These questions are a compilation of questions asked on previous exams or homework assignments. These problems are merely a representation of the type of questions that could be covered on the current exam. It is **not** an exact replication of what will be asked on the current exam nor should it be the only study reference you use to prepare for the exam. **For practice, I have provided far more questions than what will be on the actual exam.**
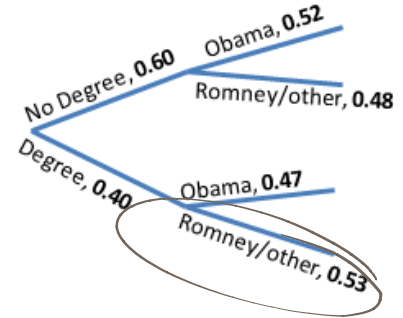
True and False: Questions 1-5.

1. Under a standard normal density 50% of the data falls below the mean.
   a. True
   b. False

2. For data that are right skewed, the mean is less than the median.
   a. True
   b. False

3. For any two events A and B, the probability of the intersection of the two events can be found by computing the product of the probabilities of each individual event. That is, $P(A \cap B) = P(A)P(B)$.
   a. True
   b. False

4. An observational study imposes treatments onto subjects or units for the purpose of measuring a response.
   a. True
   b. False

5. A confidence interval is *less precise* as the sample size, n, *increases*.
   a. True
   b. False

Use the following for the next two questions. In the 2012 presidential election, exit polls from the critical state of Ohio is summarized by the following tree diagram. "Degree" implies the voter had earned a college degree.

6. Given a randomly selected voter had earned a college degree, what is the probability they did not vote for Obama
   a. 0.40 $P(O'|D) = 0.53$
   b. 0.21
   c. 0.53
   d. 0.48

(Tree diagram)
No Degree, 0.60 — Obama, 0.52 ; Romney/other, 0.48
Degree, 0.40 — Obama, 0.47 ; Romney/other, 0.53

7. What is the probability a randomly selected individual voted for Obama?
   a. 0.52
   b. 0.99
   c. 0.47
   d. 0.50

$$P(O) = P(O \cap D) + P(O \cap D')$$
$$= P(O|D)P(D) + P(O|D')P(D')$$
$$= 0.47(0.4) + 0.52(0.6)$$
$$= 0.5$$

Use the following for the next two questions. An injection molding process for making detergent bottles uses three different machines the table gives the distribution for the number of machines operating at any given time.

y_i	0	1	2	3
$P(y_i)$	0.020	0.030	0.200	0.750

8. Find the probability that 2 or fewer machines are running at the same time.
   a. 0.200
   b. 0.050 $0.2 + 0.03 + 0.02$
   c. 0.250
   d. 0.75

9. Find the standard deviation of machines running.
   a. 0.398 $E(Y) = .03 + 2(0.2) + 3(.75) = 2.68$
   b. 0.631
   c. 2.68 $SD(Y) = \sqrt{\sum (y_i - E(Y))^2 P(y = y_i)}$
   d. 7.98

$$= \sqrt{\begin{array}{l}(2.68)^2(0.02) + (1-2.68)^2(0.03) \\ + (2-2.68)^2(0.2) + (3-2.68)^2(0.75)\end{array}} = 0.631$$

10. Solar-heat installations successfully reduce the utility bill 60% of the time. What is the probability that at least 9 out of 10 solar-heat installations are *successful* and will reduce the utility bill?
    a. 0.0464 $p = 0.6 \quad n = 10$
    b. 0.9432
    c. 0.0403
    d. 0.8429 $P(x=9) + P(x=10) = \binom{10}{9} 0.6^9 (0.4) + \binom{10}{10} 0.6^{10}$

11. The number of cars entering a roadway at a specific exit during the morning rush hour traffic (6AM to 9AM, M-F) is a random variable modeled by a Poisson distribution, with an average of 6 cars per minute. What is the probability the number of cars entering the roadway is between 4 and 6 cars per minute (inclusive)?

