EXAM 2 SOLUTIONS

STAT 511

SPRING 2012

POINTS WERE ASSIGNED AS FOLLOWS.

1) 14
2a) 6
2b) 6
2c) 16
2di) 6
2dii) 6
2diii) 6
3a) 12
3b) 3
4a) 10
4b) 7
4c) 8

1. FROM THE INFORMATION GIVEN, WE KNOW THAT THE BLUE OF μ BASED ON Y_1, Y_2, Y_3, Y_4 IS $\dfrac{Y_1 + Y_2 + Y_3 + Y_4}{4}$.

THE VARIANCE OF THIS ESTIMATOR IS $\left(\tfrac{1}{4}\underset{\sim}{1}\right)' \begin{bmatrix} 5 & 1 & 1 & 1 \\ 1 & 5 & 1 & 1 \\ 1 & 1 & 5 & 1 \\ 1 & 1 & 1 & 5 \end{bmatrix} \left(\tfrac{1}{4}\underset{\sim}{1}\right) = 2.$

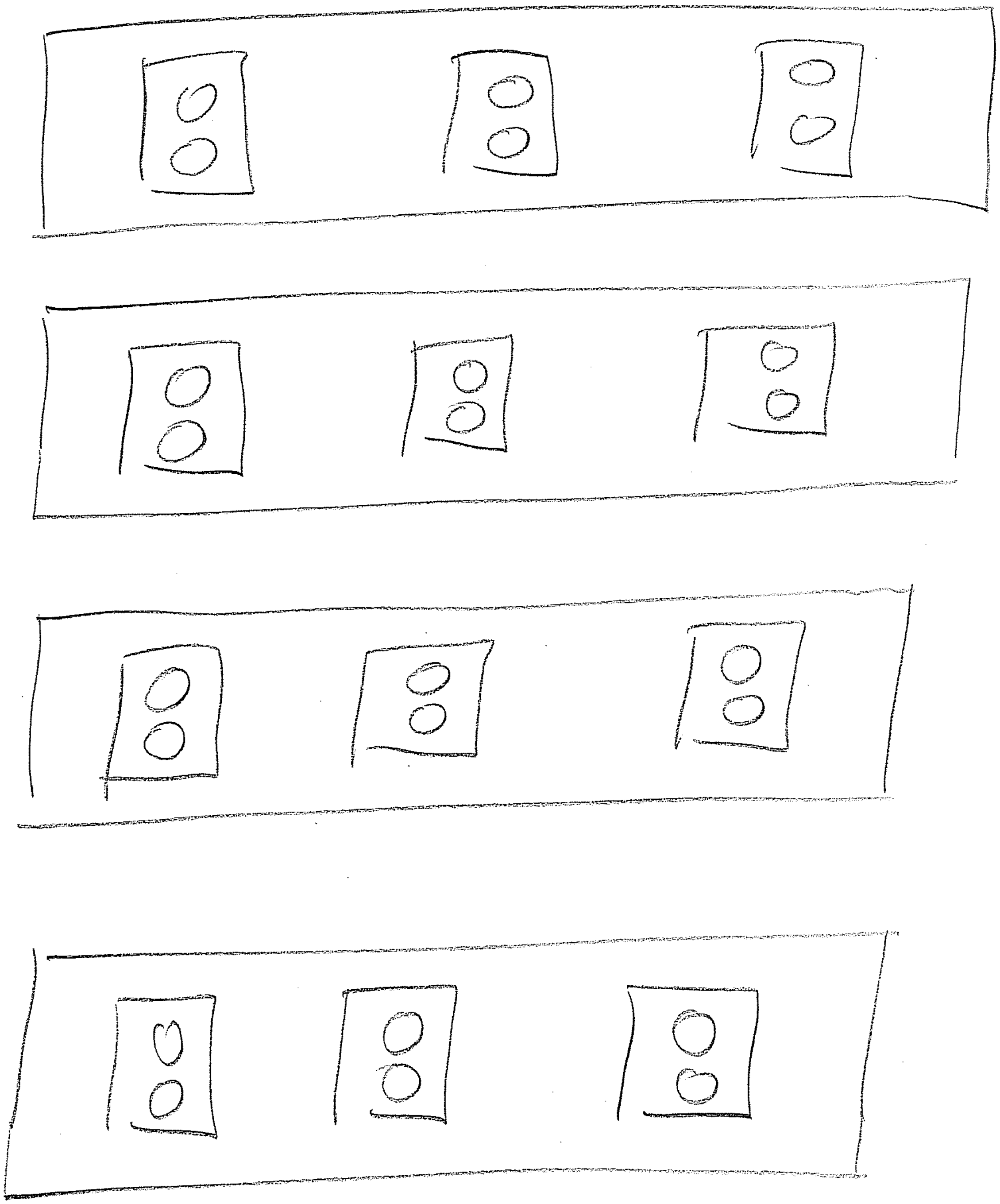
THE BLUE OF μ BASED ON Y_5 IS Y_5 WITH VARIANCE 4.

