UNSW PSYCHSOC

PSYC2001
Practice Examination
2021

Questions

1. **What does the Central Limit Theorem state about the sampling distribution of the mean?**

2. **What is the difference between using X/*S* and *chi/sigma* when respectively expressing the mean and standard deviation?**

3. **What does a z-score of -1.68 represent?**

4. **What is the value of alpha for a 99% confidence interval?**

5. **What is a Type II error?**

6. **State the two types of alternative hypotheses for one-tailed t-tests with a single mean**

7. **State three similarities between normal distribution and t-distribution**

8. **What is the difference between standard deviation and standard error?**

9. **When is the null hypothesis retained when dealing with confidence intervals?**

10. **What is a repeated measures design with two levels? What does it mean for the data collected?**

11. **What is the null hypothesis for a research design with independent means?**

12. **When should we use a t-test, and when should we use a z-test?**

13. **Describe what an extraneous variable is, and give examples.**

14. **Distinguish between a dependent and independent samples design.**

15. **Describe random sampling and random allocation and their respective impact on validity.**
16. **Explain the issue of rank order effects and how it can be addressed?**

17. **What are the five factors that affect statistical power?**

18. **What is the difference between delta and gamma? Why would one use delta instead of gamma?**

19. **Which is the only factor under experimenters' control that can manipulate statistical power?**

20. **What is wrong with rejecting the null hypothesis only in a statistical sense?**

21. **Gareth wanted to measure a correlation between whether listening to math rock improves students' statistics marks. The |t| value that he measured was larger than t_c (at the .05 level of significance). Give an example of a conclusion statement about the correlation that Gareth could make.**

22. **Describe the sampling distribution of a correlation (r_{XY}) when rho=0.**

23. **Which value of r is stronger: .326 or -.479? Explain why.**

24. **Calgary decides to examine whether her cohort's PSYC2001 assignment marks are improved when students receive lectures in person. Identify the predictor and criterion variables in her examination.**

25. **What are the three assumptions for linear regression? Explain them.**

26. **Jacen believes that his appointment with his psychologist will cause him to never have depression again. What is the error that he has made, and what is an alternative statement that he could make to fix his error?**

27. **Wai-Yee wants to find a causal inference between seasons and their effect on mood. What type of design should she use to measure this causal relationship?**

28. **Yisha designs an experiment where she examines the effect of three different coffee brands on the concentration of students from UNSW and USYD. What is the factorial design used?**

29. **What are the three assumptions that need to be met in an independent means t tests and why?**

30. If there are 3 classes in a primary school that are tested together rather than separately, is a t-test considered robust? Why or why not?

31. What are the four factors that make a quality journal?

32. Peer review is the current gold-standard for evaluating the quality of research, what is peer review and why is peer review criticised?

33. What is the issue with reporting exploratory analyses as confirmatory?

34. What are the differences between exploratory and confirmatory analysis?

35. What is the experiment-wise error rate if there are 4 hypothesis tests conducted?

36. What is the main reason that only a small percentage of replications are significant, when most of these original significant studies were found to be significant? Explain.

Answers

1. The sampling distribution of the mean becomes more normal as *n* increases, regardless of population distribution.

2. Aramaic letters (*X/S*) are used to represent data from the sample, whilst Greek letters (*chi/sigma*) represent data of the population.

3. The observed value is 1.68 standard deviations below the mean; z-scores are interpreted as how far away an observed score is from the mean in standard deviations. A negative z-score indicates that the observed score is a certain number of standard deviations *below* the mean, whilst a positive z-score indicates that the observed score is a certain number of standard deviations *above* the mean.

4. Alpha is .01. Alpha is the error rate in the form of a decimal probability - since the confidence interval is stated at 99%, this leaves room for 1% of error, which translates to .01 as a decimal (the alpha value).

5. This error occurs when researchers incorrectly retain a false null hypothesis (when it should be rejected). This often occurs when the alpha level is set lower (e.g. from .05 to .01) as the level of error is reduced, and more evidence is needed to reject the null hypothesis.

6. H_1: mu > mu_0 and H_1: mu < mu_0. The alternative hypothesis for two-tailed tests (which is the focus of this course) is H_1: mu=mu_0.

7. The mean is 0, the distribution is symmetric and is unimodal.

8. Standard deviation considers variation within a sample, whilst standard error indicates variance from the sample mean in relation to the population mean.

9. When mu_0 is contained within the interval (for a single mean), since the value proposed by the null hypothesis (which we want to reject) is considered a plausible true value for the population mean within the confidence interval. When a researcher is comparing two means, if a value of 0 is contained within the confidence interval (indicating a mean difference of 0, in which two mean are not significantly different), the null hypothesis is retained.

10. Each participant undergoes the two different levels of the independent variable (e.g. a treatment versus control condition for taking medication) - this means that researchers

will get two scores for each participant, being reduced to one score for each participant that states the *difference* between their behaviours in the two different levels.

11. $mu_1=mu_2$ since there would be no behavioural difference observed between the two levels (i.e. $mu_1-mu_2=0$).

12. t-tests are used when sigma (standard deviation) is unknown, and z-tests are used when sigma is known. Sigma is the true standard deviation of the population.

13. Extraneous variables are uncontrolled variables which may affect the dependent variables in ways that are unaccounted for from independent variables. For example, individual differences such as mood, motivation, intelligence, ability, or environmental factors such as measurement error.

14. An independent samples design conducts random allocation of participants to distinct groups, whilst a dependent samples design measures the same participants across all conditions.

