4. A 20-year study of 5000 British adults noted four bad habits: smoking, drinking, inactivity, and poor diet. The study looked to show that there is a higher death rate (proportion who die in a 20-year period) among people with all four bad habits than among people with none of the four bad habits.

   (a) Was this an experiment or observational study? Explain.

   (b) What are the null and alternative hypotheses?

   (c) What would be the result of a Type I error?

   (d) What would be the result of a Type II error?

   Of the 314 people who had all four bad habits, 91 died during the study, while of the 387 people with none of the four bad habits, 32 died during the study.

   (e) Calculate and interpret the *P*-value in context of this study.

5. It is estimated that 17.4% of all U.S. households own a Roth IRA. The American Association of University Professors (AAUP) believes this figure is higher among their members and commissions a study. If 150 out of a random sample of 750 AAUP members own Roth IRAs, is this sufficient evidence to support the AAUP belief?

6. A long accepted measure of the discharge rate (in 1000 $ft^3$/sec) at the mouth of the Mississippi River is 593. To test if this has changed, ten measurements at random times are taken: 590, 596, 592, 588, 589, 594, 590, 586, 591, 589. Is there statistical evidence of a change?

7. A behavior study of high school students looked at whether a higher proportion of boys than girls met a recommended level of physical activity (increased heart rate for 60 minutes/day for at least 5 days during the 7 days before the survey). What is the proper conclusion if 370 out of a random sample of 850 boys and 218 out of an independent random sample of 580 girls met the recommended level of activity?

8. In a random sample of 35 NFL games the average attendance was 68,729 with a standard deviation of 6,110, while in a random sample of 30 Big 10 Conference football games the average attendance was 70,358 with a standard deviation of 9,139. Is there evidence that the average attendance at Big 10 Conference football games is greater than that at NFL games?

9. A study is proposed to compare two treatments for patients with significantly narrowed neck arteries. Some patients will be treated with surgery to remove built-up plaque, while others will be treated with stents to improve circulation. The response variable will be the proportion of patients who suffer a major complication such as a stroke or heart attack within one month of the treatment. The researchers decide to block on whether or not a patient has had a mini stroke in the previous year.

   (a) There are 1000 patients available for this study, half of whom have had a mini stroke in the previous year. Explain a block design to assign patients to treatments.

   (b) Give two methods other than blocking to increase the power of detecting a difference between using surgery versus stents for patients with this medical condition. Explain your choice of methods.

10. A car simulator was used to compare effect on reaction time between DWI (driving while intoxicated) and DWT (driving while texting). Ten volunteers were instructed to drive at 50 mph and then hit the brakes in response to the sudden image of a child darting into the road. A baseline stopping distance was established for each driver. Then one day each driver was tested for stopping distance while driving while texting, and another day the driver was tested after consuming a quantity of alcohol. For each driver, which test was done on the first day was decided by coin toss. The following table gives the extra number of feet necessary to stop at 50 mph for each driver for DWI and DWT.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| DWI | 30 | 26 | 28 | 35 | 42 | 33 | 36 | 28 | 27 | 37 |
| DWT | 30 | 31 | 25 | 39 | 45 | 32 | 38 | 30 | 28 | 38 |

The sample means are $\bar{x}_{DWI} = 32.2$, $\bar{x}_{DWT} = 33.6$, and a two-sample $t$-test, $H_0$: $\mu_{DWI} = \mu_{DWT}$, $H_a$: $\mu_{DWI} \neq \mu_{DWT}$, gives a $P$-value of 0.590, and a conclusion that there is no evidence of a difference between the effect on reaction time between DWI and DWT. Explain why this is not the proper hypothesis test, and then perform the proper test.

## SIX INVESTIGATIVE TASKS

1. An exercise electrocardiogram (EKG) checks for changes in your heart during exercise and is useful in diagnosing coronary artery disease. An EKG has fewer potential side effects but is much less precise than thallium tomography. In one EKG study, 500 volunteers with known coronary artery disease and 500 volunteers with healthy arteries underwent EKG checks. The physicians administering and evaluating the tests did not know the physical condition of any volunteer. The following table gives the numbers of volunteers whom the physicians evaluated as "positive" for coronary disease.

| | Test for coronary disease | |
|---|---|---|
| | Positive | Negative |
| Healthy volunteers | 100 | 400 |
| Volunteers with disease | 305 | 195 |

(a) *Sensitivity* is defined as the probability of a positive test given that the subject has disease. What was the sensitivity of this study?

(b) *Specificity* is defined as the probability of a negative test given that the subject is healthy. What was the specificity of this study?

(c) A valuable tool for assessing the accuracy of such studies is the *positive diagnostic likelihood ratio* ($LR^+$) which gives the ratio of the probability a positive test result will be observed in a diseased person compared to the probability that the same result will be observed in a healthy person.

$$LR^+ = \frac{\text{sensitivity}}{\text{1-specificity}}$$

What was $LR^+$ in this study, and explain why the larger the value of $LR^+$, the more useful the test.

Healthy volunteers

Volunteers with disea

| | coronary disease | |
|---|---|---|
| | Positive | Negative |
| | 100 | 400 |
| se | 305 | 195 |