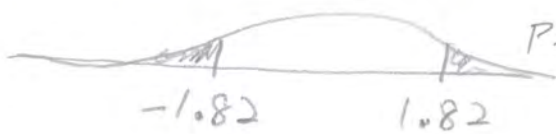
Assignment 7

(1) $Z = -1.82$

a) $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$

i)



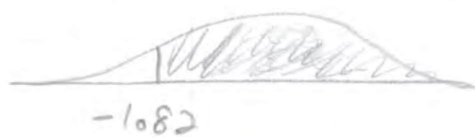
$$P = 2 * \Phi(-1.82) = 2 * 0.0344 = 0.0688$$

ii) No

iii) $\theta = \theta_0$

b) $H_0: \theta \leq \theta_0$ vs $H_1: \theta > \theta_0$

i)



$$p = 1 - \Phi(-1.82) = 1 - 0.0344 = 0.9656$$

ii) No

iii) $\theta \leq \theta_0$

c) $H_0: \theta \geq \theta_0$ vs $H_1: \theta < \theta_0$

i)

$$p = \Phi(-1.82) = 0.0344$$

(ii) Yes

iii) $\theta < \theta_0$

2

a) $Y_i \sim N(\beta x_i, \sigma^2)$, independent

b) $\ell(\beta) = \ln \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i - \beta x_i)^2}$

$$= \ln\left[\sigma^{-n}(2\pi)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\beta x_i)^2}\right]$$

$$= \ln\left(\sigma^{-n}(2\pi)^{-n/2}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2$$

$$\ell'(\beta) = 0 - \frac{1}{\cancel{2}\sigma^2}\sum_{i=1}^{n}\cancel{2}(y_i - \beta x_i)(-x_i)$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i y_i - \beta x_i^2) = \frac{1}{\sigma^2}\left(\sum_{i=1}^{n} x_i y_i - \beta\sum_{i=1}^{n} x_i^2\right) \stackrel{\text{set}}{=} 0$$

$$\Longrightarrow \sum_{i=1}^{n} x_i y_i = \beta \sum_{i=1}^{n} x_i^2 \Longrightarrow \hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

Also the least-squares estimate

c) $\hat{\beta}$ is a linear combination of normals so it is normal

$$E(\hat{\beta}) = E\left(\frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}\right) = \frac{\sum_{i=1}^{n} x_i E(Y_i)}{\sum_{i=1}^{n} x_i^2}$$

$$= \frac{\sum_{i=1}^{n} x_i x_i \beta}{\sum_{i=1}^{n} x_i^2} = \beta\,\frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^2} = \beta$$

unbiased

2c continued

$$Var(\hat{\beta}) = Var\left(\frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}\right)$$

$$= \left(\frac{1}{\sum_{i=1}^{n} x_i^2}\right)^2 Var \sum_{i=1}^{n} x_i Y_i \overset{\overset{ind}{\downarrow}}{=} \frac{1}{(\sum x_i^2)^2} \sum_{i=1}^{n} Var(x_i Y_i)$$

$$= \frac{1}{(\sum x_i^2)^2} \sum_{i=1}^{n} x_i^2 Var(Y_i) = \frac{1}{(\sum x_i^2)^2} \sum_{i=1}^{n} x_i^2 \sigma^2$$

$$= \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}, \text{ so } \hat{\beta} \sim N\left(\beta, \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}\right)$$

d) $$Z = \frac{\hat{\beta} - \beta_0}{\sqrt{\sigma^2 / \sum x_i^2}} = \frac{\sqrt{\sum_{i=1}^{n} x_i^2}\,(\hat{\beta} - \beta_0)}{\sigma}$$

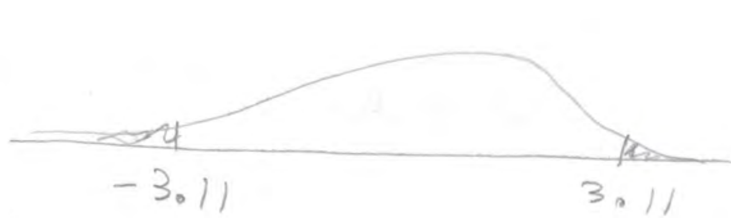
$\sim N(0,1)$ if $H_0: \beta = \beta_0$ is true

e) $\sum_{i=1}^{n} x_i y_i = 32.88$, $\sum_{i=1}^{n} x_i^2 = 28$, $\hat{\beta} = \frac{32.88}{28} = 1.174$

$$Z = \frac{\sqrt{\sum x_i^2}\,(\hat{\beta} - \beta_0)}{\sigma} = \frac{\sqrt{28}\,(1.174 - 0)}{2} = 3.11$$

