Statistic and data handling

This exam contains 38 multiple choice questions. A question and the possible answer choices appear. Select the choice you find to be the correct answer of the question. Clicking on a choice saves it as your answer for the question.

Circle the correct response for each question. Make sure that your answer is clearly marked. You will not receive partial credit for any work done.

This is a closed book, closed notes examination. You may use a calculator if you wish. However, cell phones are not permitted for use in any way.

Any discussion or otherwise inappropriate communication between examinees, as well as the appearance of any unnecessary material or cell-phone usage, will be dealt with severely. Violations may result in an "F" for this exam, "F" for the class, suspension, or expulsion.

Print your name at the top of this page in the upper right hand corner.

Good Luck!!

HONOR PLEDGE FOR THIS EXAM:

After you have finished the exam, please read the following statement and sign your name below it.

I promise that I did not discuss any aspect of this exam with anyone other than the instructor, that I neither gave nor received any unauthorized assistance on this exam, and that the work presented herein is entirely my own.

1. Which of the following statements is FALSE?
   (a) The analysis of variance is an extension of the *t*-test.
   (b) When a study has only one independent variable with two groups, F and t^2 are identical.
   (c) A direction of differences between groups must be specified in critical F.
   (d) Degrees of freedom are the number of scores free to vary once the means are known.

2. What is statistical inference?
   (a) The practice of using \cherry-picked" statistics to thrust one's opinion onto others.
   (b) The process by which sample results are used to make statements about population parameters.
   (c) The manner in which ethical issues are addressed in an experiment involving human subjects.
   (d) None of the above.

3. What does the margin of error quantify?
   (a) the chance of making a correct statement
   (b) the uncertainty arising from using a statistic to estimate a population parameter
   (c) the amount of information gained from using an experiment when compared to using an observational study
   (d) None of the above

4. Suppose that we take a random sample of size *n* from a large population (e.g., UNIEVORA students) and compute **p**, the mean of individuals who are left handed. Which of the following statements is true?
   (a) The larger the sample size, the smaller the margin of error.
   (b) The larger the sample size, the smaller the bias.
   (c) The larger the sample size, the larger the length of a 95 percent confidence interval for the proportion of left-handed individuals.

(d)   All of the above are true.

5. Consider a random sample of 100 females and 100 males. Suppose 15 of the females are left-handed and 12 of the males are left-handed. What is the estimated difference between population proportions of females and males who are left-handed (females – males)? Select the choice with the correct notation and numerical value.

(a) $p_1 - p_2 = 3$
(b) $p_1 - p_2 = 0.03$
(c) $\hat{p}_1 - \hat{p}_2 = 3$
(d) $\hat{p}_1 - \hat{p}_2 = 0.03$

6. In any normal distribution, the proportion of observations that are within 2 standard deviations of the mean is closest to

(a)   0.03
(b)   0.68
(c)   0.95
(d)   0.98

7. True or False. Outliers can greatly affect the value of the sample mean and sample standard deviation.

(a) True
(b) False

8. A random sample of 5000 students were asked whether they prefer a 10 week quarter system or a week semester system. Of the 5000 students asked, 500 students responded. The results of this survey _____

(a) can be generalized to the entire student body because the sampling was random.
(b) can be generalized to the entire student body because the margin of error was 4.5%.
(c) should not be generalized to the entire student body because the non-response rate was 90%.
(d) should not be generalized to the entire student body because the margin of error was 4.5%.What is a sampling distribution?

9. In the bird-oxygen rate example we did in class, we examined how much oxygen birds consumed at different temperatures. We computed the least-squares regression equation to be
$$O_2{}^{rate} = 3.5 - 0.1 \, (Temperature)$$
What oxygen rate (measured in ml/g/hr) would you predict for a bird subjected to a temperature of 8 degrees?

(a)   -0.1
(b)   2.7
(c)   3.5

(d)   4.3

10. A simple random sample of size $n = 25$ is drawn from a population with mean 50 and standard deviation 5. What is the standard deviation of the sample mean x?

