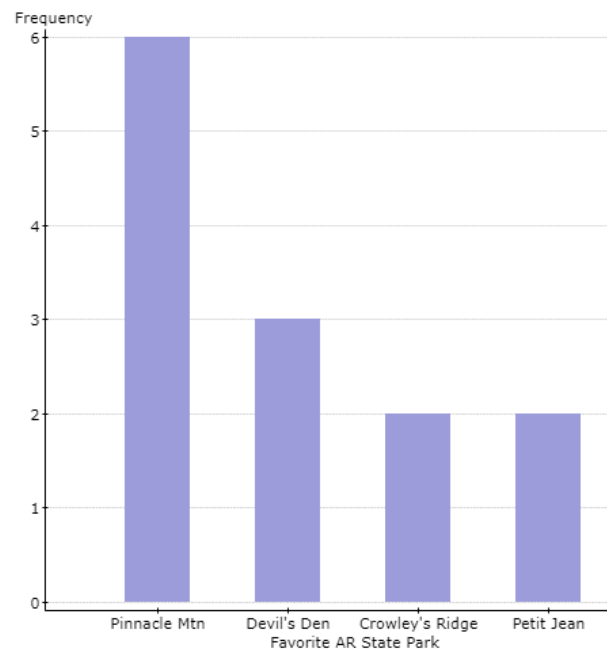


-Directions-

This exam 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 Question #1 first, then Question #2, and then Question #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. What percentage of observed values in a sample fall between the lower and upper limits of a 99% confidence interval for the population mean?
 - a. 99%
 - b. 50%
 - c. 100%
 - d. Cannot Be Determined

2. The graph to the right represents the reported favorite Arkansas state park for a sample of students at a state university. Noting the measurement scale for the data, what would be the most appropriate description of the shape of the data distribution?
 - a. Normal
 - b. Uniform
 - c. Skewed
 - d. Describing the shape of the data distribution is not appropriate in this context



3. Refer to the graph in Question 2. Observations in the sample data will be numerically coded as follows: 1 = Pinnacle Mtn., 2 = Devil's Den, 3 = Crowley's Ridge, and 4 = Petit Jean. Provide the mean, median, mode, and sample size for the sample data.
 - a. $\bar{X} = 3.25$, Median = 2.5, Mode = 1, $n = 13$
 - b. $\bar{X} = 2$, Median = 2, Mode = 1, $n = 13$
 - c. $\bar{X} = 3.25$, Median = 2.5, Mode = 1, $n = 4$
 - d. $\bar{X} = 2$, Median = 2, Mode = 1, $n = 4$

For Questions 4–6, refer to the setting and data provided below.

According to 2014 data from the Census Bureau, women make 77 cents for every dollar men make. Moreover, data released by the American Association of University Professors (AAUP) also suggests a gender wage gap in academia. A random sample of associate professor 9-month salaries (in thousands of dollars) obtained by the National Center for Education Statistics will be used to evaluate if salaries differ by gender ('Male Salary' vs. 'Female Salary'). The sample salary data for associate professors are provided. Other variables in the dataset include the name of the university ('School').

School	Male Salary	Female Salary
Stanford	144.4	129.9
Columbia	137.6	120.9
Stevenson	129.8	89.4
Duke	120.6	97.7
UC Berkely	113.6	100.9
UT Dallas	113.1	102.6
South Florida	111.8	74.1
UCLA	110.7	105.3

4. Which of the following statistical procedures would be most appropriate to test the claim that the mean male salary is higher than the mean female salary?
Assume that any necessary normality requirements hold.
- a. Two-tailed two-sample paired/dependent t-test of means
 - b. Two-tailed two-sample independent t-test of means
 - c. One-tailed two-sample paired/dependent t-test of means
 - d. One-tailed two-sample independent t-test of means
5. Referring to the setting and data provided above, what is the appropriate p-value for testing the claim that the mean male salary is higher than the mean female salary? Round to 3 decimal places.
- a. 0.003
 - b. 0.001
 - c. 0.011
 - d. 0.021

