Laying the Foundation®
Advanced Programs Division

AP* Statistics Review

Probability

Teacher Packet

Probability Rules

A sample space contains all the possible outcomes observed in a trial of an experiment, a survey, or some random phenomenon.

- The sum of the probabilities for all possible outcomes in a sample space is 1.
- The probability of an outcome is a number between 0 and 1 inclusive. An outcome that always happens has probability 1. An outcome that never happens has probability 0.
- The probability of an outcome occurring equals 1 minus the probability that it doesn't occur.
- The probability that two mutually exclusive (disjoint) events occur is 0.

Strategies for solving probability problems:
Draw a picture of the situation

- Use a chart, table, tree diagram, Venn Diagram, normal curve

When is a binomial distribution appropriate?

- If there are exactly 2 outcomes, usually designated success and failure, for each trial.
- If the number of trials is fixed.
- If the trials are independent.
- If the probability of success is the same for each trial.

When is a geometric distribution appropriate?

- If there are exactly 2 outcomes for each trial.
- If the trials are independent.
- If the probability of success is the same for each trial.
- If there is not a fixed number of trials. The trials continue until a success/failure is achieved.

When is a normal distribution appropriate?

- If the data is modeled by a continuous distribution and is given as normal or the sample size is large enough (second semester topic).
- If the data is modeled by a binomial distribution and np and $n(1-p)$ are large enough.

Is there a formula on the AP formula sheet that applies?

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$
- $\mu_x = E(X) = \sum x_i P(x_i)$
- $\sigma_x^{\ 2} = \sum (x_i - \mu_x)^2 \cdot P(x_i)$
- $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
-

Is there a formula/idea that is not on the formula sheet that applies?

- If events are disjoint, then $P(A \cap B) = 0$
- If events are independent, $P(A|B) = P(A)$ or $P(A \cap B) = P(A) \cdot P(B)$
- For any 2 random variables X and Y, $\mu_{x \pm y} = \mu_x \pm \mu_y$.
- For any 2 independent random variables X and Y, $\sigma^2_{\ x \pm y} = \sigma^2_{\ x} + \sigma^2_{\ y}$
- Z-score = $\dfrac{\text{Value of interest - mean}}{\text{standard deviation}}$

What steps are needed if a simulation is appropriate?

- Model the component of interest in the problem with some chance mechanism.
- State any assumptions being made (usually independent trials and constant probability).
- Describe how the simulation will be run. If using random digits, be sure to state whether duplicates are allowed. Be sure to give a stopping rule.
- Conduct the simulation with a reasonable number of replications.
- State the conclusion reached in the context of the problem.

Multiple Choice Questions on Probability

Questions 1 and 2 refer to the following situation. The class of 1968 and 1998 held a joint reunion in 2008 at the local high school. Attendees were asked to complete a survey to determine what they did after graduation. Here is the information obtained.

	College	Job	Military	Other
1968	56	73	85	7
1998	173	62	37	20

1. What is the probability that a randomly selected attendee graduated in 1998 and went into the military?

   (A) 0.072
   (B) 0.127
   (C) 0.303
   (D) 0.596
   (E) 0.669

2. What is the probability that a randomly selected 1968 graduate went to college after graduation?

   (A) 0.245
   (B) 0.253
   (C) 0.560
   (D) 0.592
   (E) 0.755

3. A fair die is rolled 3 times. The first 2 rolls resulted in 2 fives. What is the probability of not rolling 5 on the next roll?

   (A) 1
   (B) $\frac{5}{6}$
   (C) $\binom{3}{1}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)$
   (D) $\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)$
   (E) 0

4. In a game, a spinner with five equal-sized spaces is labeled from A to E. If a player spins an A they win 15 points. If any other letter is spun the player loses 4 points. What is the expected gain or loss from playing 40 games?

(A) Gain of 360 points
(B) Gain of 55 points
(C) Gain of 8 points
(D) Loss of 1 point
(E) Loss of 8 points

5. Let X be a random variable whose distribution is normal with mean 30 and standard deviation 4. Which of the following is equivalent to $P(X \geq 26)$?