a. 0.8925
b. 0.4551
c. 0.2945
d. 0.1606

$$P(X=4) + P(X=5) + P(X=6)$$

$$= \frac{6^4 e^{-6}}{4!} + \frac{6^5 e^{-6}}{5!} + \frac{6^6 e^{-6}}{6!}$$

12. Based on the definition of a probability density function, why is this function invalid?

$$f(y) = \begin{cases} \frac{1}{15}(4-y), & for\ 1 \le y \le 5 \\ 0, & otherwise \end{cases}$$

a. Probability density functions can't be piece-wise functions.
b. The variable should be x not y.
c. The range of the random variable can only be from $-\infty$ to $+\infty$.
d. The function does not integrate to 1 over the sample space.

$$\int_1^5 \frac{1}{15}(4-y)\,dy = \int_1^5 \frac{4}{15}\,dy - \int_1^5 \frac{y}{15}\,dy = \frac{4}{15}y\Big|_1^5 - \frac{y^2}{30}\Big|_1^5$$

$$= 0.2667$$

Use the following for the next three questions. The random variable X is described by the following pdf:

$$f(x) = \begin{cases} 3x^{-4}, & x > 1 \\ 0, & otherwise \end{cases}$$

13. What is the **cumulative** density function?

a. $F(x) = 1$
b. $F(x) = -x^{-3}$
c. $F(x) = 1 - x^{-3}$
d. $F(x) = 3x^{-4}$

$$F(x) = \int_1^x 3t^{-4}\,dt = -t^{-3}\Big|_1^x$$

$$= -x^{-3} + 1$$

14. What is the probability X will be less than 1.2?

a. 0.634
b. 0.421
c. 0.579
d. 0.113

$$P(x < 1.2) = F(1.2) = 0.421$$

15. What value represents the 95th percentile?

a. 2.71
b. 1.65
c. 1.96
d. 1.17

$$F(x) = 0.95$$

$$0.95 = -x^{-3} + 1$$
$$-0.05 = -x^{-3}$$
$$0.05 = x^{-3}$$
$$x = 0.05^{-1/3}$$

Use the following for the next two questions. The length of time for one individual to be served at a restaurant is a random variable, X, with the following probability density function:

$$f(x) = \frac{1}{4}e^{-x/4} \text{ for } x \ge 0$$

16. Find the probability that an individual would wait longer than 10 minutes to be served?

a. ≈ 0.00
b. **0.08**
c. 0.94
d. None of the above.

$$P(x>10) = 1 - P(x<10) = 1 - \int_0^{10} \tfrac{1}{4} e^{-x/4}\, dx$$

$$= 1 + \left(-e^{-\frac{x}{4}} \Big|_0^{10}\right) = 0.08$$

17. Short answer. What value is the top 10% of times longer than? (90th percentile) Show work. Round answer to two decimal places.

$$0.9 = \int_0^x \tfrac{1}{4} e^{-t/4}\, dt$$

$$0.9 = -e^{-\frac{t}{4}} \Big|_0^x$$

$$0.9 = -e^{-\frac{x}{4}} + 1$$

$$-0.1 = -e^{-x/4}$$

$$0.1 = e^{-x/4}$$

$$(0.1)^{-4} = e^{x}$$

$$\ln(0.1^{-4}) = x$$

$$x = 9.21$$

Use the following for the next three questions. The Lightning Bolt Electric Car company claims their new car "The Buzzer" can travel on average 170 miles on one charge ($\mu = 170\ miles$). It is known that the population distribution associated with the distance per charge in normally distributed with standard deviation of 30 miles.

18. Draw the distribution. On the axis indicate the mean and $\pm 1, \pm 2, \pm 3$ standard deviations away from the mean.

80 110 140 170 200 230 260
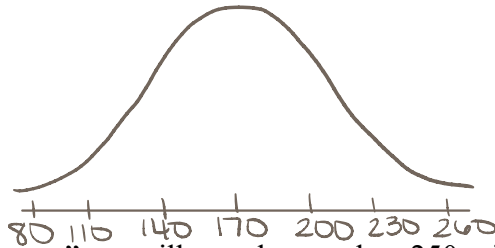
19. What is the probability a "Buzzer" car will travel more than 250 miles on a single charge?

   a. 0.0267
   b. 0.9733
   c. 0.0038
   d. 0.9962

$P(X>250) = 1 - pnorm(250, 170, 30)$

$= 0.0038$

20. Find the 95th percentile for buzzer car distances.

   a. 230 miles
   b. 219.4 miles
   c. 309.7 miles
   d. 211.2 miles.

