

1. Consider a 3×3 factorial analysis with factors A and B under an ordinary (fixed effects) linear model. The effects model under the R baseline restriction has parameter vector for mean responses

$$\gamma = (\mu^*, \alpha_2^*, \alpha_3^*, \beta_2^*, \beta_3^*, \alpha\beta_{22}^*, \alpha\beta_{23}^*, \alpha\beta_{32}^*, \alpha\beta_{33}^*)'$$

a) Write out all 9 cell means in terms of the entries of γ in the table below. Let rows correspond to levels of A (1 to 3 top to bottom) and columns correspond to levels of B (1 to 3 left to right).

b) Give below a matrix C so that the *testable* hypothesis $H_0 : C\gamma = \mathbf{0}$ is the hypothesis $H_0 : \beta_j = 0 \forall j$. (As always, $\beta_j = \mu_j - \mu_{..}$.)

c) Give below a matrix C so that the *testable* hypothesis $H_0 : C\gamma = \mathbf{0}$ is the hypothesis that considering only levels 2 and 3 of A and levels 2 and 3 of B (the lower right 4 cells in the table) “there are no interactions” (considering only these 4 cells an interaction plot of means would have the “parallelism” property).

d) Suppose that all cell sample sizes are 1 and write $\mathbf{Y}' = (y_{11}, y_{12}, y_{13}, y_{21}, y_{31}, y_{22}, y_{23}, y_{32}, y_{33})$. The hypothesis in c) is that $E\mathbf{Y}$ is in some subspace of \mathfrak{R}^9 of dimension 8. Write below a 9×8 matrix \mathbf{X}_0 so that $C(\mathbf{X}_0)$ is this subspace. (Hint: Note that this hypothesis says nothing at all about 5 of the 9 mean responses, and be careful to pay attention to the order of the entries of \mathbf{Y} listed above.)

2. A data set of Kaplan, et al. (1972) on the metabolism of sulfisoxazole can be found on page 273 of Bates and Watts. This substance was administered intravenously to a subject and blood samples were taken over time. Concentration of sulfisoxazole in the plasma, y (in $\mu\text{g/ml}$), was measured. Bates and Watts suggest an analysis of the data based on a model for concentration as a function of time, t (in minutes from injection),

$$y_i = \theta_1 \exp(-\theta_2 t_i) + \theta_3 \exp(-\theta_4 t_i) + \varepsilon_i \quad (*)$$

Attached to this exam is an R printout that you should use to answer the following questions.

a) There are actually two different non-linear regression models represented on the R printout. The first is model (*) and the second is the less complex model $y_i = \theta_1 \exp(-\theta_2 t_i) + \varepsilon_i$. What **quantitative** support does the printout provide for the use of model (*) rather than the simpler second model?

b) In fitting model (*), what one gets as parameter estimates depends upon one's starting values for $(\theta_1, \theta_2, \theta_3, \theta_4)$. In fact, it is possible to produce exactly the same predicted values and SSE as shown on the printout for a different set of parameter estimates $(\hat{\theta}_1^*, \hat{\theta}_2^*, \hat{\theta}_3^*, \hat{\theta}_4^*)$. How is this already obvious from (*)?

Henceforth base your responses on model (*).

c) Based on model (*), an initial concentration is $\theta_1 + \theta_3$. Give approximate 95% confidence limits for this value. (You don't need to do arithmetic, but YOU MUST PLUG IN NUMBERS.)

d) It is potentially of interest to know the value of

$$t_{.5} = \text{the time at which concentration is } .5(\theta_1 + \theta_3)$$

$t_{.5}$ is some function of $(\theta_1, \theta_2, \theta_3, \theta_4)$, say $\tau(\theta_1, \theta_2, \theta_3, \theta_4)$. It is possible to show (don't try to do so) that here

$$\tau(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4) = 2.06 \quad \text{and} \quad \nabla \tau = \left(\frac{\partial \tau}{\partial \theta_i} \Big|_{(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4)} \right) = (-.01664, -.4352, .08311, 240.4)$$

Give approximate 95% confidence limits for $t_{.5}$. (Again, don't do arithmetic, BUT DO PLUG IN.)

