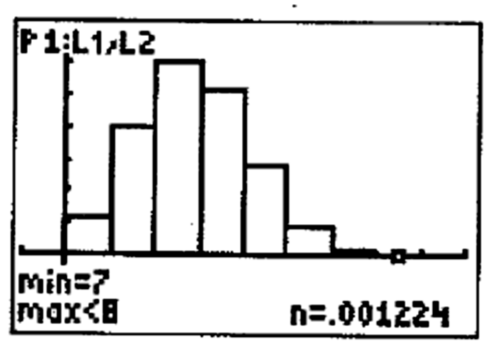
#17. Which of the following random variables is geometric?
The number of times I have to roll a die to get TWO 6's.
The number of cards I deal from a well-shuffled deck of 52 cards until I get a heart.
The number of digits I read in a randomly selected row of the random digits table until I find a 7.
The number of 7's in a row of 40 random digits.
The number of 6's I get if I roll a die 10 times.

#18. Seventeen people have been exposed to a particular disease. Each one independently has a 40% chance of contracting the disease. A hospital has the capacity to handle 10 cases of the disease. What is the probability that the hospital's capacity will be exceeded?

a. 0.011
b. 0.035
c. 0.092
d. 0.965
e. 0.989

#19. The figure shows the probability distribution of a discrete random variable X. Which of the following best describes this random variable?

a. Binomial with n = 8, p = 0.1
b. Binomial with n = 8, p = 0.3
c. Binomial with n = 8, p = 0.8
d. Geometric with p = 0.1
e. Geometric with p = 0.2



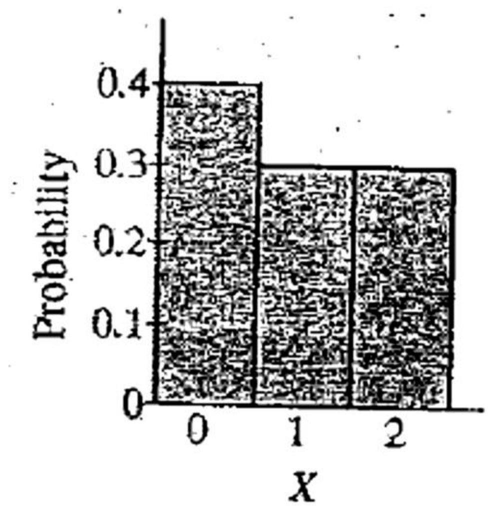
#20. A test for extrasensory perception (ESP) involves asking a person to tell which of 5 shapes – a circle, star, triangle, diamond, or heart – appears on a hidden computer screen. On each trial, the computer is equally likely to select any of the 5 shapes. Suppose researchers are testing a person who does not have ESP and so is just guessing on each trial. What is the probability that the person guess the first 4 shapes incorrectly but gets the fifth correct?

a. 1/5
b. $\left(\frac{4}{5}\right)^4$
c. $\left(\frac{4}{5}\right)^4 * \left(\frac{1}{5}\right)$
d. $\binom{5}{1}\left(\frac{4}{5}\right)^4 * \left(\frac{1}{5}\right)$
e. 4/5

#21. The probability of obtaining a head when a certain coin is flipped is about 0.65. Which of the following is closest to the probability that heads would be obtained 15 or fewer times when this coin is flipped 25 times?

a. 0.14
b. 0.37
c. 0.39
d. 0.60
e. 0.65

#22. The number of points, X, scored in a game has the probability distribution below:



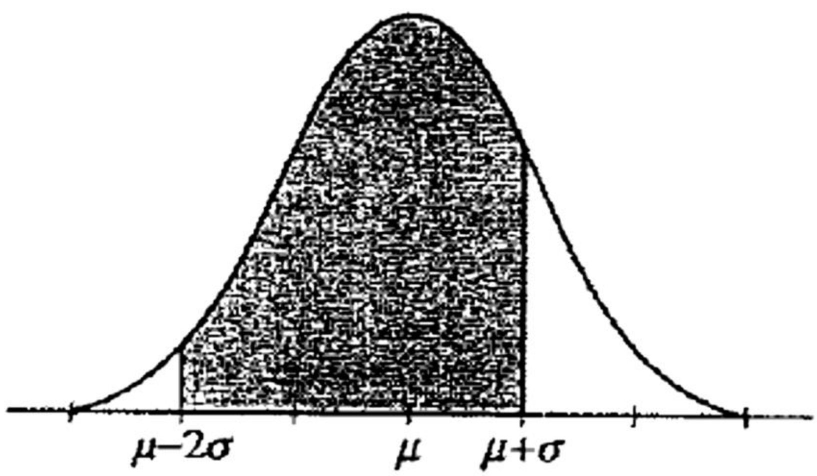
The number of points obtained in one game is independent of the number of points obtained in a second game. When the game is played twice, the sum of the number of points for both times could be 0, 1, 2, 3, or 4.
If Y represents the sampling distribution of the sum of the scores when the game is played twice, for which value of Y will the probability be greatest?

