HW7 Solutions

1) Sincich, 6.86

6.86 a. To determine if the mean size of California homes exceeds the national average, we test:

H_0: $\mu = 2230$
H_a: $\mu > 2230$

The test statistic is $z = \dfrac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} = \dfrac{2347 - 2230}{257 / \sqrt{100}} = 4.55$

The rejection region requires $\alpha = .01$ in the upper tail of the z-distribution. From Table IV, Appendix B, $z_{.05} = 2.33$. The rejection region is $z > 2.33$.

Since the observed value of the test statistic falls in the rejection region ($z = 4.55 > 2.33$), H_0 is rejected. There is sufficient evidence to indicate the mean size of California homes exceeds the national average at $\alpha = .01$.

b. To compute the power, we must first set up the rejection regions in terms of .

$$\bar{x}_0 = \mu_0 + z_\alpha \sigma_{\bar{x}} \approx \mu_0 + 2.33\left(\frac{s}{\sqrt{n}}\right) = 2,230 + 2.33\left(\frac{257}{\sqrt{100}}\right) = 2,289.88$$

We would reject H_0 if $\bar{x} > 2,289.88$

The power of the test when $\mu = 2,330$ would be:

$$\text{Power} = P(\bar{x} > 2289.88 \mid \mu = 2,330) = P\left(z > \frac{\bar{x}_0 - \mu_a}{\sigma_{\bar{x}}}\right) = P\left(z > \frac{2,289.88 - 2,330}{257 / \sqrt{100}}\right)$$

$$= P(z > -1.56) = .5 + .4406 = .9406$$

c. The power of the test when $\mu = 2,280$ would be:

$$\text{Power} = P(> 2289.88 \mid \mu = 2,280) = P\left(z > \frac{\bar{x}_0 - \mu_a}{\sigma_{\bar{x}}}\right) = P\left(z > \frac{2,289.88 - 2,280}{257 / \sqrt{100}}\right)$$

$$= P(z > 0.38) = .5 - .1480 = .3520$$

2) Sincich, 7.1

7.1 a. $\mu_1 \pm 2\sigma_{\bar{x}_1} \Rightarrow \mu_1\ 2\frac{\sigma_1}{\sqrt{n_1}} \Rightarrow 150 \pm 2\frac{\sqrt{900}}{\sqrt{100}} \Rightarrow 150 \pm 6 \Rightarrow (144, 156)$

b. $\mu_2 \pm 2\sigma_{\bar{x}_2} \Rightarrow \mu_2 \pm 2\frac{\sigma_2}{\sqrt{n_2}} \Rightarrow 150 \pm 2\frac{\sqrt{1600}}{\sqrt{100}} \Rightarrow 150 \pm 8 \Rightarrow (142, 158)$

c. $\mu_{\bar{x}_1-\bar{x}_2} = \mu_1 - \mu_2 = 150 - 150 = 0$

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} = \sqrt{\frac{900}{100}+\frac{1600}{100}} = \sqrt{\frac{2500}{100}} = 5$$

d. $(\mu_1 - \mu_2) \pm 2\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \Rightarrow (150 - 150) \pm 2\sqrt{\frac{900}{100}+\frac{1600}{100}} \Rightarrow 0 \pm 10 \Rightarrow (-10, 10)$

e. The variability of the difference between the sample means is greater than the variability of the individual sample means.

3) Sincich, 7.2

7.2 a. $\mu_{\bar{x}_1} = \mu_1 = 12$ $\sigma_{\bar{x}_1} = \frac{\sigma_1}{\sqrt{n_1}} = \frac{4}{\sqrt{64}} = .5$

b. $\mu_{\bar{x}_2} = \mu_2 = 10$ $\sigma_{\bar{x}_2} = \frac{\sigma_2}{\sqrt{n_2}} = \frac{3}{\sqrt{64}} = .375$

c. $\mu_{\bar{x}_1-\bar{x}_2} = \mu_1 - \mu_2 = 12 - 10 = 2$

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} = \sqrt{\frac{4^2}{64}+\frac{3^2}{64}} = \sqrt{\frac{25}{64}} = .625$$

d. Since $n_1 \geq 30$ and $n_2 \geq 30$, the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is approximately normal by the Central Limit Theorem.

4) Sincich, 7.3 parts (a), (b), and (e)

7.3 a. For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .025$. From Table IV, Appendix B, $z_{.025} = 1.96$. The confidence interval is:

$$(\bar{x}_1 - \bar{x}_2) \pm z_{.025}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \Rightarrow (5{,}275 - 5{,}240) \pm 1.96\sqrt{\frac{150^2}{400}+\frac{200^2}{400}}$$
$$\Rightarrow 35 \pm 24.5 \Rightarrow (10.5, 59.5)$$

We are 95% confident that the difference between the population means is between 10.5 and 59.5.

b. The test statistic is $z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(5275 - 5240) - 0}{\sqrt{\frac{150^2}{400} + \frac{200^2}{400}}} = 2.8$

