-Directions-

This competition includes 25 multiple-choice questions and three open-response questions that might be used as tie breakers. For questions 1 through 25 (the multiple-choice questions), mark your answer choice in the appropriate location on the sheet provided. After completing questions 1 through 25, answer each tie breaker question in sequential order (i.e., complete Tie Breaker #1 first, then Tie Breaker #2, and then Tie Breaker #3 last). Be sure that your name is printed on each of the tie break questions. When time is called, you will be asked to turn in your multiple-choice question answer sheet and your written responses to the tie breaker questions.

1. All of the following increase the width of a confidence interval **except**:

   a. Increased Confidence Level
   b. Increased Variability
   c. Increased Sample Size
   d. Decreased Sample Size

2. A 99% t-based confidence interval for the mean price for a drug (dollars per tablet) is calculated using a simple random sample of drug prices for 50 prescription drugs. Given that the 99% confidence interval is $3.01 < μ < $4.29, what is the sample mean price for a drug (dollars per tablet)?
   Please select the best answer of those provided below.

   a. $0.64
   b. $3.65
   c. Not Enough Information; we would need the prices for all 50 prescription drugs
   d. Not Enough Information; we would need the variation in the population of prices

3. The primary goal of a confidence interval is to provide a range of possible values for a **sample statistic**.

   a. True
   b. False
   c. Sometime True, Sometimes False
   d. None of the Above

4. During a two-week period, a surgeon has a mean completion time for three different surgeries of 4 hours and 32 minutes (i.e. 272 minutes) with a variance of 0 minutes. The next surgery performed by the surgeon was completed in 4 hours and 14 minutes (i.e. 254 minutes). What is the standard deviation for the four surgery completion times?

   a. 81 minutes
   b. 0 minutes
   c. 9 minutes
   d. 18 minutes

5. Provide the probabilistic definition for the significance level, α.
   a. $P(Reject\ \mathrm{H_0} |\ \mathrm{H_0}\ \text{True})$
   b. $P(Reject\ \mathrm{H_0} |\ \mathrm{H_0}\ \text{False})$
   c. $P(Fail\ to\ Reject\ \mathrm{H_0} |\ \mathrm{H_0}\ \text{True})$
   d. $P(Fail\ to\ Reject\ \mathrm{H_0} |\ \mathrm{H_0}\ \text{False})$

6. The Global Assessment of Functioning (GAF) is a numeric scale used by mental health clinicians to rate the social, occupational, and psychological functioning of an individual. Scores range from 100 (extremely high functioning) to 1 (severely impaired functioning). The distribution of observed GAF scores for several individuals is negatively skewed with a mean of 72 and a standard deviation of 15. If the set of all observed GAF scores are 'centered' (i.e., the mean for the variable is subtracted from each raw score), which of the following describes the shape, center, and spread of the new distribution of centered scores?

   a. Normally Distributed; with a mean of 0 and a standard deviation of 1
   b. Normally Distributed; with a mean of 0 and a standard deviation of 15
   c. Negatively Skewed; with a mean of 0 and a standard deviation of 1
   d. Negatively Skewed; with a mean of 0 and a standard deviation of 15

7. The statement "If there is sufficient evidence to reject a null hypothesis at the 10% significance level, then there is sufficient evidence to reject it at the 5% significance level" is:

   a. Always True
   b. Never True
   c. Sometimes True; the p-value for the statistical test needs to be provided for a conclusion

8. Suppose that X is a continuous random variable with a probability density function (pdf) $f(x) = \frac{1}{4}x^3, 0 \leq x \leq 2$. Is it more likely for X to be closer to two or to zero?

   a. Closer to Two
   b. Closer to Zero
   c. Neither
   d. Both have Equal Density

Assume the following probabilities for two events, A and B: $P(A) = 0.30, P(B) = 0.45$.
Use this information to answer questions 9 - 12 below.