a. 0
b. 1
c. 2
d. 3
e. 4

#23. A recent study was conducted to investigate the duration of time required to complete a certain manual dexterity task. The reported mean was 10.2 seconds with a standard deviation of 16.0 seconds. Suppose the reported values are the true mean and standard deviation for the population of subjects in the study. If a random sample of 144 subjects is selected from the population, what is the approximate probability that the mean of the sample will be more than 11.0 seconds?

a. 0.1151
b. 0.2743
c. 0.7257
d. 0.8849
e. Based on the values of the true mean and true standard deviation, it can be concluded that the population distribution is not normal and therefore the probability cannot be calculated.

#24. A certain type of remote-control car has a fully charged battery at the time of purchase. The distribution of running times of cars of this type, before they require recharging of the battery for the first time after its period of initial use, is approximately normal with a mean of 80 minutes and a standard deviation of 2.5 minutes. The shaded area in the figure below represents which of the following probabilities?



a. The probability that the running time of a randomly selected car of this type, before it requires recharging of the battery for the first time after its period of initial use, is between 75 minutes and 82.5 minutes.
b. The probability that the running time of a randomly selected car of this type, before it requires recharging of the battery for the first time after its period of initial use, is between 75 minutes and 85 minutes.
c. The probability that the running time of a randomly selected car of this type, before it requires recharging of the battery for the first time after its period of initial use, is between 77.5 minutes and 82.5 minutes.
d. The probability that the running time of a randomly selected car of this type before it requires recharging of the battery for the first time after its period of initial use is between 77.5 minutes and 85 minutes.

#25. Ten percent of all Dynamite Mints candies are orange and 45 percent of Holiday Mints candies are orange. Two independent random samples, each of size 25, are selected – one from Dynamite Mints candies and the other from Holiday Mints candies. The total number of orange candies in the two samples is observed. What are the expected total number of orange candies and the standard deviation for the total number of orange candies, respectively, in the two samples?

a. 7 and 2.905
b. 7 and 3.987
c. 13.75 and 2.233
d. 13.75 and 2.905
e. 13.75 and 3.987

#26. Traffic data revealed that 35 percent of automobiles traveling along a portion of an interstate highway were exceeding the legal speed limit. Using highway cameras and license plate registrations, it was also determined that 52 percent of sports cars were also speeding along the same portion of the highway. What is the probability that a randomly selected car along the same portion of the highway was a speeding sports car?

a. 0.870
b. 0.673
c. 0.182
d. 0.170
e. It cannot be determined from the information given

#27. A magazine has 1,620,000 subscribers, of whom 640,000 are women and 980,000 are men. Thirty percent of the women read the advertisements in the magazine and 50 percent of the men read the advertisements in the magazine. A random sample of 100 subscribers is selected. What is the expected number of subscribers in the sample who read the advertisements?

a. 30
b. 40
c. 42
d. 50
e. 80

#28. Joey and Matthew plan to visit a bookstore. Based on their previous visits to this bookstore, the probability distributions of the number of books they will buy are given below:

Number of books Joe will buy	0	1	2
Probability	0.50	0.25	0.25

Number of books Matthew will buy	0	1	2
Probability	0.25	0.50	0.25

Assuming that Joe and Matthew make their decisions independently, what is the probability that they will purchase no books on this visit to the bookstore?

a. 0.0625
b. 0.1250
c. 0.1875
d. 0.2500
e. 0.7500

<u>Refer to the following to answer the next TWO questions:</u>

Every Thursday, Matt and Dave's Video Venture has "roll-the-dice" day. A customer may choose to roll two fair dice and rent a second movie for an amount (in cents) equal to the numbers uppermost on the dice, with the larger number first. For example, if the customer rolls a two and a four, a second movie may be rented for $0.42. If a two and a two are rolled, a second movie may be rented for $0.22. Let X represent the amount paid for a second movie on roll-the-dice day. The expected value of X is $0.47 and the standard deviation of X is $0.15

