

*Caveat:* The appearance of a problem on this review sheet doesn't guarantee that a similar problem will appear on the exam. Similarly, the non-appearance of a kind of a problem doesn't imply that it won't appear on the exam. Last, the number of problems on this review and their difficulty is not to be taken as a statement about the exam. Enjoy!

1. A sociologist believes there is a strong association between combined parental income and college GPA. A large random and independent sample of students is taken, and both the relevant income and the GPA are measured. **What statistics procedure will the sociologist do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test?**
2. A crotchety professor of statistics at Harvard is worried about grade inflation in the United States. The professor takes a random sample of 51 public and private schools across the United States and finds the GPAs of all the students of all the schools for each the years 1983, 1988, 1993, 1998, 2003, 2008, and 2013. **What statistics procedure will the professor do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test?**
3. A large dealer of farm and construction equipment is deciding whether or not to stock John Deere earth-movers or Caterpillar earth-movers (they can't both be sold at the same location). The dealer tests the John Deere bulldozer for 7 days and finds the mean of pounds of dirt moved per day is 27,400 pounds with a standard deviation of 822 pounds. The dealer tests the Caterpillar bulldozer for 11 days and finds the mean of pounds of dirt moved per day is 33,900 pounds with a standard deviation of 551 pounds. Should the dealer stock the Caterpillar? **What statistics procedure will the dealer do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test?**
4. In part because of the tradition of marriages taking place in June, a team of wedding planners believe that people celebrate their nuptials differently in different seasons. The planners think that 50% of weddings take place during the summer, 30% take place during the fall, 10% during the winter, and 10% during the spring. The planners check the number of marriages taking place in a local county for the previous year and assume that it constitutes a random sample for the distribution of all weddings in the state. (They will proceed cautiously and carefully.) **What statistics procedure will the planners do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test?**
5. (Real data reported in the Boston Globe on 4/11/14) A study by New Jersey-based TD Bank found that out of a sample of 233 couples in Greater Boston, 41 percent maintain separate checking accounts so they can indulge themselves without dipping into the household accounts. The survey found that couples here and across the United States maintain a joint checking account for the obvious household expenses: mortgages, groceries, utilities, and other bills. But when they want to splurge on themselves, for a spa treatment, a daily dose of the venti caramel macchiato at Starbucks, or behind-home-plate seats at a Red Sox game, they spend from their individual accounts. An economist has thought for years that the true proportion of couples keeping separate accounts is actually 35%. **What statistics procedure will the economist do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test? The test produces a P-value of .00001—write a conclusion for the economist.**

6. An animal rights advocate is curious about the distribution of pets in some of the various neighborhoods and surrounding cities comprising the greater Boston area. The advocate creates four pet categories, [Dog, cat, other, none], and divides the region into [Back Bay, Beacon Hill, South End, North End, Roxbury, West Roxbury, East Boston, Allston, Brighton, Quincy]. The advocate takes a random sample of households in each of the areas and collects data. **What statistics procedure will the advocate do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test?**

7. A scientist is convinced that there is a cure for baldness. The scientist creates a three-month protocol consisting of daily electrical stimulation of the affected area of the scalp, followed by rubbing a proprietary cream on the affected area. (Quack! Quack!) The scientist collects a random sample of 118 bald or balding men and measures the number of square millimeters of ‘bald scalp’ on each of their heads. The scientist then runs the protocol for the supposed cure for the three months and then measures again the number of square millimeters of ‘bald scalp’ on each of their heads. **What statistics procedure will the scientist do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test? The test produces a P-value of .8899—write a conclusion for the scientist.**

8. A hair-product marketer believes that it is still the case that more women over the age of 40 dye their hair than do men over the age of 40. The marketer takes a random sample of 14 unisex salons at different price points and finds that from a sample of 346 women, 62% color their hair. From a sample of 291 men, 55% dye their hair. **What statistics procedure will the marketer do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test?**

9. A travel industry expert believes that at any given time, the price of airline tickets from New York to London depends on the cost of airplane fuel that day, the unemployment rate that day, the peak temperature of that day in New York, and the exchange rate of the dollar to the British pound on that day. The expert takes a random sample of 27 days from 2013 and collects the relevant data. **What statistics procedure will the expert do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test?**

10. An educational researcher believes that there is an association between the highest level of schooling [high school diploma, bachelor’s degree, graduate or professional degree] and whether or not a person is currently [employed, voluntarily unemployed, or involuntarily unemployed]. The researcher carefully selects a large random sample of people across the state and collects the data. **What statistics procedure will the researcher do with the data? What are the  $H_0$  and the  $H_a$  for the appropriate hypothesis test?**

11. A botanist in Portland, Oregon is interested in the mean height of *all* conifers west of the Mississippi River. (There are many kinds of conifers, including—but not limited to—all the different kinds of pine trees.) The botanist strolls down the street measuring the heights of 12 conifers in the front yards of his neighbors on the block by using a super-sweet laser-range device. The botanist plans on making a confidence interval using Minitab. **Name at least one assumption the botanist failed to satisfy.**