**PSY 350 Exam #1 KEY Spring 2018**

1. If you find an **interesting outlier,** you should:
2. always delete it.
3. delete it if it is more than +/- 2 SDs from the mean.
4. revise your hypotheses to include it.
5. compare your results with and without including it.

*The most common incorrect answer here was B; this can help you identify a possible outlier, but if you have reason to believe it’s an* ***interesting*** *outlier and not an error, you shouldn’t automatically delete it. You need to try to understand what might be unusual about it.* ***Lecture 4.7***

1. You calculate the Pearson correlation between hours spent watching television on the evening before an exam and mental alertness during the exam, and you find a value of

***r* = -.89.** What can you conclude from this correlation?

1. There is a very weak, negative, linear relationship.
2. There is a very strong, negative, linear relationship.
3. There is a moderately strong, positive, linear relationship.
4. There is a very strong, positive, nonlinear relationship.

*95% of the class got this answer correct – I’m pretty sure that’s a record. Well done!* ***Lecture 3.6; G&W 15.1-15.3***

1. Suzie is using a popular measure of emotional adjustment. Following the instructions in the test manual, she converts each participant’s score to a z-score, and then she multiplies each z-score by 10 and adds 50 so that the new scores look like those in the manual. What should Suzie expect to happen to the **distribution** of her participants’ scores on this variable?
2. The distribution will change in unpredictable ways; Suzie should not do this.
3. The mean and standard deviation will stay the same, but the shape of the distribution will change.
4. The mean will stay the same, but the standard deviation and the shape of the distribution will change.
5. The mean and standard deviation will change, but the shape of the distribution will stay the same.

*This is an important property of distributions; we can add, subtract, multiply, and divide constants without changing the shape of the distribution or the rank order of scores. We use this property all the time to convert measures so that they are easier to compare and interpret.* ***Lecture 3.2; G& W 5.4-5.7***

1. Summer is conducting a study in which she manipulates participants’ moods (positive or neutral) and then has them play a board game with three other participants. She records whether each participant placed first, second, third, or fourth in the game and uses that as her dependent variable. What type or level of variable is this?
2. Nominal
3. Ordinal
4. Interval
5. Ratio

*This is an ordinal variable, because Summer is keeping track of participants’ rank order, not their absolute scores. We don’t know whether the difference in scores between the person who came first and the person who came second is at all similar to the difference in scores between the person who came second and the person who came first.* ***Lecture 2.1; G&W 1.3***

1. The point of a measure of **central tendency** is to:
2. assess the representativeness of a distribution.
3. evaluate the consistency of a distribution.
4. summarize a distribution in just one number.
5. Q6

*Measures of central tendency are first and foremost summary statistics; they wrap up a bunch of observations into a single number. The most common incorrect answer here was A; I think perhaps some of you were thinking that we could compare the mean of a sample to the mean of a population, but we can’t do that with* ***only*** *a measure of central tendency, and even that is not really an indicator of the representativeness of a distribution (there is not a single statistic that allows us to do that).* ***Lecture 2.4; G&W 3.1-3.5***

1. Which type(s) of reliability does Cronbach’s alpha assess?
2. Interrater reliability.
3. Test-retest reliability.
4. Internal consistency reliability.
5. Alpha is appropriate for any type of reliability.

*Each type of reliability assesses a different type of error and thus requires a different analysis; they are not interchangeable. Cronbach’s alpha is an index of internal consistency reliability.* ***Lecture 4.3.***

1. Randall is conducting a study in which he wants participants to indicate whether they recognize a series of faces, over 300 trials. What should Randall do to **reduce** the likelihood of inattentive responding?
2. Use the long strings approach.
3. Use the psychometric antonym approach.
4. Explain to participants why his study is important.
5. Nothing; inattentive responding is only a concern in surveys.