6. Refer to Questions 4-5. Using a 0.05 significance level, which of the following is the most appropriate conclusion for the hypothesis test given the results?
- Reject the null hypothesis; there is sufficient evidence to suggest that the mean male salary is significantly higher than the mean female salary.
 - Reject the null hypothesis; there is not sufficient evidence to suggest that the mean male salary is significantly higher than the mean female salary.
 - Fail to reject the null hypothesis; there is sufficient evidence to suggest that the mean male salary is significantly higher than the mean female salary.
 - Accept the null hypothesis; there is not sufficient evidence to suggest that the mean male salary is significantly higher than the mean female salary.
7. In which of the following would we **not** observe sampling error?
- Convenience Sample
 - Quota Sample
 - Census
 - Simple Random Sample
 - Stratified Sample
 - Cluster Sample
- All of the Above
 - I and II only
 - III, IV, V, and VI only
 - III only
8. Observed prices for a particular coffee maker on Black Friday at different retailers are negatively skewed with a mean of \$29 and a standard deviation of \$15. If the set of all observed Black Friday prices for the coffee maker are converted to a z-score scale, which of the following describes the shape and center of the new distribution of these z-scores?
- Normally distributed with a mean of 0
 - Normally distributed with a mean of \$29
 - Negatively skewed with a mean of 0
 - Cannot Be Determined

9. A study of 24 coffee-drinking mathletes was conducted to determine if math competition exam scores are associated with morning caffeine intake levels (little/no caffeine, moderate caffeine, or high caffeine). The mathletes were randomly assigned to one of the three different caffeine intake level groups. On the morning of the competition, 8 mathletes were given a cup of coffee that was actually decaffeinated, 8 were given a cup of coffee containing 200 mg of caffeine, and 8 were given a cup of coffee containing 400 mg of caffeine. The mathletes were not told which type of coffee they were given. After drinking their cup of coffee, all 24 mathletes sat for the competition exam and their scores were recorded. What type of study is described above? Please select the best answer of those provided below.
- a. Survey
 - b. Observational Study
 - c. Experimental Study
 - d. Single-Blind Experimental Study
10. The mean marathon completion time for three friends was 4 hours and 32 minutes with a standard deviation of 0 minutes. At the finish line these three friends discovered that another friend in their training group had completed the marathon in 4 hours and 14 minutes. Using this information provide the population standard deviation for the four friends marathon completion times. Round to 3 decimal places.
- a. 81 minutes
 - b. 0 minutes
 - c. 9 minutes
 - d. Cannot Be Determined
11. It was reported that 14% of residents in a local community had the flu during the week of January 16. The indoor soccer season for the Blue Crickets, a youth soccer team from this local community, began this same week (i.e. the week of January 16). Assuming there are 18 youth players on the Blue Crickets roster, what is the probability that at least one player has the flu during the week of January 16? Round to 3 decimal places.
- a. 0.066
 - b. 0.194
 - c. 0.934
 - d. 0.740

For Questions 12-14, refer to the table, which summarizes data for a sample of 105 individuals.

Health Outcome	Vaccination Status	
	Not Vaccinated (No)	Vaccinated (Yes)
Sick (Chickenpox)	38	12
Healthy (No Chickenpox)	4	51

12. What is the probability of one randomly selected individual becoming sick with Chickenpox given the individual had been vaccinated? Round to 3 decimal places.
- a. 0.240
 - b. 0.114
 - c. 0.962
 - d. 0.190
13. If *two* of the 105 individuals are randomly selected without replacement, what is the probability of both selected individuals becoming sick with Chickenpox?
- a. 0.476
 - b. 0.227
 - c. 0.224
 - d. 0.036
14. Assume simple random sampling for the data summarized in the table above.
Let p_V represent the proportion of individuals that are vaccinated.
What is the 99% confidence interval to estimate p_V , the population proportions of individuals that are vaccinated? Round to 3 decimal places.
- a. (0.034, 0.194)
 - b. (0.506, 0.694)
 - c. (0.521, 0.679)
 - d. (0.477, 0.723)

15. Which measures of central tendency would be equivalent in a perfectly normal distribution?

- I. The mean
- II. The median
- III. The mode
- IV. The range

- a. All of the Above
- b. I and II only
- c. I, II, and III
- d. None of the Above

16. If large simple random samples are taken over and over again from the same population and 95% confidence intervals for the population mean are constructed for each, about 95% of the constructed confidence intervals contain the _____.

