

Name: _____

Score: _____ / _____

Homework 4

1

A sample of 100 young women of Chinese descent in a US college had average BMI 21.5 kg/m² and standard deviation 1.9 kg/m². Suppose you wish to test the hypothesis that the mean BMI in this population is equal to the value of 21.1 kg/m² (consistent with the population of young women in China) against the alternative that the mean BMI is inconsistent with that in a population with mean 21.1 kg/m². The t-statistic you obtain from these data is _____.

Answer Point Value: 3.0 points

Answer Key: 2.1|2.12

2

A sample of 100 young women of Chinese descent in a US college had average BMI 21.5 kg/m² and standard deviation 1.9 kg/m². Suppose you wish to test the hypothesis that the mean BMI in this population is equal to the value of 21.1 kg/m² (consistent with the population of young women in China) against the alternative that the mean BMI is inconsistent with that in a population with mean 21.1 kg/m². The hypothesis test you use will involve a t distribution with _____ degrees of freedom.

Answer Point Value: 3.0 points

Answer Key: 99

3

A sample of 100 young women of Chinese descent in a US college had average BMI 21.5 kg/m² and standard deviation 1.9 kg/m². Suppose you wish to test the hypothesis that the mean BMI in this population is equal to the value of 21.1 kg/m² (consistent with the population of young women in China) against the alternative that the mean BMI is inconsistent with that in a population with mean 21.1 kg/m². Carrying out this test, you obtain the p-value p=_____.

Answer Point Value: 3.0 points

Answer Key: 0.03|0.04

4

Researchers evaluated two versions of a cash transfer program in Zomba, Malawi, which gave money to young girls attending schools and measured their academic performance. One version of the program gave the girls money unconditionally, while the other version gave the girls money only when they had good attendance records. When testing the hypothesis that academic performance of the girls was unrelated to the version of the cash transfer program received against the alternative that there was a difference in academic performance, researchers obtained $p=0.04$. Which interpretations below are incorrect?

☐☐☐☐☐

Answer Point Value: 4.0 points

Answer Key: A,B,C,E

5

This question involves data from the 1991 and 2009 waves of the China Health and Nutrition Study, contained in the file `chinalong.dta`. All participants in this file were 65 or older at the time of the 2009 wave.

Generate a new variable, X , that is the increase in BMI between 1991 and 2009 for each subject. Assume that the population increases in BMI follow an approximate normal distribution with mean μ . The estimated mean increase in BMI in this sample over the 18 year period is ____ kg/m^2 .

Answer Point Value: 3.0 points

Answer Key: 0.84|0.85

6

This question involves data from the 1991 and 2009 waves of the China Health and Nutrition Study, contained in the file `chinalong.dta`. All participants in this file were 65 or older at the time of the 2009 wave.

Which of the attached options describes the appropriate hypothesis testing setup if the researcher would like to test whether the 2009 BMI is the same or lower in 1991 versus the alternative hypothesis that BMI has increased during the time under study?

- ☐
- ☐
- ☐
- ☐

Answer Point Value: 3.0 points

Answer Key: C

7

This question involves data from the 1991 and 2009 waves of the China Health and Nutrition Study, contained in the file `chinalong.dta`. All participants in this file were 65 or older at the time of the 2009 wave.

Carry out a two-sided test of the null hypothesis that there is no change in BMI versus the alternative that there was some sort of change in the BMI. The t-statistic you obtain is _____.

Answer Point Value: 3.0 points

Answer Key: 9.4|9.5

8 This question involves data from the 1991 and 2009 waves of the China Health and Nutrition Study, contained in the file chinalong.dta. All participants in this file were 65 or older at the time of the 2009 wave.

Which is the correct interpretation of the results of the t-test you conducted in the previous problem?

- ☐
- ☐
- ☐
- ☐

Answer Point Value: 3.0 points

Answer Key: B

9 Investigators also wish to know whether changes in BMI depend on smoking status. (Recall smoker=1 for smokers and smoker=0 for nonsmokers.) Conduct a two-sample t-test to compare BMI changes of smokers and nonsmokers. The p-value you obtain from this test is ____.