*The long strings and psychometric antonyms approaches are tools to* ***detect*** *inattentive responding after it has occurred, not to reduce it before it happens. The psychometric antonyms approach wouldn’t be appropriate here because Randall is only asking his participants to indicate recognition or not – there are no terms that could be antonyms. As discussed in class, getting participants on board with the importance of your study* ***does*** *help improve their motivation to pay attention. Inattentive responding can be a concern in all types of research.* ***Lecture 4.5***

1. Which of these is a **random sample**?
2. Obtaining a complete list of the population and using a random number table to select individuals.
3. Choosing a public location and rolling a die to select which individuals to approach.
4. Choosing several different public locations and approaching every 5th person.
5. Obtaining a complete list of the population and selecting the first 100 people.

*To obtain a random sample, the population must be* ***known*** *and you must have a statistically random method of selecting them. This is the only way we can be sure that each individual in the sample has a truly equal chance of being selected.* ***Lecture 5.4; mentioned in G&W 7.1.***

1. Alec measures the number of calories consumed by each of his participants over a three-day healthy eating intervention. To visualize his data, Alec should use:
2. a bar chart, because his data is categorical.
3. a bar chart, because his data is continuous.
4. a histogram, because his data is categorical.
5. a histogram, because his data is continuous.

*I ended up dropping this item because I think several of you were confused about which variable I meant – the calories or the intervention. As I gave you almost no information about the intervention, I had assumed you would focus on the calories. Remember that a frequency distribution summarizes* ***just one variable****, not one variable in relation to another, and that calories are a* ***continuous variable*.** *A histogram is the most appropriate choice here.* ***Lecture 2.3, G & W 2.1-2.3***

1. The Law of Large Numbers states that:
2. the more independent samples we take, the more likely we are to get significant results.
3. the larger the sample, the more likely it is that the sample mean will be close to the population mean.
4. the larger the population mean, the more likely it is that we will estimate it correctly in our sample.
5. the more variability we have in a distribution, the less likely we are to experience sampling error.

*95% correct, well done!*

1. Which approach to handling missing data is most likely to **bias** your analyses in undesirable ways?
2. Pairwise deletion.
3. Mean imputation.
4. Multiple imputation.
5. Maximum likelihood.

*As discussed in class, all of these are strategies that a researcher could use to handle missing data. Pairwise deletion is superior to listwise deletion because you lose less data. Imputation is a way to get around deleting data at all; researchers* ***used*** *to use mean imputation until they realized that it* ***biased their data*** *by reducing the variability (you end up with a lot of values close to the mean). Multiple imputation and maximum likelihood both solve this problem and appear to work quite well.*

1. When we calculate the variance for a sample rather than a population, we use a slightly different formula because:
2. the sample mean is different from the population mean.
3. the sample will always be more variable than the population.
4. the population will always be more variable than any sample.
5. we need to take the representativeness of the sample into account.

*I dropped this item because it was extremely difficult; I think a lot of you were pulled by the distractors. Choice A is a true statement, but it doesn’t have anything to do with how we calculate the variance of a sample. Choice D looked appealing because it is important to use representative samples, but (as mentioned earlier) there is not a single statistic that assesses representativeness. We use the corrected formula (dividing by n-1 rather than n) to calculate the variance for a sample because otherwise we will always underestimate the true variance; no sample is as variable as the entire population. That’s going to be important soon, when we start discussing the difference between a z-distribution and a t-distribution.* ***Lecture 2.6; G & W 4.1-4.6***

1. A correlation is **defined as:**
2. the distance of each variable from the population mean.
3. the probability that a third variable causes both variables.
4. the significance of the difference between the variances of two variables.
5. the ratio of the shared variance between two variables to their total pooled variance.

*The most common incorrect answer here was C, which was a nonsense answer – we have not even discussed significance yet in any context, and we don’t do find a difference between variances when we calculate a correlation. Perhaps some of you were thrown off by the term “pooled variance,” but I am not sure why you wouldn’t read that as a synonym for “vary together” or why you’d prefer C here.* ***Lecture 3.7, G & W 15.2***

1. Kristin finds that most people who score above the mean for one variable in her study tend to score below the mean for another variable. What can we expect Kristin to find in her analyses?
2. these two variables will be unrelated.
3. these two variables will have a positive covariance.
4. these two variables will have a negative covariance.
5. these two variables will have a significant mean difference.

