Week 3 Quiz

Quiz, 14 questions

13/14 points (92.85%)

Congratulations! You passed!

Next Item

1 / 1 point

1.

Which of the following is **false** about bootstrapping?

- ( ) A bootstrap confidence interval constructed based on a biased sample will still yield an unbiased estimate for the population parameter of interest.

Correct

This question refers to the following learning objective(s):

Construct bootstrap confidence intervals using one of the following methods:

- Percentile method: XX% confidence level is the middle XX% of the bootstrap distribution.
- Standard error method: If the standard error of the bootstrap distribution is known, and the distribution is nearly normal, the bootstrap interval can also be calculated as $\bar{x}_{boot} \pm z^{\star} SE_{boot}$.

Recognize that when the bootstrap distribution is extremely skewed and sparse, the bootstrap confidence interval may not be appropriate.

If the original sample is biased then the bootstrap distribution will yield a **biased** estimated for the population of interest, bootstrapping is not a remedy for bad data collection.

- ( ) Bootstrap distributions are constructed by sampling with replacement from the original sample, while sampling distributions are constructed by sampling with replacement from the population.
- ( ) Bootstrap distributions that are extremely skewed or have isolated clumps of values may yield unreliable confidence intervals.
- ( ) The endpoints of a 95% bootstrap confidence interval are the cutoff values for the top and bottom 2.5% of the bootstrap distribution.

0 / 1

point

Week 3 Quiz

Quiz, 14 questions

13/14 points (92.85%)

2.

Suppose we wanted to compare the rates of return for two stocks: the technology company Intel and the U.S. airline Southwest Airlines. To compare the rates of return, we take a random sample of 50 days of Intel's stock returns and another random sample of 50 days for Southwest's stock returns (not necessarily the same days). These data should not be treated as paired. Why would these data not be considered paired data?

- 50 observations is not enough to be able to consider the data as paired.
- When random sampling is involved, data can't be treated as paired.
- The data can't be considered paired data because the two companies are in different industries.

This should not be selected

This question refers to the following learning objective(s):

- Define observations as paired if each observation in one dataset has a special correspondence or connection with exactly one observation in the other data set.
- Carry out inference for paired data by first subtracting the paired observations from each other, and then treating the set of differences as a new numerical variable on which to do inference (such as a confidence interval or hypothesis test for the average difference).

If the data were collected on the same day, the industry would not matter.

- The data can't be considered paired data because the days for which we have Intel data may be different from the days for which we have Southwest Airlines data.

---

1 / 1 point

3.

Which of the following is **false** about bootstrap and sampling distributions?

- Both distributions get narrower as the standard deviation decreases.
- Both distributions are created by sampling with replacement from the population.

Correct

This question refers to the following learning objective(s): Describe how bootstrap distributions are constructed, and recognize how they are different from sampling distributions.

The bootstrap distribution is created by sampling with replacement from the original sample, not the population.

Bootstrap distributions are centered at the sample statistic, sampling distributions are centered at the population parameter.

Both distributions are comprised of sample statistics.

---

1 / 1 point

4.

Your friend, who took statistics a few years ago, recently read a study which examined whether there is any difference between the average birth weights of babies born to smoking mothers vs. non-smoking mothers. Your friend asked you to remind him what it means when the study says ``a 95% confidence interval for the difference between the average birth weight from non-smoking mothers and smoking mothers ($\mu_{non} - \mu_{smoke}$) is 0.2 to 0.9 pounds." Of the following possible responses to your friend's question, which is **true** according to the study?

- The study data does not provide convincing evidence (at 10% significance level) of a difference between the average birth weight from smoking mothers and non-smoking mothers.
- The study data does not provide convincing evidence (at 5% significance level) of a difference between the average birth weight from smoking mothers and non-smoking mothers.
- The study's authors are 95% confident that babies born to non-smoking mothers are on average 0.2 to 0.9 pounds lighter than babies born to smoking mothers.
- The study's authors are 95% confident that babies born to non-smoking mothers are on average 0.2 to 0.9 pounds heavier than babies born to smoking mothers.

Correct

This question refers to the following learning objective(s):

- Recognize that a good interpretation of a confidence interval for the difference between two parameters includes a comparative statement (mentioning which group has the larger parameter).
- Recognize that a confidence interval for the difference between two parameters that doesn't include 0 is in agreement with a hypothesis test where the null hypothesis that sets the two parameters equal to each other is rejected.

---

1 / 1 point

5.

When doing inference on a single mean, which of the following is the **correct** justification for using the t-distribution rather than the normal distribution?

Because the standard error estimate may not be accurate.

