

Test Directions: Do not round intermediate steps unless asked to do so. Rounding instructions for final answers are specified in each question. The option E) NOTA means *None of the Above* answers are correct. This is a test heavily dependent on supplementary information provided in the test, so be sure to thoroughly read any extra directions given.

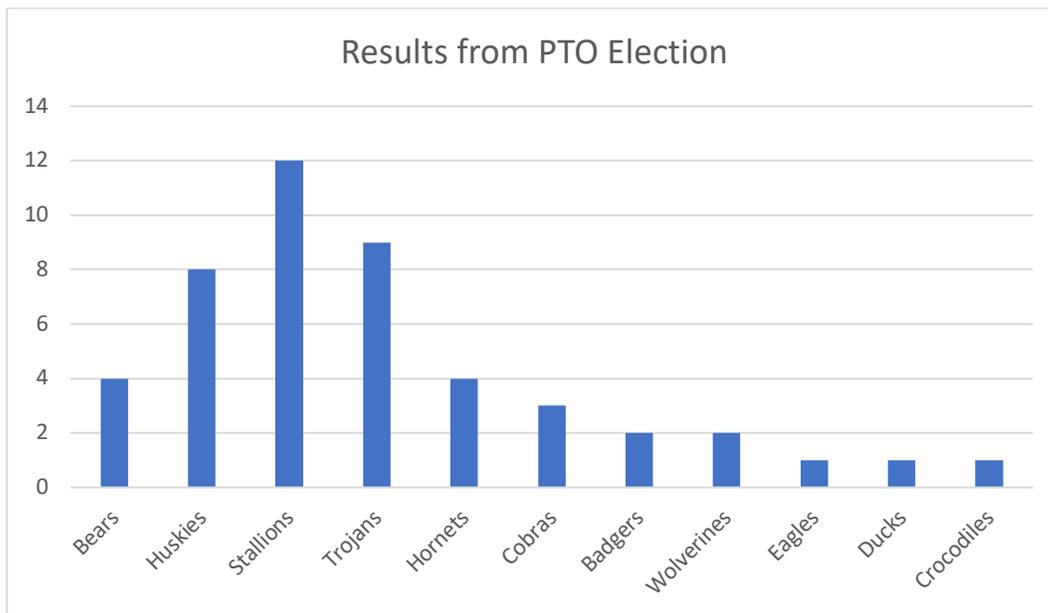
1. Jerry is interested in finding the distribution of birth month based upon a random sample of students at his very large middle school. He splits his school into groups by grade (6th grade, 7th grade, and 8th grade) and then takes an SRS of 3 classrooms for each grade. Within each randomly selected classroom, Jerry then takes an SRS of 5 students. Which of the following types of sampling methods has Jerry just performed?

- A) Stratified B) Cluster C) Convenience D) Multistage E) NOTA

2. The data obtained for the variable of interest in the previous question (namely, birth month) is best classified as having which of the following levels of measurement?

- A) Ratio B) Interval C) Nominal D) Ordinal E) NOTA

3. A local high school wishes to switch their mascot. The PTO holds an election where a simple plurality vote is required to win. The results are documented in the chart below. The vertical scale is the number of votes cast for each mascot choice.



The distribution above is best described as which of the following?

- A) Skewed Left B) Skewed Right C) Symmetric D) Center-Skewed E) NOTA

4. If we consider a random variable X , such that X follows an exactly normal distribution with a mean of 1 and standard deviation of 2, and we desire to create a graph of the distribution such that the X-Axis represent the domain of random variable X and the Y-Axis represent the proportion of observations at or below each value in the domain of random variable X , which of the following best describes what we have just created?

- A) A graph of the cumulative distribution function of X C) A line graph of X E) NOTA
 B) A graph of the probability density function of X D) A histogram of X