$= 3.11$

(2f)



$$P = 2 * \Phi(-3.11) = 2 * 0.0009$$
$$= 0.0018$$

g) Critical value is 2.576; Yes

h) $\beta > 0$

(3) a) $\text{Power} = P_\theta\left(|Z_n| \geq z_{1-\alpha/2}\right)$

$$= 1 - P_\theta\left(-z_{1-\alpha/2} < \frac{\sqrt{n}(\bar{X}_n - \theta_0)}{\sqrt{\theta_0(1-\theta_0)}} < z_{1-\alpha/2}\right)$$

$$= 1 - P_\theta\left(-z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta_0(1-\theta_0)}{n}} < \bar{X}_n - \theta_0 < z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta_0(1-\theta_0)}{n}}\right)$$

$$= 1 - P_\theta\left(\theta_0 - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta_0(1-\theta_0)}{n}} < \bar{X}_n < \theta_0 + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta_0(1-\theta_0)}{n}}\right)$$

$$= 1 - P_\theta\left[\theta_0 - \theta - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta_0(1-\theta_0)}{n}} < \bar{X}_n - \theta < \theta_0 - \theta + z_{1-\alpha/2}\sqrt{\frac{\theta_0(1-\theta_0)}{n}}\right]$$

3a continued

$$= 1 - P_\theta\left[\sqrt{n}(\theta_0 - \theta) - z_{1-\frac{\alpha}{2}}\sqrt{\theta_0(1-\theta_0)} < \sqrt{n}(\bar{X}_n - \theta) < \sqrt{n}(\theta_0 - \theta) + z_{1-\frac{\alpha}{2}}\sqrt{\theta_0(1-\theta_0)}\right]$$

$$= 1 - P_\theta\left[\frac{\sqrt{n}(\theta_0 - \theta)}{\sqrt{\theta(1-\theta)}} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta_0(1-\theta_0)}{\theta(1-\theta)}} < \frac{\sqrt{n}(\bar{X}_n - \theta)}{\sqrt{\theta(1-\theta)}} < \frac{\sqrt{n}(\theta_0 - \theta)}{\sqrt{\theta(1-\theta)}} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta_0(1-\theta_0)}{\theta(1-\theta)}}\right]$$

$Z \sim N(0,1)$

$$\overset{n\to\infty}{\approx} 1 - \left(\underbrace{\Phi\left(\frac{\sqrt{n}(\theta_0 - \theta)}{\sqrt{\theta(1-\theta)}} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{\theta_0(1-\theta_0)}{\theta(1-\theta)}}\right)}_{A} - \underbrace{\Phi\left(\frac{\sqrt{n}(\theta_0 - \theta)}{\sqrt{\theta(1-\theta)}} - z_{1-\alpha/2}\sqrt{\frac{\theta_0(1-\theta_0)}{\theta(1-\theta)}}\right)}_{B}\right)$$

$$= 1 + B - A$$

3b When $\theta = \theta_0$, $A = \Phi\left(0 + z_{1-\frac{\alpha}{2}}(1)\right)$

$$= \Phi\left(z_{1-\alpha/2}\right) = 1-\alpha/2$$

And $B = \Phi\left(0 - z_{1-\alpha/2}(1)\right) = \Phi\left(-z_{1-\alpha/2}\right) = \alpha/2$

And $1 + B - A = 1 + \frac{\alpha}{2} - (1-\alpha/2)$

$$= 1 + \frac{\alpha}{2} - 1 + \frac{\alpha}{2} = \alpha$$

c) $A = \Phi\left(\frac{\sqrt{100}\,(0.5-0.45)}{\sqrt{0.45*0.55}} + 1.96\sqrt{\frac{0.5*0.5}{0.45*0.55}}\right)$

$$= \Phi(2.97) = 1 - \Phi(-2.97)$$

$$= 1 - 0.0015 = 0.9985$$

$B = \Phi\left(\frac{\sqrt{100}\,(0.5-0.45)}{\sqrt{0.45*0.55}} - 1.96\sqrt{\frac{.5*.5}{0.45*0.55}}\right)$

$$= \Phi(-0.96) = 0.1685$$, and

Power $= 1 + B - A = 1 + 0.1685 - 0.9985$

$= 0.17$ which is pathetic.

Do we like this probability of a correct decision?