#29. If a customer rolls the dice and rents a second movie every Thursday for 20 consecutive weeks, what is the total amount that the customer would expect to pay for these second movies?

a. $0.45
b. $0.47
c. $0.67
d. $3.00
e. $9.40

#30. If a customer rolls the dice and rents a second movie every Thursday for 30 consecutive weeks, what is the approximate probability that the total amount paid for these second movies will exceed $15.00?

a. 0
b. 0.09
c. 0.14
d. 0.86
e. 0.91

#31. A manufacturer makes light bulbs and claims that their reliability is 98 percent. Reliability is defined to be the proportion of non-defective that are produced over the long term. If the company's claim is correct, what is the expected number of non-defective light bulbs in a random sample of 1,000 bulbs?

a. 20
b. 200
c. 960
d. 980
e. 1,000

#32. The XYZ Office Supplies Company sells calculators in bulk at wholesale prices, as well as individuals at retail prices. Next year's sales depend on market conditions, but executives use probability to find estimates of sales for the coming year. The following tables are estimates for next year's sales:

WHOLESALE SALES				
Number Sold	2,000	5,000	10,000	20,000
Probability	0.1	0.3	0.4	0.2

RETAIL SALES			
Number Sold	600	1,000	1,500
Probability	0.4	0.5	0.1

What profit does XYZ Office Supplies Company expect to make for the next year if the profit from each calculator sold is $20 at wholesale and $30 at retail?

a. $10,590
b. $220,700
c. $264,750
d. $833,100
e. $1,002,500

#33. Circuit boards are assembled by selecting 4 computer chips at random from a large batch of chips. In this batch of chips, 90 percent of the chips are acceptable. Let X denote the number of acceptable chips out of a sample of 4 chips from this batch. What is the least probable value of X?

a. 0
b. 1
c. 2
d. 3
e. 4

#34. At Kennett High School, 5% of athletes play both football and some other contact sport, 30% play football, and 40% play other contact sports. If there are 200 athletes, how many play neither football nor any other contact sport?

a. 20
b. 70
c. 80
d. 100
e. 130

#35. A fair coin is flipped 10 times and the number of heads is counted. This procedure of 10 coin flips is repeated 100 times and the results are placed in a frequency table. Which of the frequency tables below is most likely to contain the results from these 100 trials?

(A)

Number of Heads	Frequency
0	19
1	12
2	9
3	6
4	2
5	1
6	3
7	5
8	8
9	14
10	21

(B)

Number of Heads	Frequency
0	9
1	9
2	9
3	9
4	9
5	10
6	9
7	9
8	9
9	9
10	9

(C)

Number of Heads	Frequency
0	0
1	0
2	6
3	9
4	22
5	24
6	18
7	12
8	7
9	2
10	0

(D)

Number of Heads	Frequency
0	7
1	10
2	6
3	11
4	8
5	10
6	9
7	12
8	7
9	11
10	9

(E)

Number of Heads	Frequency
0	0
1	0
2	0
3	2
4	24
5	51
6	22
7	1
8	0
9	0
10	0

ANSWERS:

#1. C	#2. D	#3. C	#4. B	#5. B
#6. C	#7. E	#8. E	#9. B	#10. A
#11. B	#12. D	#13. E	#14. D	#15. D
#16. D	#17. C	#18. B	#19. B	#20. C
#21. B	#22. C	#23. B	#24. A	#25. D
#26. C	#27. C	#28. B	#29. E	#30. Omit
#31. D	#32. B	#33. A	#34. B	#35. C

AP Statistics –Probability Review A

#1. A coin is flipped three times.

- ***Event A consists of observing exactly 2 heads***.
- ***Event B consists of observing one or more tails.***

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

a) List the members of event A.

b) List the members of event B.

c) Find the union of event A and event B.

d) Find the intersection of event A and event B.

e) Find P(A|B)

#2. A player rolls two dice. The players receive $1.00 for each dot on the face of the two dice. For example, he will be paid $9.00 if the dice show a 4 and a 5. Let *X* = the number of dollars won each game.

a) Is *X* a ***discrete or continuous*** random variable?

b) How many possible outcomes are there?

c) Show the ***probability distribution*** of *X*.

X	2	3	4	5	6	7	8	9	10	11	12
P(X)											

d) Find the ***expected value*** of *X*.