3. Consider a scenario in which 12 samples of a large lot of material are sent 3 apiece to 4 different labs for hardness testing. At the labs, each specimen is tested twice. For

y_{ijk} = the hardness measured on the k th test of the j th specimen at the i th lab

where $i = 1, 2, 3, 4$ and $j = 1, 2, 3$ and $k = 1, 2$, suppose that

$$y_{ijk} = \mu_i + \phi_{ij} + \varepsilon_{ijk} \quad (**)$$

for constants $\mu_1, \mu_2, \mu_3, \mu_4$, the ϕ_{ij} iid $N(0, \sigma_\phi^2)$ independent of the iid $N(0, \sigma^2)$ random variables ε_{ijk} .

a) For the y_{ijk} written in dictionary order in \mathbf{Y} , what are $\mathbf{X}, \boldsymbol{\beta}, \mathbf{Z}$, and \mathbf{u} so the model can be written in standard mixed linear model form?

Stat 511 Exam II Spring 2004 Printouts

Problem 2

```
> time<-c(.25,.50,.75,1.00,1.50,2.00,3.00,4.00,6.00,12.00,24.00,48.00)
> conc<-c(215.6,189.2,176.0,162.8,138.6,121.0,101.2,88.0,61.6,22.0,4.4,.01)
```

```
> blood.fm1<-nls(formula=conc~(theta1*exp(-theta2*time)+theta3*exp(-
theta4*time)),start=c(theta1=200,theta2=5,theta3=70,theta4=.1),trace=T)
```

```
55161.36 : 200.0 5.0 70.0 0.1
19963.2 : 52.4343646 2.3213531 120.7827395 0.1637390
7107.025 : 66.8313947 0.2547202 168.7018220 0.1692898
7100.616 : 62.4548374 0.2588085 173.0702195 0.1700836
7093.994 : 58.4335743 0.2630338 177.0836029 0.1708158
7093.283 : 51.0640427 0.2717472 184.4379396 0.1721668
7090.37 : 44.8794949 0.2810323 190.6084239 0.1733210
7084.891 : 39.7259944 0.2908481 195.7490068 0.1743080
7076.77 : 35.4470720 0.3011472 200.0161628 0.1751539
7066.12 : 31.8979404 0.3118771 203.5546740 0.1758805
7053.153 : 28.9518490 0.3229817 206.4912762 0.1765066
7046.573 : 24.0509948 0.3458234 211.3753644 0.1775888
7027.508 : 20.6117921 0.3700385 214.8019914 0.1784065
7018.247 : 15.7654693 0.4202012 219.6310298 0.1796486
7005.513 : 10.8169763 0.5269206 224.5724600 0.1811158
6793.44 : 8.5514174 0.7518429 226.9114329 0.1821330
5964.866 : 11.2819924 1.1022659 224.5584007 0.1818996
4472.041 : 19.9510024 1.2816124 216.8710602 0.1799516
2510.098 : 35.5643008 1.2857303 202.9706711 0.1760439
627.6562 : 58.7612699 1.2944981 182.3755254 0.1696706
34.79365 : 81.4693395 1.3049686 162.3408378 0.1621803
34.37898 : 81.2428056 1.3060937 162.5976318 0.1617880
34.37898 : 81.2416105 1.3060546 162.5975340 0.1617901
```

```
> blood.fm2<-nls(formula=conc~theta1*exp(-
theta2*time),start=c(theta1=240,theta2=.5),trace=T)
```

```
10440.87 : 240.0 0.5
1835.811 : 206.0920489 0.1812607
788.3232 : 209.0524358 0.2227155
751.5445 : 211.4521261 0.2337930
750.7822 : 211.8584747 0.2354533
750.769 : 211.9122833 0.2356727
750.7688 : 211.9192618 0.2357012
750.7688 : 211.9201658 0.2357049
```

```
> summary(blood.fm2)
```

```
Formula: conc ~ theta1 * exp(-theta2 * time)
```

```
Parameters:
```

	Estimate	Std. Error	t value	Pr(> t)	
theta1	211.9202	5.9877	35.39	7.69e-12	***
theta2	0.2357	0.0176	13.39	1.04e-07	***

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 8.665 on 10 degrees of freedom

Correlation of Parameter Estimates:

```
      theta1
theta2 0.7198
```

```
> summary(blood.fm1)
```

Formula: conc ~ (theta1 * exp(-theta2 * time) + theta3 * exp(-theta4 * time))

Parameters:

	Estimate	Std. Error	t value	Pr(> t)	
theta1	81.24161	6.10743	13.302	9.74e-07	***
theta2	1.30606	0.19757	6.611	0.000168	***
theta3	162.59753	6.94309	23.419	1.18e-08	***
theta4	0.16179	0.00878	18.427	7.75e-08	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.073 on 8 degrees of freedom

Correlation of Parameter Estimates:

```
      theta1 theta2 theta3
theta2 -0.4135
theta3 -0.7648 0.8792
theta4 -0.7743 0.7351 0.9291
```

```
> vcov(blood.fm1)
```

	theta1	theta2	theta3	theta4
theta1	37.30071480	-0.498973414	-32.42998263	-4.152238e-02
theta2	-0.49897341	0.039032249	1.20604785	1.275166e-03
theta3	-32.42998263	1.206047855	48.20643196	5.663934e-02
theta4	-0.04152238	0.001275166	0.05663934	7.708909e-05

```
> predict(blood.fm1)
```

```
[1] 214.76259903 192.24579905 174.52247123 160.31594516 139.01485001
[6] 123.60956647 101.68867320 85.56219164 61.62444317 23.33134279
[11] 3.34784268 0.06893125
```

```
> x<-seq(0,50,.5)
```

```
> yhat.fm1<-(coef(blood.fm1)[1]*exp(-
coef(blood.fm1)[2]*x)+coef(blood.fm1)[3]*exp(-coef(blood.fm1)[4]*x))
```

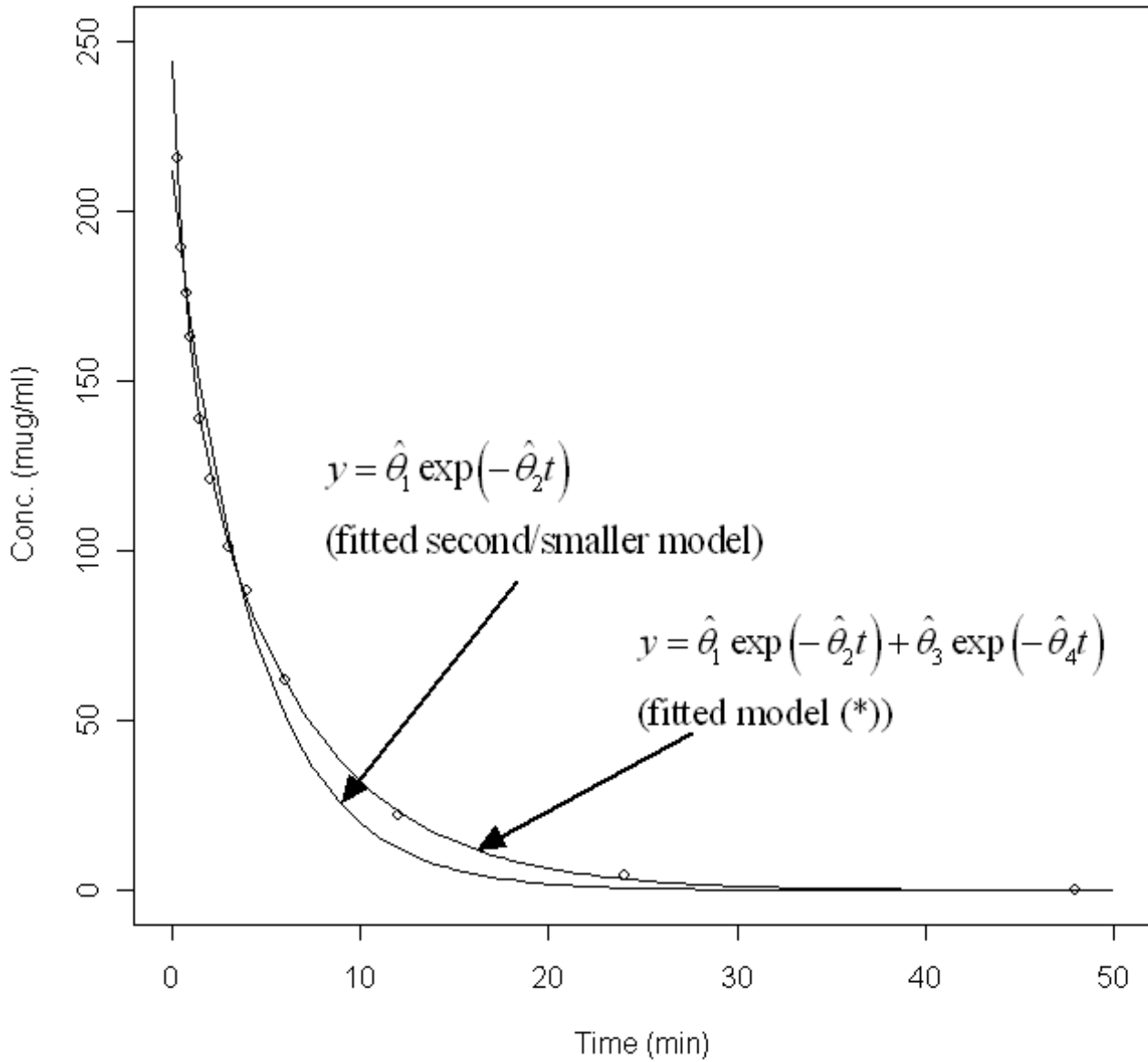
```
> yhat.fm2<-(coef(blood.fm2)[1]*exp(-coef(blood.fm2)[2]*x))
```

```
> plot(c(0,50),c(0,250),type="n",xlab="Time (min)",ylab="Conc. (mug/ml)")
```

```
> points(time,conc)
```

```
> lines(x,yhat.fm1)
```

```
> lines(x,yhat.fm2)
```