1. (CONTINUED)

THESE ARE INDEPENDENT BLUEs OF M. THUS, THEY CAN BE OPTIMALLY COMBINED BY INVERSE VARIANCE WEIGHTING. WE HAVE

$$\left(\frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{4}}\right)\frac{5+3+8+8}{4}+\left(\frac{\frac{1}{4}}{\frac{1}{2}+\frac{1}{4}}\right)9=7.$$

2.



2 a) $\mathrm{COV}(y_{ij1}, y_{ij2})$

$$= \mathrm{COV}(g_i + t_{ij} + e_{ij1},\ g_i + t_{ij} + e_{ij2})$$

$$= \mathrm{COV}(g_i, g_i) + \mathrm{COV}(t_{ij}, t_{ij})$$

$$= \sigma_g^2 + \sigma_t^2$$

$$\text{CORRELATION} = \frac{\mathrm{COV}(y_{ij1}, y_{ij2})}{\sqrt{\mathrm{VAR}(y_{ij1})\,\mathrm{VAR}(y_{ij2})}}$$

$$= (\sigma_g^2 + \sigma_t^2) / (\sigma_g^2 + \sigma_t^2 + \sigma_e^2).$$

2b)

$H_0: \mu + \omega_j + \overline{\gamma_{\cdot}} + \overline{\phi_{j\cdot}}$ SAME FOR $j = 1, 2, 3$

$$\Longleftrightarrow H_0: \left(\omega_1 + \overline{\phi_{1\cdot}}\right) - \left(\omega_2 + \overline{\phi_{2\cdot}}\right) = 0$$

$$\text{AND } \left(\omega_1 + \overline{\phi_{1\cdot}}\right) - \left(\omega_3 + \overline{\phi_{3\cdot}}\right) = 0.$$

THE ABOVE ARE STATED IN TERMS OF TESTABLE QUANTITIES. I ALSO ACCEPTED

$H_0: \omega_1 + \overline{\phi_{1\cdot}} = \omega_2 + \overline{\phi_{2\cdot}} = \omega_3 + \overline{\phi_{3\cdot}}$

2 b) (CONTINUED) A COMMON <u>INCORRECT</u> ANSWER WAS $H_0: W_1 = W_2 = W_3$.

THIS DOES NOT IMPLY NO WATERING LEVEL MAIN EFFECTS. TO SEE THIS, START WITH THE TABLE OF MEANS BELOW.

		GENO 1	GENO 2
WL	1	$\mu + W_1 + \gamma_1 + \phi_{11}$	$\mu + W_1 + \gamma_2 + \phi_{12}$
	2	$\mu + W_2 + \gamma_1 + \phi_{21}$	$\mu + W_2 + \gamma_2 + \phi_{22}$
	3	$\mu + W_3 + \gamma_1 + \phi_{31}$	$\mu + W_3 + \gamma_2 + \phi_{32}$

THERE ARE WATERING LEVEL MAIN EFFECTS IF AND ONLY IF THE 3 ROW AVERAGES ARE NOT ALL THE SAME.

IT IS EASY TO MAKE THESE ROW AVERAGES DIFFERENT EVEN IF $W_1 = W_2 = W_3$.

