

BASIC STATISTICS SELF TEST

1. A researcher is interested in the travel time of Utrecht University students to college. A group of 50 students is interviewed. Their mean travel time is 16.7 minutes. For this study the mean of 16.7 minutes is an example of a(n)
 - A. Parameter
 - B. Statistic
 - C. Population
 - D. Sample

2. A researcher is curious about the IQ of students at the Utrecht University. The entire group of students is an example of a:
 - A. Parameter
 - B. Statistic
 - C. Population
 - D. Sample

3. Statistical techniques that summarize and organize the data are classified as:
 - A. Population statistics
 - B. Sample statistics
 - C. Descriptive statistics
 - D. Inferential statistics

4. A sports psychologist was interested in the effects of a six-week imagery intervention on an athlete's ability to execute sport-specific skills such as penalty taking in football. How might you define the imagery variable?
 - A. Independent variable
 - B. Dependent variable
 - C. Outcome variable
 - D. Resultant variable

5. Five-point Likert scales (*strongly disagree, disagree, neutral, agree, strongly agree*) are frequently used to measure motivations and attitudes. A Likert scale is a:
 - A. Discrete variable.
 - B. Ordinal variable.
 - C. Categorical variable.
 - D. All of the above options (A, B and C)

6. In a 500m speed skating race, time results would be considered an example of which level of measurement?
 - A. Nominal
 - B. Ordinal
 - C. Interval
 - D. Ratio

7. IQ tests are standardized so that the mean score is 100 for the entire group of people who take the test. However, if you select a group of 50 who took the test, you probably would not get 100. What statistical concept explains the difference between the two means?
- Statistical error
 - Inferential error
 - Residual error
 - Sampling error
8. A researcher studies the factors that determine the number of children future couples decide to have. The variable 'number of children' is a :
- Discrete variable
 - Continuous variable
 - Categorical variable
 - Ordinal variable
9. A teacher asks students to identify their favorite reality television show. What type of measurement scale do the different television shows make up?
- Nominal
 - Ordinal
 - Interval
 - Ratio
10. The median is always:
- The most frequently occurring score in a data set
 - The middle score when results are ranked in order of magnitude
 - The same as the mean
 - The difference between the maximum and minimum scores.
11. The seminar rooms in the library are identified by the letters A to H. A researcher records the number of classes held in each room during the first semester. What kind of graph would be appropriate to present the frequency distributions of these data?
- Histogram
 - Scatterplot
 - Bar chart
 - Box plot
12. A set of scores is presented in a histogram. The histogram shows a series of bars that tend to decrease in height from left to right. What is the shape of the distribution?
- Leptokurtic
 - Positively skewed
 - Negativity skewed
 - Normal

13. What is the mean for the following scores: 2, 5, 4, 1, 8?

- A. 3
- B. 4
- C. 5
- D. 20

14. What is the mean for the scores shown in the frequency distribution?

- A. 1.5
- B. 3.0
- C. 2.9
- D. 5.8

Value	f
1	1
2	3
3	3
4	2
5	1

15. What is the median for the following scores: 2, 5, 4, 1, 8?

- A. 3.5
- B. 4
- C. 4.5
- D. 7

16. A teacher gave a statistics test to a class of Geography students and computed the measures of central tendency for the test scores. Which of the following statements cannot be an accurate description of the scores?

- A. The majority of students had scores above the mean.
- B. The majority of students had scores above the median.
- C. The majority of students had scores above the mode.
- D. All of the above options (A, B and C) are false statements.

17. Which of the following sets of scores has the greatest variability (range)?

- A. 2, 5, 8, 11
- B. 13, 13, 13, 13
- C. 20, 25, 26 ,27
- D. 42, 43, 44, 45

18. Which of the following statements is the most accurate description for the concept of standard deviation?

- A. The total distance from the smallest score to the highest score.
- B. The square root of the total distance from the smallest score to the highest score.
- C. The squared average distance between all scores and the mean.
- D. The average distance between a score and the mean.