3 d

$$A = \Phi\left(\frac{\sqrt{783}\,(0.5-0.45)}{\sqrt{0.45 \times 0.55}} + 1.96\sqrt{\frac{0.5 \times 0.5}{0.45 \times 0.55}} \right)$$

$= \Phi(4.78) \simeq 1$ off the charts

$$B = \Phi\left(\frac{\sqrt{783}\,(0.5-0.45)}{\sqrt{0.45 \times 0.55}} - 1.96\sqrt{\frac{0.5 \times 0.5}{0.45 \times 0.55}} \right)$$

$= \Phi(0.84) = 1 - \Phi(-0.84) = 1 - 0.2005$

$= 0.7995$

Power $= 1 + B - A = 1 + 0.7995 - 1$

$= 0.7995$ That's better, almost 80%

But it's a lot of people for a taste test.

(4) Power $= P_\lambda(Z_n > z_{1-\alpha})$

$$= P_\lambda\left(\frac{\sqrt{n}(\bar{X}_n - \lambda_0)}{\sqrt{\lambda_0}} \geq z_{1-\alpha}\right)$$

$$= P_\lambda\left(\bar{X}_n - \lambda_0 \geq \frac{\sqrt{\lambda_0}}{\sqrt{n}} z_{1-\alpha}\right)$$

$$= P_\lambda\left(\bar{X}_n \geq \lambda_0 + \frac{\sqrt{\lambda_0}}{\sqrt{n}} z_{1-\alpha}\right)$$

$$= P_\lambda\left(\underbrace{\frac{\sqrt{n}(\bar{X}_n - \lambda)}{\sqrt{\lambda}}}_{Z \dot\sim N(0,1)} \geq \frac{\sqrt{n}(\lambda_0 - \lambda)}{\sqrt{\lambda}} + \frac{\sqrt{\lambda_0}}{\sqrt{\lambda}} z_{1-\alpha}\right)$$

$$\approx 1 - \Phi\left(\frac{\sqrt{n}(\lambda_0 - \lambda)}{\sqrt{\lambda}} + \frac{\sqrt{\lambda_0}}{\sqrt{\lambda}} z_{1-\alpha}\right)$$

with $\alpha = 0.05$, $\lambda_0 = 8$ & $\lambda = 9$, equals

$$1 - \Phi\left(\frac{\sqrt{n}(8-9)}{\sqrt{9}} + \sqrt{\frac{8}{9}} \times 1.645\right)$$

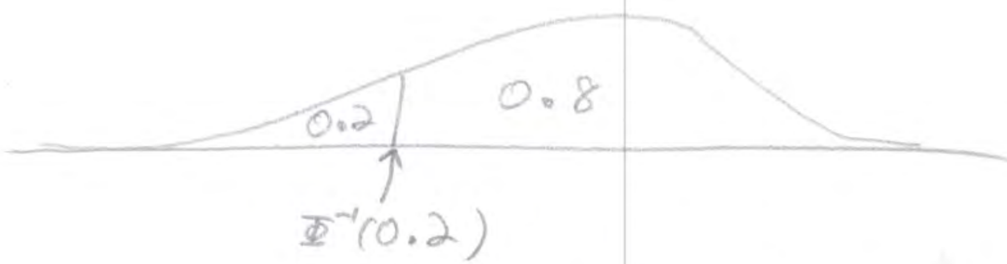
4 continued

$$= 1 - \Phi\left(-\frac{\sqrt{n}}{3} + 1.55\right) \stackrel{set}{=} 0.80$$

Desired power

$$\iff \Phi\left(1.55 - \frac{\sqrt{n}}{3}\right) = 0.2$$

$$\iff 1.55 - \frac{\sqrt{n}}{3} = \Phi^{-1}(0.2)$$



Looking in the normal table, find

$\Phi(-0.84) = 0.2005$ That's close enough

$$1.55 - \frac{\sqrt{n}}{3} = -0.84 \iff \frac{\sqrt{n}}{3} = 2.39$$

$$\iff \sqrt{n} = 7.17 \iff n = 7.17^2 = 51.4$$

So choose $n = 52$

Cross-checking,

$$1 - \Phi\left(\frac{\sqrt{52}(8-9)}{\sqrt{9}} + \sqrt{\frac{8}{9}} \times 1.645\right) = 1 - \Phi(-.85)$$

$$= 1 - 0.1977 = 0.8023 \checkmark$$

(5) $\text{Power} = P(Z^* \le z_\alpha) = P\left(\frac{\sqrt{n}(\bar{X} - 0)}{1} \le -1.645 \right)$

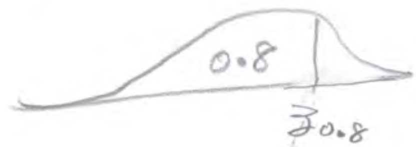
$= P\left(\bar{X} \le \frac{-1.645}{\sqrt{n}} \right)$

