Homework 3

8.6 (a) $E(\hat{\theta_3}) = aE(\hat{\theta_1}) + (1-a)E(\hat{\theta_2}) = a\theta + (1-a)\theta = \theta$

(b) $V(\hat{\theta_3}) = a^2V(\hat{\theta_1}) + (1-a)^2V(\hat{\theta_2}) = a^2\sigma_1^2 + (1-a)^2\sigma_2^2$, since it was assumed that $\hat{\theta_1}$ and $\hat{\theta_2}$ are independent. To minimize $V(\hat{\theta_3})$, we can take the first derivative (with respect to a), set it equal to zero, to find

$$a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

(One should verify that the second derivative test shows that this is indeed a minimum.)

8.17 It is given that $\hat{p}_1$ is unbiased, and since $E(Y) = np$, $E(\hat{p}_2) = \frac{np+1}{n+2}$.

(a) $B(\hat{p}_2) = \frac{np+1}{n+2} - p = \frac{1-2p}{n+2}$

(b) Since $\hat{p}_1$ is unbiased, $MSE(\hat{p}_1) = V(\hat{p}_1) = \frac{p(1-p)}{n}$. $MSE(\hat{p}_2) = V(\hat{p}_2) + B(\hat{p}_2) = \frac{np(1-p)+(1-2p)^2}{(n+2)^2}$.

(c) Considering the inequality

$$\frac{np(1-p)+(1-2p)^2}{(n+2)^2} > \frac{p(1-p)}{n},$$

this can be written as

$$(8n+4)p^2 - (8n+4)p + n > 0.$$

Solving for p using the quadratic formula, p will be away from .5 by at least $\sqrt{\frac{n+1}{8n+4}}$.

8.28 The point estimate is given by the difference of the sample proportions: $.70 - .54 = .16$ and an error bound is $2\sqrt{\frac{.7(.3)}{180} + \frac{.54(.46)}{100}} = .121$.

8.90 (a) For the 95% CI for the difference in mean verbal scores, the pooled sample variance is $s_p^2 = \frac{14(42)^2+14(45)^2}{28} = 1894.5$ and thus

$$446 - 534 \pm 2.048\sqrt{1894.5(\frac{2}{15})} = -88 \pm 32.55 \quad or \quad (-120.55, -55.45)$$

.

(b) For the 95% CI for the difference in mean math scores, the pooled sample variance is $s_p^2 = \frac{14(57)^2+14(52)^2}{28} = 2976.5$ and thus

$$548 - 517 \pm 2.048\sqrt{2976.5(\frac{2}{15})} = 31 \pm 40.80 \quad or \quad (-9.8, 71.8)$$

.

(c) At the 95% confidence level, there appears to be a difference in the two mean verbal SAT scores achieved by the two groups. However, a difference is not seen in the math SAT scores.

(d) We assumed that the sample measurements were independently drawn from normal populations with $\sigma_1 = \sigma_2$.

8.93 (a) Since the two random samples are assumed to be independent and normally distributed, the quantity $2\bar{X} + \bar{Y}$ is normally distributed with mean $2\mu_1 + \mu_2$ and variance $(\frac{4}{n} + \frac{3}{m})\sigma^2$. Thus, is as σ^2 is known, then $2\bar{X} + \bar{Y} \pm 1.96\sigma\sqrt{\frac{4}{n} + \frac{3}{m}}$ is a 95% CI for $2\mu_1 + \mu_2$.

(b) Recall the $\frac{1}{\sigma^2}\Sigma_{i=1}^{n}(X_i - \bar{X})^2$ has a chi-square distribution with $n-1$ degrees of freedom. Thus, $[\frac{1}{3\sigma^2}]\Sigma_{i=1}^{m}(Y_i - \bar{Y})^2$ is chi-square with $(m-1)$ degrees of freedom and the sum of these is chi-square with $n+m-2$ degrees of freedom. Then , by using Definition 7.2, the quantity

$$T = \frac{2\bar{X} + \bar{Y} - (2\mu_1 + \mu_2)}{\sigma\sqrt{\frac{4}{n} + \frac{3}{m}}},$$

where

$$\hat{\sigma}^2 = \frac{\Sigma_{i=1}^{n}(X_i - \bar{X})^2 + \frac{1}{3}\Sigma_{i=1}^{m}(Y_i - \bar{Y})^2}{n+m-2}.$$

Then , the 95% CI is given by $2\bar{X} + \bar{Y} \pm t_{.025}\sigma\sqrt{\frac{4}{n} + \frac{3}{m}}$.

Question 6 Denote Y as the number of phone calls professor receive each night, and suppose professor has friends $1, 2, \ldots, K$, each has probability $p_1, p_2, \ldots, p_k$ (which are very small and sum to 1) to make a phone call to professor. Then,

$$\begin{aligned} P(Y = 1) &= p_1(1-p_2)(1-p_3)\cdots(1-p_k) \\ &\quad + p_2(1-p_1)\cdots(1-p_k) \\ &\quad + \vdots \\ &\quad + (1-p_1)(1-p_2)\cdots(1-p_{k-1})p_k \\ &\approx p_1 e^{-1} + \ldots + p_k e^{-1} \\ &\approx e^{-1} \end{aligned}$$