Answer Point Value: 4.0 points

Answer Key: 0.003|0.004

10 Interpret the results of this t-test in language suitable for publication in a journal, being sure describe the changes in both groups and the estimated difference (and accompanying 95% CI) between changes in both groups in easy to understand terms.

Answer Point Value: 10.0 points

Model Short Answer: -----

11 Suppose the China Health and Nutrition Survey (CHNS) investigators wish to design a follow-up study of participants, measuring BMI again in 2014. They wish to test the hypothesis that there is no change in BMI of participants in the time period 2009-2014.

First, they use the CHNS 1991-2009 data to estimate the standard deviation of change in BMI. Using data from all 821 subjects, their estimated standard deviation is ____.

Answer Point Value: 3.0 points

Answer Key: 2.5|2.6

12

Suppose the China Health and Nutrition Survey investigators wish to design a follow-up study of participants, measuring BMI again in 2014. They wish to test the hypothesis that there is no change in BMI of participants in the time period 2009-2014.

When designing this follow-up study, they assume the standard deviation of the change in BMI over the 5 year period of interest will be 2.5 kg/m^2 . They would ideally like to be able to detect a difference of 0.1 standard deviations, which corresponds to ____ kg/m^2 .

Answer Point Value: 3.0 points

Answer Key: 0.249|0.251

13

Suppose the China Health and Nutrition Survey investigators wish to design a follow-up study of participants, measuring BMI again in 2014. They wish to test the hypothesis that there is no change in BMI of participants in the time period 2009-2014.

Use Stata to complete the following sample size calculation. The investigators need at least ____ participants to test this hypothesis with a minimum detectable difference of 0.1 standard deviations in BMI, assuming a type I error rate of 0.05, 80% power, and a standard deviation of 2.5 kg/m^2 .

NOTE: Fractions of participants are not allowed in sample size calculations! You don't want to be the 0.5 of a person in any study! :) For sample size calculations, always round up to the nearest person (so if your calculation says you need 23.4 people, round up to 24 people...Stata will do this for you).

Answer Point Value: 3.0 points

Answer Key: 784.9|785.1

14

Suppose the China Health and Nutrition Survey investigators wish to design a follow-up study of participants, measuring BMI again in 2014. They wish to test the hypothesis that there is no change in BMI of participants in the time period 2009-2014.

Use Stata to conduct the following sample size calculation. The investigators need at least ____ participants to test this hypothesis with a minimum detectable difference of 0.1 standard deviations in BMI, assuming a type I error rate of 0.05, 90% power, and a standard deviation of 2.5 kg/m²

Answer Point Value: 3.0 points

Answer Key: 1050.9|1051.1

15

Suppose the China Health and Nutrition Survey investigators wish to design a follow-up study of participants, measuring BMI again in 2014. They wish to test the hypothesis that there is no change in BMI of participants in the time period 2009-2014.

The investigators have a limited budget and can only afford to follow up 400 participants. The minimum detectable difference they can detect with $\alpha=0.05$ and 80% power is ____ kg/m².

Answer Point Value: 3.0 points

Answer Key: 0.35|0.36

16

Suppose the China Health and Nutrition Survey investigators wish to design a follow-up study of participants, measuring BMI again in 2014. They wish to test the hypothesis that there is no change in BMI of participants in the time period 2009-2014.

The investigators have a limited budget and can only afford to follow up 400 participants. The minimum detectable difference they can detect with $\alpha=0.05$ and 80% power is ____ standard deviations.

Answer Point Value: 3.0 points

Answer Key: 0.14|0.15

17

The UNC Atherosclerosis Risk in Communities (ARIC) study collected longitudinal data on the level of physical activity among people aged 45-64, summarized in an activity score variable taking integer values from 2 (low activity) to 10 (high activity). These scores were ascertained at baseline (year 0 of the study) and follow-up (year 6), and a change score for each participant was constructed by subtracting the year 0 score from the year 6 score.