(a)   1
(b)   2
(c)   5
(d)   10

11. A random sample of *n* = 1000 portfolio rate returns (measured as a percentage) is taken; here are the data:
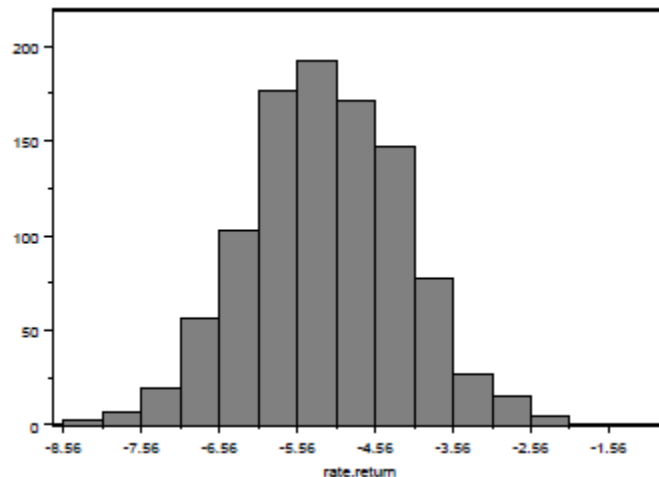

The distribution of the rate of returns is best described as
(a)   skewed left
(b)   symmetric
(c)   skewed right

12. In Question 11, the sample mean rate of return is closest to
   (a)   1
   (b)   -1
   (c)   5
   (d)   -5

13. In Question 11, the sample standard deviation is closest to
   (a)   -1
   (b)   1
   (c)   4
   (d)   11

14. In a simple random sample (SRS) of *n* = 100 Columbia residents, a 99 percent confidence interval for the mean age of Columbia residents was computed to be (29.8; 38.5). What is the interpretation attached to this interval?
   (a)   We are 99 percent confident that the mean age of Columbia residents is between 29.8 and 38.5.
   (b)   Ninety-nine percent of the residents in our sample had ages between 29.8 and 38.5.
   (c)   We are 99 percent confident that the mean age of the residents in our sample is between 29.8 and 38.5.
   (d)   All of the above are valid interpretations.

15. In Question 14, what is one way to decrease the length of the confidence interval?
   (a) Increase the sample size
   (b) Use a smaller confidence level
   (c) Both (a) and (b) are correct
   (d) Neither (a) nor (b) are correct

16. Which statistic provides a measure of variability (or spread)?
   (a)   mean

(b)   median
(c)   maximum
(d)   standard deviation

17. In the scatterplot below, the correlation is closest to
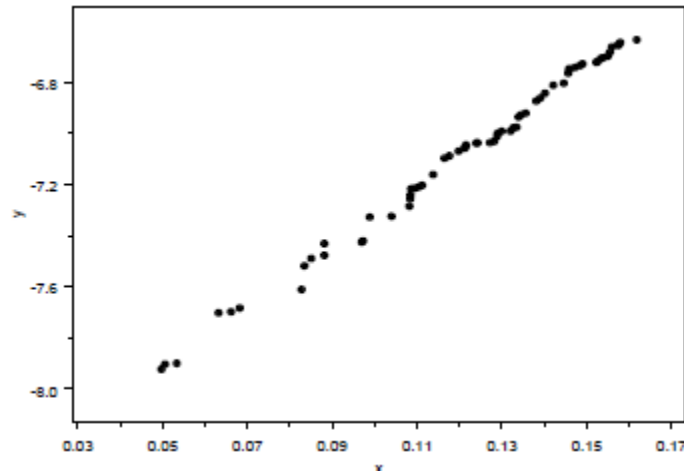
(a) $r = -0.87$
(b) $r = 0.25$
(c) $r = 0.99$
(d) $r = 1.03$

18. _____ are used to infer that the results from a sample are reflective of the true population scores.

(a)   Descriptive statistics
(b)   Regression statistics
(c)   Correlated statistics
(d)   Inferential statistics

19. In the t-Student test, the _____ states the means are _____.

(a)   null hypothesis; not equal
(b)   research hypothesis; equal
(c)   null hypothesis; equal
(d)   research hypothesis; not equal

20. The F statistic is a ratio of two types of variance: variance_____ and within groups.

(a)   random
(b)   individual
(c)   between
(d)   systematic

21. A Type I [1] error occurs when the null hypothesis is

(a)   rejected and the research hypothesis is actually false.
(b)   accepted but and research hypothesis is actually true.
(c)   rejected and null hypothesis is actually true.
(d)   accepted and null hypothesis is actually true.

22. Which of the following statements is TRUE?

(a) A very low significance level increases the chances of a Type I error.

---

[1] In statistical hypothesis testing, a **type I error** is the incorrect rejection of a true 'null hypothesis' (a "false positive")

(b)   If the effect size is small, a Type II[2] error is unlikely.
(c)   When the null hypothesis is rejected, the population means are equal.
(d)   True differences are more likely to be detected if the sample size is large.

23. If a mechanic looks at your car engine and says there is nothing wrong with it and your car breaks down when you leave the garage, what type of error did the mechanic make?
(a)   Type I.
(b)   Type II
(c)   Systematic error
(d)   Matrix error

24. If the null hypothesis was rejected and there was 1 chance out of 100 that the decision was wrong, what was the alpha level in the study?
(a)   .01
(b)   .10
(c)   .001
(d)   .100

25. The probability of a Type II error is related to _____.
(a)   sample size
(b)   significance level (alpha)
(c)   effect size
(d)   All of the above.