FOR EXAMPLE, TAKE

$$\mu = W_1 = W_2 = W_3 = \gamma_1 = \gamma_2 = \phi_{11} = \phi_{12} = \phi_{21} = \phi_{22} = 0$$

AND $\phi_{31} = \phi_{32} = 1$. THEN THE ROW AVERAGES ARE 0, 0, AND 1, WHICH IMPLIES THAT THERE ARE WATERING LEVEL MAIN EFFECTS.

IS THERE INTERACTION IN THIS EXAMPLE? IF YOU THINK THERE IS INTERACTION BECAUSE SOME

OF THE ϕ_{ij} PARAMETERS ARE NOT 0 OR
BECAUSE NOT ALL ϕ_{ij} PARAMETERS ARE
EQUAL, THEN YOU ARE STILL NOT UNDERSTANDING
AN IMPORTANT AND FUNDAMENTAL S11
CONCEPT. YOU NEED TO KEEP YOUR EYE ON
THE CELL MEANS IN THE TABLE RATHER
THAN NONESTIMABLE ELEMENTS OF β.
THERE IS NO INTERACTION IN MY EXAMPLE
BECAUSE THE GENOTYPE DIFFERENCE BETWEEN CELL
MEANS IS THE SAME FOR EVERY ROW OF THE TABLE.

2c) $X = \left[\underset{24\times 1}{\underset{\sim}{1}},\ \underset{4\times 1}{\underset{\sim}{1}} \otimes \underset{3\times 3}{I} \otimes \underset{2\times 1}{\underset{\sim}{1}},\ \underset{12\times 1}{\underset{\sim}{1}} \otimes \underset{2\times 2}{I},\ \underset{4\times 1}{\underset{\sim}{1}} \otimes \underset{6\times 6}{I} \right]$

$\underset{\sim}{\beta} = (\mu, w_1, w_2, w_3, \gamma_1, \gamma_2, \phi_{11}, \phi_{12}, \phi_{21}, \phi_{22}, \phi_{31}, \phi_{32})'$

$Z = \left[\underset{4\times 4}{I} \otimes \underset{6\times 1}{\underset{\sim}{1}},\ \underset{12\times 12}{I} \otimes \underset{2\times 1}{\underset{\sim}{1}} \right]$

$\underset{\sim}{u} = (g_1, g_2, g_3, g_4, t_{11}, t_{12}, t_{13}, t_{21}, t_{22}, t_{23},$
$t_{31}, t_{32}, t_{33}, t_{41}, t_{42}, t_{43})'.$

2d) THIS IS A SPLIT-PLOT DESIGN THAT IS COMPLETELY ANALOGOUS TO THE FIELD EXPERIMENT WITH GENOTYPES AND FERTILIZER LEVELS DISCUSSED IN COURSE NOTES. HERE GREENHOUSES ARE BLOCKS, WATERING LEVEL IS THE WHOLE-PLOT TREATMENT FACTOR, AND GENOTYPE IS THE SPLIT-PLOT TREATMENT FACTOR. IN CLASS, WE

LEARNED THAT BLOCK-BY-WHOLE PLOT
TREATMENT FACTOR INTERACTION EFFECTS
CORRESPOND TO WHOLE-PLOT EXPERIMENTAL
UNITS. THUS, GREENHOUSE-BY-WATER LEVEL
MEAN SQUARE IS THE DENOMINATOR FOR
THE F-TEST OF WATERING LEVEL MAIN
EFFECTS. WE ALSO LEARNED THAT THE
ERROR FOR THE TEST OF SPLIT-PLOT MAIN
EFFECTS AND W.P.-BY-S.P. INTERACTION
BOTH USE AN ERROR THAT IS DETERMINED

BY POOLING

BLOCK X S.P. = GH X GENO WITH

BLOCK X W.P. X S.P. = GH X WL X GENO

THUS, F-STATISTICS ARE

AS FOLLOWS.

