Solution to Practice Midterm, Oct. 4, 2003,

October 3, 2003

1

a) $\hat{p} = \frac{29}{200} = .145$. The 95% confidence interval for p is:

$$\frac{\hat{p} + \frac{z_{\alpha/2}}{2n} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n} + \frac{z^2_{\alpha/2}}{4n^2}}}{1 + (z_{\alpha/2})^2/n} = (.103, .200)$$

b) Approximately, $n = \frac{4\hat{p}\hat{q}z^2_{\alpha/2}}{d^2}$, to halve the interval, they need 4(200) people in total. So they need to sample 600 more people.
c)$\hat{p_1} = .145$, $\hat{p_2} = .32$, $m = 200$, $n = 300$,
a $\alpha = 0.01$, so 99% confidence interval for the difference is:

$$\hat{p_1} - \hat{p_2} \pm z_{\alpha/2}\sqrt{\frac{\hat{p_1}\hat{q_1}}{m} + \frac{\hat{p_2}\hat{q_2}}{n}} = (-.270, -.080)$$

b $\alpha = .05$, $\hat{p} = .25$, to test $H_0 : p_1 - p_2 \geq 0$ vs. $H_a : p_1 - p_2 < 0$.

$$z = \frac{\hat{p_1} - \hat{p_2}}{\sqrt{\hat{p}\hat{q}(1/m + 1/n)}} = -4.43 < -1.65$$

so there is significant evidence that the TV advertising campaign is effective. We carried out one-sided test.
c Since $z = -4.43$, so p-value$= P(Z < -4.43) < .0001$.

2

i) $\bar{x}_D = 362.8 - 392.9 = -30.1$, $n_D = 20$, $\alpha = .01$;
Test $H_0 : \mu_D = 0$ vs. $H_a : \mu_D \neq 0$.

$$t = |\frac{\bar{x}_d - \mu_D}{s_D/\sqrt{n_D}}| = 2.890 > t_{19,.005} = 2.861$$

Conclusion: There is significant difference in the nutritive value of these two varieties of corn.
ii) $t = 2.890$, p-value$= P(|T_{19}| > 2.89) \approx .01$.

3

a) $\alpha = .01$, $\bar{x} = 611.8$, $\bar{y} = 565.0$, $s_1 = 84.0$, $s_2 = 82.9$, $m = 145$ and $n = 79$;
By assuming $\sigma_1 = \sigma_2 = \sigma$, pooled sample variance:

$$s_p^2 = (\frac{m-1}{m+n-2})s_1^2 + (\frac{n-1}{m+n-2})s_2^2 = (83.6)^2$$

Test: $H_0 : \mu_1 - \mu_2 = 0$ vs. $H_a : \mu_1 - \mu_2 \neq 0$.

$$z = |\frac{\bar{x} - \bar{y}}{s_p\sqrt{\frac{1}{m} + \frac{1}{n}}}| = 4.00 > z_{.005} = 2.58$$

Conclusion: There is significant difference between mean SAT score of male and female students.

b) Test: $H_0 : \mu_1 - \mu_2 \leq 10$ vs. $H_a : \mu_1 - \mu_2 > 10$.

$$z = \frac{\bar{x} - \bar{y} - \Delta}{s_p\sqrt{\frac{1}{m} + \frac{1}{n}}} = 3.15 > z_{.01} = 2.33$$

Conclusion: There is significant evidence that mean SAT score of male students is at least 10 points higher than female students.

c) $\sigma_1 = 84$, $\sigma_2 = 83$

Test: $H_0 : \mu_2 \leq .89\mu_1$ vs. $H_0 : \mu_2 > .89\mu_1$.

Since $\bar{X} \sim N(\mu_1, \frac{84}{\sqrt{145}} = 6.98)$, $\bar{Y} \sim N(\mu_2, \frac{83}{\sqrt{79}} = 9.34)$, so $\bar{Y} - .89\bar{X} \sim N(\mu_2 - .89\mu_1, \sqrt{9.34^2 + (6.98 * .89)^2} = 11.2)$.

$$z = \frac{\bar{x} - .89\bar{y}}{11.2} = 1.894 > z_{.05} = 1.65$$

Conclusion: There is significant evidence that the mean SAT score of female students is at least 89% of the mean score of male students.

4

a) $H_0 : p = .6$ vs. $H_a : p > .6$; $\alpha = .05$ with rejection region $X > K$.

$$P(X > K) = P(\frac{X - np}{\sqrt{npq}} > \frac{K - np}{\sqrt{npq}}) = P(Z > \frac{K - np}{\sqrt{npq}}) = .05$$

$$K = 1.65\sqrt{100(.6)(.4)} + 100(.6) = 68$$

b) Type I error:

$$\alpha = P(X > 63|p = .6) = P(Z > \frac{63 - 100(.6)}{\sqrt{100(.6)(.4)}}) = P(Z > .612) = .271$$

Type II error:

$$\beta = P(X < 63|p = .7) = P(Z < \frac{63 - 100(.7)}{\sqrt{100(.7)(.3)}}) = P(Z < -1.53) = .063$$

5

a) $X_i \sim N(0, \sigma^2), i = 1, ..., n; (i.i.d.)$

$$\begin{aligned} E(\hat{\sigma}^2) &= E(\frac{1}{n}\sum_{i=1}^{n} X_i^2) = \frac{1}{n}E(\sum_{i=1}^{n} X_i^2) \\ &= EX_1^2 = \sigma^2 \end{aligned}$$

b) Since $\frac{X_i}{\sigma} \sim N(0,1)$, so $\sum_{i=1}^{n}(\frac{X_i}{\sigma})^2 \sim \chi_n^2$, i.e. $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi_n^2$. Then 95% confidence interval for σ^2 is:

$$(\frac{n\hat{\sigma}^2}{\chi^2_{n,.975}}, \frac{n\hat{\sigma}^2}{\chi^2_{n,.025}})$$