Stat 511 Exam 2

April 11, 2003 Prof. Vardeman

1. The first printout attached to this exam concerns the analysis of some data from a tire grip study. Drag, $x$ (measured in %), is applied to a tire and the (grip) force, $y$ (measured in lbs), with which it holds the pavement is measured. Physical theory suggests that

$$y \approx a \exp(b x)$$

We'll consider two analyses of $n = 19$ data pairs $(x_i, y_i)$. The first is one based on the model

$$\ln y_i = \ln a + b x_i + e_i \qquad (*)$$

and the second is based on the model

$$y_i = a \exp(b x_i) + e_i \qquad (**)$$

(where in both cases the errors $e_i$ are iid $\mathrm{N}(0, s^2)$). (Actually, simple plots of $\ln y$ vs $x$ and of $y$ vs $x$ call into question the constant variance assumptions. But we will ignore this concern for the narrow purposes of this exam.)

a) Based on the linear model in (*), give 95% confidence limits for $a$ (not $\ln a$ ) and 95% limits for $b$ . (Plug in, but you need not do arithmetic.)

b) Give approximate 95% confidence limits for $a$ and $b$ based on the nonlinear model (**).

c) Give (any set of) approximate 90% confidence limits for $s$ in the nonlinear model (**). (Plug in numbers, but you need not do arithmetic.)

d) Find approximate 95% prediction limits for the next grip force if $x = .5$, based on the nonlinear model (**). (There is enough information on the printout to allow you to apply a bit of calculus and some simple matrix calculations to get this. When you have reduced your answer to a completely numerical expression, you may stop without doing arithmetic.)

e) The graph on the printout is a contour plot for the error sum of squares. Based on this plot, how plausible is the hypothesis $H_0\colon a = 650$ and $b = -.007$? Make some quantitative statement if you can. (It's possible, for example, to get a rough idea of a $p$-value from the plot.)

f) A model is linear in its (mean) parameters exactly when the error sum of squares is quadratic in those parameters (and produces exactly elliptical level contours). In light of this fact, what does the substantial agreement between answers to parts a) and b) and the appearance of the contour plot indicate about model (**) and the inferences based on it?

2. Consider the small mixed model

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = \underset{6\times 1}{\mathbf{1}} m + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} + \underset{6\times 1}{\mathbf{e}}$$

for $\mathbf{G} = s_u^2 \underset{4\times 4}{\mathbf{I}}$ and $\mathbf{R} = s^2 \underset{6\times 6}{\mathbf{I}}$ .

a) What is the variance-covariance matrix for $\mathbf{Y}$ in this model?

b) REML estimation of variance components in this model is maximum likelihood estimation of variance components based on $\underset{5\times 6}{\mathbf{B}}\mathbf{Y}$ for an appropriate $\mathbf{B}$. Give one such $\mathbf{B}$.

The second printout attached to this exam enables a mixed effects analysis for the observation vector $\mathbf{Y} = (.7,.6,2.1,2.3,.1,1.5)'$.

c) The numerical maximization of the restricted likelihood produces $\hat{s} = .1118$. There is another route to producing essentially this value. Show an elementary calculation that produces this value.

d) Give 95% exact confidence limits for $s$ . (Show some work. Plug in numerical values, but you don't need to do arithmetic.)

e) What are 95% approximate confidence limits for $s_u$?

f) Observations $y_1$ and $y_2$ are from "group 1." How do you suggest predicting another (unobserved) value from this group? (Give a numerical value and say what it is in terms used in lecture.) Would you use the same numerical value for predicting if instead of the mixed model we used a fixed effects model with 4 unknown group means and $s^2\mathbf{I}$ for a variance-covariance matrix? (If not, what value would you use?)

3. In the study of the precision of a measuring device, each of $a = 2$ widgets (call "widgets" levels of Factor A) was measured $m = 2$ times by each of $b = 5$ different technicians (call "technicians" levels of Factor B). The resulting data can be thought of as having $2 \times 5$ (complete, balanced, replicated) factorial structure. With

$$y_{ijk} = \text{measurement } k \text{ by technician } j \text{ on widget } i$$

model as

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}$$

where $\mu$ is the only fixed effect, all the random effects are independent, the $\alpha_i \sim \mathrm{N}\left(0, \sigma_\alpha^2\right)$, the $\beta_j \sim \mathrm{N}\left(0, \sigma_\beta^2\right)$, the $\alpha\beta_{ij} \sim \mathrm{N}\left(0, \sigma_{\alpha\beta}^2\right)$, and the $\epsilon_{ijk} \sim \mathrm{N}\left(0, \sigma^2\right)$. Standard 2-way factorial ANOVA calculations were done and produced the following ANOVA table.

| **Source** | ***df*** | ***MS*** | **E*MS*** |
|---|---|---|---|
| A | 1 | 20.0 | $\sigma^2 + 2\sigma_{\alpha\beta}^2 + 10\sigma_\alpha^2$ |
| B | 4 | 10.0 | $\sigma^2 + 2\sigma_{\alpha\beta}^2 + 4\sigma_\beta^2$ |
| AB | 4 | 6.0 | $\sigma^2 + 2\sigma_{\alpha\beta}^2$ |
| Error | 10 | 2.0 | $\sigma^2$ |

a) A quantity of serious interest in this context is $\sigma_\beta^2 + \sigma_{\alpha\beta}^2$ (which is called a measure of measurement "reproducibility"). Find a sensible point estimate of this quantity.