2d)

i) $$\frac{321.8/2}{116.4/6}$$

ii) $$\frac{2.5/1}{(11.7+14.5)/(3+6)}$$

iii) $$\frac{75.1/2}{(11.7+14.5)/(3+6)}$$

3 a)

SOURCE	DF
H	1
D	1
H×D	1
FARM(H, D)	16
SOW (FARM, H, D)	140
V	2
H×V	2
D×V	2
H×D×V	2
ERROR	312 ← V × FARM(H, D) + V × SOW(FARM, H, D)
C. TOTAL	479

b) SEE CIRCLED TERMS ABOVE

4. LET Y_{ij} DENOTE THE MEASUREMENT OF CHEMICAL CONCENTRATION FOR THE j^{TH} PLANT THAT RECEIVED TREATMENT i ($i = 1, 2$; $j = 1, 2, 3, 4$). FOR EXAMPLE, Y_{24} IS THE MEASUREMENT FOR THE PLANT WITH ID NUMBER 8.

IF WE ASSUME

$$Y_{ij} = \mu_i + P_{ij} + e_{ij}, \quad \text{WHERE}$$

$P_{11}, \ldots, P_{24} \overset{iid}{\sim} N(0, \sigma_P^2)$ INDEPENDENT OF

$e_{11}, \ldots, e_{24} \overset{iid}{\sim} N(0, \sigma_e^2)$, THEN THE

MEASUREMENTS HAVE THE JOINT

DISTRIBUTION SPECIFIED IN THE

PROBLEM STATEMENT. WE CAN

WRITE THIS MODEL AS

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + Z\underset{\sim}{u} + \underset{\sim}{e}, \quad \text{WHERE}$$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{24} \end{bmatrix} = \underset{\sim}{y}, \quad \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = X, \quad \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \underset{\sim}{\beta},$$

$$Z = \underset{8 \times 8}{I},$$

$$\underset{\sim}{u} = \begin{bmatrix} p_{11} \\ p_{12} \\ \vdots \\ p_{24} \end{bmatrix}, \quad \underset{\sim}{e} = \begin{bmatrix} e_{11} \\ e_{12} \\ \vdots \\ e_{24} \end{bmatrix}.$$

$$G = Var(\underset{\sim}{u}) = \sigma_p^2 \, \underset{8 \times 8}{I}$$

$$R = Var(\underset{\sim}{e}) = \sigma_e^2 \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & 0 & & \\ & & 1/2 & & & & & \\ & & & 1/2 & & & & \\ & & & & 1 & & & \\ & 0 & & & & 1 & & \\ & & & & & & 1/2 & \\ & & & & & & & 1/2 \end{bmatrix}$$

$$\mathrm{VAR}(\underset{\sim}{y}) = \Sigma = ZGZ' + R$$

$$= \sigma_p^2 \underset{8\times 8}{\underline{I}} + \sigma_e^2 \,\mathrm{DIAG}(1,1,\tfrac{1}{2},\tfrac{1}{2},1,1,\tfrac{1}{2},\tfrac{1}{2})$$

$$= \sigma_e^2 \left(\sigma_p^2/\sigma_e^2 \underset{8\times 8}{\underline{I}} + \mathrm{DIAG}(1,1,\tfrac{1}{2},\tfrac{1}{2},1,1,\tfrac{1}{2},\tfrac{1}{2}) \right)$$

$$= \sigma_e^2 \,\mathrm{DIAG}(4,4,3.5,3.5,4,4,3.5,3.5)$$

WHEN $\sigma_p^2/\sigma_e^2 = 3$.

a) THE AITKEN MODEL HOLDS WITH

$\sigma^2 = \sigma_e^2$ AND $V = \text{DIAG}(4, 4, 3.5, 3.5, 4, 4, 3.5, 3.5)$

WE CAN USE WEIGHTED LEAST SQUARES WITH WEIGHTS $1/4,\ 1/4,\ 1/3.5,\ 1/3.5,\ 1/4,\ 1/4,\ 1/3.5,\ 1/3.5$

THE BLUE OF μ_2 IS SIMPLY A WEIGHTED AVERAGE:

$$\frac{1/4\ 82 + 1/4\ 76 + 1/3.5\ 77 + 1/3.5\ 70}{1/4 + 1/4 + 1/3.5 + 1/3.5} = \hat{\mu}_2$$

