*For all questions, answer choice (E) NOTA means that none of the given answers is correct. Good Luck!*

1. A and B are two correlated random variables. Given that $\mu_A = 16$, $\sigma_A = 5$, $\mu_B = 12$, and $\sigma_B = 2$, what is the probability that $A + B$ is between 25 and 30.
   (A) 0.3561   (B) 0.6342   (C) 0.5205   (D) 0.4956   (E) NOTA

2. Rahul and Anirudh go to the DoubleTree gym and find a bunch of Indians dancing. Due to this distraction, the amount they can bench press is significantly altered. They bench press once each. The sum of the z-scores of their individual bench presses relative to the population distribution of bench presses (lbs)($\mu = 123.86, \sigma = 38.6$) is 3.86. Rahul bench presses twice as much as Anirudh. What is the average of their bench presses? (round only the final answer to the nearest whole number)
   (A) 189 lbs   (B) 203 lbs   (C) 196 lbs   (D) 198 lbs   (E) NOTA

3. Siddharth, the god of chemistry, always gets bored during chemistry class. To make the best use of his time and reduce the boredom, Siddharth decided to draw a 3×3 square grid (it consists of 9 unit squares), and color it in randomly using his 3 favorite colors: razzmatazz, xanadu, and caput mortuum. Assuming that every square is colored, and each one only has one color on it, what is the probability of the grid not having a 2 by 2 square completely filled with one color? Express your answer as a decimal rounded to the nearest thousandth.
   (A) 0.913   (B) 0.869   (C) 0.814   (D) 0.836   (E) NOTA

4. In his quest to achieve the perfect body, Rahul conducts an experiment to find the most effective conditioning method. He gathers a group of 100 high school seniors from Florida. The students are first split based on their gender. Then, for no particular reason, Rahul splits the students based on hair lengths of $0 - 10$ inches, $10 - 20$ inches, and greater than 20 inches. All the students were put through intensive training for a week. Students ran on either a $5, 10$, or 20 mile trail, and ran on the same trail for either $3, 5$, or 7 days a week. There was also a control group which did not run at all. Based on the given information, how many treatments did the experiment have?
   (A) 18   (B) 6   (C) 27   (D) 54   (E) NOTA

5. A troupe of Indian girls masquerading as clowns decide to go bowling with the hope of being accepted into the PBL. In a round of bowling, the six, of presumably equal skill, post scores of $98, 94, 88, 82, 76, 72$. To be recognized as a legitimate team, the average of the team's score must be at least 100. Tanvi is quite enraged because she believes the players' average score is exactly 110. Find the power of the alternative of the significance test that the given round of bowling is a significant deviation from the team's supposed average score of 110 at the 0.05 level. (Round intermediate steps to three decimal places)
   (A) 0.89   (B) 0.69   (C) 0.99   (D) 0.79   (E) NOTA

6. A hospital wishes to test the hypotheses for paramedic response times:
   H_0: $\mu = 6.4$ minutes
   H_A: $\mu > 6.4$ minutes
   The hospital decides to carry out the test using a 5% significance level. If they reject the null hypothesis, the hospital will implement new policies to help lower the response times. If they fail to reject the null hypothesis, the policies will remain the same. For the patients, which is the more serious error: a Type I error or a Type II error? Select the answer choice with the correct explanation corresponding with the correct error.
   (A) Type I error: The hospital will unnecessarily have to implement new policies.
   (B) Type II error: The hospital will unnecessarily have to implement new policies.
   (C) Type I error: The policies will remain the same and result in response times greater than 6.4 minutes.
   (D) Type II error: The policies will remain the same and result in response times greater than 6.4 minutes.
   (E) NOTA
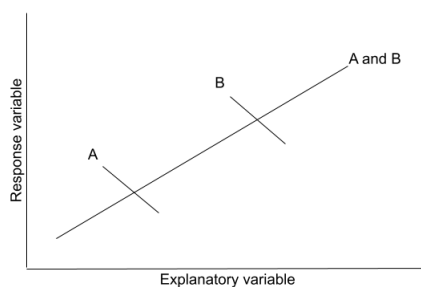
7. What phenomenon does the graph above illustrate? (all lines are data-modeled linear regression lines of the sets)

