BIOSTATISTICS QUIZ

ANSWERS

- 1. When you read scientific literature, do you know whether the statistical tests that were used were appropriate and why they were used?
 - a. Always
 - b. Mostly
 - c. Rarely
 - d. Never

Answer: Yes and mostly are acceptable. While students are encouraged to continue on with this quiz, if an answer c or d was given it is highly recommended that the student pursue the biostatistics refresher course.

- 2. Which of the following statements are true?
 - a. The p-value is the probability of the sample data arising by chance.
 - b. The p-value is an arbitrary value, designated as the significance level.
 - c. The p-value is the chance of getting an observed effect if the null hypothesis was false.
 - d. The p-value is the chance of getting an observed effect if the null hypothesis was true.
 - e. A very small p-value allows us to say that there is enough evidence to accept the null hypothesis.

Answer: b and d

3. Answer true or false for the following statements:

The 95% confidence interval for the mean:

- a. Contains the sample mean with 95% certainty.
- b. Is less likely to contain the population mean than the 99% confidence interval.
- c. Contains 95% of the observations in the population.
- d. Is approximately equal to the sample mean plus and minus two standard deviations
- e. Can be used to give an indication of whether the sample mean is a precise estimate of the population mean.

Answers:

- a. False: it contains the population mean with 95% certainty. It always contains the sample mean.
- b. True
- c. False: In repeated samples, around 95% of the 95% confidence intervals (CI) will contain the population mean. Another way to think about 95% CI is if the same study were repeated 100 times then the mean of 95 of these 100 studies would lie somewhere within the 95% CI.
- d. False: is approximately equal to \pm two standard errors about the sample mean.
- e. True (Narrow confidence intervals indicate the sample mean is a precise estimate.)

- 4. A study was conducted into the influence of spaying of bitches on their subsequent development of urinary incontinence. Young adult bitches presenting for spaying were randomly allocated to immediate ovariohysterectomy or to a deferred operation 6 months later. The bitches were followed over the 6 months. What type of variable is 'development of urinary incontinence'?
 - a. Qualitative variable
 - b. Quantitative variable
 - c. Categorical variable
 - d. Binary variable
 - e. Continuous variable

Answer: a, c and d.

5. How would you approach the analysis of this study?

Answer:

- Specify the null hypothesis.
- Display data in 2×2 table, cross-classifying the 'spaying status with the ;'incontinence'.
- State and calculate the test statistic to be used (use computer or perform by hand using appropriate formula).
- Specify degrees of freedom and p value.
- Decide whether or not to reject null hypothesis.
- Calculate the confidence interval (computer or by hand using appropriate formula).
- Review whether or not assumptions of the test were satisfied.
- Draw some clinical conclusions from the analysis.
- 6. What statistical analysis would you use to answer the question in the above study?
 - a. Two-sample t-test?
 - b. Correlation?
 - c. Chi-squared test?
 - d. Paired t-test?
 - e. Linear regression?

Answer: c

7. Answer true or false for the following statements:

The paired t-test:

- a. Tests the null hypothesis that the two population means are equal
- b. Must have equally sized numbers of observations in each group.
- c. Assumes that the data in each group are normally distributed.
- d. Is appropriate for comparing the means of independent groups of observations.
- e. When appropriately used, is more powerful when the sample size is large.

Answers:

- a. False: Tests the null hypothesis that the mean of the differences in the population is zero.
- b. True (This is required in order to calculate the difference for each pair of observations.)
- c. False: the paired t-test makes the assumption that the differences between the pairs are normally distributed.
- d. False: the two-sample t-test should be used for comparing the means of independent groups of observations.
- e. True
- 8. Answer true or false for the following statements:

A correlation coefficient:

- a. Should not be calculated when there is an underlying relationship between the two variables but it is not linear.
- b. Does not provide evidence of a causal relationship between two variables.
- c. Should not be used to judge the biological importance of the relationship between two variables.
- d. Should be performed only when certain assumptions are satisfied (e.g. variables measured on a random sample of individuals, both the variables are quantitative and at least one of the two variables need to be normally distributed).

Answers: all true.

9. Review the follow 4 figures and answer the questions below.



Figure 1



Figure 2



Figure 3



Figure 4

Answer true or false for the following statements:

- a. The horizontal line within the box in a box-and-whisker plot in Figure 1 indicates the median value.
- b. The scatter diagram in Figure 2 is useful to present data when we are interested in examining the relationship between two quantitative variables.
- c. The histogram in Figure 3 is skewed to the right (or positively skewed) and can be transformed into a Normal distribution using a logarithmic transformation of the data.
- d. The box plots in Figure 4 demonstrate that the data for 3 groups of dogs are normally distributed.
- e. Figure 1 is an appropriate distribution for a continuous variable.

Answers: all true

- 10. A study was conducted to investigate the relationship between sheep live weight (kg) and its chest girth (cm). A random sample of 66 sheep was weighed and simultaneously had their chest girth measured. Answer true or false for the following statements:
 - a. Analysis of the data from this study could be performed using a two-sample t-test.
 - b. A scatter diagram should be used to present the data.
 - c. Chest girth measurement is a categorical variable.
 - d. Simple linear regression could be used to describe the straight line relationship between sheep live weight and chest girth.
 - e. Sheep live weight can be predicted by measuring chest girth as long as the relationship between these variables is linear and chest girth is measured without error.

Answers:

- a. False. Simple linear regression or correlation analysis should be used to analyse the data.
- b. True.
- c. False. This variable is a continuous variable.
- d. True.
- e. True.