$= P\left[\frac{\sqrt{n}(\bar{X} - -1)}{1} \le \sqrt{n}\left(\frac{-1.645}{\sqrt{n}} + 1 \right) \right]$

$= P[Z \le \sqrt{n} - 1.645] = \Phi(\sqrt{n} - 1.645)$

$\overset{set}{=} 0.8 \implies \Phi^{-1}\left(\Phi(\sqrt{n} - 1.645)\right) = \Phi^{-1}(0.8)$

$\implies \sqrt{n} - 1.645 = \text{qnorm}(0.8)$



$= 0.8416$

$\implies \sqrt{n} = 0.8416 + 1.645 = 2.49$

$\implies n = 2.49^2 = 6.2$, so let $n = 7$

(6) H_0 will <u>not</u> be rejected if & only if

$$\chi^2_{\alpha/2}(n-1) < \frac{(n-1)S^2}{\sigma_0^2} < \chi^2_{1-\frac{\alpha}{2}}(n-1)$$

$$\iff \frac{1}{\chi^2_{\alpha/2}(n-1)} > \frac{\sigma_0^2}{(n-1)S^2} > \frac{1}{\chi^2_{1-\alpha/2}(n-1)}$$

$$\iff \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} < \sigma_0^2 < \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}$$

Re-deriving the confidence interval,

$$1-\alpha = P\left(\chi^2_{\alpha/2}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{1-\alpha/2}(n-1)\right)$$

$$= P\left(\frac{1}{\chi^2_{\alpha/2}(n-1)} > \frac{\sigma^2}{(n-1)S^2} > \frac{1}{\chi^2_{1-\alpha/2}(n-1)}\right)$$

$$= P\left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right)$$

QED

7 (a) $P = \overbrace{1 - F(T)}^{\text{Between } 0 \,\&\, 1}$

(b) If H_0 is true so F is cdf of T,

For $0 < x < 1$,

$$f_P(x) = \frac{d}{dx} \Pr(P \le x) = \frac{d}{dx} \Pr(1 - F(T) \le x)$$

$$= \frac{d}{dx} \Pr(F(T) \ge 1 - x)$$

$$= \frac{d}{dx} \Pr\left(F^{-1}(F(T)) \ge F^{-1}(1-x)\right)$$

$$= \frac{d}{dx} \Pr\left(T \ge F^{-1}(1-x)\right)$$

$$= \frac{d}{dx} \left(1 - F\left(F^{-1}(1-x)\right)\right)$$

$$= \frac{d}{dx} \left(1 - (1-x)\right) = \frac{d}{dx} x = 1$$

for $0 < x < 1$! <u>uniform</u>

8 (a) $P_1 \ldots P_m \overset{iid}{\sim} U(0,1)$

(b) $f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} P(Y \le y)$

$= \frac{d}{dy} P_r\left(-2 \ln P \le y\right) = \frac{d}{dy} P_r\left(\ln P \ge -\frac{y}{2}\right)$

$= \frac{d}{dy} P_r\left(P \ge e^{-y/2}\right)$

$= \frac{d}{dy}\left(1 - F_p\left(e^{-y/2}\right)\right) = -f_p\left(e^{-y/2}\right) \cdot e^{-y/2}\left(-\frac{1}{2}\right)$

$= \frac{1}{2} e^{-y/2} I\left(0 < e^{-y/2} < 1\right)$

$= \frac{1}{2} e^{-y/2} I\left(-\infty < -\frac{y}{2} < 0\right)$

$= \frac{1}{2} e^{-y/2} I\left(-\infty < -y < 0\right)$

$= \frac{1}{2} e^{-y/2} I\left(\infty > y > 0\right) = \frac{1}{2} e^{-y/2} I(y > 0)$

so it's exponential with $\lambda = \frac{1}{2}$,

MGF $(1-2t)^{-1} = (1-2t)^{-\frac{2}{2}}$

Also $\chi^2(2)$

(8c) $W = \sum_{i=1}^{n} Y_i \sim \chi^2(2m)$

(d) Small p-values are surprising if the individual null hypotheses are true

small P_i yields large negative $\ln P_i$, large positive $-2\ln P_i$, large sum.

(e) $\chi^2_{1-\alpha}(2m)$

(f) $W = -2(\ln 0.255 + \ln 0.065 + \ln 0.268 + \ln 0.044 + \ln(0.08) + \ln(0.135))$
$= 26.14$

Critical value is $\chi^2_{0.95}(12) = 21.03$

Reject H_0, conclude not all null hypotheses were true. The phenomenon being investigated could be real.