    (A) Lagrange Error Bound
    (B) Residual failure
    (C) Type I and/or type II error
    (D) Simpson's paradox
    (E) NOTA

8. Martin would like to increase the output of eggs from his chicken coop. He randomly selects 10 of the hens from his coop to use in an experiment. He gives each of the 10 chickens their normal diet for one month. For the next month he gives them a special, homemade diet. Below are the results for how many eggs each hen laid for each month.

Hen Name	Eggs laid in month of normal diet	Eggs laid in month of special, homemade diet
Firestar	21	29
Flitz	27	35
Meat	17	21
Joey	35	50
Princess Peach	19	25
Red	24	23
Rainbow	15	17
SharQeisha	22	25
Seandavius	26	33
Tiana	25	25

What is the $p-$value of the test to determine if there is significant evidence that the special, homemade diet increases the amount of eggs the hens lay? Assume that all the conditions for the test have been met. Round your answer to 4 decimal places.

   (A) 0.0029   (B) 0.0033   (C) 0.0066   (D) 0.0766   (E) NOTA

9. The writers of this test are permitted to have exactly 2 or 3 easy questions, $6, 7, 8,$ or 9 medium questions, $13, 14, 15,$ or 16 challenging questions, and 5 hard questions for a total of exactly 30 questions. Suppose this test has 4 easy questions, 6 medium questions, 16 challenging questions, and 4 hard questions. Tanvi wants to incriminate the test-writers for breaking the rules, so she finds the lowest possible p-value in a χ^2 goodness of fit test of the distribution of question difficulty using the designated parameters against the actual test. Assuming she did so correctly, what approximate p-value did she calculate? (Hint: Tanvi manipulates the expected counts while still complying with the rules)

   (A) 0.27   (B) 0.53   (C) 0.32   (D) 0.69   (E) NOTA

10. For a legitimate uniform probability density function of X with domain $[0, 5]$, which of the following are true statements:

   I. $P(X = 3.5|X > 2) = P(X < 0||X > 5)$
   II. The range of the function is constant as the domain grows.
   III. The probability density function and the cumulative distribution function of a uniform distribution are the same.

(A) I and II only   (B) II only   (C) II and III only   (D) I, II, and III   (E) NOTA

11. Chief reporters Yang and Hu conducted a survey of all 418 undergraduate students enrolled at the Indian Institute of Science in Bangalore, Karnataka to obtain information on the student preferences of 3 celebrities: Shahrukh Khan (SRK), Justin Bieber (JB), and Virat Kohli (VK). Every undergraduate student was required to reply to the survey. Given that 237 students said that they liked SRK and VK, 138 students said that they liked JB and SRK, 120 students said that they liked all 3 celebrities, 285 students said that they liked SRK, 133 students said that they liked VK and JB, 270 students said that they liked VK, and there were 3 times as many students who did not like any of the celebrities than who liked only JB, what is the expected value of students who like JB, if the results of the survey are extrapolated to the entire student body of 3743 students at the institute? The answer should be rounded to the nearest whole person.

     (A) 1576   (B) 3562   (C) 2552   (D) 1781   (E) NOTA

12. In his collection of data from Lake Munson, Anirudh finds a pronounced pattern in the residuals of his data in the relationship between variables X (independent) and Y (dependent). He further investigates this, finding that, for his raw data, the coefficient of determination using an exponential model is greater than that of a power model. For his upcoming presentation, how should Anirudh plot the explanatory variable against the response variable of his original data to best achieve approximate linearity?

