HW4

1. $Y_{ij} = \mu_i + \varepsilon_{ij}$, $i = 1, \dots, I$, $j = 1, \dots, J$
$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$

a.) $\mathcal{L}(\mu_i | \sigma^2, Y_{ij}) = (2\pi\sigma^2)^{-IJ/2} \exp\left\{\frac{-1}{2\sigma^2} \sum_{i=1}^{I}\sum_{j=1}^{J} (Y_{ij} - \mu_i)^2\right\}$

$\ell(\mu_i) = -\frac{IJ}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{I}\sum_{j=1}^{J}(Y_{ij} - \mu_i)^2$

$\max(\ell(\mu)) \Rightarrow \max\left(-\frac{1}{2\sigma^2}\sum\sum(Y_{ij} - \mu_i)^2\right)$

$\Rightarrow \min\left(\sum\sum(Y_{ij} - \mu_i)^2\right)$

b.) $\frac{\partial \ell}{\partial \mu_i} = +\frac{1}{\sigma^2}\sum_{j=1}^{J}(Y_{ij} - \mu_i)$

$0 = \sum_{j=1}^{J}(Y_{ij} - \hat{\mu}_i) \Rightarrow \hat{\mu}_i = \bar{Y}_{i\cdot}$

c.) ~~$\mathcal{L}_0(\mu_i, \sigma^2 | Y) =$~~

$\max_{\sigma^2 > 0, \mu_i = \mu} \mathcal{L}_0(\sigma^2 | Y) = (2\pi\hat{\sigma}_0^2)^{-n/2}\exp\left\{\frac{-1}{2\hat{\sigma}_0^2}\sum\sum(Y_{ij} - \mu)^2\right\}$, where $\hat{\sigma}_0^2 = \frac{1}{IJK}\sum\sum(Y_{ij} - \mu)^2$

$= (2\pi\hat{\sigma}_0^2)^{-n/2}\exp\left\{\frac{-IJK}{2}\right\}$

$\max_{\sigma^2 > 0, \mu_i \in \mathbb{R}} \mathcal{L}(\mu, \sigma^2 | Y) = (2\pi\hat{\sigma}^2)^{-n/2}\exp\left\{-\frac{1}{2\hat{\sigma}^2}\sum\sum(Y_{ij} - \hat{\mu}_i)^2\right\}$, where $\hat{\sigma}^2 = \frac{1}{IJK}\sum\sum(Y_{ij} - \hat{\mu}_i)^2$

$= (2\pi\hat{\sigma}^2)^{-n/2}$

$\Lambda(Y) = \frac{\max_{\Omega_0} \mathcal{L}(\mu, \sigma^2 | Y)}{\max_{\Omega} \mathcal{L}(\mu, \sigma^2 | Y)} = \frac{(2\pi\hat{\sigma}_0^2)^{-IJK/2}}{(2\pi\hat{\sigma}^2)^{-IJK/2}} = \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_0^2}\right)^{\frac{IJK}{2}}$

d) $\sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2 = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.} + \bar{Y}_{i.} - \bar{Y}_{..})^2$

$$= \sum \sum \left[(Y_{ij} - \bar{Y}_{i.})^2 + 2(Y_{ij} - \bar{Y}_{i.})(\bar{Y}_{i.} - \bar{Y}_{..}) + (\bar{Y}_{i.} - \bar{Y}_{..}) \right]^2$$

$$= \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})^2 + 2 \sum_i (\bar{Y}_{i.} - \bar{Y}_{..}) \underbrace{\sum_j (Y_{ij} - \bar{Y}_{i.})}_{=0 \text{ for all } i} + J \sum_i (\bar{Y}_{i.} - \bar{Y}_{..})^2$$

$$\sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2 = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})^2 + J \sum_i (\bar{Y}_{i.} - \bar{Y}_{..})^2$$

SSTotal = SSW + SSB

e) $\bar{Y}_{i.} \overset{iid}{\sim} N(\mu, \frac{\sigma^2}{J})$ where $\mu_1 = \mu_2 = \cdots = \mu_I = \mu$

f) $\frac{1}{\sigma^2} SSB = \frac{1}{\sigma^2} J \sum_{i=1}^{I} (\bar{Y}_{i.} - \bar{Y}_{..})^2$