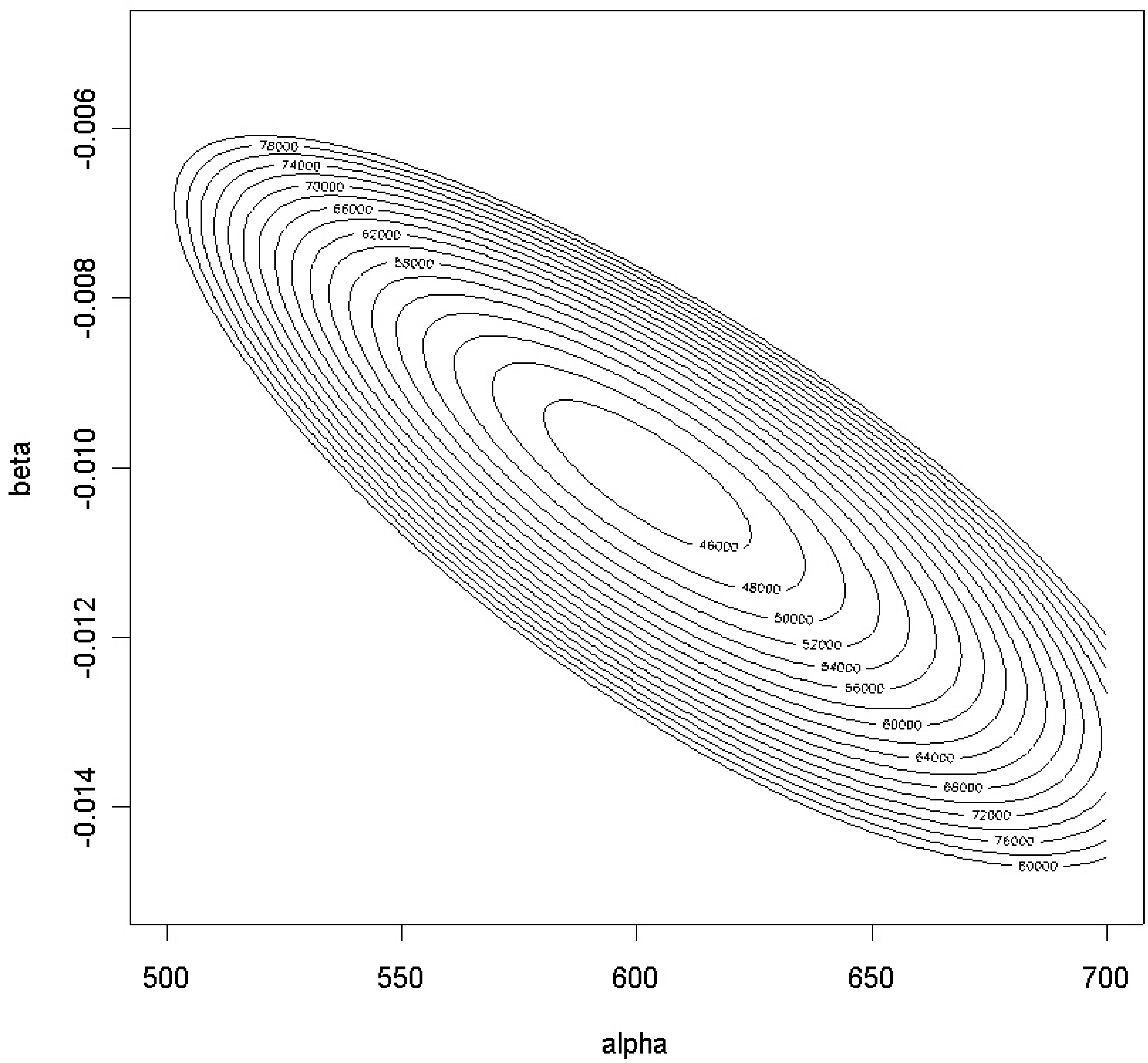
b) Make approximate 90% confidence limits for $\sigma_\beta^2 + \sigma_{\alpha\beta}^2$. (As usual, you should plug in numbers, but you don't need to do arithmetic.)

