

Self Assessment Questions & Answers

Correlation

Question 1

- If Spearman's correlation coefficient is equal to 1, then:
 - a. The rankings of the 2 variables is vastly different.
 - b. The rankings of the 2 variables partially agree
 - c. All of the “total variation” is “explained” by the regression line
 - d. The rankings of the 2 variables totally agree

Answer 1

- If Spearman's correlation coefficient is equal to 1, then:
 - a. The rankings of the 2 variables is vastly different.
 - b. The rankings of the 2 variables partially agree
 - c. All of the “total variation” is “explained” by the regression line
 - d. The rankings of the 2 variables totally agree

Question 2

- Which is true?
 - a. The Spearman correlation coefficient is the nonparametric analog of the Pearson correlation coefficient.
 - b. The Spearman correlation ranges between 0 and 1.
 - c. The Spearman correlation assesses the association between two nominal variables.
 - d. Some information is lost from the original data when using the Pearson

Answer 2

- Which is true?
 - a. The Spearman correlation coefficient is the nonparametric analog of the Pearson correlation coefficient.
 - b. The Spearman correlation ranges between 0 and 1.
 - c. The Spearman correlation assesses the association between two nominal variables.
 - d. Some information is lost from the original data when using the Pearson

Question 3

- The appropriate null hypothesis for testing no association between two ordinal variables in the population is
 - a. There is no correlation between X and Y in the population
 - b. There is a correlation between X and Y in the population
 - c. There is no correlation between the ranks of X and Y in the population
 - d. There is a correlation between the ranks of X and Y in the population

Answer 3

- The appropriate null hypothesis for testing no association between two ordinal variables in the population is
 - a. There is no correlation between X and Y in the population
 - b. There is a correlation between X and Y in the population
 - c. There is no correlation between the ranks of X and Y in the population
 - d. There is a correlation between the ranks of X and Y in the population

Question 4

- When is it appropriate to use the Spearman correlation to assess the relationship between 2 continuous variables?
 - a. When there are outliers in the data
 - b. When the data is skewed
 - c. When the observations are ranked or ordinal
 - d. All of the above

Answer 4

- When is it appropriate to use the Spearman correlation to assess the relationship between 2 continuous variables?
 - a. When there are outliers in the data
 - b. When the data is skewed
 - c. When the observations are ranked or ordinal
 - d. All of the above

Question 5

- You come across a data set in which the Spearman correlation between 2 variables is 1.0. A friend asks you whether that means that the data, if plotted, would fall on a straight line. You answer
 - a. Yes, since the interpretation is the same as for a Pearson correlation
 - b. No, for a Spearman correlation, 2.0 is perfect association, 1.0 is only modest
 - c. No, but the means are necessarily the same
 - d. No, but the subjects should be in the same order on the two variables
 - e. No, the Spearman correlation is looking for non-linear associations

Answer 5

- You come across a data set in which the Spearman correlation between 2 variables is 1.0. A friend asks you whether that means that the data, if plotted, would fall on a straight line. You answer
 - a. Yes, since the interpretation is the same as for a Pearson correlation
 - b. No, for a Spearman correlation, 2.0 is perfect association, 1.0 is only modest
 - c. No, but the means are necessarily the same
 - d. No, but the subjects should be in the same order on the two variables
 - e. No, the Spearman correlation is looking for non-linear associations

Question 6

- Which of the following would most likely exhibit a correlation of 1.0? (What would you expect the other to be: positive, negative, 1.0, -1.0, or 0)
 - a. Fahrenheit temperature with Celsius temperature
 - b. Age in years with remaining time till death
 - c. Height in years with time to take an exam
 - d. Height in inches with weight in kilograms
 - e. Systolic blood pressure with diastolic pressure

Answer 6

- Which of the following would most likely exhibit a correlation of 1.0? (What would you expect the other to be: positive, negative, 1.0, -1.0, or 0)
 - a. Fahrenheit temperature with Celsius temperature **1.0**
 - b. Age in years with remaining time till death **negative**
 - c. Height in years with time to take an exam **near 0**
 - d. Height in inches with weight in kilograms **positive**
 - e. Systolic blood pressure with diastolic pressure **positive**

