Biost/Stat 533
Spring 2008

Homework Assignment #9
KEY

1. A manufacturer of locknuts found unwanted differences in the torque values of its product. (Torqe is the work (force $\times$ distance) required to tighten the nut.) The manufacturer did an experiment to study two factors that might affect torque values. The first factor was the type of manufacturing process. The second factor was the medium onto which the locknut would be threaded (a bolt or a mandrel). The following table gives the experimental data.

	Manufacturing Process		
	A	B	C
bolt	20,16,17,18,15, 16,19,14,15,24	26,40,28,38,38, 30,26,38,45,38	25,40,30,17,16, 45,49,33,30,20
mandrel	24,18,17,17,15, 23,14,18,12,11	32,22,30,35,32, 28,27,28,30,30	10,13,17,16,15, 14,11,14,15,16

(a) Give the ANOVA table for the two-way ANOVA model with interactions.
(b) Is there an interaction between these two factors for torque value?
(c) Regardless of your answer to (b), give the ANOVA table for the two-way ANOVA model without interactions.
(d) Descriptively, does manufacturing process or medium appear to be the more important factor? (*Hint:* consider the mean square.)

Solution: Table 1 shows the ANOVA table for the two-way model with interactions.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
type	1	821.40	821.40	12.78	0.0007
process	2	1054.23	527.12	8.20	0.0008
type$\times$ process	2	406.90	203.45	3.17	0.0501
Residuals	54	3469.80	64.26		

Table 1: Problem 1 (a)

(b) At the 5% level, we would fail to reject the null hypothesis of no interaction effect.

(c) Table 2 shows the ANOVA table for the two-way model without interactions.
(d) Descriptively, it appears that more of the variation in the outcome is due to the type of bolt rather than the process because $MS(\text{type}) > MS(\text{process})$.

2. In class we did an example about orthogonal contrasts in one-way ANOVA. The same ideas can be used with continuous variables. Suppose

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
type	1	821.40	821.40	11.87	0.0011
process	2	1054.23	527.12	7.61	0.0012
Residuals	56	3876.70	69.23		

Table 2: Problem 1 (c)

we model $E[Y]$ as a function of x using data from a planned experiment where x takes on exactly 3 values. The data matrix is

x	y
40	25.66
50	29.15
60	35.73
40	28.00
50	35.09
60	39.56
40	20.65
50	29.79
60	35.66

Consider the following design matrix for a linear model with parameters β_0, L and Q. (The rows correspond to the same order of the data as in the data table above.)

$$X = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}$$

(a) Interpret the parameters β_0, L and Q.

(b) A common practice would be to use the same design matrix but to divide the second column by $\sqrt{2}$ and divide the third column by $\sqrt{6}$. What is the purpose of doing this?

(c) Find the least-squares estimates of β_0, L and Q using the re-scaled design matrix described in (b).

(d) Here is a different model: $E[y] = \beta_0 + \beta_1 X + \beta_2 x^2$.
Consider the relationship between the two models. What is an advantage of using the first model compared to using $E[y] = \beta_0 + \beta_1 X + \beta_2 x^2$?

Solution: (a) We can rewrite the model as

$$E[Y|X=x] = \begin{cases} \beta_0 - L + Q & \text{if} \quad x = 40 \\ \beta_0 - 2Q & \text{if} \quad x = 50 \\ \beta_0 + L + Q & \text{if} \quad x = 60 \end{cases}$$

Then β_0 can be interpreted as the the overall mean, since $\frac{1}{3}\{E[Y|X=40] + E[Y|X=50] + E[Y|X=60]\} = \beta_0$. L can be interpreted as

the linear effect, since it is a contrast between the groups with $X = 60$ and $X = 40$. Note that $(E[Y|X=40] - E[Y|X=50]) - (E[Y|X=50] - E[Y|X=60]) \propto Q$. So Q can be interpreted as the difference in the linear trends between groups 1 and 2 and groups 2 and 3. That is, Q can be interpreted as the quadratic trend in the response as a function of x.

(b) The purpose of scaling the design matrix is so that the parameter estimates have equal variances. That way, the size of the estimates can be compared.

(c) Table 3 shows the resulting estimates.

	$\hat{\beta}_0$	$\hat{L}$	$\hat{Q}$
est.	31.03	6.11	−0.156

Table 3: Problem 2 (c)

(d) Both models can measure a linear and quadrataic trend in the response.
But in the model with L and Q the linear and quadratic effect estimates
are uncorrelated.

3. Suppose an experiment is done to compare two varieties of tomatoes in plots that can hold two plants each. There are two such plots available for the study, so that the experimental design is

1	2
1	2

(plants will be randomized within a plot).

