

Hypothesis Testing Practice

1. The Centers for Disease Control (CDC) reported on trends in weight, height, and body mass index from the 1960's through 2002.¹ The general trend was that Americans were much heavier and slightly taller in 2002 as compared to 1960; both men and women gained approximately 24 pounds, on average, between 1960 and 2002.

In 2002, the mean weight for men was reported at 191 lbs. Suppose that an investigator hypothesizes that weights are even higher in 2006 (i.e., that the trend continued over the subsequent 4 years). In order to test the hypotheses at a significance level of 0.01, we select a random sample of American males in 2006 and measure their weights. Suppose we have resources available to recruit $n = 100$ men into our sample. We weigh each participant and find the average to be 197.1 lbs with a standard deviation of 25.6 lbs.

- a) What are the null and alternative hypotheses for this problem?
 - b) What is the test statistic for this problem?
 - c) What is the p -value here?
 - d) At a significance level of 0.01, what would your decision be?
2. The National Center for Health Statistics (NCHS) published a report in 2005 entitled *Health, United States*, containing extensive information on major trends in the health of Americans. Data are provided for the US population as a whole and for specific ages, sexes, and races.

The NCHS report indicated that in 2002 Americans paid an average of \$3,302 per year on health care and prescription drugs. An investigator hypothesizes that in 2005 expenditures have decreased primarily due to the availability of generic drugs. To test the hypothesis at a significance level of 0.05, a sample of 100 Americans are selected and their expenditures on health care and prescription drugs in 2005 are measured. The sample data has a mean of \$3190 and a standard deviation of \$890. Is there statistical evidence of a reduction in expenditures on health care and prescription drugs in 2005?

- a) What are the null and alternative hypotheses for this problem?
 - b) Explain what both Type I and Type II errors mean for this situation.
 - c) What is the test statistic for this problem?
 - d) What conclusion could you draw and why?
 - e) What is the p -value here?
3. The NCHS reported that the mean total cholesterol level in 2002 for all adults was 203. Total cholesterol levels in participants who attended the seventh examination of the offspring in the Framingham Heart Study are summarized as follows: $n = 3,310$, $\bar{x} = 200.3$, and $s = 36.8$. Is there statistical evidence, at $\alpha = 0.01$, of a difference in mean cholesterol levels in the Framingham Offspring?
 - a) What are the null and alternative hypotheses for this problem?
 - b) What is the test statistic for this problem and what is your conclusion?
 - c) What is the p -value here?

4. Consider again the NCHS-reported mean total cholesterol level in 2002 for all adults of 203. Suppose a new drug is proposed to lower total cholesterol. A study is designed to evaluate the efficacy of the drug in lowering cholesterol. Fifteen patients are enrolled in the study and asked to take the new drug for 6 weeks. At the end of 6 weeks, each patient's total cholesterol level is measured, and the sample statistics are as follows: $n = 15$, $\bar{x} = 195.9$ and $s = 28.7$. Is there statistical evidence of a reduction in mean total cholesterol in patients after using the new drug for 6 weeks?
- What are the null and alternative hypotheses for this problem?
 - What is the test statistic for this problem and what is your conclusion?
 - Explain what a Type II error means for this situation.
5. A PGA (Professional Golf Association) tournament organizer is attempting to determine whether hole (pin) placement has a significant impact on the average number of strokes for the 13th hole on a given golf course. Historically, the pin has been placed in the front right corner of the green, and the historical mean number of strokes for the hole has been 4.25, with a standard deviation of 1.6 strokes. On a particular day during the most recent golf tournament, the pin was placed toward the back of the green and 64 golfers played the hole with the pin in a different place.
- What are the null and alternative hypotheses for this test?
 - On this day, this sample average was 4.75 strokes and the standard deviation was still 1.6. What is the test statistic here?
 - Would you reject the null hypothesis at a significance level of 0.01 or not? Why?
 - What is the p-value for this sample?
 - Based on your p-value calculation above, what would you conclude about the test?
6. A local chamber of commerce claims that the mean family income in a city is \$32250. A bright statistics student suspects this is too high and decides to run a hypothesis test to verify this. A sample of 105 families is selected and the mean is \$31,710 with a standard deviation of \$3480.
- What are the null and alternative hypotheses for this situation?
 - What is your test statistic?
 - At a significance level of 0.05, what is your decision?
7. A health-care actuary has been investigating the possibility that the cost of maintaining the cancer patients within its plan is significantly higher than other types of patients. The non-cancer patients have historically submitted costs of about \$1080 per month. A sample of 36 cancer cases for is taken with an average cost of \$1165 and a standard deviation of \$180. Is there any evidence of a significant change?
- What are the null and alternative hypotheses?
 - What is your test statistic?
 - At a significance level of 0.01, what is your threshold, or cutoff value?
 - What is the decision here?
8. A college bookstore tells prospective students that the average cost of its textbooks is \$52. A group of smart statistics students thinks that the average cost is higher. In order to test the bookstore's claim

against their alternative, the students select a random sample of size 100. Assume that the mean from their random sample is \$52.80 and the standard deviation is \$4.50. Perform a hypothesis test at the 5% level of significance and state your decision.

9. A credit card company is interested in the average balance that is carried by its college-aged cardholders. They are particularly interested in whether the average balance of the college-aged cardholders is significantly different from that of the cardholders who are in their early 20s, who have an average balance of \$1385. A sample of 40 college-aged cardholder accounts reveals an average balance of \$1,250 and a standard deviation of \$350.
 - a) Find a 95% confidence interval for the mean account balance for the college-aged users on their credit cards.
 - b) What would the null and alternative hypotheses be for this?
 - c) What is your decision at $\alpha = 0.01$? Justify your answer.

10. The lifetime of a disk drive head is normally distributed with a population mean of 1000 hours. A new process has been developed and the head engineer for the disk drive processing believes the drives should be able to last longer. A sample of 64 drives is taken and the average of these is 1060 hours with a standard deviation of 120 hours.
 - a) What are the null and alternative hypotheses for this problem?
 - b) What is the test statistic for this problem?
 - c) What is the p-value here?
 - d) At a significance level of 0.05, what would your decision be? How about at a 0.10 level?