One researcher studied a sample of African-American women and found that the estimated mean change score was 0.29 (95% CI=(0.17, 0.41)) among the 295 women who were retired at year 6. The corresponding estimate among the 841 women still working at year 6 was 0.15 (95% CI=(0.05, 0.25)).

Because a t distribution is quite similar to the normal distribution for large samples, we will use critical values for the normal (not t) distribution for this question.

The estimated standard deviation of the change scores among retired women is ____.

Answer Point Value: 4.0 points

Answer Key: 1.05|1.1

18

The UNC Atherosclerosis Risk in Communities (ARIC) study collected longitudinal data on the level of physical activity among people aged 45-64, summarized in an activity score variable taking integer values from 2 (low activity) to 10 (high activity). These scores were ascertained at baseline (year 0 of the study) and follow-up (year 6), and a change score for each participant was constructed by subtracting the year 0 score from the year 6 score.

One researcher studied a sample of African-American women and found that the estimated mean change score was 0.29 (95% CI=(0.17, 0.41)) among the 295 women who were retired at year 6. The corresponding estimate among the 841 women still working at year 6 was 0.15 (95% CI=(0.05, 0.25)).

Because a t distribution is quite similar to the normal distribution for large samples, we will use critical values for the normal (not t) distribution to answer this particular question.

The estimated standard deviation of the mean change score (the standard error of the mean) among retired women is ____.

Answer Point Value: 4.0 points

Answer Key: 0.06|0.07

19

Which test can be used to assess whether the activity levels of retired women are the same at follow-up as they were at baseline?

- ☐
- ☐
- ☐
- ☐

Answer Point Value: 3.0 points

Answer Key: B

20

Suppose you wish to test the hypothesis that there is no change in physical activity level from baseline to follow-up among African-American women who are retired at follow-up, against the alternative that there is some change in activity. Suppose the population mean change is given by μ . Which of the following describes the appropriate hypothesis testing framework?

- ☐
- ☐
- ☐
- ☐

Answer Point Value: 3.0 points

Answer Key: A

21

Evaluate this hypothesis using a critical value of $\alpha=0.05$. In reporting your evaluation, be sure to provide (1) the point estimate of interest, (2) a 95% interval estimate, and (3) the t statistic, degrees of freedom, and p-value. Report the results of your statistical test in language suitable for publication in a scientific journal and be sure to relate the statistical results to the subject matter (examples: physical activity does not change in the 6-year time period; or physical activity does change over time, in fact we estimated an average decrease of 1 point (95% CI= (0.8, 1.2)) in the activity score).

Answer Point Value: 10.0 points

Model Short Answer: -----

22

The investigators next wish to test whether the change is the same among retired and working women. In order to carry out the test, they need to know whether the standard deviations among working and retired African-American women are the same. You test this hypothesis; the correct p-value from this test is....

- ☐
- ☐
- ☐

Answer Point Value: 3.0 points

Answer Key: C

23

Suppose you obtain $p > 0.05$. How do you go about testing the hypothesis that the mean changes among working and retired African-American women are the same?

- ☐
- ☐
- ☐
- ☐

Answer Point Value: 3.0 points

Answer Key: B,C

24

Based on the test you conducted concerning the variances, which version of the t-test should you use?

- ☐
- ☐

Answer Point Value: 3.0 points

Answer Key: B

Test whether the change is the same among retired and working women. Evaluate this hypothesis using a critical value of $\alpha=0.05$. In reporting your evaluation, be sure to provide (1) the point estimate of interest, (2) a 95% interval estimate, and (3) the test statistic, degrees of freedom, and p-value. Report the results of your statistical test in language suitable for publication in a scientific journal and be sure to relate the statistical results to the subject matter (i.e., be sure to say whether the groups are the same, and if they're different, say which one has the biggest change and whether that change represents an increase or a decrease in activity over the six-year time period).

Answer Point Value: 10.0 points

Model Short Answer: -----