*This is the definition of a negative covariance – individuals with deviations in one direction on one variable tend to have deviations in the other direction on the other variable. The means are not at all relevant here; there’s no reason that the means couldn’t be precisely the same.* ***Lecture 3.6; G & W 15.2-15.3***

1. Which statistic would be most helpful for evaluating whether two raters show high levels of **agreement**?
2. Cohen’s kappa
3. Cronbach’s alpha
4. Pearson correlation
5. Spearman correlation

*Kappa is* ***designed*** *to assess agreement. The most common incorrect answer was C, which is a* ***good*** *index of interrater* ***reliability****, but isn’t enough to tell you about agreement. As discussed in class, two raters can have high levels of reliability (a high correlation) without actually giving the same ratings (agreement).* ***Lecture 4.2.***

1. The Central Limit Theorem states that the distribution of sample means:
2. will have a normal shape only if the original variable is normally distributed.
3. will have a mean equal to the population mean.
4. will have a larger standard deviation than the population standard deviation.
5. will be uncorrelated to the population mean.

*According to the Central Limit Theorem, for sufficiently large samples, the distribution of the sample means will have a mean equal to the population mean and a* ***smaller*** *standard deviation – this is why we use the* ***standard error*** *rather than the SD to evaluate sample means. The CLT also states that because sampling error is random, the distribution of means will be normal even if the original distribution was not normally distributed.* ***Lecture 5.5; G&W 7.2 (p. 200)***

1. If a variable has a large variance, you can confidently conclude that:
2. you have at least one substantial outlier in the data.
3. on average, the individual values in the data are spread out from the mean.
4. your study contains a lot of error and should be viewed with caution.
5. that variable is probably unrelated to any other variable in your study.

*This is the conceptual definition of the variance – on average, how far are the individual observations from the mean? A large variance* ***might*** *be caused by a large outlier, or by a lot of error, but not necessarily – you can’t tell either of those things by looking at the variance.* ***Lecture 2.6. G & W 4.2 – 4.6.***

1. Jodie wants to study people with mild, possibly undiagnosed levels of depression. Jodie needs to collect:
2. a probability sample, because a random sample is necessary for statistical analysis.
3. a probability sample, because otherwise she cannot generalize her conclusions at all.
4. a nonprobability sample, because she cannot obtain a sampling frame for this population.
5. a nonprobability sample, because she is not interested in the probability of depression.

*As discussed in class, the distinction between a probability sample and a nonprobability sample is* ***whether or not you have a sampling frame*** *(which is a comprehensive list of all of the individuals in the population). We very often use nonprobability samples in psychology research for the reason illustrated here: often, our populations are unknown and unknowable. So although the theory of inferential statistics assumes random samples, we perform analyses on nonprobability samples all the time, and we can generalize our conclusions, we just have to be very careful about how we do so. The probability/ nonprobability nature of a sample and the probability of an outcome (depression) have nothing to do with one another.* ***Lecture 5.3.***

1. It would be more appropriate to report the **median** than the mean when:
2. your variable is measured at the nominal level.
3. your variable is distributed symmetrically.
4. your variable has very little variance.
5. your variable is skewed.

*The median is most useful when describing a distribution that is* ***not*** *symmetrical. In a symmetrical distribution, the mean and the median are the same, so there’s no advantage to reporting the median. However, when the distribution is skewed (e.g., household income in the US), the mean can be misleading because it is distorted by the extreme values. In those cases, the median is a more accurate representation of the “typical” value in a distribution. Neither the mean nor the median is appropriate for nominal variables, and whether your variable has much variance is not relevant.* ***Lecture 2.4 & 2.5; G & W 3.1 – 3.6.***

1. Ryan summarizes his analyses by organizing his numeric results logically into rows and columns. He should refer to this summary in his writing as:
2. a table.
3. a figure.
4. an appendix.
5. any of these – they are interchangeable.

*93% correct. Good job, make sure you remember this as you write your papers!*

1. As discussed in lecture, the **validity** of a measure is best **defined** as the degree to which:
2. the measure is widely used and accepted within the scientific community.
3. the measure correlates with other measures of the same underlying variable.
4. we obtain consistent results with the measure across different studies and samples.
5. we have evidence to support the interpretations we want to make about that measure.