(A) $P(X < 34)$
(B) $P(X \leq 26)$
(C) $P(26 \leq X \leq 34)$
(D) $1 - P(X \leq 34)$
(E) $P(X \geq 34)$

6. The distribution of heights of male high school students has a mean of 68 inches and variance of 1.52 square inches. The distribution of female high school students has a mean of 66 inches and a variance of 1.64 square inches. If the heights of the male and female students are independent, what is the standard deviation of the difference in their heights?

(A) 0.12 inches
(B) 0.35 inches
(C) 1.48 inches
(D) 1.78 inches
(E) 2.24 inches

7. If $P(A) = 0.34$ and $P(A \text{ or } B) = 0.71$, which of the following is false?

(A) $P(B) = 0.37$, if A and B are mutually exclusive.
(B) $P(B) = 0.561$, if A and B are independent.
(C) $P(B)$ cannot be determined if A and B are neither mutually exclusive nor independent.
(D) $P(A \text{ and } B) = 0.191$, if A and B are independent.
(E) $P(A|B) = 0.34$, if A and B are mutually exclusive.

8. In a litter of eight puppies, 5 are female. 2 of the puppies are picked at random. Which of the following is true?

(A) The probability that both puppies are female is $\left(\frac{2}{5}\right)^2$.

(B) The probability that both puppies are female is $\left(\frac{5}{8}\right)^2$

(C) The probability that both puppies are female is $\left(\frac{5}{8}\right)\left(\frac{4}{7}\right)$

(D) The expected number of female puppies is 1.25.

(E) The situation can be described by a binomial model.

9. Homes built in the suburbs typically have none to three-car garages. Let X be the number of garage stalls per home found in a sample of 200 homes in a local suburban area. From the data obtained, $P(X = 0) = 0.06$, $P(X = 1) = 0.45$, and $P(X = 2) = 0.32$. Find the mean number of garage stalls per home for the sample of homes.

(A) 1.09
(B) 1.15
(C) 1.5
(D) 1.6
(E) 2

10. The probability that a randomly chosen American is a Republican is 0.35. What is the probability that in a sample of 10 Americans, that at least 1 will be a Republican?

(A) 0.9865
(B) 0.2275
(C) 0.0725
(D) 0.0135
(E) 0.0072

Free Response Questions on Probability

1. A one-mile relay race has 4 horses running a quarter of a mile each. The riders of each horse must pass a ribbon to the next rider at the end of each leg. Each of these horses has participated in many quarter-mile races before and the table below summarizes the mean and standard deviation of their previous times.

Horses	Mean	Standard Deviation
Morning Spark	32.6 sec	2.5 sec.
Dew on the Meadow	29.4 sec	2.1 sec
April Shower	37.7 sec	3.9 sec
Night Sky	26.8 sec	1.9 sec

(a) Assuming the times follow a normal model, what is the probability that Night Sky can run a quarter mile in under 24.5 seconds?

(b) What is the probability that in the next twelve races that Night Sky enters, that he will run under 24.5 seconds in 5 of those races?

(c) If the times of the horses are independent, what is the mean and standard deviation of the combined team score for a one-mile relay race?

2. Anna and Paulina are playing video games. Anna has a 60% chance of winning the first game that they play, but this grows to 70% for winning the second game in a row. If Anna loses the first game, she only has a 30% chance of winning the second game.

(a) What is the probability that they each win one of two games played?

(b) What is the probability distribution of the number of games that Anna wins?

(c) What is the mean and standard deviation of the number of games Anna wins?

3. A survey at a local college asked a random sample of faculty and a random sample of students the color of the car that they would like to drive. The results are given in the table.

	Faculty	Students
Silver	40	10
Black	20	147
Red	35	86
Other	25	17

(a) If a person is chosen at random from all those surveyed, what is the probability that they would like a black car?

(b) If the person chosen at random is a faculty member, what is the probability that they would prefer a black car? Show your work.

(c) Based on your answers in part (a) and part (b), is car color choice independent of college role (faculty, student) for those in this sample?

