Hypothesis Testing Practice Solutions

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?
   Soln: $H_0\colon \mu = 191$ vs $H_A\colon \mu > 191$.

   b) What is the test statistic for this problem?
   Soln: $Z = \frac{197.1-191}{25.6/\sqrt{100}} = 2.38$.

   c) What is the *p*-value here?
   Soln: $P(Z > 2.38) = 1 - 0.9913 = 0.0087$.

   d) At a significance level of 0.01, what would your decision be?
   Soln: At the 0.01 level, I WOULD be able to reject the null hypothesis because the *p*-value = 0.0087 < 0.01 = α. Therefore, I would be inclined to believe the mean weight has increased.

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?
   Soln: $H_0\colon \mu = 3302$ vs $H_A\colon \mu < 3302$.

   b) Explain what both Type I and Type II errors mean for this situation.
   Soln: A Type I error here would mean we conclude that the expenditures ARE lower when they

actually are not. A Type II error would mean we do NOT reject the null hypothesis, concluding the mean expenditures have not lowered, when they actually have.

c) What is the test statistic for this problem?
Soln: $Z = \frac{3190-3302}{890/\sqrt{100}} = -1.26$.

d) What conclusion could you draw and why?
Soln: At 0.05, the one-sided cutoff value is $Z = -1.645$. Since our test statistic is –1.26 > our threshold value of –1.645, we should *not* reject the null hypothesis here. We do not have statistically significant evidence (at $\alpha = 0.05$) that the mean expenditures are lower.

e) What is the *p*-value here?
Soln: $P(Z < -1.26) = 0.1038$, leading to the same conclusion because 0.1038 > 0.05.

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?
Soln: $H_0\text{: } \mu = 203$ vs $H_A\text{: } \mu \neq 203$.

b) What is the test statistic for this problem and what is your conclusion?
Soln: $Z = \frac{200.3-203}{36.8/\sqrt{3310}} = -4.22$. Since –4.22 is below the two-sided cutoff at $\alpha = 0.01$ of –2.575, we can reject the null hypothesis and conclude there IS a difference in mean cholesterol levels.

c) What is the *p*-value here?
Soln: $2 \times P(Z < -4.22) < 0.0001$, leading to the same conclusion as above.

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?

a) What are the null and alternative hypotheses for this problem?
Soln: $H_0\text{: } \mu = 203$ vs $H_A\text{: } \mu < 203$.

b) What is the test statistic for this problem and what is your conclusion?
Soln: $t = \frac{195.9-203}{28.7/\sqrt{15}} = -0.96$. Since –0.96 is NOT below the one-sided cutoff at $\alpha = 0.01$ at 14

degrees of freedom of –2.624, we can NOT reject the null hypothesis and should conclude there is not a lowered average cholesterol level.

c) Explain what a Type II error means for this situation.
Soln: A Type II error means we do NOT reject the null hypothesis, concluding the mean cholesterol levels have not lowered when the drug actually works.

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 13^{th} 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.

a) What are the null and alternative hypotheses for this test?
Soln: Since there is no specific *direction* indicated here, this should be a two-sided test. $H_0\colon \mu = 4.25$ vs $H_A\colon \mu \neq 4.25$.

b) On this day, this sample average was 4.75 strokes and the standard deviation was still 1.6. What is the test statistic here?
Soln: $Z = \frac{4.75-4.25}{1.6/\sqrt{64}} = 2.5.$

c) Would you reject the null hypothesis at a significance level of 0.01 or not? Why?
Soln: The two-sided cutoff at a significance level of 0.01 is 2.575. Since the test statistic is *less than* this threshold, we cannot reject the null hypothesis.

d) What is the p-value for this sample?
Soln: The area beyond $Z = 2.5$ is 1.00 – 0.9938 = 0.0062. Since this was a two-sided test, this area needs to be doubled, giving a p-value of 0.0124.

e) Based on your p-value calculation above, what would you conclude about the test?
Soln: As before, this result indicates that we should *not* reject the null hypothesis, this time because the p-value of 0.0124 > 0.01 = significance level.

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.

a) What are the null and alternative hypotheses for this situation?
Soln: Since there is a suggestion that the quoted average might be *too high,* this is a one-sided test. $H_0\colon \mu = 32250$ vs $H_A\colon \mu < 32250$.

b) What is your test statistic?

Soln: $Z = \frac{31710-32250}{3480/\sqrt{105}} = -1.59.$

c) At a significance level of 0.05, what is your decision?
Soln: The one-sided cutoff at a significance level of 0.05 is –1.645. Since the test statistic does *not* exceed this threshold, we cannot reject the null hypothesis.

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?

   a) What are the null and alternative hypotheses?
   Soln: Since there is a suggestion that the cancer patient costs are *higher,* this is a one-sided test. $H_0\colon \mu = 1080$ vs $H_A\colon \mu > 1080$.

   b) What is your test statistic?
   Soln: Our Z-statistic is $Z = \frac{1165-1080}{180/\sqrt{36}} = 2.833$.

   c) At a significance level of 0.01, what is your threshold, or cutoff value?
   Soln: The one-sided cutoff at a significance level of 0.01 is 2.33.

   d) What is the decision here?
   Soln: Since the test statistic of 2.833 *does* exceed the threshold of 2.33, we can reject the null hypothesis and conclude the cancer patient average is probably higher than the non-cancer average.

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.

   Soln: Since there is a suggestion that the textbook costs are *higher,* this is a one-sided test. $H_0\colon \mu = 52$ vs $H_A\colon \mu > 52$. For the sample of 100, $Z = \frac{52.80-52}{4.50/\sqrt{100}} = 1.778$. The cutoff for a 5% level one-sided test is 1.645, and the test value exceeds this. Therefore, there is enough evidence to suggest the textbook average is higher than $52.

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.
Soln: $95\%CI: 1250 \pm (1.96)\frac{350}{\sqrt{40}} \rightarrow [1141.53,\ 1358.47]$.

b) What would the null and alternative hypotheses be for this?
Soln: Since the question pivots on comparing to some known average for the customers in their early 20s, they should be $H_0: \mu = 1385$ vs $H_A: \mu \neq 1385$.

c) What is your decision at α = 0.01? Justify your answer.
Soln: At 0.01, the two-sided cutoff value is Z = ±2.575. Here, the test statistic is $Z = \frac{1250-1385}{350/\sqrt{40}} = -2.439$. Since this is *not* beyond our threshold, we should *not* reject the null hypothesis here. We should conclude that the college-aged users do not behave significantly differently than those in their early 20s.

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?
Soln: $H_0: \mu \leq 1000$ vs $H_A: \mu > 1000$.

b) What is the test statistic for this problem?
Soln: $Z = \frac{1060-1000}{120/\sqrt{64}} = 4.00$.

c) What is the p-value here?
Soln: $p - \text{value} = P(Z > 4.00)$. From Z table, this value is actually OFF the chart (for the one I gave you), so the area above 4.00 is definitely less than the area above 3.09 (the highest entry on the table), which is 0.001. Therefore, the area above 4.0 will be smaller than 0.001, or approximately 0. So the p-value is essentially 0.

d) At a significance level of 0.05, what would your decision be? How about at a 0.10 level?
Soln: At α = 0.05, I WOULD reject the null hypothesis. I would also reject the null hypothesis at α = 0.01. (p-value < 0.05 AND 0.01, so both levels lead to rejecting the null hypothesis.)