19. What is the variance for the following set of scores: 2, 2, 2, 2, 2.

- A. 0
- B. 2
- C. 4
- D. 25

20. Normally distributed data are normally referred to as:
- A. Bell-shaped
 - B. Asymmetrical
 - C. Skewed
 - D. Peaked
21. Take the formula $Z = (X - \mu)/\sigma$, where μ is the mean of the population, X is the value of the element, Z is the z-score and σ is the standard deviation. What does this formula calculate?
- A. Confidence interval.
 - B. Standard score.
 - C. Standard error of the mean.
 - D. Variance.
22. A population has a mean of $\mu=35$ and a standard deviation of $\sigma=5$. After 3 points are added to every score in the population, what are the new values for the mean and standard deviation?
- A. $\mu=35$ and $\sigma=5$
 - B. $\mu=35$ and $\sigma=8$
 - C. $\mu=38$ and $\sigma=5$
 - D. $\mu=38$ and $\sigma=8$
23. Of the following Z-score values, which one represents the location closest to the mean?
- A. $Z=+0.5$
 - B. $Z=+1,0$
 - C. $Z=-1.5$
 - D. $Z=-0.3$
24. If the scores on a test have a mean of 26 and a standard deviation of 4, what is the z-score for a score of 18?
- A. 2
 - B. 11
 - C. -2
 - D. -1.41
25. A population has a $\mu=50$ and $\sigma=10$. If these scores are transformed into z-scores, the population of z-scores will have a mean and standard deviation of:
- A. $\mu=50$ and $\sigma=10$
 - B. $\mu=50$ and $\sigma=1,96$
 - C. $\mu=1$ and $\sigma=0$
 - D. $\mu=0$ and $\sigma=1$

26. If all possible samples of size $n=30$ are selected from a population with $\mu=80$ and $\sigma=10$ and the mean is computed for each sample, then what shape is expected for the distribution of sample means?
- A. The sample means tend to form a normal-shaped distribution whether the population is normal or not.
 - B. The sample means tend to form a normal-shaped distribution only if the population distribution is normal.
 - C. The sample size of $n=30$ is too small to predict the shape of the distribution.
 - D. The mean of each sample will be very close to 80, hence the distribution of means will have little variability.
27. What is a definition of the standard error?
- A. Standard deviation of the sample.
 - B. Squared standard deviation.
 - C. Standard deviation of sample means.
 - D. Standard deviation of the population mean.
28. If a researcher sets a level of significance at .05 (i.e. 5%), what does this mean?
- A. Five times out of 100, a significant result will be found that is due to chance alone and not to true relationship.
 - B. Ninety-five times out of 100, a significant result will be found that is due to chance alone and not to true relationship.
 - C. Five times out of 100, a significant result will be found that is not due to chance, but to true relationship.
 - D. None of the above.
29. When does a researcher risk a Type I error?
- A. Anytime the decision is 'fail to reject'.
 - B. Anytime H_0 is rejected.
 - C. Anytime H_1 is rejected.
 - D. All of the above options.
30. Which of the following assumptions are required if an independent t-test is to be used?
- A. Samples are drawn from a normally distributed population.
 - B. Homogeneity of variances (equal variances).
 - C. The data are either interval or ratio scales.
 - D. All the above assumptions (A, B and C) are required.
31. What is the correct decision in a hypothesis if the data produce a t-statistic that is in the critical region?
- A. Reject H_0
 - B. Fail to reject H_0
 - C. Reject H_1
 - D. Fail to reject H_1

32. How does the shape of the t distribution compare to the normal distribution?
- A. The t distribution is taller and less spread out, especially when n is large.
 - B. The t distribution is taller and less spread out, especially when n is small.
 - C. The t distribution is flatter and more spread out, especially when n is large.
 - D. The t distribution is flatter and more spread out, especially when n is small.
33. Suppose you wanted to apply a one-tailed test as opposed to a two-tailed test. How would you convert a significance of $p = .284$?
- A. $.284 / 2 = .142$
 - B. $.284 \times 2 = .568$
 - C. $.284 / 0.05 = 5.68$
 - D. $.284 \times 0.05 = .0142$
34. A research report summarizes the results of a t -test by stating: $t(35)=5.2, p<0.05$. Which of the following is a correct interpretation of this report?
- A. The H_0 was not rejected and the probability of a Type I error is less than .05.
 - B. The H_0 was not rejected and the probability of a Type II error is less than .05.
 - C. The H_0 was rejected and the probability of a Type I error is less than .05.
 - D. The H_0 was rejected and the probability of a Type II error is less than .05.
35. Which of the following is true about a 95% confidence interval of the mean of a given sample:
- A. 95 out of 100 sample means will fall within the limits of the confidence interval.
 - B. There is a 95% chance that the population mean will fall within the limits of the confidence interval.
 - C. 95 out of 100 population means will fall within the limits of the confidence interval.
 - D. There is a .05 probability that the population mean falls within the limits of the confidence interval.
36. What effect would increasing the sample size have on a confidence interval?
- A. The confidence interval would increase in size.
 - B. The confidence interval would decrease in size.
 - C. The confidence interval is unaffected by sample size.
 - D. The confidence interval could either increase or decrease in size.
37. In an independent t -test output of SPSS, the Levene's test result is $p = .006$. What can we infer from this number?
- A. The means of both groups are assumed to be unequal.
 - B. The means of both groups are assumed to be equal.
 - C. The variances of both groups are assumed to be unequal.
 - D. The variances of both groups are assumed to be equal.