The p-value of the test is $P(z \le -2.8) + P(z \ge 2.8) = 2P(z \ge 2.8) = 2(.5 - .4974)$
$= 2(.0026) = .0052$

Since the p-value is so small, there is evidence to reject H_0. There is evidence to indicate the two population means are different for $\alpha > .0052$.

c. The p-value would be half of the p-value in part **b**. The p-value $= P(z \ge 2.8) = .5 - .4974 = .0026$. Since the p-value is so small, there is evidence to reject H_0. There is evidence to indicate the mean for population 1 is larger than the mean for population 2 for $\alpha > .0026$.

d. The test statistic is $z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(5275 - 5240) - 25}{\sqrt{\frac{150^2}{400} + \frac{200^2}{400}}} = .8$

The p-value of the test is $P(z \le -.8) + P(z \ge .8) = 2P(z \ge .8) = 2(.5 - .2881) = 2(.2119) = .4238$

Since the p-value is so large, there is no evidence to reject H_0. There is no evidence to indicate that the difference in the 2 population means is different from 25 for $\alpha \le .10$.

e. We must assume that we have two independent random samples.

5) Sincich, 7.15

7.15 a. Let μ_1 = mean forecast error of buy-side analysts and μ_2 = mean forecast error of sell-side analysts. For confidence coefficient 0.95, $\alpha = .05$ and $\alpha/2 = .05/2 = .025$. From Table IV, Appendix B, $z_{.025} = 1.96$. The 95% confidence interval is:

$$(\bar{x}_1 - \bar{x}_2) \pm z_{.025}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \Rightarrow (.85 - (-.05)) \pm 1.96\sqrt{\frac{1.93^2}{3{,}526} + \frac{.85^2}{58{,}562}} \Rightarrow .90 \pm .064 \Rightarrow (.836,\ .964)$$

We are 95% confident that the difference in the mean forecast error of buy-side analysts and sell-side analysts is between .836 and .964.

b. Based on 95% confidence interval in part **a**, the buy-side analysts has the greater mean forecast error because our interval contains positive numbers.

c. The assumptions about the underlying populations of forecast errors that are necessary for the validity of the inference are:

1. The samples are randomly and independently sampled.
2. The sample sizes are sufficiently large.

6) Sincich, 7.20

7.20 a. The first population is the set of responses for all business students who have access to lecture notes and the second population is the set of responses for all business students not having access to lecture notes.

b. To determine if there is a difference in the mean response of the two groups, we test:

H_0: $\mu_1 - \mu_2 = 0$
H_a: $\mu_1 - \mu_2 \neq 0$

The test statistic is $z = \dfrac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \dfrac{(8.48 - 7.80) - 0}{\sqrt{\dfrac{.94}{86} + \dfrac{2.99}{35}}} = 2.19$

The rejection region requires $\alpha/2 = .01/2 = .005$ in each tail of the z-distribution. From Table IV, Appendix B, $z_{.005} = 2.58$. The rejection region is $z < -2.58$ or $z > 2.58$.

Since the observed value of the test statistic does not fall in the rejection region ($z = 2.19 \not> 2.58$), H_0 is not rejected. There is insufficient evidence to indicate a difference in the mean response of the two groups at $\alpha = .01$.

c. For confidence coefficient .99, $\alpha = .01$ and $\alpha/2 = .01/2 = .005$. From Table IV, Appendix B, $z_{.005} =$ 2.58. The confidence interval is:

$$(\bar{x}_1 - \bar{x}_2) \pm z_{.005}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \Rightarrow (8.48 - 7.80) \pm 2.58\sqrt{\frac{.94}{86} + \frac{2.99}{35}}$$
$$\Rightarrow .68 \pm .801 \Rightarrow (-.121, 1.481)$$

We are 99% confident that the difference in the mean response between the two groups is between $-.121$ and 1.481.

d. A 95% confidence interval would be smaller than the 99% confidence interval. The z value used in the 95% confidence interval is $z_{.025} = 1.96$ compared with the z value used in the 99% confidence interval of $z_{.005} = 2.58$.

7) Sincich, 7.26

7.26 a.

Pair	Difference
1	3
2	2
3	2
4	4
5	0
6	1

$$\bar{d} = \frac{\sum_{i=1}^{n_d} d_i}{n_d} = \frac{12}{6} = 2$$

$$s_d^2 = \frac{\sum_{i=1}^{n_d} d_i^2 - \frac{\left(\sum_{i=1}^{n_d} d_i\right)^2}{n_d}}{n_d - 1} = \frac{\left(34 - \frac{(12)^2}{6}\right)}{5} = 2$$

b. $\mu_d = \mu_1 - \mu_2$

c. For confidence coefficient .95, $\alpha = .05$ and $\alpha/2 = .025$. From Table V, Appendix B, with df $= n_D - 1 = 6 - 1 = 5$, $t_{.025} = 2.571$. The confidence interval is:

$$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n_d}} = 2.571 \frac{\sqrt{2}}{\sqrt{6}} \Rightarrow 2 \pm 1.484 \Rightarrow (.516, 3.484)$$

d. H_0: $\mu_d = 0$
H_a: $\mu_d \neq 0$

The test statistic is $t = t = \frac{\bar{d}}{s_d / \sqrt{n_d}} = \frac{2}{\sqrt{2}/\sqrt{6}} = 3.46$

The rejection region requires $\alpha/2 = .05/2 = .025$ in each tail of the t-distribution with df $= n_D - 1 = 6 - 1 = 5$. From Table V, Appendix B, $t_{.025} = 2.571$. The rejection region is $t < -2.571$ or $t > 2.571$.

