INDEPENDENT SAMPLES OR DEPENDENT / PAIRED SAMPLES?

For each of the scenarios given below, identify whether the two samples being used to compare the two "groups" are independent or dependent. Also give the two variables of interest in the problem, write down the appropriate hypotheses for answering the question of interest and define the parameters in each case.

1. Last year a random sample of 500 U.S. adults were asked whether or not they or someone in their family owned a small business. These data were then compared to similar data collected on a random sample of 1200 U.S. adults ten years ago. We'd like to use the data to see if there is a difference between the proportions of all U.S. adults who own (or someone in their family owns) a small business for the two time periods (last year versus ten years ago). ***Independent***

   Variables: Own a small business? (comparison variable)
   Year (grouping variable)

   H_o: $p_1 = p_2$ ***H_a: $p_1 \neq p_2$***

   Parameters:
   p_1 = proportion of all U.S. adults who own (or family owns) a small business last year
   p_2 = proportion of all U.S. adults who own (or family owns) a small business ten years ago

2. To compare two programs for training industrial workers to perform a skilled job, 20 workers were randomly selected. 10 were randomly chosen and trained under Method 1 while the remaining 10 were trained using Method 2. Method 1 is the new, experimental method believed to improve performance. After completion of the training, each worker was subjected to a test that measures task performance time. Is there evidence that Method 1 training improved worker efficiency over that of Method 2, on average? ***Independent***

   Variables: Task performance time (comparison variable)
   Method (grouping variable)

   H_o: $\mu_1 = \mu_2$ ***H_a: $\mu_1 < \mu_2$***

   Parameters: ***μ_1 = average performance time for all workers trained under Method 1***
   μ_2 = average performance time for all workers trained under Method 2

3. A manufacturer claims that their boot waterproofing is better than the major brand. In an effort to demonstrate this, the company randomly selected 20 people and randomly assigned a new boot treated with their waterproofing product to one foot and a new boot with the competitor's waterproofing product to the other foot. The study subjects were to wear the boots in bad weather for 1 year and at the end of the year the amount of water damage in each boot was measured (high numbers mean more damage). Is there evidence that the manufacturer's waterproofing product is better than the competition's? ***Dependent***

   Variables: Water damage (comparison variable)
   Brand (grouping variable)

   H_o: $\mu_{manuf} = \mu_{comp}$ ***H_a: $\mu_{manuf} < \mu_{comp}$***

   Parameters: ***μ_{manuf} = mean water damage for all boots treated with manufacturer's product***
   μ_{comp} = mean water damage for all boots treated with competitor's product

4. An analyst for Educational Testing Service wants to compare the mean GMAT scores of students before and after taking a GMAT review course. To do this, ETS randomly selects 50 students who signed up for the course and administers the GMAT to each student twice, before and after the review course. Is there evidence that the review course improves scores for all GMAT takers, on average? ***Dependent***

   Variables: GMAT score (comparison variable)
   Time (grouping variable)

   H_o: $\mu_{before} = \mu_{after}$ H_a: $\mu_{before} < \mu_{after}$

   Parameters: μ_{before} ***= mean score for all GMAT takers before taking the review course***
   μ_{after} ***= mean score for all GMAT takers after taking the review course***

5. Is there a difference in average family spending on health care for households in low and high socioeconomic groups? An economist gathers data on 100 households in each group to try to answer this question. ***Independent***

   Variables: Health care expenditure (comparison variable)
   Socioeconomic group (grouping variable)

   H_o: $\mu_{low} = \mu_{high}$ H_a: $\mu_{low} \neq \mu_{high}$

   Parameters:
   μ_{low} ***= average family spending on health care for all families in low socioeconomic group***
   μ_{high} ***= average family spending on health care for all families in high socioeconomic group***

6. An independent research organization conducts a study to see if there is a difference in lifetime variability of incandescent and LED light bulbs. To do this they randomly select 25 different light intensities ("60W", "40W", etc.) and test both an incandescent and LED bulb of that intensity. Is there enough evidence of a difference? ***Dependent***

   Variables: Lifetime (comparison variable)
   Bulb type (incandescent, LED) (grouping variable

   H_o: $\sigma_I = \sigma_L$ H_a: $\sigma_I \neq \sigma_L$

   Parameters: σ_I ***= lifetime standard deviation of all incandescent bulbs***
   σ_L ***= lifetime standard deviation of all LED bulbs***

7. Is one brand of motorcycle tires better than another? To answer this question a study is conducted where 100 motorcycles are randomly selected, brand A is mounted on the front and brand B is mounted on the back. Treadwear is measured after a year of driving. ***Dependent***

   Variables: treadwear; brand

   H_o: $\mu_A = \mu_B$ H_a: $\mu_A \neq \mu_B$

   Parameters: μ_A ***= mean treadwear for all brand A tires***
   μ_B ***= mean treadwear for all brand B tires***

8. A random sample of consumers was asked both before and after an advertising campaign for Brand X about which brand of soft drink they purchased (Brand X or not). We'd like to use the data to determine if there is an increase in the proportion of consumers who purchase Brand X after the advertising campaign. ***Dependent***

   Variables: Purchase Brand X? (comparison variable)
   Time (grouping variable)

   H_o: $p_{before} = p_{after}$ H_a: $p_{before} < p_{after}$

   Parameters:
   p_{before} ***= proportion of all consumers who purchased Brand X before the advertising campaign***
   p_{after} ***= proportion of all consumers who purchased Brand X after the advertising campaign***

9. Par, Inc., a manufacturer of golf equipment, wants to test a new cut-resistant golf ball. The company hopes the new cut-resistant coating will not only improve durability, but also improve driving distance consistency. The company randomly selects 40 balls of both the new and current models and subjects them to driving tests performed by an automated hitting machine. Are the new balls more consistent? ***Independent***

   Variables: Distance (comparison variable)
   Ball type (new, current) (grouping variable)

   H_o: $\sigma_{new} = \sigma_{current}$ H_a: $\sigma_{new} < \sigma_{current}$

   Parameters: σ_{new} ***= distance standard deviation of all new balls***
   $\sigma_{current}$ ***= distance standard deviation of all current balls***

10. Mutual funds are classified as *load* funds, requiring the investor to pay a fee up front when the initial investment is made, or *no-load* funds which do not require this fee. A random sample of 30 load funds and 30 no-load funds were collected to see if load funds are worth the extra fees by generating a higher return on investment, on average, than no-load funds. Are load funds worth it? ***Independent***

    Variables: Return (comparison variable)
    Type of fund (grouping variable)

    H_o: $\mu_{load} = \mu_{noload}$ H_a: $\mu_{load} > \mu_{noload}$

    Parameters: μ_{load} ***= mean rate of return for all load funds***
    μ_{noload} ***= mean rate of return for all no-load funds***

11. Because the Asian economy faltered in the last few months of 1997, investors anticipated a negative effect on the average earnings of US companies in the fourth quarter of 1997. To see if there was evidence of this downturn, 14 companies were randomly sampled and their fourth-quarter earnings per share numbers in 1996 and 1997 were observed. Do the data show enough evidence of a downturn? ***Dependent***

    Variables: 4^{th} quarter earnings per share (comparison variable)
    Year (grouping variable)

    H_o: $\mu_{1996} = \mu_{1997}$ H_a: $\mu_{1996} > \mu_{1997}$

    Parameters: μ_{1996} ***= mean 4Q 1996 EPS for all US companies***
    μ_{1997} ***= mean 4Q 1997 EPS for all US companies***