Problem 3

```
> lab<-c(1,1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3,4,4,4,4,4,4)
> LAB<-as.factor(lab)
> specimen<-c(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10,11,11,12,12)
> SPECIMEN<-as.factor(specimen)
> y
 [1] 101.68 101.32 101.36 101.06 101.36 101.83 102.89 102.83 103.63 103.64
[11] 106.60 106.58 103.55 103.48 103.44 103.74 101.62 101.99 107.12 106.94
[21] 108.26 108.33 106.84 106.83

> fml.out<-lme(y~LAB-1,random=~1|SPECIMEN)
```

```

> summary(fm1.out)
Linear mixed-effects model fit by REML
Data: NULL
      AIC      BIC    logLik
41.92476 47.89915 -14.96238

Random effects:
Formula: ~1 | SPECIMEN
      (Intercept)  Residual
StdDev:    1.172893 0.1724569

Fixed effects: y ~ LAB - 1
      Value Std.Error DF  t-value p-value
LAB1 101.4350 0.6808201  8 148.9894    0
LAB2 104.3617 0.6808201  8 153.2882    0
LAB3 102.9700 0.6808201  8 151.2441    0
LAB4 107.3867 0.6808201  8 157.7313    0
Correlation:
      LAB1 LAB2 LAB3
LAB2 0
LAB3 0    0
LAB4 0    0    0

Standardized Within-Group Residuals:
      Min      Q1      Med      Q3      Max
-1.35273732 -0.33630320 -0.01064225  0.31939459  1.37258065

Number of Observations: 24
Number of Groups: 12

> fixed.effects(fm1.out)
      LAB1      LAB2      LAB3      LAB4
101.4350 104.3617 102.9700 107.3867

> vcov(fm1.out)
      LAB1      LAB2      LAB3      LAB4
LAB1 0.463516 0.000000 0.000000 0.000000
LAB2 0.000000 0.463516 0.000000 0.000000
LAB3 0.000000 0.000000 0.463516 0.000000
LAB4 0.000000 0.000000 0.000000 0.463516

> random.effects(fm1.out)
(Intercept)
1  0.06430488
2 -0.22259382
3  0.15828894
4 -1.48560765
5 -0.71889560
6  2.20450325
7  0.53917170
8  0.61336964
9 -1.15254134
10 -0.35285243
11  0.89861950
12 -0.54576707

```



```
> intervals(fml.out)
Approximate 95% confidence intervals
```

```
Fixed effects:
```

	lower	est.	upper
LAB1	99.86503	101.4350	103.0050
LAB2	102.79169	104.3617	105.9316
LAB3	101.40003	102.9700	104.5400
LAB4	105.81669	107.3867	108.9566

```
attr("label")
```

```
[1] "Fixed effects:"
```

```
Random Effects:
```

```
Level: SPECIMEN
```

	lower	est.	upper
sd((Intercept))	0.7161808	1.172893	1.920852

```
Within-group standard error:
```

	lower	est.	upper
	0.1155930	0.1724569	0.2572941