Please select the best answer of those provided below.

- a. Sample mean
- b. Population mean
- c. Population standard deviation
- d. None of the above

17. Players on a university rugby team have normally distributed heights with a mean of 70 inches and a standard deviation of 2 inches. The team captain has a height of 73 inches. Given that there are 30 players on the team, how many players would be expected to be shorter than the captain? Round to the nearest whole number.

- a. 28
- b. 2
- c. 15
- d. 30

18. Provide the probabilistic definition for the significance level, α .

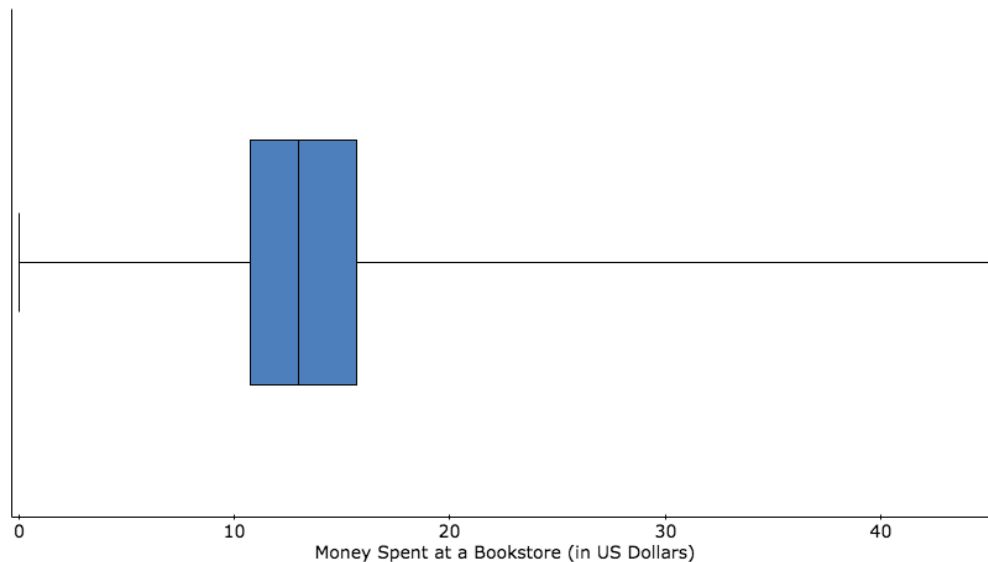
- a. $P(\text{Reject } H_0 \mid H_0 \text{ True})$
- b. $P(\text{Reject } H_0 \mid H_0 \text{ False})$
- c. $P(\text{Fail to Reject } H_0 \mid H_0 \text{ True})$
- d. $P(\text{Fail to Reject } H_0 \mid H_0 \text{ False})$

19. A researcher is attempting to predict the number of art museum visits individuals make during a year ('Art Museum') using the number of books the individual reports reading during the same year ('Books'). A small sample is collected and data is provided.

Books	12	28	2	0	12	0	3
Art Museum	3	10	1	5	1	0	1

Referring to the setting and data provided above, what proportion of the variability in the number of art museum visits an individual makes during a year can be accounted for by the number of books an individual reads during a year? Assume there is a linear relationship between the variables. Round to 3 decimal places.

- a. 0.719
 - b. 0.756
 - c. 0.571
 - d. -0.756
20. The boxplot below represents the amount of money spent (in dollars) by individuals during a single trip to a local bookstore. Determine the relationship between the mean and the median.



- a. Mean = Median
- b. Mean \approx Median
- c. Mean < Median
- d. Mean > Median

21. Refer to the boxplot provided in Question 20. What is the approximate value for the mean amount of money spent (in dollars) during a single trip to a local bookstore?
Please select the best answer of those provided below.

- a. \$45
- b. \$13
- c. \$10
- d. Cannot Be Determined

22. A researcher is investigating job satisfaction in adults as it relates to salary. Job satisfaction scores are measured on a 0 to 9 point scale with higher scores representing higher levels of job satisfaction. Salaries are recorded in thousands of dollars.