9. Calculate and report $P(A \mid B)$ if the given events, A and B, are known to be independent.

   a. 0.135
   b. 0.300
   c. 0.450
   d. 0.615

10. Calculate and report $P(A \cap B)$ if the given events, A and B, are known to be independent.

    a. 0.135
    b. 0.300
    c. 0.450
    d. 0.615

11. Calculate and report $P(A \cup B)$ if the given events, A and B, are known to be independent.

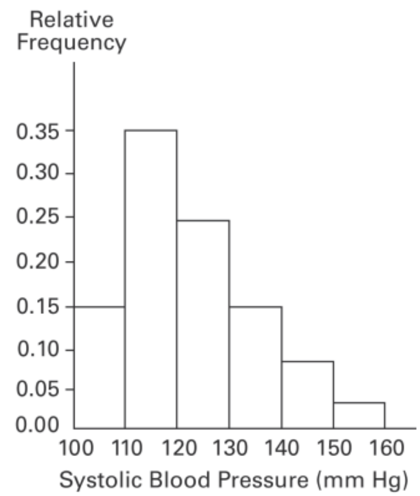
    a. 0.135
    b. 0.300
    c. 0.450
    d. 0.615

12. Are the events, A and B, disjoint in this situation?

    a. Yes
    b. No
    c. Cannot Be Determined
    d. None of the Above

For Questions 13 - 16, refer to the histogram, which shows relative frequency of systolic blood pressure for people aged 25 – 40 years.

Systolic Blood Pressure for People Aged 25 – 40 Years



13. How many individuals were included in this sample?
    a. 100
    b. It cannot be determined from the graph
    c. 160
    d. 1
    e. 6

14. What percentage of people aged 25 – 40 years have systolic blood pressure less than 110?
    a. 15%
    b. 0.15%
    c. 85%
    d. 0.85%
    e. 1.5%

15. Does this histogram appear to show a sample that follows a Normal distribution?
    a. No, the frequencies start off low, increase to a maximum, and then decrease to low again, and are roughly symmetric.
    b. Yes, the frequencies start off low, increase to a maximum, and then decrease to low again, and are roughly symmetric.
    c. Yes, the distribution appears to be skewed to the right.
    d. No, the distribution appears to be skewed to the right.
    e. It cannot be determined from the graph.

16. Is it unlikely to find someone who has a systolic blood pressure of at least 150? (Consider an event to be unlikely if it occurs less than 5% of the time)
    a. Yes, the appropriate probability is less than 0.05
    b. Yes, the appropriate probability is greater than 0.05
    c. No, the appropriate probability is less than 0.05
    d. No, the appropriate probability is greater than 0.05
    e. It cannot be determined from the graph

17. Suppose that we have a sample space, $\Omega = \{1,2,3,4,5,6,7,8,9\}$, and events, $A = \{2,4,6,9\}$, and $B = \{1,3,7,8\}$. Which of the following represents A ∩ B, the intersection of A and B ?
    a. $\{1,2,3,4,5,6,7,8,9\}$
    b. $\{1,3,5,7,8\}$
    c. $\{2,4,5,6,9\}$
    d. $\{5\}$
    e. $\emptyset$

18. A quiz consists of 10 Multiple Choice questions. Assume that the questions are independent. In addition, assume that each question has 4 choices which are equally likely to be correct on each question. What is the probability of guessing on each of the 10 quiz questions and getting at least 7 questions correct? Round to 4 decimal places.
    a. 0.0004
    b. 0.0031
    c. 0.0035
    d. 0.9965
    e. 0.9969

19. What percentage of measurements in a population are above the median?
    a. 0%
    b. 0.5%
    c. 50%
    d. 100%
    e. Cannot Be Determined

20. What percentage of measurements in a skewed population are above the mean?
    a. 0%
    b. 0.5%
    c. 50%
    d. 100%
    e. Cannot Be Determined

For Questions 21–23, refer to the relevant results from a regression analysis provided below.

> A simple random sample of 5k race times for 32 competitive female runners aged 15-24 years old resulted in a mean 5k race time of 16.79 minutes. The simple linear regression equation that fit the sample data was obtained and found to be $\hat{y} = 21.506 - 0.276x$ where x represents the age of the runner in years and $\hat{y}$ represents the 5k race time in minutes. When testing the claim that there is a linear correlation between age and 5k race times of competitive female runners, an observed test statistic of $t = -7.87$ resulted in an approximate p-value of 0.0001.