## Printout #1

```
> drag<-c(10,10,10,20,20,20,30,30,30,50,50,50,70,70,70,100,100,100,100)
> grip<-c(550,460,610,510,410,580,470,360,480,390,310,400,300,280,340,250,200,
200,200)
> d<-data.frame(drag,grip)
> lgrip<-log(grip)
> ld<-data.frame(drag,lgrip)

> lm.out<-lm(lgrip~drag,data=ld)

> summary(lm.out)

Call:
lm(formula = lgrip ~ drag, data = ld)

Residuals:
     Min       1Q   Median       3Q      Max
-0.20591 -0.07680  0.02142  0.09303  0.16860

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.3992499  0.0517216  123.72  < 2e-16 ***
drag        -0.0102413  0.0008749  -11.71 1.47e-09 ***
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Residual standard error: 0.1234 on 17 degrees of freedom
Multiple R-Squared: 0.8896,     Adjusted R-squared: 0.8831
F-statistic:   137 on 1 and 17 DF,  p-value: 1.470e-09

> nlm.out<-nls(formula=grip~alpha*exp(beta*drag),start=c(alpha=exp(6.3992499),
beta=-0.0102413),trace=T)
44569.64 :  601.3937632  -0.0102413
44419.53 :  601.94401292  -0.01008628
44419.51 :  601.98037841  -0.01008891

> summary(nlm.out)

Formula: grip ~ alpha * exp(beta * drag)

Parameters:
        Estimate Std. Error t value Pr(>|t|)
alpha 601.980378  28.603057   21.05 1.30e-13 ***
beta   -0.010089   0.001161   -8.69 1.16e-07 ***
---
Signif. codes:  0 `***' 0.001 `**' 0.01 `*' 0.05 `.' 0.1 ` ' 1

Residual standard error: 51.12 on 17 degrees of freedom

Correlation of Parameter Estimates:
       alpha
beta -0.7887
```

```
> confint.nls(nlm.out,level=.95)
Waiting for profiling to be done...

.
.
.

2.5%          97.5%
alpha 543.39335246 663.907579508
beta   -0.01263842  -0.007740446

> ss<-function(a,b)
+ (550-a*exp(10*b))^2+
+ (460-a*exp(10*b))^2+
+ (610-a*exp(10*b))^2+
+ (510-a*exp(20*b))^2+
+ (410-a*exp(20*b))^2+
+ (580-a*exp(20*b))^2+
+ (470-a*exp(30*b))^2+
+ (360-a*exp(30*b))^2+
+ (480-a*exp(30*b))^2+
+ (390-a*exp(50*b))^2+
+ (310-a*exp(50*b))^2+
+ (400-a*exp(50*b))^2+
+ (300-a*exp(70*b))^2+
+ (280-a*exp(70*b))^2+
+ (340-a*exp(70*b))^2+
+ (250-a*exp(100*b))^2+
+ (200-a*exp(100*b))^2+
+ (200-a*exp(100*b))^2+
+ (200-a*exp(100*b))^2
+ }

> ss(601.980378,-0.010089)
[1] 44419.51

> alpha<-500:700
> beta<-seq(-.015,-.005,.0001)
> SumofSquares<-outer(alpha,beta,FUN=ss)
```

```
> contour(alpha,beta,SumofSquares,levels=seq(40000,80000,2000),xlab="alpha",
ylab="beta")
```



## Printout #2

```
> y<-c(.7,.6,2.1,2.3,.1,1.5)
> group<-c(1,1,2,2,3,4)
> gd<-groupedData(y~1|group)
> fm1<-lme(y~1,random=~1|group)
> summary(fm1)
Linear mixed-effects model fit by REML
 Data: NULL
       AIC      BIC    logLik
  13.73695 12.56526 -3.868475

Random effects:
 Formula: ~1 | group
        (Intercept)  Residual
StdDev:   0.9211296 0.1117760

Fixed effects: y ~ 1
               Value Std.Error DF  t-value p-value
(Intercept) 1.113638 0.4630979  4 2.404757   0.074

Standardized Within-Group Residuals:
        Min          Q1         Med          Q3         Max
-0.82361234 -0.39112820 -0.04071808  0.32529533  0.96568027

Number of Observations: 6
Number of Groups: 4

> intervals(fm1)

Approximate 95% confidence intervals
 Fixed effects:
                 lower     est.    upper
(Intercept) -0.1721281 1.113638 2.399404

 Random Effects:
  Level: group
                    lower      est.    upper
sd((Intercept)) 0.4117237 0.9211296 2.060798

 Within-group standard error:
     lower       est.      upper
0.04197268 0.11177602 0.29766689

> fixed.effects(fm1)
(Intercept)
   1.113638
> random.effects(fm1)
  (Intercept)
1  -0.4602492
2   1.0784223
3  -0.9989286
4   0.3807555

> predict.nlme(fm1)
        1         1         2         2         3         4
0.6533886 0.6533886 2.1920601 2.1920601 0.1147092 1.4943934
attr(,"label")
[1] "Fitted values"
```