26. Which of the following is NOT a reason for a Type II error?
(a)   incomprehensive instructions to participants
(b)   a very weak manipulation of the independent variable
(c)   using a dependent measure that is unreliable and insensitive
(d)   increasing sample size

27. Dr. Probability is using a *t*-test to compare the means of two groups. There are 25 participants in each group. How many degrees of freedom are there in this test?
(a)   23
(b)   24
(c)   48
(d)   49

28. Before the 2004 Presidential election, the confidence statement was given: "We are 95 percent confident that between 45.6% and 54.4% of likely voters in Iowa will vote for John Kerry." Assuming that this information was based on a random sample, what is the margin of error associated with this statement?
(a)   2.2 percent
(b)   4.4 percent
(c)   8.8 percent
(d)   17.6 percent

29. What does it mean for a measurement to have predictive validity?
(a)   The measurement must be easy to observe.
(b)   The measurement must always be correct.
(c)   The measurement must be able to predict success on tasks that are related to the property being measured.

---

[2] A **type II error** is the failure to reject a false null hypothesis (a "false negative")

(d)   None of the above.

30. For a random sample of 9 women, the average resting pulse rate is x = 76 beats per minute, and the sample standard deviation is s = 5. The standard error of the sample mean is
(a)   0.557
(b)   0.745
(c)   1.667
(d)   2.778

31. Recently, I found the following question in an online poll (at www.foxnews.com):
     'Should states fund tuition for illegal aliens?'
What is wrong with this question?
(a)   Political questions should never be asked.
(b)   All online questions should be worded in the past tense.
(c)   A negative connotation is attached to the word "illegal."
(d)   Everyone agrees that illegal aliens should be deported.

32. In carrying out a survey, which of the following would be best classified as a sampling error?
(a)   individual nonresponse
(b)   the use of poorly-worded questions
(c)   responses which are not truthful
(d)   undercoverage

33. A randomly selected sample of 1000 college students was asked whether they had ever used the drug Ecstasy. Sixteen percent (16% or 0.16) of the 1000 students surveyed said they had. Which one of the following statements about the number 0.16 is correct?
(a)   It is a sample proportion (mean value).
(b)   It is a population proportion (mean value).
(c)   It is a margin of error.
(d)   It is a randomly chosen number.

34. A simple random sample of n = 400 Birmingham residents was recently asked the question: *Have you ever been the victim of a hate crime?* From this sample, 100 residents responded "yes." From this information, what is the value of the sample proportion of hate-crime victims?
(a)   0.05
(b)   0.10
(c)   0.20
(d)   0.25

35. I read an advertisement recently in which a credit card company promised that I could reduce my debt by 150 percent. Which of the following statements is (are) true?
(a)   This is possible if my debt is more than 150 dollars.
(b)   This is possible if my debt has recently increased by at least 150 percent.
(c)   The company's claim makes no sense.
(d)   Both (a) and (b).

36. In a past General Social Survey, a random sample of men and women answered the question "Are you a member of any sports clubs?" Based on the sample data, 95% confidence intervals for the population proportion (mean) who would answer "yes" are 0.13 to 0.19 for women and 0.247 to 0.33 for men. Based on these results, you can reasonably conclude that
    (a) At least 25% of American men and American women belong to sports clubs.
    (b) At least 16% of American women belong to sports clubs.
    (c) There is a difference between the proportions of American men and American women who belong to sports clubs.
    (d) There is no conclusive evidence of a gender difference in the proportion belonging to sports clubs.

37. Suppose a 95% confidence interval for the proportion of Americans who exercise regularly is 0.29 to 0.37. Which one of the following statements is FALSE?
    (a) It is reasonable to say that more than 25% of Americans exercise regularly.
    (b) It is reasonable to say that more than 40% of Americans exercise regularly.
    (c) The hypothesis that 33% of Americans exercise regularly cannot be rejected.
    (d) It is reasonable to say that fewer than 40% of Americans exercise regularly.

Solutions:

1	c	15.	b	29.	c
2	b	16.	b	30.	c
3	a	17.	d	31.	c
4	a	18.	c	32.	c
5	d	19.	d	33.	d
6	c	20.	c	34.	a
7	a	21.	c	35.	d
8	c	22.	c	36.	c
9	d	23.	c	37.	c
10	a	24.	d	38.	b
11	b	25.	a		
12	c	26.	c		
13	b	27.	b		
14	b	28.	a		