Under $H_0: \mu_1 = \mu_2 = \cdots = \mu_I = \mu$, each $\bar{Y}_{i.}$ is iid $N(\mu, \frac{\sigma^2}{J})$. By the result in the one-sample normal case that $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$, it follows that

$\frac{1}{\sigma^2} SSB \overset{H_0}{\sim} \chi^2_{I-1}$.

g) $\frac{1}{\sigma^2} SSW = \frac{1}{\sigma^2} \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})^2$

Since $\frac{1}{\sigma^2} \sum_j (Y_{ij} - \bar{Y}_{i.})^2 \sim \chi^2_{J-1}$ (again by appealing to one-sample normal case), and since the summands are independent for different values of i, the summation property of independent chi-squared variables gives

$\frac{1}{\sigma^2} SSW \sim \chi^2_{I(J-1)}$

h) The results of (f) and (g) are not sufficient. We also need independence of SSW and SSB, which is true since SSW can be written as the sum of group sample variances, and SSB can be written as a function of the group sample means. $SSW \perp SSB$ then follows from the independence of s^2 and the sample mean.

i) $F^* = \frac{SSB/(I-1)}{SSW/[I(J-1)]} \overset{H_0}{\sim} F_{I-1, I(J-1)}$

(2)

a.)

	Df	Sum Sq.	Mean Sq.	F
Type	2	754.4	377.2	6.65
Residuals	45	2551.5	56.7	

b.) Estimate σ^2 by $MSE = \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})^2 = 56.7$

c.) $SSTot = 754.4 + 2551.5 = 3305.9$

d.) $F^* = 6.65 \overset{H_0}{\sim} F_{2,45}$

Critical value for $\alpha = 0.05$: $F_{2,45}(0.95) = 3.204$ from F-table.

$F^* > 3.204 \Rightarrow$ Reject H_0: No difference in mean cycles until failure.

That is, at least one watch type has a different average number of cycles until failure, at significance level 0.05.

(3) a) $I = 3,\ J = 3,\ K = 6$

b) $\varepsilon_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$

c) μ : overall pop. mean

α_i: Difference of group i mean with overall pop. mean, marginal across drive type

β_j: Diff. of group j mean with overall pop. mean, marginal across technician

δ_{ij}: Additional effect of type-tech. combination ij on mean response relative to the main effects of drive type i and tech. j.

d) See ANOVA notes on course website.

e)

Source	df	Source	MS	F
Drive Type	2	200	100	22.5
Technician	2	10	5	1.13
Interaction	4	90	22.5	5.1
Error	45	200	4.44	
Total	53	500		

f.) $MSE = 4.44$

(3) g.) Test for effect of types: $F^* = 22.5 \overset{H_0}{\sim} F_{2,45}$

Reject Ho if $F^* > 3.204$

$\Rightarrow$ Conclusion: At least one drive type has a different mean repair time, at significance $\alpha = 0.05$

h.) Test for interaction of drive type and tech: $F^* = 5.1 \overset{H_0}{\sim} F_{4,45}$

Reject Ho if $F^* > 2.579$.

$\Rightarrow$ Conclusion: Reject Ho, at least one combination of drive type and technician has a different mean repair time than the marginal effects of that drive type and technician, at significance level $\alpha = 0.05$.

(4.) a) $\dfrac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$

b.) Use the fact that group sample variances are unbiased for σ^2:

$$\mathbb{E}[S_p^2] = \frac{(J_1 - 1)\mathbb{E}[S_1^2] + (J_2 - 1)\mathbb{E}[S_2^2]}{J_1 + J_2 - 2} = \frac{(J_1 - 1)\sigma^2 + (J_2 - 1)\sigma^2}{J_1 + J_2 - 2}$$

$$= \sigma^2$$

c) i) Constant variance assumption is violated.

ii) This violation could possibly inflate or deflate the type I error, depending on the relationship of unbalanced group sizes and the difference in the variances across the groups.

See HW 3 #1.

d) For testing K hypotheses simultaneously, the Bonferroni correction uses significance level α/K for each individual hypothesis test to control the overall family-wise type I error to be less than α.