Exam 2 STAT 511 Solutions

1a) $\mathrm{Var}(y_i) = \mathrm{Var}(|x_i|\,\varepsilon_i)$

$$= |x_i|^2\,\mathrm{Var}(\varepsilon_i)$$

$$= x_i^2\,\sigma^2$$

$$\mathrm{cov}(|x_i|\,y_i, |x_j|\,y_j) = |x_i||x_j|\,\mathrm{cov}(y_i, y_j)$$

$$= |x_i||x_j|\;0$$

$$= 0$$

Thus, $\mathrm{Var}(\underset{\sim}{y}) = \sigma^2\,\mathrm{diag}(x_1^2, \ldots, x_n^2)$, i.e.,

$$\mathrm{Var}(\underset{\sim}{y}) = \sigma^2 \begin{bmatrix} x_1^2 & 0 & \cdots & \cdots & 0 \\ 0 & x_2^2 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & x_{n-1}^2 & 0 \\ 0 & \cdots & \cdots & 0 & x_n^2 \end{bmatrix}.$$

1b) $\operatorname{Var}(\underset{\sim}{y}) = \sigma^2 V$, where

$$V = \operatorname{diag}(x_1^2, x_2^2, \ldots, x_n^2).$$

Because $x_1, x_2, \ldots, x_n$ are known, V is known. Thus, this is a special case of the Aitken model.

The design matrix is

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \equiv \underset{\sim}{x}$$ and the parameter vector is $\underset{\sim}{\beta} = [\beta]$ $(p=1)$.

The BLUE is

$$(X'V^{-1}X)^{-1}X'V^{-1}\underset{\sim}{y} = (\underset{\sim}{x}'V^{-1}\underset{\sim}{x})^{-1}\underset{\sim}{x}'V^{-1}\underset{\sim}{y}$$

$$= \frac{\sum_{i=1}^{n} x_i y_i / x_i^2}{\sum_{i=1}^{n} x_i^2 / x_i^2} = \frac{1}{n}\sum_{i=1}^{n} \frac{y_i}{x_i}.$$

2a) This is a split-plot experiment. The whole-plot experimental units are pots. The Split-plot experimental units are seedlings.

b) Seedlings

c)

Source	DF
Watering level	3-1 = 2
pot(wat. lev.)	(10-1)(3) = 27
injection	2-1 = 1
wat. lev. x injection	(3-1)(2-1) = 2
error	87
c. total	120-1 = 119

Note that "error" is a combination of

injection x pot(wat. lev.) (2-1)(27) = 27

and

seedling (injection, pot, wat. lev.) = (2-1)(60) = 60.

3a) Let $\varepsilon \sim N(0,1)$ and independent of W_1.

Then W_2 has the same distribution as $W_1 + \varepsilon$.

$$\Rightarrow \begin{bmatrix} W_1 \\ W_2 \end{bmatrix} \stackrel{d}{=} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} W_1 \\ \varepsilon \end{bmatrix}$$

$$\sim N\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} I \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \right)$$

$$\stackrel{d}{=} N\left(\underset{\sim}{0}, \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \right) \equiv N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$

b) $$E(W_1 \mid W_2) = \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (W_2 - \mu_2)$$

$$= 0 + (1)(1/2)(1.7 - 0)$$

$$= 0.85.$$

4a) $Y_{1j} - Y_{2j} = \mu_1 + u_j + e_{1j} - (\mu_2 + u_j + e_{2j})$

$$= \mu_1 - \mu_2 + e_{1j} - e_{2j}$$

$$E(Y_{1j} - Y_{2j}) = \mu_1 - \mu_2 + E(e_{1j}) - E(e_{2j})$$

$$= \mu_1 - \mu_2$$

$$Var(Y_{1j} - Y_{2j}) = Var(e_{1j} - e_{2j}) = Var(e_{1j}) + Var(e_{2j})$$

$$= 2\sigma_e^2.$$

Thus, $d_1, \ldots, d_{20} \overset{iid}{\sim} N(\mu_1 - \mu_2, 2\sigma_e^2)$.