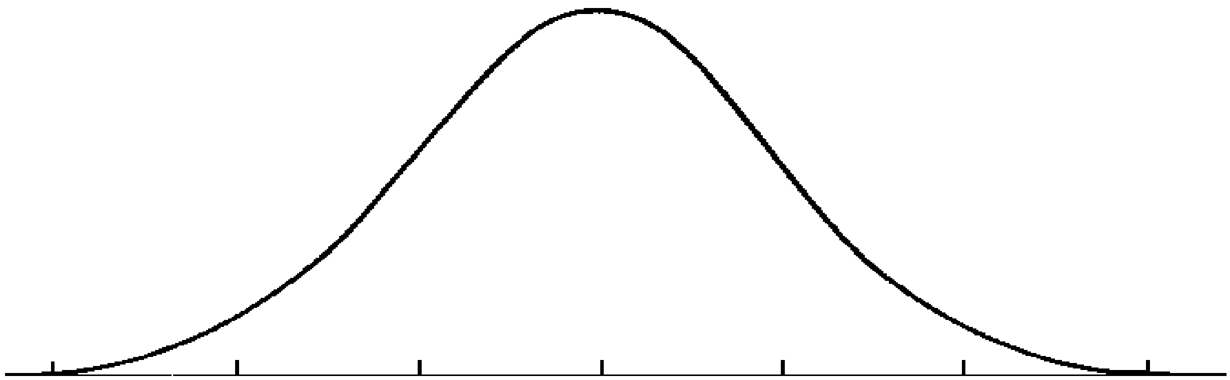
$$E(x) = \sum_{i=1} x_i p_i = x_1 p_1 + x_2 p_2 + \cdots$$

e) Find the ***variance*** of the pay off.

$$\sigma_X^2 = \sum (x_i - \mu_x)^2 p_i$$

#3. Let X = the actual weight of a 2-lb. bag of rice. The mean amount of rice is 32 ounces, with standard deviation is 0.25 ounces.

a) Is X a ***discrete or continuous variable***?



b) What ***percentage*** of the population will be within 31.5 and 32.5 ounces?

c) P(X = 32)

d) $P(31.5 \le X \le 32.5)$

e) $P(X > 32.21)$

f) P(X < 32.08)

g) P(31.6 < X < 32.33)

#4. A researcher surveyed students majoring in business and asked what type of car they own.

	Bought New	**Bought Used**	**TOTAL**
Male	**21**	**32**	**53**
Female	**44**	**15**	**59**
TOTAL	**65**	**47**	**112**

a) What is the probability that a female student purchased a new car?

b) What is the probability that a male student purchased a used car?

c) What is the probability that a student purchased a new car?

d) What is the probability that a student is female?

e) Given that a student is female what is the probability that she purchased a used car?

f) Given that a car is new what is the probability that the owner is a male?

g) Are the events ‘student is female’ and ‘bought used car’ independent?

#5. A new detergent is found to remove excess dirt and stains satisfactorily on 88% of the items washed. Assume that 3 items are to be washed with the new detergent.

a) What type of distribution does this follow? What are n and p?

b) What is the probability of satisfactory results on all 3 items?

c) What is the probability that exactly two items are cleaned satisfactorily?

d) What is the probability that none are cleaned satisfactorily?

e) What is the probability that at least one item is cleaned satisfactorily?

f) What is the probability that less than 2 items are cleaned satisfactorily?

#6. Given three outcomes E_1, E_2, and E_3 with probabilities:

$P(E_1) = 0.05$, $P(E_2) = 0.55$, and $P(E_3) = 0.4$

Let event $A=\{E_1,E_3\}$, and $B=\{E_2, E_3\}$

a) Find the probability of A ***union*** B.

b) Find the probability of A ***intersect*** B.

c) Find the probability of A ***given*** B.

d) Find the probability of B ***given*** A.

e) Are A and B ***independent***?

#7. You can insure a $25,000 diamond for its total value by paying a premium of *C* dollars. If the probability of theft in a given year is 0.20, what premium should the insurance company charge if it wants the expected gain to be equal to $1000?

#8. Microcomputers are shipped to the University bookstore from three factories A, B, and C. You know that factory A produces 20% defective microcomputers, whereas B produces 10% defectives and C only 5% defectives. The manager in the store receives a new shipment of microcomputers and discovers that 40% are from factory C, 40% are from factory B, and 20% are from factory A. (Hint: make a tree diagram)

a) What is the probability of finding a defective microcomputer in this shipment?

b) Are the events “microcomputer comes from factory A” and “microcomputer comes from factory B” mutually exclusive? Are they independent?

c) Suppose the manager randomly selects one microcomputer, and discovers that it is defective. What is the probability that it came from factory A?