A simple random sample of 64 adults was used to construct the simple linear regression equation $\hat{y} = 8.245 - 0.098x$ where x represents the salary of the adult in thousands of dollars and y represents the job satisfaction score.

One individual adult has an observed job satisfaction score of 5 and a residual of 0.871. What was the predicted job satisfaction score of this individual adult? Round to 3 decimal places.

- a. 5.871
- b. 4.129
- c. 8.626
- d. 6.884

23. A sample of 362 college bound, high school students was utilized in a study to determine the effect of school type (Private, Public, or Online/Virtual) on overall SAT scores. Partial output, applying the methods of one-way analysis of variance, is provided below. Provide 'FIND ME #1' (i.e. DF_W) and 'FIND ME #2' (i.e. F) as the answer to this question.

Source	DF	Sum of Squares	Mean Square	F
Between	2	324269	162134	FIND ME #2
Within	FIND ME #1	4986078	13889	

- a. $DF_W = 361; F = 11.674$
- b. $DF_W = 359; F = 11.674$
- c. $DF_W = 359; F = 0.065$
- d. $DF_W = 361; F = 0.065$

24. The mean of a distribution is 8.3, the variance is 6.76, the range is 9, and the median is 7. What is the standard deviation for the distribution? Round to 3 decimal places.

- a. 2.600
- b. 1.500
- c. 4.500
- d. Cannot Be Determined

25. Assume that student waiting times at bus stops are uniformly distributed between 12 and 28 minutes. What is the probability that a randomly selected student has a waiting time between 20 and 25 minutes? Round to 3 decimal places.

- a. 0.500
- b. 0.313
- c. 0.188
- d. 0.200

-Tie Breaker Question 1-

Measles is a highly contagious, serious disease caused by a virus. Previously known to result in major epidemics and millions of deaths each and every year, there was a sharp decline in measles-related deaths with the introduction of the measles vaccine in 1963. However, despite the availability of this effective vaccine, the disease still remains one of the leading causes of death among young children globally. Both vaccine efficacy and a critical vaccination coverage is needed to prevent large outbreaks. The measles vaccine is highly effective in preventing disease but vaccination coverage is insufficient in some areas. The World Health Assembly recommends that 90% of individuals in an area be vaccinated in order to achieve herd immunity, a form of indirect protection from infectious disease that occurs when a large percentage of a population has become immune to an infection and, consequently, individuals that are not immune are still partially protected from disease.

In a particular area, 99.8% of individuals are either vaccinated or female or both, 44.7% of individuals are vaccinated and female, and 51.1% of individuals are female. Based upon this information, has the critical vaccination level of 90% been achieved? What percentage of the individuals in this particular area are vaccinated?

Round only your final answer to 3 decimal places. You must provide reasoning for your answer.

-Tie Breaker Question 2-

The table below represents data from the General Social Survey and provides the level of health and the highest degree earned for a random sample of individuals. The goal of the study was to determine if there was a statistically significant association between education level and the healthiness of an individual.

Education Level	Health			
	Excellent	Good	Fair	Poor
High School or Less	537	1079	557	170
Associate or 2-Year Degree	80	138	49	11
Bachelor's Degree	229	276	64	12
Graduate Degree	130	147	32	2

Do the data indicate an association between education level and the healthiness of an individual? 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 Question 3-

The 115th Congress has an all-time high of 22 female senators. Note that there are 100 total senators. Six senators will be randomly selected to form a working group for a major piece of legislation. What is the probability that there will be *more* women than men selected for the six-member group? Round to 3 decimal places.

You must provide reasoning for your answer.

Multiple Choice Key

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

-Tie Breaker Question 1-

ABBREVIATED Question: In a particular area, 99.8% of individuals are either vaccinated or female or both, 44.7% of individuals are vaccinated and female, and 51.1% of individuals are female. Based upon this information, has the critical vaccination level of 90% been achieved? What percentage of the individuals in this particular area are vaccinated?