21. The proportion of variation in 5k race times that can be explained by the variation in the age of competitive female runners was approximately 0.81. What is the value of the sample linear correlation coefficient?

    a. -0.9
    b. -0.81
    c. 0.405
    d. 0.81
    e. 0.9

22. Using all of the results provided, is it reasonable to predict the 5k race time (minutes) of a competitive female runner 6 years of age?

    a. Yes; linear correlation between age and 5k race times is statistically significant
    b. No; the age provided is beyond the scope of our available sample data
    c. Yes; both the sample linear regression equation and an age in years is provided
    d. No; linear correlation between age and 5k race times is not statistically significant
    e. Yes; the age provided is beyond the scope of our available sample data

23. The statement “Since the p-value for the linear correlation is so small, we may conclude that getting older causes the female runners to have lower 5k race times,” is:

    a. True
    b. False
    c. Sometimes true and sometimes false

24. A 95% confidence interval for a population proportion was calculated to be $0.256 < p < 0.321$. We want to know if $p = 0.25$.
    a. Based on the Confidence Interval we conclude that p is significantly less than 0.25.
    b. Based on the Confidence Interval we conclude that p is significantly greater than 0.25
    c. Based on the CI we conclude that p is does not significantly differ from 0.25.
    d. Not enough information is given to make any conclusion

25. Assuming weights of male athletes are normally distributed with a mean of 180 lbs and a standard deviation of 10 lbs, what is the probability that a randomly selected male athlete weighs more than 200 lbs? Round to 4 decimal places.
    Also, is this probability the same as the probability that a randomly selected sample of size n (where $n > 1$) has a mean weight more than 200 lbs?

    a. 0.0228; yes, these two probabilities would be the same
    b. 0.0228; no, these two probabilities would not be the same
    c. 0.9772; yes, these two probabilities would be the same
    d. 0.9772; no these two probabilities would not be the same
    e. It cannot be determined with the information provided

<u>Tie Breaker 1</u>

Assume the following probabilities for two events, A and B:

$$P(A) = 0.50, P(B) = 0.60, P(A \cup B) = 0.86$$

Are the events, A and B, independent in this situation?
You must provide reasoning for your answer.

Tie Breaker 2

The table below relates to a study where the length of hands and feet are measured for a group of college freshmen. The goal of the study was to determine if there was a statistically significant association between foot and hand length.

	Right foot longer	Left foot longer	Both feet same	Row sample size
Right hand longer	50.00%	13.64%	36.36%	22
Left hand longer	8.00%	36.00%	56.00%	25
Both hands same	22.64%	24.53%	52.83%	53

The table above provides row percentages and sample sizes. For example, 50.00% of students with 'Right hand longer' have their 'Right foot longer' and there are 22 total 'Right hand longer' students.

1) Based upon the row percentages provided, fill in the contingency table below with the approximate cell counts, rounding to the nearest whole number.

	Right foot longer	Left foot longer	Both feet same
Right hand longer			
Left hand longer			
Both hands same			

2) Do the data indicate an association between foot and hand length? Conduct an appropriate hypothesis test to answer this question using a 0.05 significance level. Provide the hypotheses, test statistic(s), p-value(s), and a formal conclusion.

Tie Breaker 3

For some $n \in \mathbb{N}$ I have calculated a probability as:

$$\frac{1}{(n-1)!} - \frac{1}{n!} = \frac{1}{8}$$

What is the value of n? You must provide reasoning for your answer.

Multiple Choice Key

1. c
2. b
3. b
4. c
5. a
6. d
7. c
8. a
9. b
10.a
11.d
12.b
13.b
14.a
15.d
16.a
17.e
18.c
19.c
20.e
21.a
22.b
23.b
24.b
25.b

-Tie Breaker Question 1-

Assume the following probabilities for two events, A and B:

$P(A) = 0.50$, $P(B) = 0.60$, $P(A \cup B) = 0.86$

Are the events, A and B, independent in this situation?

You must provide reasoning for your answer.