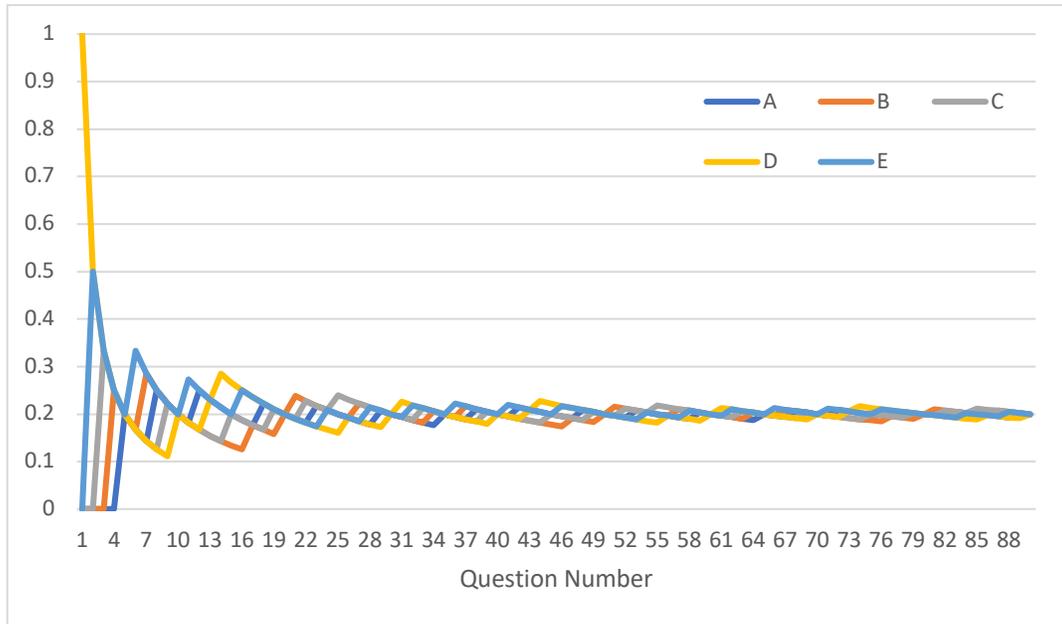
5. Consider the graph of the probability density function for random variable X from the previous problem. Jeff wants to find how the graph of random variable X is affected by transformations. If he were to multiply each individual value in the domain of X by 2 and then add 1, what is now the area under the transformed density function?

- A) 3 B) 1 C) 0 D) 2 E) NOTA

6. I am practicing for my baseball league’s Home Run Derby by hitting pitches thrown by a fellow teammate. My goal for the competition is to hit at least 5 home runs in a single batting practice (a set of 10 pitches). If the probability that I hit a home run on any given, independent pitch is 0.4, what is the probability that I reach my goal in my next practice, to four decimal places?

- A) 0.3669 B) 0.1662 C) 0.8338 D) 0.6331 E) NOTA

Questions 7-8 Refer to the Following Chart and Description:



This chart above shows the proportion of each correct answer (on the y-axis) for all questions up to and including a certain question number (on the x-axis) for a 90 question English test. The test has five answer choices for each question and for a given question; any of the 5 choices are equally likely to be the correct answer.

7. Considering both the provided chart and given description above, what additional information is still needed to properly interpret the chart?

- A) Vertical Gridlines B) Y-Axis Scale C) A Significance Level (α) D) Legend E) NOTA

8. The behavior of the relative frequencies of each answer choice as question number increases, as displayed in the provided chart, is a consequence of which of the following?

- A) Bayes' Theorem B) Law of Averages C) Law of Large Numbers D) Central Limit Theorem E) NOTA

9. Joseph knows that some Statistics FAMAT tests will ask about the names of various theorems and paradoxes, so he begins researching. In his research, he discovers this generalized definition on Wikipedia:

“_____ paradox is a phenomenon in probability and statistics, in which a trend appears in several different groups of data but disappears or reverses when these groups are combined.”

Identify the last name of the individual that correctly completes the Wikipedia definition, and then find the number of distinct permutations of the letters in this individual’s last name. For this question, assume an uppercase letter is indistinguishable from a lowercase one. Additionally, do not include the possessive apostrophe and s (’s) in your count of letters.

- A) 720 B) 1260 C) 2520 D) 6720 E) NOTA

Questions 10-13 Refer to the Following Information:

The playoff bracket for my basketball league has just been finalized! The single-elimination bracket is provided on this page, with my team being the first-seeded Alligators. All teams are labeled on the bracket by their seed number followed by their team name. All series are the best of five games, meaning that the first team to three wins in a series moves on to the next round. An interesting point about the teams and games in this tournament is that the odds against a team winning for any particular game is represented by the seed numbers as a ratio. For example, in my team’s first series against the fourth-seeded “Dragons”, the odds against us winning a game in the series are 1:4, meaning the odds in our favor are 4:1, and thus the probability that my team wins any particular game in that series is 4/5. This is true for all games in every series, regardless of the previous or future outcomes in the tournament. Thus, all games are independent of each other.



10. What is the probability, to three decimal places, that the Dragons win any given game against the Cheetahs, given that both teams each win their first series?
- A) 0.429 B) 0.750 C) 0.571 D) 1.333 E) NOTA
11. What is the probability, to four decimal places, that the Cheetahs win 2 games or fewer in their first 4 games against the Bears? That is, they have played exactly 4 games out of a possible total of 5 games with a potential 5th game pending and so either the series ends at 3-1 in favor of the bears or is tied at 2-2 with the series deciding 5th game yet to be played.
- A) 0.8208 B) 0.6912 C) 0.5248 D) 0.6048 E) NOTA
12. In the first series of games pitting the Alligators against the Dragons, what is the expected number of games played until a winner is decided to four decimal places?
- A) 3.5871 B) 4.3928 C) 3.6336 D) 3.7820 E) NOTA
13. What is the probability, to three decimal places, that the Alligators will beat the Bears in the Championship series in 4 games or fewer?
- A) 0.191 B) 0.593 C) 0.381 D) 0.296 E) NOTA
14. Consider the discrete, uniform random variable X defined on a finite set of n distinct integer values. Which of the following answer choices is equivalent to the expression $\frac{(y-\mu_x)\sqrt{n}}{\sqrt{\sum(x_i-\mu_x)^2}}$ for some outcome y in the domain of X ?
- A) μ_x B) Z_y C) \sqrt{n} D) σ_x E) NOTA