     (A) X vs. Y (B) $\log(X)$ vs. Y (C) X vs. $\log(Y)$ (D) $\log(X)$ vs. $\log(Y)$ (E) NOTA

13. A valid probability distribution on the interval $[-a, a]$ can be modeled by the function $f(x) = \sqrt{a^2 - x^2}$. Find a.

     (A) $\sqrt{\dfrac{2}{\pi}}$ (B) $\sqrt{\dfrac{\pi}{2}}$ (C) $\sqrt{\pi}$ (D) $\sqrt{2}$ (E) NOTA

14. Suppose a random variable x takes only values of 0 and 1 at an equal rate. What is the standard deviation of x?

     (A) 0   (B) 0.5   (C) $\sqrt{0.5}$ (D) $\dfrac{1}{12}$ (E) NOTA

15. Rohan is playing basketball. He makes 750 out of his last 1000 free throw attempts. Assuming every free throw attempt is independent and we can view the 1000 attempts as a SRS, which of the following is the best point estimate for the probability that he makes the very next free throw attempt and why?

     (A) $\dfrac{1}{2}$ - because he either makes it, or he doesn't   (B) $\dfrac{3}{4}$ - because of the Central Limit Theorem

     (C) $\dfrac{1}{2}$ - because of the binomial setting   (D) $\dfrac{3}{4}$ - because of the law of large numbers   (E) NOTA

16. Consider a Markov chain that has 3 states, $S = 1, 2, 3$, and a transition matrix of $P = \begin{bmatrix} \frac{1}{8} & \frac{3}{8} & \frac{1}{2} \\ \frac{1}{7} & \frac{3}{7} & \frac{3}{7} \\ \frac{4}{5} & \frac{1}{20} & \frac{3}{20} \end{bmatrix}$

What is the stationary distribution for this Markov chain? Round each component of the vector to the nearest ten-thousandth.

     (A) $(0.2458, 0.4218, 0.3324)$ (B) $(0.3701, 0.2740, 0.3559)$ (C) $(0.1765, 0.3872, 0.4363)$
     (D) $(0.3913, 0.2817, 0.3270)$ (E) NOTA

17. Three distributions walk into a bar: a normal distribution centered at 5 with some positive s, a binomial distribution with $p = 0.75$ and some integer n greater than 25, and a geometric distribution with $p = 0.001$. The bartender, a bimodal distribution with a complex, decides not to serve one of the three distributions because its general shape is not like the others based off of skewness. Who is left parched?

     (A) The normal distribution   (B) The binomial distribution   (C) The geometric distribution
     (D) Neither of the three   (E) NOTA

18. The Cleveland Cavaliers are in the 2020 NBA Finals. Down by 6 points with 1 minute left in the game, Cleveland needs to win this game. Right now, Kevin Porter Jr. of the Cavaliers is at the free throw line to shoot one free throw (he makes 52% of all free throws he takes). When the game resumes, exactly one of the following possibilities will occur within each 30 second increment until the conclusion of the game: neither team scores (40% chance), Cleveland scores 3 points (20% chance), the opposing team scores 2 points (25% chance), Cleveland scores 2 points and Porter Jr. attempts a free throw (15% chance). If Cleveland leads the game at any point, they will win the game. What is the probability that Cleveland wins? (in the case of a drawn game, regard it as a loss for Cleveland; round only the final answer)
    (A) 0.008   (B) 0.02   (C) 0.08   (D) 0.04   (E) NOTA

19. Rayyan wants to compare the effectiveness of his tutoring service with that of Siddhi's. He took a random sample of 459 students from his service and another random sample of 554 students from Siddhi's service. At the end of the year, only 40 of the students from Rayyan's sample passed the final exam, while 400 of the students from Siddhi's sample passed the final exam. Compute the standard error for the significance test. Round your answer to the nearest ten thousandth place.
    (A) 0.0231   (B) 0.6591   (C) 0.0469   (D) 0.0313   (E) NOTA

20. What is the standard deviation of the set which consists of A, B, C, and D? Round the final answer and each of the values in the set to the nearest hundredth place.