(We have normality because linear combinations of normals are normal. We have independence because all e_{ij}'s are independent.)

4b) We should conduct a paired-data t-test. If you forgot the expression for the test statistic that you should have learned in a first statistics course, it is fortunately easy to derive using what we have learned this semester.

Take

$$\underset{\sim}{y} = \underset{\sim}{d}, \quad X = \underset{\sim}{1}, \quad \underset{\sim}{\beta} = \mu_1 - \mu_2, \quad \sigma^2 = 2\sigma_e^2$$

$$\hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y} = (\underset{\sim}{1}'\underset{\sim}{1})^{-1}\underset{\sim}{1}'\underset{\sim}{y} = \bar{y}_. = \bar{d}_.$$

$$Var(\hat{\underset{\sim}{\beta}}) = \sigma^2(X'X)^{-1} = 2\sigma_e^2/n = \sigma_e^2/10$$

$$\hat{\sigma}^2 = \widehat{2\sigma_e^2} = \frac{(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})'(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})}{n - \text{rank}(X)}$$

$$= \frac{\sum_{i=1}^{20}(d_i - \bar{d}_.)^2}{19}.$$

Thus, to test $H_0: \mu_1 = \mu_2 \iff H_0: \mu_1 - \mu_2 = 0$,

we use $t = \dfrac{\bar{d}.}{\sqrt{\hat{\sigma}^2/n}} = \dfrac{\bar{d}.}{\sqrt{\dfrac{\sum_{i=1}^{20}(d_i - \bar{d}.)^2}{19} \Big/ 20}}$

4c) Noncentral t with d.f. $= 19$

and noncentrality parameter

$$\frac{\mu_1 - \mu_2}{\sqrt{\sigma^2/20}} = \frac{\mu_1 - \mu_2}{\sqrt{2\sigma_e^2/20}} = \frac{\mu_1 - \mu_2}{\sqrt{\sigma_e^2/10}}.$$

This is true because

$$t = \frac{\dfrac{\bar{d}. - (\mu_1 - \mu_2)}{\sqrt{\sigma^2/20}} + \dfrac{\mu_1 - \mu_2}{\sqrt{\sigma^2/20}}}{\sqrt{\hat{\sigma}^2/\sigma^2}},$$

$\dfrac{\bar{d}. - (\mu_1 - \mu_2)}{\sqrt{\sigma^2/20}} \sim N(0,1)$ independent of $\dfrac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{19}/19$.

4d) This is the classic case of two independent normal samples. If you forgot the formulas, they are easy to derive. Take

$$\underset{\sim}{y} = \begin{bmatrix} \underset{\sim}{a} \\ \underset{\sim}{b} \end{bmatrix}, \quad X = \begin{bmatrix} \underset{20\times1}{\underset{\sim}{1}} & \underset{20\times1}{\underset{\sim}{0}} \\ \underset{20\times1}{\underset{\sim}{0}} & \underset{20\times1}{\underset{\sim}{1}} \end{bmatrix}, \quad \underset{\sim}{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$\sigma^2 = Var(u_j + e_{ij}) = \sigma_u^2 + \sigma_e^2.$$

$$\hat{\underset{\sim}{\beta}} = (X'X)^{-1}X'\underset{\sim}{y} = \begin{bmatrix} 20 & 0 \\ 0 & 20 \end{bmatrix}^{-1} \begin{bmatrix} a_\cdot \\ b_\cdot \end{bmatrix} = \begin{bmatrix} \bar{a}_\cdot \\ \bar{b}_\cdot \end{bmatrix}$$

$$\underset{\sim}{c}'\hat{\underset{\sim}{\beta}} = [1, -1] \begin{bmatrix} \bar{a}_\cdot \\ \bar{b}_\cdot \end{bmatrix} = \bar{a}_\cdot - \bar{b}_\cdot$$

$$Var(\underset{\sim}{c}'\hat{\underset{\sim}{\beta}}) = \underset{\sim}{c}' \sigma^2 (X'X)^{-1} \underset{\sim}{c} = \sigma^2 [1, -1] \begin{bmatrix} 1/20 & 0 \\ 0 & 1/20 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= \sigma^2/10$$