*As we discussed in lecture, early textbooks tend to define validity as “whether something measures what it claims to measure.” This is a pretty big oversimplification; the modern view, as presented in class, argues that a measure isn’t “valid” or “invalid” in its own right – we need to think about how we are using that measure and what we are interpreting it to mean. Some interpretations of a measure might be more valid than others; for example, the MMPI has quite a bit of validity evidence for identifying some kinds of mental illness, but very little evidence of validity as a predictor of job performance (yet it is used that way more than you’d think!). In the same way, a count of aggressive behaviors among children in a lab setting might be reasonably interpreted as a measure of aggression, but not so appropriate as an indicator of developmental delays. It’s critically important that researchers not just say “oh, this measure is valid” without thinking about what it is valid* ***for****. Choice B is one of the* ***types*** *of evidence we might collect about validity, but remember that validity is* ***accumulated evidence*** *– not just one type* ***–*** *so**that**answer isn’t a good* ***definition*** *of what validity means.* ***Lecture 4.4***

1. Which of the following is true about the relationship between correlation and causation?
2. Correlation is necessary, but not sufficient, for causation.
3. Causation is necessary, but not sufficient, for correlation.
4. Correlation usually means that causation is *not* present.
5. Causation and correlation are totally unrelated.

*93% correct – again, well done!*

1. Test-retest reliability is **not** very informative when you are:
2. measuring something that should be stable over time.
3. measuring something you expect will change a lot over time.
4. using a well-established, published test.
5. aggregating observations across items or raters.

*Test-retest reliability is designed to measure the amount of error due to time, or short-term temporary fluctuations. It only makes sense to consider those as* ***error*** *if we are pretty confident that the thing we are measuring is stable. If we’re measuring something that ought to change, we* ***expect*** *the test-retest correlation to be low, and so it’s not very informative. We absolutely need to verify the reliability of any measure we use, no matter how often it’s been used before, and we are very often interested in both test-retest and internal consistency reliability at the same time. The two types of reliability address different types of error, which may or may not both be relevant in any particular context.* ***Lecture 4.3***

1. According to the APA style manual, which of these is the **best** way to write about your analyses?
2. “I calculated the mean and standard deviation for each variable.”
3. “The mean and standard deviation for each variable were calculated.”
4. “The researchers calculated the mean and standard deviation for each variable.”
5. “The mean and the standard deviation for each variable were revealed to be…”

*As discussed in lecture and in the APA style manual, APA recommends using* ***first person active voice*** *to write about things you actually did. It is not inherently subjective to use “I” – you are reporting an objective fact,* ***you*** *as the researcher did this. The other options are either passive voice (you can’t tell* ***who*** *performed the action in the sentence) or untrue (you are responsible for your analyses, not some nonexistent “researchers”).* ***Lecture 3.3; APA Manual Sections 3.09 and 3.18****.*

1. On a memory test, Evan obtains a z-score of -1.87 and Lina obtains a z-score of 1.12. Whose z-score is **less likely** (probable)?
2. Evan’s.
3. Lina’s.
4. Both are equally unlikely.
5. It is impossible to tell without more information.

*The size of the z-score – its distance from zero, whether positive or negative – tells us how far from the mean it is. Z-scores farther from the mean are* ***less likely*** *than scores close to the mean. A z-score with a larger absolute value is less likely than a z-score with a smaller absolute value.* ***Lecture 3.1 & 5.2; G & W 5.1-5.3 and 6.2-6.3.***

1. If the distribution of a variable is positively skewed, it means that:
2. there are a few very high values, so that the mean is larger than the median.
3. there are a few very low values, so that the mean is smaller than the median.
4. the mean of the data is higher than we would expect to see by chance.
5. it would be inappropriate to report the median for this variable.

*The direction of skew is defined by* ***where the extreme observations are****, not by where the majority of the data is. A skewed distribution generally has a few extreme values, and those values distort the mean in their own direction (high extreme values make the mean higher, low extreme values make the mean lower). If we didn’t have those values, the rest of the data would be normal. Therefore, we name the skew for the values that cause the skew. A positively skewed distribution has a few extremely high values, and a negatively skewed distribution has a few extremely low values.*