Key to Probability Multiple Choice

1. A $\frac{37}{513}$
2. B P(College given 1968 graduation) = 56/221
3. B What happens on the first two rolls does not affect the chance of getting a 5 again.
4. E $E(onegame) = \frac{1}{5}(15) + \frac{4}{5}(-4) = -\frac{1}{5}$. There is an average loss of 8 points over 40 games.
5. A The symmetry of the graph and the continuity give this result.
6. D $\sqrt{1.52 + 1.64} = 1.78$ (the variances are already squared)
7. E Statement is true if mutually exclusive is replaced by independent. Using the formula for $P(A \cup B)$ shows the others are true.
8. C The situation is not binomial because the probability of success is not the same for each trial. A (using an incorrect probability), B, D, and E are based on binomial models.
9. D $E(X) = 0(.06) + 1(.45) + 2(.32) + 3(.17)$
10. A $1 - P$(no Republicans in the group of 10) = 1- $(.65)^{10}$. Here because of the large population from which the sample is drawn, the situation can be modeled with a binomial model. (the 10% rule applies)

Rubric for Probability Free Response

1. **Solution**

Part (a)

For Night Sky

$$P(time < 24.5) = P(z < \frac{24.5 - 26.8}{1.9}) = 0.1131$$

The probability that Night Sky runs a quarter mile in less than 24.5 seconds is .1131

Part (b)

X = the number of races that Night sky runs in less than 24.5 seconds

X is binomial with n = 12 races and p = .1131

$$P(X = 5) = \binom{12}{5}(.1131)^5(.8869)^7 = .0063$$

Part (c)

$$\mu_{total} = \mu_{MS} + \mu_{DM} + \mu_{AS} + \mu_{NS} = 32.6 + 29.4 + 37.7 + 26.8 = 126.5 \text{ seconds}$$

$$\sigma_{Total} = \sqrt{\mu_{MS}{}^2 + \mu_{DM}{}^2 + \mu_{AS}{}^2 + \mu_{NS}{}^2}$$

$$= \sqrt{2.5^2 + 2.1^2 + 3.9^2 + 1.9^2} = 5.43$$

Scoring

Each part is *essentially correct* (E), *partially correct* (P), or *incorrect* (I).

Part (a) is *essentially correct* if the probability is correctly calculated and work is shown to support this result.

Part (a) is *partially correct* if:

the probability is correctly calculated but no work is shown
OR
reasonable work is shown (including calculator output) but the answer is not correct. Score part(a) as incorrect if the probability is not between 0 and .5.

Part (b) is *essentially correct* if:

1. the student recognizes the setting as binomial
2. the probability in part (a) is used for p
3. work is shown – the correct values for n and x are *given* and the desired probability calculated, or the binomial formula is correctly evaluated.

Part (b) is *partially correct* if:

the student recognizes the situation as binomial and identifies p from part (a) but does not compute the desired probability

OR

the student computes the probability as either $(.1131)^5(.8869)^7$ or $\binom{12}{5}(.1131)^5$

OR

the student gives the correct probability, but no work is shown.

Part (c) is *essentially correct* if both the mean and standard deviation of the team time are correctly calculated (except for arithmetic mistakes)

Part (c) is *partially correct* if only one of these is correctly computed(except for arithmetic mistakes)

4 **Complete Response** (3E)
All three parts essentially correct

3 **Substantial Response** (2E 1P)
Two parts essentially correct and one part partially correct

2 **Developing Response** (2E 0P or 1E 2P or 3P)
Two parts essentially correct and no parts partially correct
OR
One part essentially correct and two parts partially correct
OR
Three parts partially correct

1 **Minimal Response** (1E 1P or 1E 0P or 2P)
No parts essentially correct and either zero or one parts partially correct
OR
No parts essentially correct and two parts partially correct

2. **Solution**

Part (a)

Each girl wins one game if Anna wins first and then Paulina OR if Paulina wins and then Anna. The probability that they each win one game is
$P(AP\ or\ \ PA) = (.6)(.3) + (.4)(.3) = .18 + .12 = .3$

Part (b)

Let X be the number of games Anna wins. She wins 0, 1, or 2 games. The probability that she wins no games is (.4)(.7) = .28. The probability that she wins 2 games is (.6)(.7) = .42. The probability distribution is given in the table:

X	0	1	2
P(X)	.28	.3	.42

Part (c)

$\mu_x = (0)(.28) + 1(.3) + 2(.42) = 1.14$ games

$\sigma_x = \sqrt{(0-1.14)^2(.28) + (1-1.14)^2(.3) + (2-1.14)^2(.42)} = .825$ games

Scoring

Each part is *essentially correct* (E), *partially correct* (P), or *incorrect* (I).

