

QUESTIONS

1. Explain the difference between ratio, interval, ordinal, and nominal scales of measurement. Which are used for parametric tests, which are used for nonparametric tests. How can you convert between categorical and continuous data?
2. Explain how skew and kurtosis can affect measurements of central tendency and/or measurement of spread.
3. What is the difference between standard deviation and standard error? Make sure to explain what a sampling distribution is in your answer.
4. Why does sample variance underestimate the true population variance when uncorrected for? What is the appropriate correction?
5. Explain the logic of null hypothesis significance testing.
6. How are the critical values defined for t versus z ? Which test is more powerful: z or t ? Why?
7. What is a “ p value”? What does the p value tell you? How do p values differ from effect size? Does change in sample size affect either p values or effect sizes?
8. How can outliers cause a Type I or a Type II error?
9. Explain the difference between parametric and nonparametric tests. When is it appropriate versus inappropriate to use either type of test and why?
10. In which test(s) do we have the assumption of linearity? Why is this important?
11. What type of chi square can be converted into Pearson’s r ? How would you do this?
12. What is the purpose of calculating adjusted r ?
13. Compare the distributions for Chi square, F , and t . What are the similarities? What are the differences?
14. What is the logic of ANOVA?
15. Explain four issues of machine learning that prevent it from being a golden ticket solution to large scale data analysis.