Week 3 Quiz

Quiz, 14 questions

13/14 points (92.85%)

Correct

With a small sample size our estimate of the standard error as $\frac{s}{\sqrt{n}}$ is not reliable, since the sample standard deviation, s, may not be a reliable estimate for the population standard deviation σ when the sample size is low. We make up for this by using the t instead of the normal distribution.

Because the t-distribution is not symmetric.

All of the above.

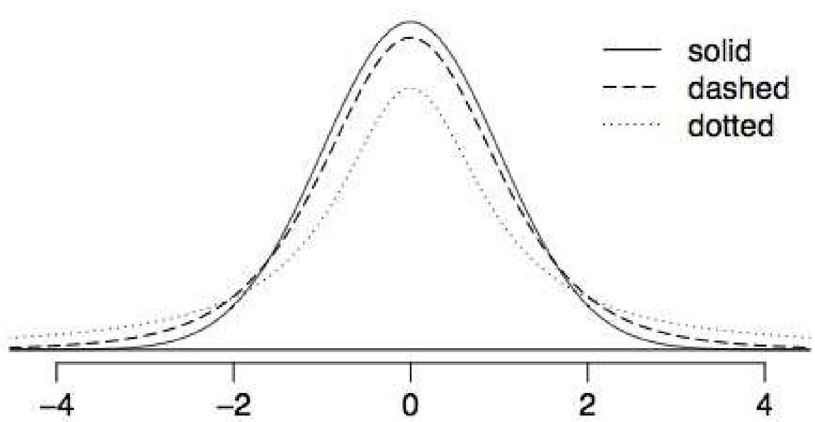
None of the above.

---

1 / 1 point

6.

How does the shape of the t-distribution change as the sample size increases?



It becomes skewed

It becomes more normal looking

Correct

This question refers to the following learning objective(s):

- Describe how the t-distribution is different from the normal distribution, and what ``heavy tail" means in this context.
- Note that the t-distribution has a single parameter, degrees of freedom, and as the degrees of freedom increases this distribution approaches the normal distribution.

As the degrees of freedom increases the t-distribution starts approaching the normal distribution.

- It becomes flatter
- It becomes wider

---

1 / 1 point

7.
Air quality measurements were collected in a random sample of 25 country capitals in 2013, and then again in the same cities in 2014. We would like to use these data to compare average air quality between the two years. Which of the following tests is the **most** appropriate?

- independent samples t-test with one-sided alternative hypothesis
- paired t-test with one-sided alternative hypothesis
- paired t-test with two-sided alternative hypothesis

Correct
We're looking for a difference between the two years' averages, and the data are collected in the same cities in both years, creating dependency requiring paired analysis.

- independent samples t-test with two-sided alternative hypothesis

---

1 / 1 point

8.
We are testing the following hypotheses:

$H_0 : \mu_1 = \mu_2$

$H_A : \mu_1 \neq \mu_2$

The sample sizes are $n_1 = 20$ and $n_2 = 40$. Your friend who is working on this hypothesis test calculates a Z statistic, Z = 2.5. Which of the following is true?

- She should be using an F statistic instead of a Z statistic.
- p-value is approximately 0.006.

She should be using a T statistic instead of a Z statistic.

Week 3 Quiz

Quiz, 14 questions

13/14 points (92.85%)

Correct

This question refers to the following learning objective(s):

- Use a T-statistic, with degrees of freedom $df = n - 1$ for inference for a population mean.
- Use a T-statistic, with degrees of freedom $df = min(n_1 - 1, n_2 - 1)$ for inference for difference between means of two population means using data from two small samples.
- Describe how to obtain a p-value and a critical t-score ($t^\star_{df}$) for a confidence interval.

Since one of the samples is small, she should be using a T statistic instead.

p-value is approximately 0.012.

---

1 / 1 point

9\.
We would like to test if students who are in the social sciences, natural sciences, arts \& humanities, and other fields spend the same amount of time studying for this course. What type of test should we use?

z-test

F-test (ANOVA)

Correct

There are many groups, and we're comparing averages.

t-test for two independent groups

t-test for two dependent groups

---

1 / 1 point

10\.
Which of the following is **most** useful for checking the equal variance across groups condition for ANOVA?

Summary statistics suggesting that the means of each group are roughly equal.

Side-by-side box plots showing roughly equally sized boxes for each group.

Correct
This question refers to the following learning objective(s):

List the conditions necessary for performing ANOVA

- the observations should be independent within and across groups
- the data within each group are nearly normal
- the variability across the groups is about equal

and use graphical diagnostics to check if these conditions are met.

Equal variability across groups is one of the required conditions for ANOVA.

- Summary statistics suggesting roughly equal ranges for each group.
- Histograms suggesting nearly normal distributions of data in each group.