Solution:

No. A and B are NOT independent.

Reasoning (One Possible Answer):

Notice that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

So, $0.86 = 0.50 + 0.60 - P(A \cap B)$

Solving for $P(A \cap B)$, we find that $P(A \cap B) = 0.24$

If A and B are independent, $P(A \cap B) = P(A) * P(B)$

$P(A \cap B) = 0.24 \neq 0.30 = 0.50 * 0.60 = P(A) * P(B)$

Thus, A and B are NOT independent

Rubric:

0 pts – Answer that A and B are independent

1 pt – Answer that A and B are NOT independent but with no reasoning or improper reasoning

2 pts – Answer that A and B are independent but with an arithmetic error in proper reasoning

3 pts – Answer that A and B are NOT independent with proper reasoning

*Note that 'proper reasoning' implies that the response includes a valid probabilistic definition of independent events such as $P(A \cap B) = P(A) * P(B)$ or $P(A|B) = P(A)$ or $P(B|A) = P(B)$ etc.

-Tie Breaker Question 2-

The table below relates to a study where the length of hands and feet are measured for a group of college freshmen. The goal of the study was to determine if there was a statistically significant association between foot and hand length.

	Right foot longer	Left foot longer	Both feet same	Row sample size
Right hand longer	50.00%	13.64%	36.36%	22
Left hand longer	8.00%	36.00%	56.00%	25
Both hands same	22.64%	24.53%	52.83%	53

The table above provides row percentages and sample sizes. For example, 50.00% of students with 'Right hand longer' have their 'Right foot longer' and there are 22 total 'Right hand longer' students.

Solution Provided in Red:

3) Based upon the row percentages provided, fill in the contingency table below with the approximate cell counts, rounding to the nearest whole number.

	Right foot longer	Left foot longer	Both feet same
Right hand longer	**11**	**3**	**8**
Left hand longer	**2**	**9**	**14**
Both hands same	**12**	**13**	**28**

4) Do the data indicate an association between foot and hand length? Conduct an appropriate hypothesis test to answer this question using a 0.05 significance level. Provide the hypotheses, test statistic(s), p-value(s), and a formal conclusion

• Hypotheses $\begin{cases} H_0: No\ association\ between\ foot\ and\ hand\ length\ (independent) \\ H_1: There\ is\ an\ association\ between\ foot\ and\ hand\ length\ (dependent) \end{cases}$

• Test Statistic $\chi^2 = 11.942, df = 4$

• P-Value $p - value = 0.0178$

• Formal Conclusion at $\boldsymbol{\alpha} = \mathbf{0.05}$ (in terms of $\boldsymbol{H0}$) Reject the null hypothesis at the 5% significance level. There is sufficient evidence to support the claim that foot and hand length are associated/dependent.

Rubric: 0 pts to 5 pts Possible 1 point for each completely correct answer and 0 points for each incorrect answer of the following tie breaker components: (1) approximate cell counts, (2) hypotheses, (3) test statistic, (4) p-value, and (5) formal conclusion

-Tie Breaker Question 3-

For some $n \in \mathbb{N}$ I have calculated a probability as:

$$\frac{1}{(n-1)!} - \frac{1}{n!} = \frac{1}{8}$$

What is the value of n? You must provide reasoning for your answer.

Solution:

$n = 4$

Reasoning (One Possible Arithmetic Based Proof):

$$\frac{n}{n} \cdot \frac{1}{(n-1)!} - \frac{1}{n!} = \frac{1}{8}$$

$$\frac{n}{n!} - \frac{1}{n!} = \frac{1}{8}$$

$$\frac{n-1}{n!} = \frac{1}{8}$$

$$8(n-1) = n!$$

$$8 = n \cdot (n-2)!$$

$$\frac{8}{n} = (n-2)!$$

Thus, $n \in \{4,8\}$. Checking these we see that $n = 4$

Rubric:

0 pts – Wrong Answer with no reasoning

1 pt – Correct Answer with no reasoning

2 pts – Wrong Answer but they used definitions of factorials or other proper reasoning

3 pts – Correct Answer with correct reasoning (I suppose they could just brute force the answer)