Chec 9

Q2 3/3/1/3

a. With large sample sizes, a 95% confidence interval for the difference of population means, $\mu_1 - \mu_2$, is $(\bar{x}-\bar{y}) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{m}+\frac{s_2^2}{n}} = (\bar{x}-\bar{y}) \pm 1.96\sqrt{[SE(\bar{x})]^2+[SE(\bar{y})]^2}$. Using the values provided, we get $(64.9-63.1) \pm 1.96\sqrt{(.09)^2+(.11)^2} = 1.8 \pm .28 = (1.52, 2.08)$. Therefore, we are 95% confident that the difference in the true mean heights for younger and older women (as defined in the exercise) is between 1.52 inches and 2.08 inches.

b. The null hypothesis states that the true mean height for younger women is 1 inch higher than for older women, i.e. $\mu_1 = \mu_2 + 1$. The alternative hypothesis states that the true mean height for younger women is <u>more</u> than 1 inch higher than for older women.
The test statistic, z, is given by

$$z = \frac{(\bar{x}-\bar{y})-\Delta_0}{SE(\bar{x}-\bar{y})} = \frac{1.8-1}{\sqrt{(.09)^2+(.11)^2}} = 5.63$$

The P-value is $P(Z \geq 5.63) = 1 - \Phi(5.63) \approx 1 - 1 = 0$. Hence, we reject H_0 and conclude that the true mean height for younger women is indeed more than 1 inch higher than for older women.

c. From the calculation above, P-value $= P(Z \geq 5.63) = 1 - \Phi(5.63) \approx 1 - 1 = 0$. Therefore, yes, we would reject H_0 at any reasonable significance level (since the P-value is lower than any reasonable value for α).

d. The actual hypotheses of (b) have not been changed, but the subscripts have been reversed. So, the relevant hypotheses are now H_0: $\mu_2 - \mu_1 = 1$ versus H_a: $\mu_2 - \mu_1 > 1$. Or, equivalently, H_0: $\mu_1 - \mu_2 = -1$ versus H_a: $\mu_1 - \mu_2 < -1$.

Q-n 3 is 10 pt and 5/5

Let μ_1 = the population mean pain level under the control condition and μ_2 = the population mean pain level under the treatment condition.

a. The hypotheses of interest are H_0: $\mu_1 - \mu_2 = 0$ versus H_a: $\mu_1 - \mu_2 > 0$. With the data provided, the test statistic value is $z = \frac{(5.2-3.1)-0}{\sqrt{\frac{2.3^2}{43}+\frac{2.3^2}{43}}} = 4.23$. The corresponding P-value is $P(Z \geq 4.23) = 1 - \Phi(4.23) \approx 0$.
Hence, we reject H_0 at the $\alpha = .01$ level (in fact, at any reasonable level) and conclude that the average pain experienced under treatment is less than the average pain experienced under control.

b. Now the hypotheses are H_0: $\mu_1 - \mu_2 = 1$ versus H_a: $\mu_1 - \mu_2 > 1$. The test statistic value is $z = \frac{(5.2-3.1)-1}{\sqrt{\frac{2.3^2}{43}+\frac{2.3^2}{43}}} = 2.22$, and the P-value is $P(Z \geq 2.22) = 1 - \Phi(2.22) = .0132$. Thus we would reject H_0 at the $\alpha = .05$ level and conclude that mean pain under control condition exceeds that of treatment condition by more than 1 point. However, we would not reach the same decision at the $\alpha = .01$ level (because $.0132 \leq .05$ but $.0132 > .01$).

Q-n 19 is 10 pt.

19. For the given hypotheses, the test statistic is $t = \dfrac{115.7-129.3+10}{\sqrt{\frac{5.03^2}{6}+\frac{5.38^2}{6}}} = \dfrac{-3.6}{3.007} = -1.20$, and the df is

$$v = \frac{(4.2168+4.8241)^2}{\frac{(4.2168)^2}{5}+\frac{(4.8241)^2}{5}} = 9.96$$, so use df = 9. The P-value is $P(T \le -1.20 \text{ when } T \sim t_9) \approx .130$.

Since .130 > .01, we don't reject H_0.

Q-n25 part c) 10 pts for fully done problem but take 2 pts off if the question at the end is not answered explicitly.

c. From the data provided, $\bar{x} = 110.8$, $\bar{y} = 61.7$, $s_1 = 48.7$, $s_2 = 23.8$, and $\nu \approx 15$. The resulting 95% CI for the difference of population means is $(110.8-61.7) \pm t_{.025,15}\sqrt{\frac{48.7^2}{12}+\frac{23.8^2}{14}} = (16.1, 82.0)$. That is, we are 95% confident that wines rated ≥ 93 cost, on average, between \$16.10 and \$82.00 more than wines rated ≤ 89. Since the CI does not include 0, this certainly contradicts the claim that price and quality are unrelated.

Q-n 36 is 10 pts

36. From the data provided, $\bar{d} = 7.25$ and $s_D = 11.8628$. The parameter of interest: μ_D = true average difference of breaking load for fabric in unabraded or abraded condition. The hypotheses are H_0: $\mu_D = 0$ versus H_a: $\mu_D > 0$. The calculated test statistic is $t = \dfrac{7.25-0}{11.8628/\sqrt{8}} = 1.73$; at 7 df, the P-value is roughly .065. Since .065 > .01, we fail to reject H_0 at the $\alpha = .01$ level. The data do not indicate a significant mean difference in breaking load for the two fabric load conditions.

Q-n 37 is 10 pts and 4/6

a. This exercise calls for paired analysis. First, compute the difference between indoor and outdoor concentrations of hexavalent chromium for each of the 33 houses. These 33 differences are summarized as follows: $n = 33$, $\bar{d} = -.4239$, $s_D = .3868$, where d = (indoor value – outdoor value). Then $t_{.025,32} = 2.037$, and a 95% confidence interval for the population mean difference between indoor and outdoor concentration is $-.4239 \pm (2.037)\left(\frac{.3868}{\sqrt{33}}\right) = -.4239 \pm .13715 = (-.5611, -.2868)$. We can be highly confident, at the 95% confidence level, that the true average concentration of hexavalent chromium outdoors exceeds the true average concentration indoors by between .2868 and .5611 nanograms/m^3.

b. A 95% prediction interval for the difference in concentration for the 34th house is $\bar{d} \pm t_{.025,32}\left(s_D\sqrt{1+\frac{1}{n}}\right) = -.4239 \pm (2.037)\left(.3868\sqrt{1+\frac{1}{33}}\right) = (-1.224, .3758)$. This prediction interval means that the indoor concentration may exceed the outdoor concentration by as much as .3758 nanograms/m^3 and that the outdoor concentration may exceed the indoor concentration by a much as 1.224 nanograms/m^3, for the 34th house. Clearly, this is a wide prediction interval, largely because of the amount of variation in the differences.