Since the observed value of the test statistic falls in the rejection region ($3.46 > 2.571$), H_0 is rejected. There is sufficient evidence to indicate that the mean difference is different from 0 at $\alpha = .05$.

8) Sincich, 7.31

7.31 a. Let μ_1 = the mean salary of technology professionals in 2003 and μ_2 = the mean salary of technology professionals in 2005. Let $\mu_d = \mu_1 - \mu_2$.

To determine if the mean salary of technology professionals at all U.S. metropolitan areas has increased between 2003 and 2005, we test:

H_0: $\mu_1 - \mu_2 = 0$ OR H_0: $\mu_d = 0$

H_a: $\mu_1 - \mu_2 < 0$ H_a: $\mu_d < 0$

b.

Metro Area	2003 Salary ($ thousands)	2005 Salary ($ thousands)	Difference (2003 – 2005)
Silicon Valley	87.7	85.9	1.8
New York	78.6	80.3	–1.7
Washington, D.C.	71.4	77.4	–6.0
Los Angeles	70.8	77.1	–6.3
Denver	73.0	77.1	–4.1
Boston	76.3	80.1	–3.8
Atlanta	73.6	73.2	0.4
Chicago	71.1	73.0	–1.9
Philadelphia	69.5	69.8	–0.3
San Diego	69.0	77.1	–8.1
Seattle	71.0	66.9	4.1
Dallas-Ft. Worth	73.0	71.0	2.0
Detroit	62.3	64.1	–1.8

c. $\bar{d} = \dfrac{\sum_{1}^{n_d} d_i}{n_d} = \dfrac{-25.7}{13} = -1.98$

$$s_d^2 = \frac{\sum_{1}^{n_d} d_i^2 - \frac{\left(\sum_{1}^{n_d} d_i\right)^2}{n_d}}{n_d - 1} = \frac{206.59 - \frac{(-25.7)^2}{13}}{13-1} = 12.9819$$

$$s_d = \sqrt{s_d^2} = \sqrt{12.9819} = 3.6$$

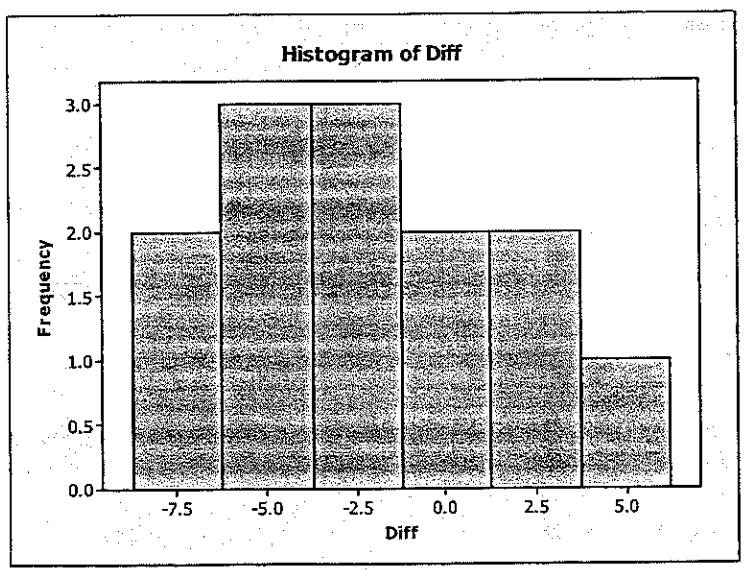
d. The test statistic is $t = \dfrac{\bar{d} - \mu_o}{s_d / \sqrt{n_d}} = \dfrac{-1.98 - 0}{3.6/\sqrt{13}} = -1.98$

e. The rejection region requires $\alpha = .10$ in the lower tail of the t-distribution with $df = n_d - 1 = 13 - 1 = 12$. From Table V, Appendix B, $t_{.10} = 1.356$. The rejection region is $t < -1.356$.

f. Since the observed value of the test statistic falls in the rejection region ($t = -1.98 < -1.356$), H_0 is rejected. There is sufficient evidence to indicate the mean salary of technology professionals at all U.S. metropolitan areas has increased between 2003 and 2005 at $\alpha = .10$.

g. In order for the inference to be valid, we must assume that the population of differences is normal and that we have a random sample.

Using MINITAB, the histogram of the differences is:



The graph is fairly mound-shaped although it is somewhat skewed to the right. Since there are only 13 observations, this graph is close enough to being mound-shaped to indicate the normal assumption is reasonable.