15. Random sampling refers to how participants are sampled from the population, and is representative of the population. This promotes external validity, as there is a wider range of participants that are accounted for, becoming more representative of the population. Random allocation refers to assigning participants randomly to groups to reduce potential extraneous variables/individual differences. This promotes internal validity, as individual differences unaccounted for by the independent variable are spread across the groups.

16. Rank order effects refer to how some stimuli/conditions being presented earlier may be responded to differently by stimuli presented later e.g. participants may simply do better on a second test because they have completed the first test, thus confounding the experiment. Rank order effects can be addressed by counterbalancing, which is assigning participants a random order of conditions.

17. The size of alpha, the size of gamma, the size of n (population), the size of sigma, the directionality of the alternative hypothesis

18. Gamma is measured in units of standard deviation (sigma), whereas delta is measured in units of standard error ($sigma_M$). Delta is used over gamma when dealing with sampling distribution, since sampling distribution requires the standard error of the mean within a sample.

19. Sample size (n) - increasing sample size decreases standard error (as we have more people, finding values closer to the true population mean; also, since sample size is the denominator (n) in the equation of standard error, as n grows, the calculated value of standard error decreases). The other factors that affect statistical power cannot be controlled.

20. All null hypotheses are likely to be rejected (since it is highly unlikely that mu_0 is exactly equal to mu) - the focus should be on rejecting the null hypothesis to an important degree, ensuring that the minimum effect size is of practical importance.

21. "There is sufficient evidence at the .05 level of significance that the correlation between math rock and performance in statistics exams in the population differs from 0."

22. The sampling distribution is skewed - the degree of skew increases as rho approaches +1 or -1. If rho=0, the sampling distribution would be symmetrical and nearly normal (resembling a t-distribution).

23. r= -.479 is stronger than .326 as -.479 is closer to -1 than .326 is closer 1. Whilst -.479 is negative, it shows a stronger negative correlation between two variables, whereas .326 shows a positive but weaker correlation between two variables. The closer r is to -1, there is a stronger negative correlation, whilst the closer r is to 1, there is a stronger positive correlation.

24. The predictor variable (independent variable) is whether students receive lectures in person or not. The criterion variable (dependent variable) is the cohort's PSYC2001 assignment marks. The predictor variable predicts scores on the criterion variable.

25. 1. X and Y form a bivariate normal distribution (both X and Y are normal).
2. X and Y are linearly related (the means of distributions of Y scores for each X fall on a straight line)
3. Variance of distributions of Y scores for each X score is the same (homoscedasticity)

26. Jacen has confused correlation with causality; there is a correlation between seeing a psychologist and lower chances of developing a depressive disorder, however, his appointment with the psychologist only *predicts* the lowered chance of developing a depressive disorder (it does not cause so). Instead, he could say "because I had an appointment with a clinical psychologist, I predict that I will have less likely chances of developing a depressive disorder".

27. Whilst an experimental design (where one independent variable is manipulated) is considered the "golden standard," seasons cannot be manipulated. A quasi-longitudinal design should be used, as behaviour is observed at different occasions over time (i.e. individuals' moods are assessed during each season, to measure a relationship between season and mood). The quasi-independent variable is season, which cannot be manipulated but allocated (i.e. participants' moods are measured when each season occurs).

28. A 3x2 factorial is used, as there are two factors [independent variables] (coffee brand and university). Each factor respectively has three and two levels, making it a 3x2 design. Concentration is the dependent variable that is measured, potentially affected by all levels and factors.

29. 1. Normal distribution of scores in each population
2. Population variances are equal
3. Observations are independent, within and between groups
These assumptions need to be met, so the actual distribution of t values follows the theoretical t distribution, and as a result the Type 1 error rate is maintained at the nominal rate of α.

30. If the participants in the 3 classes are tested together, the behaviour of one participant may influence the behaviour of the other participants (rather than participants acting as they usually would without peer influence). As a result, the t-test is not robust against the violation of independence of observations.

31. 1. Quality of editors and reviewers
2.Acceptance rate
3.Impact factor: average number of times each published paper is cited in another paper
4. Reputation (in psychology, American Psychological Association (APA) journals are often the highest rated)

32. Peer review can be defined as the evaluation of the quality of research by other researchers, it has been criticised as overly conservative, biased and time-consuming.

33. The issue is that the reader assumes a Type 1 error rate of 5%, when the Type 1 error rate is actually much higher because many tests are performed on the data set in an exploratory analyses. For example, if there are multiple hypothesis tests performed on the same data set e.g. two hypothesis tests [$1 - (1 - 0.05)^2 = 0.0975$], the type 1 error rate is 9.75%, rather than 5%.

34. Confirmatory analysis is planned before the data is collected; only tests that are directly related to a pre-specified test are performed and the type 1 error rate is maintained at 5%. In an exploratory analysis, unplanned analysis is performed on the data set (when data has been collected, but changing the hypothesis since the data did not reflect the original hypothesis), in which many tests are performed and the type 1 error rate is not maintained at 5%.

35. $1 - (1 - 0.05)^4 = 0.185$

36. Multiple comparisons: In the original significant studies, multiple hypothesis tests are conducted on the one set of data, which increases the Type I error rate beyond 5%. These are reported as confirmatory analyses (because it is more convenient for experimenters) instead of explanatory, meaning it is reported as having a Type I error rate of 5%. So when a replication of the study is conducted using confirmatory analysis, it is not found to be significant because the right analysis has not been reported in the original study.