38. What does ANOVA stand for?
- A. Analysis of values and averages.
 - B. Analysis of variance.
 - C. Analysis of variability.
 - D. Analysis of non ordinal values.
39. What kind of variables would you crosstabulate?
- A. Two or more categorical.
 - B. Two or more continuous.
 - C. One continuous and two or more categorical.
 - D. One categorical and two or more continuous.
40. Which statistical test is used to identify whether there is a relationship between two categorical variables?
- A. Student's t-test.
 - B. Spearman's correlation test.
 - C. Pearson's Chi-square test.
 - D. Mann-Whitney test.
41. What does the statistic Cramer's V indicate?
- A. The significance of the Chi-square test.
 - B. The expected frequencies in a contingency table.
 - C. The amount of common variability of two numeric variables.
 - D. The strength of association between two categorical variables..
42. What is the null hypothesis for a Chi-square test?
- A. Both variables have a significant relationship.
 - B. Both variables have equal means.
 - C. Both variables are independent.
 - D. Both variables are dependent.
43. In order for accurate measures of the linear relationship between two variables to be achieved, what type of data are required if using Pearson's correlation coefficient?
- A. Nominal
 - B. Ordinal
 - C. Interval
 - D. Ratio
44. A Pearson correlation of $r=-0.6$ indicates
- A. An increase in X is accompanied by an increase in Y; the relationship is strong.
 - B. An increase in X is accompanied by an increase in Y; the relationship is moderate.
 - C. An increase in X is accompanied by a decrease in Y; the relationship is strong.
 - D. An increase in X is accompanied by a decrease in Y; the relationship is moderate.

45. 2. A scatterplot shows:
- A. The frequency with which values appear in the data.
 - B. The average value of groups of data.
 - C. Scores on one variable plotted against scores on a second variable.
 - D. The proportion of data falling into different categories.
46. R^2 is the notation for:
- A. The coefficient of correlation.
 - B. The coefficient of determination.
 - C. The coefficient of variation.
 - D. The coefficient of regression.
47. Suppose the correlation between height and weight for adults is +0.80. What proportion of the variability in weight can be explained by the relationship with height?
- A. 20%
 - B. 36%
 - C. 64%
 - D. 80%
48. In a linear regression equation, $Y=a + bX$, what is the b denote?
- A. The regression coefficient, the slope of the line.
 - B. The intercept with the Y-axis.
 - C. The correlation coefficient, the strength of the line
 - D. The score on the variable X.
49. In a linear regression equation, what does a slope of 2.5 indicate?
- A. For every increase of 2.5 on the y-axis, there is an increase of 5.0 on the x-axis.
 - B. For every increase of 2.5 on the x-axis, there is an equivalent increase on the y-axis.
 - C. For every increase of 1.00 on the x-axis, there is an increase of 2.5 on the y-axis.
 - D. For every increase of 1.00 on the y-axis, there is an decrease of 2.5 on the x-axis.
50. Which of the following statements about the t-statistic in regression analysis is not true?
- A. The t-statistic tests whether the regression coefficient, b , is equal to 0.
 - B. The t-statistic provides some idea of how well a predictor predicts the outcome variable.
 - C. The t-statistic can be used to see whether a predictor variable makes a statistically significant contribution to the regression model.
 - D. The t-statistic is equal to the regression coefficient divided by its standard deviation.

ANSWERS

1	B	11	C	21	B	31	A	41	D
2	C	12	B	22	C	32	D	42	C
3	C	13	B	23	D	33	A	43	C
4	A	14	C	24	C	34	C	44	D
5	D	15	B	25	D	35	B	45	C
6	D	16	B	26	A	36	B	46	B
7	D	17	A	27	C	37	D	47	C
8	A	18	D	28	A	38	B	48	A
9	A	19	A	29	B	39	A	49	C
10	B	20	A	30	D	40	C	50	D

RESULTS

correct answers:

40-50	good	basic statistical knowledge meets expectations
30-40	satisfactory	acceptable level of knowledge
20-30	unsatisfactory	some improvement required; reading Field* recommended.
<20	poor	basic knowledge inadequate; reading Field* necessary.

* Andy Field, Discovering Statistics using SPSS. 4th ed. Chapt. 1-3; 7.4-7.4.2; 8.1-8.4; 9.1-9.5.