GW20 ANSWERS

1. What is the difference between two samples that are dependent and two samples that are independent? Give an example of two dependent samples and two independent samples.

Independent samples are from two distinct and unrelated populations. Example: 30 patients are given an experimental drug and 30 other patients are given a placebo.
Dependent samples are two measurements of a sample from a single population. Example: 30 students are given an assessment test. After a specialized study course, the same students take another assessment test and the scores are compared.

2. What conditions are necessary in order to use the dependent samples t-test for the mean of the difference of two populations?

The data need to be matched by a common characteristic.

In Problems 3-10, classify the two given samples as independent or dependent. Explain your reasoning.

3. Sample 1: The SAT scores for 35 high school students who did not take an SAT preparation course
   Sample 2: The SAT scores for 40 high school students who did take an SAT preparation course

   Independent – Two distinct populations students who took prep class, students who didn't

4. Sample 1: The SAT scores for 44 high school students
   Sample 2: The SAT scores for the same 44 high school students after taking an SAT preparation course

   Dependent – One population, two measurements - same students

5. Sample 1: The weights of 51 adults
   Sample 2: The weights of the same 51 adults after participating in a diet and exercise program for one month

   Dependent – One population, two measurements - same adults

6. Sample 1: The weights of 40 females
   Sample 2: The weights of 40 males

   Unclear, but probably Independent – Two distinct populations, males females

7. Sample 1: The average speed of 23 powerboats using an old hull design
   Sample 2: The average speed of 14 powerboats using a new hull design

   Independent – Two distinct populations old hull, new hull

8. Sample 1: The fuel mileage of 10 cars
   Sample 2: The fuel mileage of the same 10 cars using a fuel additive

   Dependent – One population, two measurements - same cars

9. The table shows the braking distances (in feet) for each of four different sets of tires with the car's anti-lock braking system (ABS) on and with ABS off. The tests were done on ice with cars traveling at 15 miles per hour.

Tire Set	**1**	**2**	**3**	**4**
Braking distance with ABS	**42**	**55**	**43**	**61**
Braking distance without ABS	**58**	**67**	**59**	**75**

   Dependent – One population, two measurements - same tire sets

10. The table shows the heart rates (in beats per minute) of five people before exercising and after.

Person	1	2	3	4	5
Heart Rate before Exercising	42	55	43	61	65
Heart Rate after Exercising	58	67	59	75	90

Dependent – One population, two measurements - same person

11A(a) The goal is do simulate a random sample so we can use the independent two population t- test for means. One example would be cluster sampling, which means I would sample some random hospitals in each region and then collect as much data as I can from those hospitals. I would not support sampling only one region since that could introduce bias into the experiment, simply due to the medical costs differing by region.

11(b) Stratified sampling (by region)

11(c) Bias could occur if by picking a hospital that is not typical or representative.

11B. I would use the t-test since the sample sizes are small and the population variances are unknown. The model uses independent sampling and the sample sizes and histograms assure that the difference of sample means is approximately Normal under the Central Limit Theorem. We do need to know if the population variances are equal or unequal to determine which test to run.

11C. **DESIGN**

Ho: There is no difference in the mean length of hospital stay for patients (age 17 and younger) diagnosed with pneumonia.

Ha: There is a difference in the mean length of hospital stay for patients (age 17 and younger) diagnosed with pneumonia.

Ho: $\mu_1=\mu_2$ **Ha:** $\mu_1 \neq \mu_2$

α**=.05,** the maximum probability of making Type I error, which would be incorrectly claiming waiting times are different.

Since we are assuming the population variances are equal, we should use the **pooled variance t- test** for comparing the means of two independent populations. This will be a two-tailed test.

Reject Ho if p-value < .05

DATA

From Minitab output, **pvalue is .0030 < .05**

Decision: **Reject Ho**

CONCLUSION

There is a difference in the mean length of hospital stay for patients (age 17 and younger) diagnosed with pneumonia. These patients spend less time in the hospital than they did ten years ago, which supports the department's claim.