---

1 / 1
point

11.
A study compared five different methods for teaching descriptive statistics. The five methods were traditional lecture and discussion, programmed textbook instruction, programmed text with lectures, computer instruction, and computer instruction with lectures. 45 students were randomly assigned, 9 to each method. After completing the course, students took a 1-hour exam.

Which of the following is the correct degrees of freedom for an F-test for evaluating if the average test scores are different for the different teaching methods?

- $df_G = 4, df_E = 40$

Correct
This question refers to the following learning objective(s): Recognize that the test statistic for ANOVA, the F statistic, is calculated as the ratio of the mean square between groups (MSG, variability between groups) and mean square error (MSE, variability within errors). Also recognize that the F statistic has a right skewed distribution with two different measures of degrees of freedom: one for the numerator ($df_G = k - 1$, where k is the number of groups), and one for the denominator ($df_E = n - k$, where n is the total sample size).

The group degrees of freedom is number of levels (categories) minus 1 ($k - 1 = 5 - 1 = 4$) and the error degrees of freedom is the sample size minus the number of levels ($n - k = 45 - 5 = 40$).

- $df_G = 45, df_E = 4$

$df_G = 5, df_E = 45$

Week 3 Quiz

Quiz, 14 questions

13/14 points (92.85%)

$df_G = 4, df_E = 44$

$df_G = 40, df_E = 4$

1 / 1
point

12.
A study compared five different methods for teaching descriptive statistics. The five methods were traditional lecture and discussion, programmed textbook instruction, programmed text with lectures computer instruction, and computer instruction with lectures. 45 students were randomly assigned, 9 to each method. After completing the course, students took a 1-hour exam. We are interested in finding out if the average test scores are different for the different teaching methods.

The p-value of the test is 0.0168. What is the conclusion of the test?

- Only two group means are significantly different from each other.
- All five group means are equal to each other.
- At most two group means are significantly different from each other.
- All five group means are significantly different from each other.
- At least two group means are significantly different from each other.

Correct
This question refers to the following learning objective(s): Recognize that the null hypothesis in ANOVA sets all means equal to each other, and the alternative hypothesis suggest that at least one mean is different.

- $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$
- H_A : At least one mean is different

The p-value is low so we reject the null hypothesis.

1 / 1
point

13.

For given values of the sample mean and the sample standard deviation when n = 25, you conduct a hypothesis test and obtain a p-value of 0.0667, which leads to non-rejection of the null hypothesis. What will happen to the p-value if the sample size increases (and all else stays the same)?

- Increase
- May either increase or decrease
- Stay the same
- Decrease

Correct

This question refers to the following learning objective(s):

- Use a t-statistic, with degrees of freedom $df = n - 1$ for inference for a population mean:

$$\text{CI: } \bar{x} \pm t^{\star}_{df} SE \qquad \text{HT: } T_{df} = \frac{\bar{x} - \mu}{SE}$$

where $SE = \frac{s}{\sqrt{n}}$.

- Use a t-statistic, with degrees of freedom $df = n_{diff} - 1$ for inference for the difference in two paired (dependent) means:

$$\text{CI: } \bar{x}_{diff} \pm t^{\star}_{df} SE \qquad \text{HT: } T_{df} = \frac{\bar{x}_{diff} - \mu_{diff}}{SE}$$

where $SE = \frac{s}{\sqrt{n}}$. Note that μ_{diff} is often 0, since often $H_0 : \mu_{diff} = 0$.

As the sample size increases the standard error will decrease, which increases the test statistic, and hence decreases the p-value (the tail area).

1 / 1
point

14.

A study compared five different methods for teaching descriptive statistics. The five methods were traditional lecture and discussion, programmed textbook instruction, programmed text with lectures, computer instruction, and computer instruction with lectures. 45 students were randomly assigned, 9 to each method. After completing the course, students took a 1-hour exam. We are interested in finding out if the average test scores are different for the different teaching methods.

If the original significance level for the ANOVA was 0.05, what should be the adjusted significance level for the pairwise tests to compare all pairs of means to each other?

- 0.5

Week 3 Quiz

Quiz, 14 questions

13/14 points (92.85%)

0.05

0.25

0.005

Correct

This question refers to the following learning objective(s): Describe why conducting many t-tests for differences between each pair of means leads to an increased Type 1 Error rate, and we use a corrected significance level (Bonferroni correction, $\alpha^\star = \alpha / K$, where K is the number of comparisons being considered) to combat inflating this error rate.

- Note that $K = \frac{k(k-1)}{2}$, where k is the number of groups.

$K = \frac{5x4}{2}$, ~~$\alpha\star$ = 0.05/10 = 0.005~~