    $A = \sigma$ of a binomial distribution with 12 observations and probability of success 0.3
    $B = \sigma$ of a geometric distribution with probability of success 0.7
    $C =$ Expected value of the discrete random variable

X	1	2	3	4	5	6
$P(X)$	0.1	0.15	0.05	0.4	0.2	0.1

    $D =$ The slope of a regression line with t statistic 2.05 and standard error 0.55
    (A) 9.54   (B) 0.29   (C) 1.00   (D) 0.33   (E) NOTA

21. Tim Tebow and Tom Brady, the two greatest football players of all time, want to maintain their fitness levels by running. Tebow is initially located at the point Tim Tebow and Tom Brady, the two greatest football players of all time, want to maintain their fitness levels by running. Tebow is initially located at the point $(15, 11)$ and Brady is initially located at the point $(-1, 35)$ in a standard Cartesian plane. Before running, Tebow drank a gallon of Gatorade, which energizes him to have rapid initial speed as well as having a head start. As a result, Tebow can do 3 moves that are each 2 units in length in a single direction before Brady starts to run. During the normal running, Tebow and Brady can each only run 1 unit at a time, and Tebow must always run in either the positive y direction or the negative x direction, while Brady must always run in either the negative y direction or the positive x direction. Brady wants to meet up with Tebow while running, while Tebow wants to avoid Brady at all costs and go to the point $(-1, 35)$. Each player optimizes his movements such that his goal has the highest probability of being met. Each movement of one person is independent of that of the other person as they move and decide at the same time. What is the probability that Brady succeeds in meeting up with Tebow before Tebow reaches the point $(-1, 35)$?
    (A) $\dfrac{1}{6777216}$ (B) $\dfrac{1}{65536}$ (C) $\dfrac{1}{262144}$ (D) $\dfrac{1}{16384}$ (E) NOTA

22. Samay has a set of chores to complete over the weekend. He must wash the dishes, clean his room, and repair his trusty trident. He takes an average of 37 minutes with a standard deviation of 10 minutes to wash the dishes, an average of 28 minutes with a standard deviation of 15 minutes to clean his room, and an average of 159 minutes with a standard deviation of 50 minutes to repair his trident. Each of these events are independent of each other. The time it takes to complete all three of these events is normally distributed. What is the probability that Samay finishes between 3 and 4 hours?
    (A) 0.4144   (B) 0.5643   (C) 0.6845   (D) 0.3058   (E) NOTA

23. What is the mean and standard deviation of the following χ^2 distribution? Round your answer to the nearest thousandth place.

	Pizza	Hot Dog	Hamburger	Taco	Salad
9th Grade	20	35	16	36	6
10th Grade	41	24	19	27	8
11th Grade	27	15	31	14	11
12th Grade	34	29	13	27	4

(A) $\mu = 0$, $\sigma = 0$ (B) $\mu = 437$, $\sigma = 116.766$ (C) $\mu = 12$, $\sigma = 4.899$
(D) $\mu = 21.850$, $\sigma = 10.806$ (E) NOTA

24. Which of the following is/are true?
I. Sampling distributions are always normal.
II. A procedure is robust if the calculations remain fairly accurate when a condition is violated.
III. The coefficient of determination is the fraction of the variation in the values of y that is explained by the least-squares regression line.
IV. Normal distributions are uniform, single-peaked, and bell-shaped.
(A) I, II, and III only   (B) I, II, III, and IV   (C) IV only   (D) II and III only   (E) NOTA