15. I went trick-or-treating this last Halloween, and I managed to get an extremely large amount of candy that I stored in my pillowcase. Unfortunately, I only got 12 unique types of candy, each of which has equal probability of being selected from my pillowcase. My sister starts randomly picking one piece of candy at a time from my pillowcase, keeping the candy regardless of what type it is, and will continue to do so until she has one of each type. What is the standard deviation, to three decimal places, of the number of pieces of candy my sister will have by the time she has one of each type? Once again, since the amount of candy is sufficiently large, we can treat subsequent selections of pieces of candy as independent of each other.

- A) 37.239 B) 13.716 C) 29.571 D) 6.102 E) NOTA

16. Joe and Sally are playing 5-card poker. Because Joe happens to be a fan of aces, he places a second ace of spades into his original standard deck of 52 cards, shuffles the cards, and plays like normal. The second ace of spades is identical in every way to the original ace of spades. After playing for a while, Sally becomes suspicious of Joe’s obscure gameplay, as she doesn’t know about the 53rd card in the deck. Agitated, Sally grabs the deck from Joe and draws each card one by one, without replacement, to inspect them and discover whatever Joe may be hiding. What is the expected number of cards Sally must draw from the deck until she discovers both aces of spades.

- A) 35 B) 36 C) 37 D) 38 E) NOTA

17. I’m playing a game of Dungeons and Dragons with my friends. A particularly notorious situation for me requires rolling a fair 20-sided die, numbered 1 to 20, ten times. In my turn, at least one roll of 18 or higher in the ten rolls of the die is considered a success. To three decimal places, what is the probability that I am successful?

- A) 0.803 B) 0.450 C) 0.833 D) 0.768 E) NOTA

18. I have a bowl filled with 9 blue-colored erasers, 6 green-colored erasers, and 11 red-colored erasers. I remove erasers from the bowl and put them in my pocket for later. What is the probability, to three decimal places, that I select at least two erasers of each color if I plan to remove seven erasers in total, without replacement?

- A) 0.181 B) 0.316 C) 0.135 D) 0.301 E) NOTA

19. The Theil-Sen Estimator is another method of obtaining a sample slope for linear regression as opposed to the usual method of least-squares regression. The method involves selecting all possible pairs of points, getting the slope of the line connecting each pair, and then obtaining the median of all of those obtained slopes. This is beneficial for larger sample sizes because the procedure is rather robust with respect to a few regression outliers. Find the absolute difference between the Theil-Sen Slope and the least-squares regression slope for the following values for X and Y rounded to two decimal places.

X	2	6	10	14
Y	6	10	6	18

- A) 0.27 B) 2.60 C) 0.20 D) 0.30 E) NOTA

Questions 20-21 consider the distributions for independent, discrete random variables X and Y shown below:

X	5	6	7	8	9	10	11
P(X)	.2	.15	.15	.05	.2	.15	.1
Y	11	10	9	8	7	6	5
P(Y)	.2	.1	.25	.05	.15	.15	.1

20. Find $E(2Y + 3X)$. Round to two decimal places.

- A) 39.85 B) 40.40 C) 40.13 D) 32.10 E) NOTA

21. Find the absolute difference between $Var(2Y + 3X)$ and $2Var(Y) + 3Var(X)$. Round to four decimal places.

- A) 32.4350 B) 53.7275 C) 0.0000 D) 33.1450 E) NOTA

22. In a traditional game of 501 darts, players throw their darts and obtain points that lower their score from 501 down to exactly 0 (no more, no less). To make this even harder, the final throw to hit 0 must be a “double value” such as Double-12 or Double-20. The nine-dart finish is one of the most incredible feats in the world of darts and requires scoring all 501 points on 9 darts (the minimum number of darts to do so). One such finish involves getting five Triple-20’s, two Triple-19’s, one Triple-17, and one Double-18. In how many ways is this particular nine-dart finish possible?