Solution:

$P(V) \approx 0.934$; Yes, the critical vaccination level of 90% has been achieved

Reasoning (One Possible Answer):

Define Notation: $V = \text{"Vaccinated"}$ and $F = \text{"Female"}$

We know: $P(V \cup F) = 0.998$, $P(V \cap F) = 0.447$, and $P(F) = 0.511$

[Probabilities Known from Question are Provided]

Notice that $P(V \cup F) = P(V) + P(F) - P(V \cap F)$

[Use of Probability Rule(s)]

So, $0.998 = P(V) + 0.511 - 0.447$

Solving for $P(V)$, we find that $P(V) = 0.934$

[Finding the Probability of an Individual Being Vaccinated in the Area]

Thus, the critical vaccination level of 90% has been achieved

[Comparing this Probability of Vaccination to the Critical Vaccination Level Stated]

Rubric:

0 pts – No Correct Probabilities or Work Provided

1 pts – Probabilities Known from Question are Provided but there is No Additional Correct Work

2 pts – Probability Rules are Utilized but the Final Solution is Incorrect

3 pts – The Final Solution is Correct but Question Regarding Critical Vaccination Level is unanswered

4 pts – The Final Solution is Correct and Question Regarding Critical Vaccination Level is answered

-Tie Breaker Question 2-

Education Level	Health			
	Excellent	Good	Fair	Poor
High School or Less	537	1079	557	170
Associate or 2-Year Degree	80	138	49	11
Bachelor's Degree	229	276	64	12
Graduate Degree	130	147	32	2

Do the data indicate an association between education level and the healthiness of an individual? 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.

Solution:

- **Hypotheses**

$$\begin{cases} H_0: \text{Education Level and Health of an Individual are Independent (there is no association)} \\ H_1: \text{Education Level and Health of an Individual are Dependent (there is association)} \end{cases}$$
- **Test Statistic**

$$\chi^2 = 166.536, df = 9$$
- **P-Value**

$$p\text{-value} < 0.0001 \approx 0$$
- **Formal Conclusion at $\alpha = 0.05$ (in terms of H_0)**
 Reject the null hypothesis at the 5% significance level. There is sufficient evidence to support the claim that education level and the healthiness of an individual are associated/dependent.

Rubric: 0 pts to 4 pts Possible

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

-Tie Breaker Question 3-

The 115th Congress has an all-time high of 22 female senators. Note that there are 100 total senators. Six senators will be randomly selected to form a working group for a major piece of legislation. What is the probability that there will be *more* women than men selected for the six-member group? Round to 3 decimal places.

You must provide reasoning for your answer.

Solution: $P(\text{more women than men selected}) \approx 0.020$

*** Steps**

1. Notice the situations when there will be **more** women than men selected for the group i.e. 4 or 5 or 6 women are selected for the six-member group
2. Notice that these are disjoint/non-overlapping events and so

$$P(4 \text{ or } 5 \text{ or } 6 \text{ women}) = P(4 \text{ women}) + P(5 \text{ women}) + P(6 \text{ women})$$

3. Find each probabilities using combinations and noting that there are 22 women and 78 men in the senate to choose the six-member group from:

$$P(4 \text{ women}) = \frac{\binom{22}{4}\binom{78}{2}}{\binom{100}{6}} \approx 0.0184$$

$$P(5 \text{ women}) = \frac{\binom{22}{5}\binom{78}{1}}{\binom{100}{6}} \approx 0.0017$$

$$P(6 \text{ women}) = \frac{\binom{22}{6}\binom{78}{0}}{\binom{100}{6}} \approx 0$$

4. So, the probability that there will be **more** women than men selected for the group is:

$$\begin{aligned} P(4 \text{ or } 5 \text{ or } 6 \text{ women}) &= P(4 \text{ women}) + P(5 \text{ women}) + P(6 \text{ women}) \\ &\approx 0.0184 + 0.0017 + 0 \approx 0.020 \end{aligned}$$

Rubric: 0 pts to 4 pts Possible

1 point for each correctly calculated probability of the following: $P(4 \text{ women})$, $P(5 \text{ women})$, $P(6 \text{ women})$, and $P(4 \text{ or } 5 \text{ or } 6 \text{ women})$

If the correct final solution is provided but an intermediate probability is not explicitly stated, 4 points should be awarded as long as sufficient work was shown or sufficient reasoning was provided for the answer.