#9. Let *X* be a ***discrete random variable*** with probability distribution:

X	***p(x)***
1	.15
2	.40
3	.10
4	.35

a) Find $P(X = 2)$.

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

c) Find $P(X < 2)$

d) Find $P(2 \le X \le 4)$

e) Find $P(2 < X \le 4)$

f) Find $P(2 \le X < 4)$

g) Find $P(2 < X < 4)$

#10. Let *X* be a ***discrete random variable*** with probability distribution:

X	*p(x)*
1	.10
2	.45
3	.15
4	.25

Does this assignment define a proper probability distribution? Explain.

#11. Investing $100 in a project will yield a net return of $16, $20 or $26 with respective probabilities 0.3, 0.5 and 0.2. Let *X* be the random variable that represents the net return.

a) Write down the ***probability distribution*** of the variable *X.*

b) Find the ***expected return*** for the project, and the ***standard deviation***.

c) Define $Y = 3X + 2$. Specify the ***probability distribution*** of *Y*.

d) Find the ***mean*** and ***standard deviation*** of Y.

❖ **I can, when appropriate, use the multiplication rule for independent events to compute probabilities.**

A fair die is to be rolled 8 times. The first 7 rolls all land on 'TWO'. What is the **probability** that the die will land on 'TWO' for the 8^{th} and final roll?

❖ **I can describe a probability model for a chance process.**

A person is planning to drive a ONE hour commute to work. The probability that the person will arrive to work on time is 0.90. What is the most reasonable explanation for how that **probability** could have been estimated?

❖ **I can find the probability that an event occurs using a two-way table.**

One hundred people were interviewed and classified according to their attitude toward small cars and their personality type.

The results are shown in the table below:

		Personality Type		
		Type A	Type B	Total
Attitude Toward Small Cars	Positive	25	12	37
	Neutral	11	9	20
	Negative	24	19	43
	Total	60	40	100

Determine the following:

#1. P(Negative)

#2. P(Type A)

#3. P(Type A & Neutral)

#4. P(Type B OR Negative)

#5. P(Type A | Positive)

#6. P(Neutral | Type B)

#7. P(not Type B)

#8. P(not Positive)

- **I can find the mean and standard deviation of the sum or difference of independent random variables.**

Random Variable B is ***normally distributed*** with mean 11 and standard deviation 4, and random variable C is ***normally distributed*** with mean 9 and standard deviation 2. If B and C are **independent**, what would best describe the distribution of C – B?

- **I can use the general addition rule to calculate P(A $\cup$ B)**
- **I can determine whether two events are independent.**
- **I can use basic probability rules, including the complement rule and the addition rule for mutually exclusive events.**

#1. TRUE OR FALSE: If two or more events are **MUTUALLY EXCLUSIVE**, then the events are also **INDEPENDENT**.

#2. If events D & G are '**disjoint'**, then it would be appropriate to make the following calculation: P(D or G) = P(D) + P(G).

#3. What age groups use social networking sites? A recent study produced the following data about 768 individuals who were asked their age and which of three social networking sites they used most often. (People who did not use such sites were excluded from the study).

Age Group (Years)					
Web site	***0-24***	***25-44***	***45-64***	***Over 65***	**Totals**
Facebook	77	105	114	12	**308**
Twitter	46	110	81	7	**244**
LinkedIn	15	97	95	9	**216**
Totals	**138**	**312**	**290**	**28**	**768**

Suppose one subject from this study was selected at random.

(a) Find the **probability** that the selected subject preferred Twitter.

(b) Find the **probability** that the selected subject preferred Twitter, given that he or she was in The 45 – 64 age group.

(c) Are the events "preferred Twitter" and "age group 45 – 64" ***independent***? Explain.

- **I can interpret the standard deviation of a random variable in context.**
- **I can describe the effects of transforming a random variable by adding or subtracting a constant and multiplying or dividing by a constant.**
- **I can find the mean and standard deviation of the sum or difference of independent random variables.**

A company sells concrete in batches of 5 cubic yards. The probability distribution of X, the number of cubic yards sold in a single order for concrete from this company, is shown in the table below.