The statistical model is $y_{ij} = \mu + P_i + V_j + \epsilon_{ij}$ for $i = 1, 2$ plots and $j = 1, 2$ varieties. Assume the errors ϵ are uncorrelated and have variance σ^2.

For (a) and (c) use the data vector $\mathbf{Y}$ written in the order $\begin{matrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \end{matrix}$

(a) What is $\text{cov}(\mathbf{Y})$ when the "plot effects" P_i are considered fixed effects?
(b) What is the formula for the best linear unbiased estimator of $\hat{V}_1 - \hat{V}_2$ in terms of the Y_{ij} when the P_i are considered fixed effects?
(c) What is $\text{cov}(\mathbf{Y})$ when the "plot effects" P_i are considered random effects with variance τ^2?
(d) What is the formula for the best linear unbiased estimator of $\hat{V}_1 - \hat{V}_2$ in terms of the Y_{ij} when the P_i are considered random effects with variance τ^2? Treat σ^2 and τ^2 as known.

Solution:

(a) $\text{cov}(\mathbf{Y}) = \sigma^2\mathbf{I}$

(b) We set up a design matrix $\mathbf{X}$ for the parameters μ, P_1, P_2, V_1, V_2. However, the matrix will not have full column rank, so we need to use a technique to fit models that don't have full rank. One option is to use constraints $P_1 + P_2 = V_1 + V_2 = 0$. If we construct $\mathbf{X}$ for the parameters μ, P_1, V_1 using this constraint, we get
$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{pmatrix}$. We find that $(\mathbf{X'X})^{-1}\mathbf{X'} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \end{pmatrix}$. Calculating $(\mathbf{X'X})^{-1}\mathbf{X'Y}$, we get that $\hat{V}_1 = \frac{1}{4}(y_{11} - y_{12} + y_{21} - y_{22})$. Since $\hat{V}_1 - \hat{V}_2 = 2\hat{V}_1$ for the chosen linear constraint, we get $\hat{V}_1 - \hat{V}_2 = \frac{1}{2}(y_{11} - y_{12} + y_{21} - y_{22}) = \bar{y}_{\cdot 1} - \bar{y}_{\cdot 2}$.

(c) $\text{cov}(\mathbf{Y}) = \begin{pmatrix} \sigma^2+\tau^2 & \tau^2 & 0 & 0 \\ \tau^2 & \sigma^2+\tau^2 & & \\ 0 & 0 & \sigma^2+\tau^2 & \tau^2 \\ 0 & 0 & \tau^2 & \sigma^2+\tau^2 \end{pmatrix}$

(d) We need to use generalized least squares (because $\text{cov}(\mathbf{Y}) \neq \sigma^2\mathbf{I}$) and calculate $(\mathbf{X'V^{-1}X})^{-1}\mathbf{X'V^{-1}Y}$ where $V \equiv \text{cov}(\mathbf{Y})$.

$$\mathbf{V}^{-1} = \frac{1}{(\sigma^2+\tau^2)^2-\tau^4}\begin{pmatrix} \sigma^2+\tau^2 & -\tau^2 & 0 & 0 \\ -\tau^2 & \sigma^2+\tau^2 & & \\ 0 & 0 & \sigma^2+\tau^2 & -\tau^2 \\ 0 & 0 & -\tau^2 & \sigma^2+\tau^2 \end{pmatrix}$$

$$\mathbf{X'V}^{-1} = \frac{1}{(\sigma^2(\sigma^2+2\tau^2)}\begin{pmatrix} \sigma^2 & \sigma^2 & \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 & -\sigma^2 & -\sigma^2 \\ \sigma^2+2\tau^2 & -\sigma^2-2\tau^2 & \sigma^2+2\tau^2 & \sigma^2-2\tau^2 \end{pmatrix}$$

$$(\mathbf{X'V}^{-1}\mathbf{X})^{-1} = \text{diag}\left(\frac{\sigma^2+2\tau^2}{4}, \frac{\sigma^2+2\tau^2}{4}, \frac{\sigma^2}{4}\right)$$

$$(\mathbf{X'V}^{-1}\mathbf{X})^{-1}\mathbf{X'V}^{-1} = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

$$\begin{pmatrix} \hat{\mu} \\ \hat{P}_1 \\ \hat{V}_1 \end{pmatrix} = (\mathbf{X'V}^{-1}\mathbf{X})^{-1}\mathbf{X'V}^{-1}\mathbf{Y} = \begin{pmatrix} \frac{y_{11}+y_{12}+y_{21}+y_{22}}{4} \\ \frac{y_{11}+y_{12}-y_{21}-y_{22}}{4} \\ \frac{y_{11}-y_{12}+y_{21}-y_{22}}{4} \end{pmatrix}$$

Similary to part (b) we have $\hat{V}_1 = \frac{1}{4}(y_{11} - y_{12} + y_{21} - y_{22})$.