Part (a) is *essentially correct* if the probability is correctly calculated AND a formula, chart, or tree diagram is used to support the answer.

Part (a) is *partially correct* if

the probability is correctly calculated (except for a minor arithmetic error) but no supporting work is shown

OR

the probability is not correctly calculated because students used one of (.6)(.3) or (.4)(.3) but supporting work is shown. The calculation is incorrect if the student uses the formula $P(A \cup B)$ as if the events were not disjoint.

Part (b) is *essentially correct* if the student shows the possible number of games won by Anna AND the correct associated probabilities. If the probability in part (a) was incorrect then reasonable values for the remaining probabilities can be counted as correct.

Part (b) is *partially correct* if the student shows the possible number of games won but does not have the correct probabilities (except as noted above).

Part (c) is *essentially correct* if both the mean and standard are calculated correctly AND work is shown, with the exception of minor arithmetic errors.

Part (c) is *partially correct* if either the mean or standard deviation is calculated correctly AND the work is shown. If the variance is reported instead of the standard deviation the question is scored a P.

Part (c) is *incorrect* if both the mean and standard deviation are calculated incorrectly OR no work is shown.

4 **Complete Response** (3E)
All three parts essentially correct

3 **Substantial Response** (2E 1P)
Two parts essentially correct and one part partially correct

2 **Developing Response** (2E 0P or 1E 2P or 3P)
Two parts essentially correct and no parts partially correct
OR
One part essentially correct and two parts partially correct
OR
Three parts partially correct

1 **Minimal Response** (1E 1P or 1E 0P or 2P)
No parts essentially correct and either zero or one parts partially correct
OR
No parts essentially correct and two parts partially correct

3. **Solution**

Part(a)

$$P(black\ car) = \frac{167}{380} = .4395$$

Part (b)

$$P(black\ car \mid faculty\ member) = \frac{20}{120} = .1667$$

Part (c)

If color choice and college role were independent, then the answers in part(a) and part (b) would be equal. Since these probabilities are not equal, role and color choice are not independent for those in this sample.

Part (a) is *essentially correct* (E) (may be minor arithmetic errors) or *incorrect* (I)

Part (b) is *essentially correct* if the conditional probability is correctly calculated.

Part (b) is *partially correct* if:

the student reverses the conditioning, calculating $P(facultymember \mid black\ car)$

OR

calculates the correct probability for the wrong column, e.g. $\frac{147}{260}$

Part (b) is *incorrect* if the student calculates the joint probability: $\frac{20}{380}$

Part (c) is *essentially correct* if

the student indicated that the variables are not independent

AND

the explanation is tied to the fact that the probabilities in part (a) and part (b) are not equal.

Part (c) is *partially correct* if the student indicates that the variables are not independent, but the explanation is incorrect or not based on parts (a) and (b)

Part (c) is incorrect if
the student fails to give a numerical justification to support the argument
OR
does one of the following
- Performs an incorrect additional calculation
- Says the variables are independent based on context
- Performs a chi-square test of independence
- Only states yes with no attempt at justification

4 **Complete Response** (EEE)
All three parts essentially correct

3 **Substantial Response** (EEP, EPE, EPP, IEE)
Part (a) essentially correct and parts (b) and (c) at least partially correct
OR
Part (a) incorrect and parts (b) and (c) essentially correct

2 **Developing Response** (EEI, EIE, EPI, EIP, IEP, IPE, IPP)
Part (a) essentially correct and one of parts (b) and (c) essentially correct
OR
Part (a) incorrect and both parts (b) and (c) at least partially correct

1 **Minimal Response** (EII, IPI, IIP, IEI, IIE)
Part (a) essentially correct and parts (b) and (c) incorrect
OR
Part (a) incorrect and one of parts (b) and (c) at least partially correct