Question 7

- A researcher studying multiple sclerosis patients uses an imaging technique to measure the cross-sectional area of the spinal cord at the C2 level. This area is found to have a correlation of -0.75 with duration of the illness, $p < .0001$. The negative correlation, -0.75 suggests
 - a. There is no association between area and duration
 - b. People who have had the disease longer tend to have larger cord areas
 - c. People who have had the disease longer tend to have smaller cord areas
 - d. Both long & short illness durations have small areas; the intermediate durations have the largest areas
 - e. Both long & short illness durations have large areas; the intermediate durations have the smallest areas

Answer 7

- A researcher studying multiple sclerosis patients uses an imaging technique to measure the cross-sectional area of the spinal cord at the C2 level. This area is found to have a correlation of -0.75 with duration of the illness, $p < .0001$. The negative correlation, -0.75 suggests
 - a. There is no association between area and duration
 - b. People who have had the disease longer tend to have larger cord areas
 - c. People who have had the disease longer tend to have smaller cord areas
 - d. Both long & short illness durations have small areas; the intermediate durations have the largest areas
 - e. Both long & short illness durations have large areas; the intermediate durations have the smallest areas

Question 8

- The p value of <0.001 in this data (Question 7) and the correlation provide good statistical evidence that
 - a. The population correlation is ≤ -0.75
 - b. The population correlation is less than 0
 - c. The population correlation is within 0.001 of -0.75
 - d. The population correlation is ≥ -0.75
 - e. No conclusion can be drawn because the p value is too small to be statistically significant

Answer 8

- The p value of <0.001 in this data (Question 7) and the correlation provide good statistical evidence that
 - a. The population correlation is ≤ -0.75
 - b. The population correlation is less than 0*
 - c. The population correlation is within 0.001 of -0.75
 - d. The population correlation is ≥ -0.75
 - e. No conclusion can be drawn because the p value is too small to be statistically significant

* The p-value leads you to reject H_0 that the population correlation=0. It does not allow you to conclude that the population correlation is ≤ -0.75 . Remaining choices are completely wrong

Question 9

- Which of the following characterize the Pearson correlation? (choose all that apply)
 - a. It ranges from -1 to 1
 - b. A value of 0 indicates the absence of any association
 - c. Extreme values have little impact on it
 - d. It is best suited to Gaussian (normally distributed) data
 - e. Values > 0.5 indicate positive association whereas < 0.05 values are negative associations
 - f. H_0 says that the population correlation is 1.0
 - g. H_0 says that the population correlation is 0
 - h. Values > 0.5 in positive direction or < -0.05 in the negative direction are necessary to show causality

Answer 9

- Which of the following characterize the Pearson correlation? (choose all that apply)
 - a. It ranges from -1 to 1
 - b. A value of 0 indicates the absence of any association
 - c. Extreme values have little impact on it
 - d. It is best suited to Gaussian (normally distributed) data
 - e. Values > 0.5 indicate positive association whereas < 0.05 values are negative associations
 - f. H_0 says that the population correlation is 1.0
 - g. H_0 says that the population correlation is 0
 - h. Values > 0.5 in positive direction or < -0.05 in the negative direction are necessary to show causality

* Note that choice (b) is too strong. (c), (e), and (f) are completely wrong

Question 10

- A study* of macular degeneration patients found that age was associated with difficulty in daily activities ($r=0.32$, $p=0.003$, 95% confidence interval: 0.11 to 0.5)

In the (*sample/population*) the correlation between age and difficulty was _____.

In the (*sample/population*) you are _____% sure that the correlation is between _____ and _____.

How likely were the authors to get a correlation of 0.32 or larger entirely by chance if there is truly no association between age and difficulty? _____

* Arch Ophthalmology, 1998; 116:514-520

Answer 10

- A study of macular degeneration patients found that age was associated with difficulty in daily activities ($r=0.32$, $p=0.003$, 95% confidence interval: 0.11 to 0.5)

In the (sample/population) the correlation between age and difficulty was 0.32.

In the (sample/population) you are 95% sure that the correlation is between 0.11 and 0.5.

How likely were the authors to get a correlation of 0.32 or larger entirely by chance if there is truly no association between age and difficulty? 0.003

* Note, the confidence interval will always contain the sample correlation, and we never know directly if it contains the population correlation, so you can't check for that!

"No act of
kindness
however small
is ever
wasted"
Aesop

[This Photo](#) by Unknown Author is licensed under [CC BY](#)