$qnorm(0.95, 170, 30) = 219.4$

21. A taxi company offers shuttle service to the nearest airport. You look up the online reviews for the company taxis and find that there are 17 reviews, six of which report that the taxi never showed up. Is this a biased sampling method for obtaining customer opinion on the taxi service?

   a. The sampling method is biased because those who take the time to write an online review are likely to do so because they are satisfied with the service they received.
   b. The sampling method is biased because those who take the time to write an online review are likely to do so because they are upset with the service they received.
   c. The sampling method is not biased because those who take the time to write an online review are choosing to write the review without any incentive.
   d. This sampling method is not biased because it is random sample of people given most of the people won't know each other.

Use the following for the next four questions. The Calla Tulip are large tulips that are often included in a tulip bouquet for enhancement benefits. A large nursery grows tulips for florists. The height of the tulip is important in making bouquets. The nursery samples 40 tulips out of their field and measures their height.

Height of Calla Tulips

22. The distribution of tulip heights is

   a. bimodal.
   b. Right-skewed.
   c. Left-skewed.
   d. symmetric.
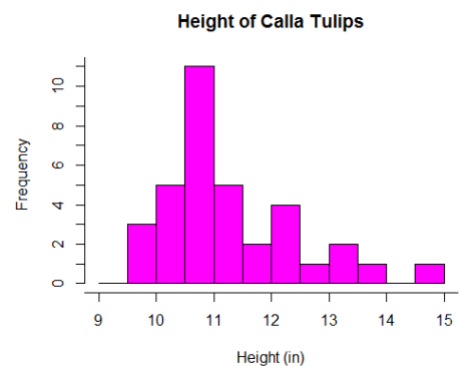
**23. Match the features of the boxplot with the corresponding letter. Shown in the plot below.**

Height of Calla Tulips
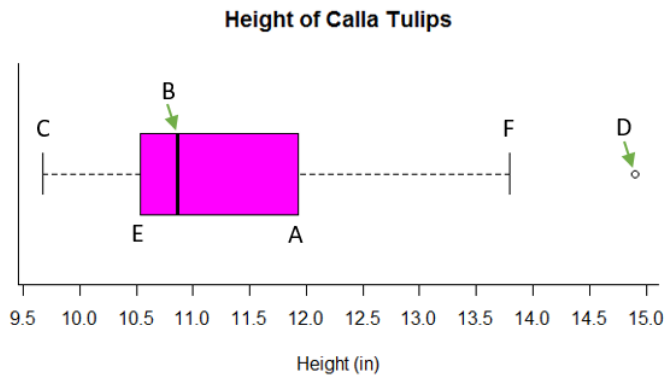
Feature	Matching Letter
Median	B
Outlier	D
First Quartile	E
Minimum (non-outlier)	C
Third Quartile	A
Maximum (non-outlier)	F

The graph depicts total fantasy football points earned by offensive players by the end of week 7 in the 2018 NFL Season by position.

NFL 2018 Offensive Players by Position

Fantasy Football Points by 10/24

24. From the plot which of the following statements is FALSE?
a. The minimum value is approximately the same for all positions.
b. There are no outliers in the quarterback (QB) position.
c. The median for running backs (RB) is approximately 25 total fantasy points.
d. Based on the distributions for each position it is safe to assume the average total number of points will be the same as the median total number of points.
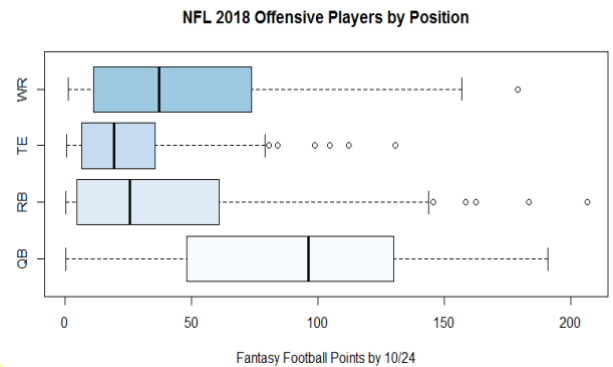
25. Which of the following statements is FALSE with respect to the standard deviation?
a. The standard deviation is the same as the variance.
b. The standard deviation is the average deviation from the mean.
c. The standard deviation is influenced by outliers in the data.
d. The standard deviation has the same units as the mean.