Since $\hat{V}_1 - \hat{V}_2 = 2\hat{V}_1$ for the chosen linear constraint, we get $\hat{V}_1 - \hat{V}_2 = \frac{1}{2}(y_{11} - y_{12} + y_{21} - y_{22}) = \bar{y}_{\cdot 1} - \bar{y}_{\cdot 2}$.

4. Suppose an experiment is done to compare three varieties of tomatoes in plots that can hold two plants each. There are three such plots available for the study, and the experimental design is

1	2
2	3
1	3

(plants will be randomized within a plot).

The statistical model is $y_{ij} = \mu + P_i + V_j + \epsilon_{ij}$ for $i = 1, 2, 3$ plots and $j = 1, 2, 3$ varieties. Assume the errors ϵ are uncorrelated and have variance σ^2.

For (a) and (c) use the data vector $\mathbf{Y}$ written in the order $\begin{matrix} y_{11} \\ y_{12} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{33} \end{matrix}$

(a) What is $\text{cov}(\mathbf{Y})$ when the "plot effects" P_i are considered fixed effects?
(b) What is the formula for the best linear unbiased estimator of $\hat{V}_1 - \hat{V}_2$ in terms of the Y_{ij} when the P_i are considered fixed effects?
(c) What is $\text{cov}(\mathbf{Y})$ when the "plot effects" P_i are considered random effects with variance τ^2?
(d) What is the formula for the best linear unbiased estimator of $\hat{V}_1 - \hat{V}_2$ in terms of the Y_{ij} when the P_i are considered random effects with variance τ^2? Treat σ^2 and τ^2 as known.

Solution:
(a) $\text{cov}(\mathbf{Y}) = \sigma^2\mathbf{I}$

(b) We go through the same mechanics as for problem 3(a). To give some insight into the result, note that $E[y_{11} - y_{12}] = V_1 - V_2$ since
all the plot effects cancel out. Similarly,
$E[y_{31} - y_{33} + y_{23} - y_{22}] = V_1 - V_2$. These are two linear unbiased estimates of $V_1 - V_2$, and they use different parts of the data so they have covariance 0. If we call these estimators $\hat{\theta}_1$ and $\hat{\theta}_2$, then any combination $c\hat{\theta}_1 + (1 - c)\hat{\theta}_2, c \in [0, 1]$ is also a linear unbiased estimate. It turns out that the BLUE gotten from the linear model and the Gauss-Markov Theorem has this form with $c = \frac{2}{3}$.

(c) $$\text{cov}(\mathbf{Y}) = \begin{pmatrix} \sigma^2+\tau^2 & \tau^2 & 0 & 0 & 0 & 0 \\ \tau^2 & \sigma^2+\tau^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2+\tau^2 & \tau^2 & 0 & 0 \\ 0 & 0 & \tau^2 & \sigma^2+\tau^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2+\tau^2 & \tau^2 \\ 0 & 0 & 0 & 0 & \tau^2 & \sigma^2+\tau^2 \end{pmatrix}$$

(d) For the design matrix $\mathbf{X}$ we use the constraint $P_3 = -P_1 - P_2$ and $V_3 = -V_1 - V_2$ (i.e., $\sum_{i=1}^3 P_i = 0$ and $\sum_{i=1}^3 V_i = 0$);

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

After many steps of computation we obtain

$$\begin{pmatrix} \hat{\mu} \\ \hat{P}_1 \\ \hat{P}_2 \\ \hat{V}_1 \\ \hat{V}_2 \end{pmatrix} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y} = \begin{pmatrix} (-y_{11}+y_{12}-y_{22}+y_{23}+y_{31}+5y_{33})/6 \\ (2y_{11}+y_{12}-y_{22}+y_{23}-2y_{31}-y_{33})/3 \\ (y_{11}-y_{12}+y_{22}+2y_{23}-y_{31}-2y_{33})/3 \\ (y_{11}-y_{12}+y_{22}-y_{23}+2y_{31}-2y_{33})/3 \\ (-y_{11}+y_{12}+2y_{22}-2y_{23}+y_{31}-y_{33})/3 \end{pmatrix}$$

Thus, the best linear unbiased estimator of $V_1 - V_2$ is

$$\hat{V}_1 - \hat{V}_2 = \frac{2y_{11} - 2y_{12} - y_{22} - 3y_{23} + y_{31} - 3y_{33}}{3}.$$

5. Random regressors. Consider the model

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \ldots, n;$$

where

$$\begin{pmatrix} x_i \\ \varepsilon_i \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_x \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{pmatrix}\right).$$

We have

$$\begin{pmatrix} Y_i \\ x_i \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix}, \begin{pmatrix} \sigma_y^2 & \sigma_{xY} \\ \sigma_{xY} & \sigma_x^2 \end{pmatrix}\right),$$

where $\sigma_y^2 = \beta_1^2\sigma_x^2 + \sigma_\varepsilon^2$, $\mu_y = \beta_0 + \beta_1\mu_x$, and $\sigma_{xY} = \beta_1\sigma_x^2$.