- A) 3024 B) 1512 C) 336 D) 168 E) NOTA

23. Many have heard of the beloved Monty Hall Problem, where you are tasked with picking one door out of three and will win whatever is behind the door you choose. Behind one of the doors is a brand-new car (considered a success), while behind the two others are goats (which is considered a failure). All the while, the game show host, Monty, knows exactly what is behind each door. You randomly select one door, and then Monty opens one of the remaining doors to reveal a goat. If you were to stay with the door you originally chose, your probability of winning the car would be 1/3. If you switch, your probability of winning dramatically rises to 2/3. Some are also familiar with such expansions of this problem like where you start with four doors, three of which have goats. Instead, consider the expansion where you start with 10 doors, 7 of which have goats while 3 have cars, and you are to select two doors at a time instead of 1. Instead of revealing 1 door, Monty will now reveal 2 doors. After Monty has revealed two of the goats, you are tasked with switching both of your chosen doors to completely different ones, switching only one of them, or not switching any doors at all. What is the expected number of cars you will win, to three decimal places, given that you will always switch both of your doors to completely different ones?

- A) 0.750 B) 0.800 C) 0.779 D) 0.834 E) NOTA

24. The Adjusted R-Squared, given by the formula $\bar{R}^2 = 1 - (1 - R^2)\left(\frac{n-1}{n-p-1}\right)$ where n is the sample size and p is the number of explanatory variables, is a statistic seen on many regression output tables and attempts to account for a phenomenon where the R-Squared fluctuates as more variables are added to a regression analysis. Considering simple, bivariate linear regression, under which of the following conditions will $\bar{R}^2 = R^2$?

- A) when $n = 1$ C) when $n = 20$ E) NOTA
 B) when $n = 10$ D) when $n = 30$

Questions 25-26 Refer to the Following Information:

In ordered datasets, quartiles are resistant measures that capture approximately 25% of the data between successive quartiles. Deciles, on the other hand, are similarly resistant measures that capture approximately 10% of the data between successive deciles. For your reference, the IDR (Inter-Decile Range) captures approximately the middle 80% of data similar to how the IQR captures the middle 50%.

25. The $1.5 * \text{IQR}$ Rule establishes upper and lower fences for which values beyond such fences are considered outliers. If, on the Standard Normal Distribution, we were to get Z-Scores for Q1 and Q3, we can establish a Z-Score cutoff for which Z-Scores beyond the fences would be considered outliers according to the $1.5 * \text{IQR}$ Rule. The outlier interval, according to this concept, can be written in the form $(-\infty, -L) \cup (L, \infty)$. What is the appropriate value for L ? Round both your critical Z-Scores for your quartiles and then your final answer each to two decimal places.

- A) 2.79 B) 3.00 C) 2.70 D) 2.68 E) NOTA

26. If our goal is to have the same Z-scores for the fences, and thus the same value for L and $-L$ for determining outliers as discussed in Question 26, but we would like to use the Inter-Decile Range and call this the $K * \text{IDR}$ Rule, what value for K should we use? Round your critical Z-Scores for your quartiles to two decimal places and round your final answer to three decimal places.

- A) 0.547 B) 1.500 C) 0.553 D) 1.089 E) NOTA

27. Below is the output table from a linear regression test for $\ln(\text{mass})$ vs. $\ln(\text{radius})$ of various Super Bouncy Balls bought from my local convenience store, whose relation shows a very strong linear pattern. Mass is measured in grams and radius is measured in decimeters.

	Coefficients	Standard Error	t-Stat	P-value
Intercept	4.49980967	0.000	-1.94308E+15	0.000
$\ln(\text{radius})$	3.0237	0.000	2.46018E+15	0.000

I pick out one of these bouncy balls and find that it has a mass of 6 grams and a radius of 0.4 decimeters. What is its residual, to the thousandths place, according to our model?

- A) 0.291 B) -0.364 C) 0.364 D) -0.291 E) NOTA

28. How many positive integers less than 14519 are divisible by 3, 11, or 19?

- A) 6922 B) 6160 C) 6183 D) 6137 E) NOTA

29. In bivariate linear regression where \hat{y}_i represents the predicted value of a response variable for a given value of an explanatory variable, *Sum of Squared Residuals* = $SS_{res} = \sum(y_i - \hat{y}_i)^2$ is the total amount of unexplained, or undetermined, variance in a response variable from its linear relationship with the explanatory variable and *Total Sum of Squares* = $SS_{tot} = \sum(y_i - \bar{y})^2$ is the total amount of variance from the mean of a response variable. Given that $SS_{tot} = 1437500.0$ and $SS_{res} = 231357.0$ for some positive, linear relationship between two variables X and Y , find the coefficient of correlation, r rounded to three decimal places.

- A) 0.401 B) 0.916 C) 0.704 D) 0.921 E) NOTA

30. What is the symbol for the equivalent population parameter for the sample statistic denoted by the letter “s”?

- A) ρ B) β_0 C) β_1 D) σ E) NOTA