25. Sanjita decided that when she is older, she wants to get a pet walrus and name it Shagun. Preparing for the future selection, Sanjita conducts a study to measure the effect of a persons age on his or her perception of a walrus′s cuteness. There are two groups that each consist of 77 randomly selected people. Each person rates pictures of walruses on their level of cuteness from a range of $1 - 12$. One group, labeled as the "young group", consists of people who are aged from $5 - 18$ years, while the other group, labeled as the "old group", consists of people who are aged from $60 - 100$ years. The mean values of the cuteness levels for the young group and the old group are 9.8 and 9.6 respectively. The standard deviations of the cuteness levels for the young group and the old group are 0.6 and 0.2 respectively. What is the $p-$value, rounded to the nearest ten-thousandth, of the significance test that can be performed to determine the probability that the young group will rate and find walruses to be cuter than the old group does?
(A) 0.0054   (B) 0.0033   (C) 0.0047   (D) 0.0072   (E) NOTA

26. Bill Li, a physics prodigy, decides that he wants to put his skills and knowledge to the test. He wants to prove his "Hot Water Postulate" using techniques derived from Lipid Combinatorics. During his proof, Bill needs to evaluate what $\sum_{k=0}^{125} \binom{500}{4k}$ is equal to. What value should Bill get?
(A) $2^{499} + 2^{249}$ (B) $2^{499} - 2^{249}$ (C) $2^{500} + 2^{250}$ (D) $2^{498} + 2^{249}$ (E) NOTA

27. Given that $A = 0.6$, $B = 0.5$, $C = 0.4$, $P(A \cup B) = 0.9$, and the events A and C are independent variables, find $P(A \cap B) + P(A \cup C)$.
(A) 1.24   (B) 0.96   (C) 0.8   (D) 0.54   (E) NOTA

28. Jason loves to ponder around and find new generalizations and patterns. He first evaluates the ceiling of the reciprocal of the correlation, $\frac{1}{r}$, which is represented by variable x, between the list of numbers n^n (mod 11) such that n is an integer from 1 to 100 inclusive and the list of numbers that are n^n (mod 11) such that n is an integer from 101 to 200 inclusive. Jason then takes x and computes the expected value of selecting one of the elements from the set of positive integers less than or equal to 20 that are congruent to x^x (mod 5), where each element has an equal chance of being selected. What value does Jason get?
(A) 11.5   (B) 10.5   (C) 8.5   (D) 9.5   (E) NOTA

29. In order to fulfill his lifelong dream of understanding the effect of a lack of sleep on intelligence, Rayyan carries out an experiment to test it so that he can further his knowledge. He took 1205 individuals and deprived them of sleep for 7 consecutive days. He then took 523 individuals who each slept exactly 11 hours for 7 consecutive days. Assume that all of the conditions and variables in the environment are held constant for both groups and all of the individuals in the sample were randomly selected. All of the subjects took an intelligence quotient exam before and after the 7 days. If the individual scored greater than 10 points less than the score they received before the experiment, then the individual was considered to have failed, which is the specific effect that Rayyan is trying to observe. 17% of the individuals who still got sleep failed, while 77% of the sleep deprived individuals failed. What is the pooled sample proportion of the test that can be performed? Round your answer to the nearest thousandth.

    (A) 0.352   (B) 0.479   (C) 0.588   (D) 0.648   (E) NOTA

30. After drinking 5 cups of coffee within an hour, Rayyan finally feels a surge of energy and has a desire to run around. Rayyan is currently in a circular room in his house and throughout the course of his running exercise, Rayyan runs to 7 distinct locations marked with a pole along the perimeter of the room. While Rayyan is running across, he is also carrying a long string with him that he connects between each pair of locations around the respective poles, which creates a line segment. As a result, the path Rayyan ran is mapped. The 7 locations were chosen optimally such that Rayyan ended up dividing the circle into the maximum number of regions possible, assuming that no more than 2 segments intersect at a point in the circle. Given that Rayyan decides to rest and sit down in one of those regions formed by the string after finishing the exercise, what is the probability that the region he sits in is a triangle? Round your answer to the nearest thousandth.

    (A) 0.593   (B) 0.547   (C) 0.614   (D) 0.531   (E) NOTA