4 b) THE BLUP OF $\underset{\sim}{u}$ IS

$G Z' \Sigma^{-1} (\underset{\sim}{y} - X \hat{\underset{\sim}{\beta}}_\Sigma)$, THE LAST ELEMENT

OF THIS VECTOR IS

$$\sigma_p^2 \left(\sigma_p^2 + \sigma_e^2/2 \right)^{-1} \left(70 - \hat{\mu}_2 \right)$$

$$= \frac{\sigma_p^2}{\sigma_p^2 + \sigma_e^2/2} \left(70 - \hat{\mu}_2 \right)$$

$$= \frac{\sigma_p^2/\sigma_e^2}{\sigma_p^2/\sigma_e^2 + 1/2} \left(70 - \hat{\mu}_2 \right) = \frac{3}{3.5} \left(70 - \hat{\mu}_2 \right).$$

THUS, THE BLUP OF THE CHEMICAL CONCENTRATION FOR PLANT 8 IS

$$\hat{\mu}_2 + \frac{3}{3.5}\left(70 - \hat{\mu}_2\right)$$

$$= \left(\frac{1}{7}\right)\hat{\mu}_2 + \left(\frac{6}{7}\right)70.$$

NOTE THAT THIS PROBLEM WAS VERY MUCH LIKE THE I.Q PREDICTION PROBLEM DISCUSSED IN CLASS.

4 c) LET $a = \sigma_p^2 + \sigma_e^2$ AND $b = \sigma_p^2 + \sigma_e^2/2$

THEN
$$\Sigma = \begin{bmatrix} a & & & & & & & \\ & a & & & & & \text{\Large 0} & \\ & & b & & & & & \\ & & & b & & & & \\ & & & & a & & & \\ & & & & & a & & \\ & \text{\Large 0} & & & & & b & \\ & & & & & & & b \end{bmatrix}$$

TAKE
$$A' = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \end{bmatrix}$$

THEN

$$E(A'\underline{y}) = \underline{0} \quad \text{AND}$$

$$VAR(A'\underline{y}) = A'\Sigma A$$

$$= \begin{bmatrix} 2a & 0 & 0 & 0 & 0 & 0 \\ 0 & 2b & 0 & 0 & 0 & 0 \\ 0 & 0 & 2a+2b & 0 & 0 & 0 \\ 0 & 0 & 0 & 2a & 0 & 0 \\ 0 & 0 & 0 & 0 & 2b & 0 \\ 0 & 0 & 0 & 0 & 0 & 2a+2b \end{bmatrix}.$$

MANY OF YOU RECOGNIZED THAT A' NEEDED TO HAVE 6 ROWS BECAUSE $n - \text{RANK}(X) = 8 - 2 = 6$. MOST OF YOU RECOGNIZED THAT EACH ROW OF A' MUST HAVE ITS FIRST 4 ELEMENTS SUM TO 0 AND ITS LAST 4 ELEMENTS SUM TO 0 BECAUSE

$$A'X = A' \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

MUST BE 0 IN ORDER FOR THE ELEMENTS OF $A'y$ TO BE ERROR CONTRASTS.

MANY OF YOU UNDERSTOOD THAT THE ROWS OF A' NEEDED TO BE LINEARLY INDEPENDENT SO THAT REML ESTIMATES OF σ_p^2 AND σ_e^2 COULD BE OBTAINED USING $A'y$. NONE OF YOU WENT THE ADDITIONAL STEP OF CHOOSING THE ROWS OF A' SO THAT THE ELEMENTS OF $A'y$ WOULD BE UNCORRELATED AS STATED IN THE PROBLEM. THE ADVANTAGE

OF UNCORRELATED ERROR CONTRASTS

IS THAT $\text{VAR}(A'\underline{y})$ IS DIAGONAL,

WHICH MAKES FINDING THE MLEs

OF σ_p^2 AND σ_e^2 BASED ON $A'\underline{y}$ A

BIT EASIER THAN WHEN $\text{VAR}(A'\underline{y})$

IS NOT DIAGONAL.