(Maybe the drug has some effect).

9 (a) i) $$S_j^2 = \frac{\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_j)^2}{n_j-1}$$

ii) χ^2, with df $\sum_{j=1}^{k}(n_j-1) = n-k$

Because sum of independent χ^2s is χ^2, with df = sum of df.

iii) $$W_2 = \sum_{j=1}^{k}\frac{(n_j-1)S_j^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{j=1}^{k}(n_j-1)\frac{\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_j)^2}{(n_j-1)}$$

$$= \frac{1}{\sigma^2}\sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_j)^2$$

(b) $$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_\circ)^2 = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j+\bar{X}_j-\bar{X}_\circ)^2$$

$$= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j)^2 + 2\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j)(X_j-\bar{X}_.)$$

$$+ \sum_{j=1}^{k}\sum_{i=1}^{n_j}(\bar{X}_j-\bar{X}_\circ)^2$$

$$= \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j)^2 + 2\sum_{j=1}^{k}(\bar{X}_j-\bar{X}_.)\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j)$$

$$+ \sum_{j=1}^{k}n_j(\bar{X}_j-\bar{X}_\circ)$$

$$= \sum_{j=1}^{k}n_j(\bar{X}_j-\bar{X}_.)^2 + \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j)^2 + \underbrace{2\sum_{j=1}^{k}(\bar{X}_j-\bar{X}_.)(n_j\bar{X}_j-n_j\bar{X}_j)}_{=0}$$

(c) $$\bar{X}_\circ = \sum_{j=1}^{k}\frac{n_j}{n}\bar{X}_j = \frac{1}{n}\sum_{j=1}^{k}\cancel{n_j}\frac{\sum_{i=1}^{n_j}X_{ij}}{\cancel{n_j}} = \frac{1}{n}\sum_{j=1}^{k}\sum_{i=1}^{n_j}X_{ij}$$

(9d) $\chi^2(n-1)$. It's just $\frac{(n-1)S^2}{\sigma^2}$ for one big sample

(e) From (b),

$$\underbrace{\frac{\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_{.})^2}{\sigma^2}}_{W} = \underbrace{\sum_{j=1}^{k} n_j(\bar{X}_j-\bar{X}_{.})^2}_{W_1} + \underbrace{\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\bar{X}_j)^2}_{W_2}$$

$W \sim \chi^2(n-1)$ by (d), $W_2 \sim \chi^2(n-k)$ by (aii)

So bec W_1 & W_2 are independent,

$$W_1 \sim \chi^2(n-1-(n-k)) = \chi^2(k-1)$$

(f) Because $\bar{X}_1, \ldots \bar{X}_k$ & $S_1^2, \ldots, S_k^2$ are all independent. W_1 is a function of $\bar{X}_1, \ldots \bar{X}_k$, W_2 is a function of $S_1^2, \ldots, S_k^2$, & functions of independent random variables are independent.

(9g)

$$F = \frac{\sum_{j=1}^{k} n_j (\bar{X}_j - \bar{X}_\cdot)^2 / (k-1)}{\sum_{j=1}^{k} \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2 / (n-k)} \sim F(k-1, n-k)$$

(h)

$$F = \frac{SSB/(k-1)}{SSW/(n-k)} = \left(\frac{n-k}{k-1}\right) \frac{SSB/SSTO}{SSW/SSTO}$$

$$= \left(\frac{n-k}{k-1}\right) \frac{R^2}{\frac{SSTO - SSB}{SSTO}} = \left(\frac{n-k}{k-1}\right) \frac{R^2}{1-R^2}$$

(i) $n = 117 + 32 + 51 = 200$

$$\bar{X}_\cdot = \frac{1}{200}(117 * 2.81 + 32 * 2.7 + 51 * 2.61)$$

$$= 2.74$$

$$SSW = (117-1) * 0.347 + (32-1) * 0.228 + (51-1) * 0.306$$

$$= 62.62$$

$$SSB = 117(2.81 - 2.74)^2 + 32(2.7 - 2.74)^2 + 51(2.61 - 2.74)^2$$

$$= 1.49$$

$$F = \frac{1.49/(3-1)}{62.62/(200-3)} = 2.34$$

(9i continued)

$F = 2.34 < 3.04$, Don't reject H_0

Conclude $\mu_1 = \mu_2 = \mu_3$. Average grades on the 3 campuses are not really different.

(ii) $R^2 = \frac{SSB}{SSB+SSW} = \frac{1.49}{1.49+62.62} = 0.02$

Not very strong at all.