X = number of cubic yards	**10**	**15**	**20**	**25**	**30**
Probability	**0.15**	**0.25**	**0.25**	**0.30**	**0.05**

The **expected value** of the probability distribution of X is 19.25 and the standard deviation is 5.76. There is a fixed cost to deliver the concrete. The profit Y, in dollars, for a particular order can be described by Y = 48X – 100. What is the **standard deviation** of Y?

- **I can use basic probability rules, including the complement rule and the addition rule for mutually exclusive events.**
- **I can use the general multiplication rule to solve probability questions.**
- **I can determine whether two events are independent.**
- **I can, when appropriate, use the multiplication rule for independent events to compute probabilities.**

A computerized device contains three components that we will call events X, Y and Z. The probabilities of failure for each component in any **ONE** year are 0.08, 0.025, and 0.06, respectively. If any one component fails, the device will fail. If the components fail ***independently*** of one another, what is the ***probability*** that the device will not fail in one year?

- **I can compute conditional probabilities.**

The probability that a new dishwasher will stop working in less than 4 years is 0.10. The probability that a dishwasher is damaged during delivery **and** stops working in less than 4 years is 0.08. The **probability** that a dishwasher is damaged during delivery is 0.18. Given that a new dishwasher is damaged during delivery, what is the **probability** that it stops working in less than 4 years?

❖ **I can calculate probabilities involving a sample mean $\overline{x}$ when the population distribution is Normal.**

The commuting time for a person to travel from home to their workplace is normally distributed with a mean of 20 minutes and a standard deviation of 4 minutes. If the person leaves home at 7:35 A.M., what is the **probability** that the person will arrive at work later than 8:04 A.M.?

Free Response Question #1: Far-sightedness:

It is generally believed that far-sightedness affects about 16% of children. A school district gives vision tests to 110 incoming kindergarten children.

a. Describe the sampling distribution (shape, center and spread) for the sample proportion by telling its mean and standard deviation. Justify your answer. ***(Be sure to check for normality and to address any other conditions/assumptions)***

b. What is the **probability** that in this group over 18% of the children will be found to be nearsighted?

<u>Free Response Question #2: Customer Care</u>

The customer care manager at a cell phone company keeps track of how long each help-line caller spends on hold before speaking to a customer service representative. He finds that the distribution of wait times for all callers has a mean of 12 minutes with a standard deviation of 5 minutes. The distribution is ***moderately skewed to the right***. Suppose the manager takes a random sample of 30 callers and calculates their mean wait time, $\bar{x}$.

a. Is it possible to calculate the standard deviation of $\bar{x}$? If it is, do the calculation. If it isn't, explain why.

b. Based on the sample size, is it safe to assume a normal approximation for the sampling distribution? Explain. If it is safe to assume normality, determine the probability that the average number of minutes of wait time will be greater than 13 minutes for the random sample of 30 callers?

AP Statistics – Unit 3 Practice Free Response Questions

The city's local video arcade is hosting a tournament which contestants play an arcade game with possible scores ranging from 0 to 20. The arcade has set up multiple game stations so that all of the contestants can play the game at the same time; thus contestant scores are independent.
Each contestant's score will be recorded as he or she finishes, and the contestant with the highest score is the winner. After practicing the game many times, PLAYER A, one of the contestants, has established the probability distribution of their scores, shown in the table below:

PLAYER A Distribution					
Scores	16	17	18	19	20
Probability	0.08	0.32	0.36	0.21	.03

PLAYER B, another contestant, has also practiced many times. The probability distribution for their scores is shown in the table below:

PLAYER B Distribution			
Scores	17	18	19
Probability	0.40	0.40	0.20

a. Calculate the **expected** score for each player.

b. Suppose that PLAYER A scores 18 and PLAYER B scores 19. The difference (PLAYER A minus PLAYER B) of their scores is -1.
   i. List **all combinations** of possible scores for PLAYER A and PLAYER B that will produce a difference (PLAYER A minus PLAYER B) of -1,

   ii. Calculate the probability for each combination.

c. Find the **probability** that the difference (PLAYER A minus PLAYER B) in their scores is -1.

d. The table below lists all the **possible differences** in the scores between PLAYER A and PLAYER B and some associated probabilities.

Distribution (PLAYER A minus PLAYER B)							
Difference	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**
Probability	**0.016**			**0.314**	**0.234**	**0.096**	**0.012**

Complete the table and calculate the probability that PLAYER B's score will be higher than PLAYER A's score.