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{n - \text{rank}(X)}$$

$$= \frac{\sum_{i=1}^{20}(a_i - \bar{a}_.)^2 + \sum_{i=1}^{20}(b_i - \bar{b}_.)^2}{38}$$

Thus, the formula for the interval is

$$\bar{a}_. - \bar{b}_. \pm t_{38}^{(.975)} \sqrt{\frac{\sum_{i=1}^{20}(a_i - \bar{a}_.)^2 + \sum_{i=1}^{20}(b_i - \bar{b}_.)^2}{(38)(10)}}$$

4e) From previous parts we have

$$E\left(\frac{\sum_{i=1}^{20}(d_i-\bar{d}_.)^2}{19}\right) = 2\sigma_e^2$$

and

$$E\left(\frac{\sum_{i=1}^{20}(a_i-\bar{a}_.)^2+\sum_{i=1}^{20}(b_i-\bar{b}_.)^2}{38}\right) = \sigma_u^2+\sigma_e^2.$$

Thus, $\hat{\sigma}_e^2 = \frac{1}{2}\,\frac{\sum_{i=1}^{20}(d_i-\bar{d}_.)^2}{19}$

and $$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^{20}(a_i-\bar{a}_.)^2+\sum_{i=1}^{20}(b_i-\bar{b}_.)^2}{38}$$

$$- \frac{1}{2}\,\frac{\sum_{i=1}^{20}(d_i-\bar{d}_.)^2}{19}.$$

4f) You could set the problem up as

$$\underset{\sim}{Y} = \begin{bmatrix} \bar{d}. \\ \bar{a}. \\ \bar{b}. \end{bmatrix}, \quad X = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \underset{\sim}{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$\underset{\sim}{\varepsilon} \sim N\left(\underset{\sim}{0}, \underbrace{\begin{bmatrix} \sigma_e^2/10 & 0 & 0 \\ 0 & \frac{\sigma_u^2 + \sigma_e^2}{20} & 0 \\ 0 & 0 & \frac{\sigma_u^2 + \sigma_e^2}{20} \end{bmatrix}}_{\Sigma} \right)$$

You could then compute

$$(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\underset{\sim}{Y}.$$

Alternatively, we know that

$\bar{d}_.$ and $\bar{a}_. - \bar{b}_.$ are independent

estimators of $\mu_1 - \mu_2$. The BLUP

will be of the form

$$\alpha \bar{d}_. + (1-\alpha)(\bar{a}_. - \bar{b}_.)$$

with weights inversely proportional

to the variances of $\bar{d}_.$ and $\bar{a}_. - \bar{b}_.$

$$Var(\bar{d}_.) = \frac{2\sigma_e^2}{20} = \sigma_e^2/10$$

$$Var(\bar{a}_. - \bar{b}_.) = \frac{\sigma_u^2 + \sigma_e^2}{10}.$$

Thus,

$$\alpha = \frac{10/\sigma_e^2}{10/\sigma_e^2 + \frac{10}{\sigma_u^2+\sigma_e^2}} = \frac{10(\sigma_u^2+\sigma_e^2)}{10(\sigma_u^2+\sigma_e^2)+10\sigma_e^2}$$

$$= \frac{\sigma_u^2+\sigma_e^2}{\sigma_u^2+2\sigma_e^2}.$$

The BLUP is, therefore,

$$\frac{\sigma_u^2+\sigma_e^2}{\sigma_u^2+2\sigma_e^2}\,\bar{d}_. + \frac{\sigma_e^2}{\sigma_u^2+2\sigma_e^2}\left(\bar{a}_. - \bar{b}_.\right).$$

Point values were as follows:

1a) 6	2a) 6	3a) 8	4a) 7	4d) 8
1b) 9	2b) 4	3b) 9	4b) 8	4e) 8
	2c) 9		4c) 8	4f) 10