(a) Derive $E[Y_i|x_i]$ and $\text{var}[Y_i|x_i]$.

Now suppose one does not observe
$x_i, i = 1, 2, \ldots, n$; but observes $w_i = x_i + u_i$, where

$$\begin{pmatrix} x_i \\ \varepsilon_i \\ u_i \end{pmatrix} \sim N_3\left(\begin{pmatrix} \mu_x \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 \end{pmatrix}\right).$$

Assume that Y is conditionally independent of w: $E[Y_i|x_i, w_i] = E[Y_i|x_i]$. Suppose the true model is $E[Y_i \mid x_i] = \beta_0 + \beta_1 x_i$ but the observed data are $(Y_i, w_i), i = 1, 2, \ldots, n$.

(b) Relate $E[Y_i \mid w_i]$ to $E[x_i \mid w_i]$.

(c) What is the joint distribution of x_i and w_i and what is $E[x_i \mid w_i]$?

(d) Combine your answers to (b) and (c) to show that
$E[Y_i|w_i] = \beta_0^* + \beta_1^* w_i$.

(e) What is the relationship between β_0^*, β_1^* and β_0, β_1?

Solution: (a)

$$\begin{aligned} E[Y_i \mid x_i] &= \mu_y + \sigma_{xY}/\sigma_x^2(x_i - \mu_x) \\ &= \beta_0 + \beta_1\mu_x + \beta_1\sigma_x^2/\sigma_x^2(x_i - \mu_x) \\ &= \beta_0 + \beta_1 x_i \end{aligned}$$

$$\begin{aligned} Var[Y_i \mid x_i] &= \sigma_y^2 - \sigma_{xY}^2/\sigma_x^2 \\ &= \beta_1^2\sigma_x^2 + \sigma_\varepsilon^2 - (\beta_1\sigma_x^2)^2/\sigma_x^2 \\ &= \sigma_\varepsilon^2. \end{aligned}$$

(b)

$$\begin{aligned} E[Y_i \mid w_i] &= E_x[E[Y_i \mid x_i, w_i] \mid w_i] \\ &= E_x[E[Y_i \mid x_i] \mid w_i] \\ &= E_x[\beta_0 + \beta_1 x_i \mid w_i] \\ &= \beta_0 + \beta_1 E[x_i \mid w_i]. \end{aligned}$$

(c)

$$\begin{pmatrix} x_i \\ w_i \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_x \\ \mu_x \end{pmatrix}, \begin{pmatrix} \sigma_x^2 & \sigma_x^2 \\ \sigma_x^2 & \sigma_x^2 + \sigma_u^2 \end{pmatrix}\right).$$

The conditional mean $E[x_i \mid w_i]$ is then

$$\begin{aligned} E[x_i \mid w_i] &= \mu_x + \frac{\sigma_{xw}}{\sigma_w^2}(w_i - \mu_x) \\ &= \mu_x + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2}(w_i - \mu_x) \\ &= \mu_x \frac{\sigma_u^2}{\sigma_x^2 + \sigma_u^2} + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_w^2} w_i \end{aligned}$$

(d)
The conditional mean $E[Y_i \mid w_i]$ (in terms of
Y and w is then

$$\begin{aligned} E[Y_i \mid w_i] &= \beta_0 + \beta_1 E[x_i \mid w_i] \\ &= \beta_0 + \beta_1[\mu_x \frac{\sigma_u^2}{\sigma_x^2 + \sigma_u^2} + \frac{\sigma_x^2}{\sigma_x^2 + \sigma_w^2} w_i] \\ &= \beta_0 + \beta_1\mu_x \frac{\sigma_u^2}{\sigma_x^2 + \sigma_u^2} + \beta_1 \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2} w_i \\ &= \beta_0^* + \beta_1^* w_i. \end{aligned}$$

(e) You can see from part (d) above that

$$\beta_0^* = \beta_0 + \beta_1 \mu_x \frac{\sigma_u^2}{\sigma_x^2 + \sigma_u^2}$$

and

$$\beta_1^* = \beta_1 \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2}$$