Use the following for the next two questions. The time it takes to heat a microwaveable burrito to an internal temperature of 160°F is normally distributed with a mean of 65 seconds and a standard deviation of 5 seconds.

26. What is the probability it will take less than 60 seconds to heat a burrito to 160°F?
a. 0.16
b. -1.00
c. 0.68
d. 0.18

$P(X < 60) = pnorm(60, 65, 5) = 0.16$

27. Suppose you have a bag of 4 burritos. Consider this bag to be a random sample of size 4. You cook each of the burritos to 160°F, what is the likelihood the *average* time will be less than 60 seconds?
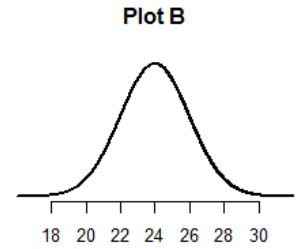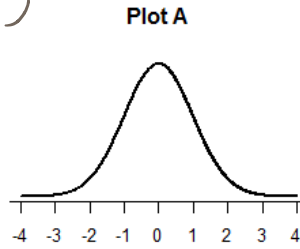a. 0.157
b. 0.029
c. 0.097
d. 0.023

$\bar{X} \sim N(65, 5/\sqrt{4})$

$P(\bar{X} < 60) = pnorm(60, 65, 5/\sqrt{4}) = 0.023$

Use the following for the next two questions. The time it takes to travel between two campuses of a university in a city via shuttle bus takes an average 24 minutes. The distribution of all times is positively skewed with a standard deviation of 12. Suppose a random sample of 36 transport times are recorded during a given week.

28. Which plot represents the distribution of sample means for a sample size of 36?
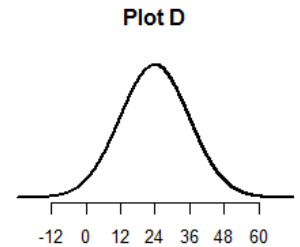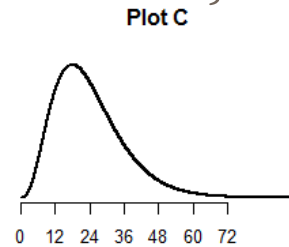
     a. Plot A $\qquad \bar{X} \sim N\left(24, \frac{12}{\sqrt{36}}\right)$

     b. **Plot B**

     c. Plot C

     d. Plot D

Plot A

Plot B

-4 -3 -2 -1 0 1 2 3 4

18 20 22 24 26 28 30

29. What is the likelihood the sample mean from the sample 36 transport times is greater than 30 minutes?

     a. 0.99865 $\quad P(\bar{X} > 30) = 1 - \text{pnorm}\left(30, 24, \frac{12}{\sqrt{36}}\right) = 0.00135$

     b. 0.50000

     c. 0.69146

     d. **0.00135**

Plot C

Plot D

0 12 24 36 48 60 72

-12 0 12 24 36 48 60

30. The **sampling distribution** of a statistic is

     a. the probability that we obtain the statistic in repeated random samples.

     b. the probability distribution of the randomly sampled statistic.

     c. the mechanism that determines whether randomization was effective.

     d. the extent to which the sample results differ systematically from the truth.

Use the following for the next three questions. A store randomly samples 603 shoppers over the course of a year and finds that 142 of them made their visit because of a coupon they received in the mail.

31. Before applying the central limit theorem and constructing a confidence interval, we need to check the sample size conditions for this particular sample. Check that the sample is large enough to proceed with confidence interval construction.

$$\hat{p}n = \frac{142}{603} \cdot 603 = 142 \geq 10 \checkmark \qquad (1-\hat{p})n = \frac{461}{603} \cdot 603 = 461 \geq 10 \checkmark$$

32. Compute the 95% confidence interval for the proportion of shoppers that visited the store because they received a coupon in the mail.

$$\frac{142}{603} \pm 1.96 \sqrt{\frac{\frac{142}{603}\left(\frac{461}{603}\right)}{603}} = (0.202, 0.269)$$

33. Interpret the confidence interval from the previous problem with context.

The 95% CI estimates the proportion of shoppers that visited the store because they received a coupon is between 0.202 and 0.269 with a point estimate of 0.235.