I have neither given nor received unauthorized assistance on this exam.

________________________________________________________
Name

________________________________________________________
Name Printed
1. An experimental data set in a set of slides due to S.A. Jenekhe found on the University of Washington Chemistry Department web site of Prof. Lawrence Ricker concerns CO$_2$ solubility in a glassy polymer. Given are the pressures, $p$, and corresponding concentrations, $c$, of CO$_2$ below.

<table>
<thead>
<tr>
<th>Pressure, $p$ (atm)</th>
<th>2.74</th>
<th>6.10</th>
<th>9.76</th>
<th>14.45</th>
<th>18.92</th>
<th>26.74</th>
<th>33.28</th>
<th>42.23</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration, $c$ (cm$^3$ (STP)/cm$^3$ polymer)</td>
<td>36.6</td>
<td>51.4</td>
<td>64.3</td>
<td>78.7</td>
<td>91.5</td>
<td>110.3</td>
<td>122.9</td>
<td>143.7</td>
</tr>
</tbody>
</table>

A standard deterministic model for gas solubility in a polymer is

$$c = H p + L_C \left( \frac{L_a p}{1 + L_a p} \right)$$

for constants $H$ (the Henry's law constant), $L_C$ (the Langmuir capacity constant), and $L_a$ (the Langmuir affinity constant). Below is a plot of these data and a fitted concentration versus pressure curve. Note that for large pressure this (fitted) curve is nearly linear with slope $H$ and intercept $L_C$, while the derivative of concentration with respect to pressure at 0 pressure is $H + L_C \cdot L_a$.

There is an R printout at the end of this exam from the session in which this plot was generated. Use it to answer the following questions about a nonlinear regression analysis of this situation based on a model

$$c_i = H p_i + L_C \left( \frac{L_a p_i}{1 + L_a p_i} \right) + \varepsilon_i$$

for iid $N(0, \sigma^2)$ errors $\varepsilon_i$, $i = 1, 2, \ldots, 8$. 
a) What are approximate 95% confidence limits for the standard deviation of concentration at any fixed pressure according to the model (*)? (If you need some percentage point(s) of a distribution that you don't have, say very carefully/completely exactly what you need.) Plug into any formula you provide.

b) Under what conditions on the parameters of model (*) is the mean concentration a simple multiple of pressure? Is there definitive evidence in these data that such a ("single mode") model is too simple and so the full complexity of model (*) is justified? Explain in terms of some measures of statistical significance.
c) What are approximate 95% confidence limits for the derivative of mean concentration with respect to pressure at 0 pressure? (Plug into an appropriate formula. You don't need to do arithmetic, but you must plug in, and if you don't have necessary percentage points of a distribution, say very carefully/completely exactly what you need.)

2. In a typical industrial "gauge R&R study," each of \( I \) different parts from some process is measured \( m \) times by each of \( J \) different operators, as a way of studying the consistency of measurement using a single gauge. We will here consider a case where the operators are "fixed" (being the only ones a company will ever use to do such measuring) while parts are "random" (representing ongoing production of such parts) and for 
\[
y_{ijk} = \text{the kth measurement obtained on part } i \text{ by operator } j
\]
model as 
\[
y_{ijk} = \mu + \alpha_i + \beta_j + \alpha \beta_{ij} + \epsilon_{ijk} \tag{**}
\]
where \( \mu \) and the \( \beta_j \) are unknown constants and the \( \alpha_i, \alpha \beta_{ij}, \) and \( \epsilon_{ijk} \) are independent random variables, with \( \alpha_i \sim \text{iid N}\left(0, \sigma_{\alpha}^2\right) \), \( \alpha \beta_{ij} \sim \text{iid N}\left(0, \sigma_{\alpha \beta}^2\right) \), and \( \epsilon_{ijk} \sim \text{iid N}\left(0, \sigma^2\right) \). (Here the \( \beta_j \) might be thought of as consistent operator biases and the \( \alpha \beta_{ij} \) might be thought of as so-called operator "nonlinearities of measurement.")

To begin, first consider a small/toy case where \( I = J = m = 2 \) (there are 2 parts, 2 operators, and each part is measured 2 times by each operator).
a) For 8 observations written down in dictionary order, show how to write out model (***) in mixed linear model matrix form (by providing the elements of \( Y = X\beta + Zu + \epsilon \) indicated below).

\[
Y = \begin{pmatrix}
y_{111} \\
y_{112} \\
y_{121} \\
y_{122} \\
y_{211} \\
y_{212} \\
y_{221} \\
y_{222}
\end{pmatrix}
\]

\[
X = \begin{pmatrix}
\beta = \begin{pmatrix}
Z = \\
u = \begin{pmatrix}
\end{pmatrix}
\end{pmatrix}
\end{pmatrix}
\]

b) Write out the following in terms of model (***) parameters.

\[
\text{Var} \, y_{111} = \text{______________________________} \\
\text{Cov} \,(y_{111}, y_{112}) = \text{______________________________}
\]

\[
\text{Cov} \,(y_{111}, y_{121}) = \text{______________________________} \\
\text{Cov} \,(y_{111}, y_{211}) = \text{______________________________}
\]

At the end of this exam, there is an R printout for an analysis based on model (***) of a modification of a real data set from an R&R study based on \( I = 4 \) parts, \( J = 3 \) operators, and \( m = 2 \) measurements per part. (These are measured heights of some steel punches in \( 10^{-3} \) inch.) Use it to answer the next two questions.
c) Based on the results on the printout
   • do you find clear evidence of differences in operator measurement "biases," and
   • do operator "nonlinearities" appear to play a large role in measurement of these punch heights?
   (Return to the parenthetical remark following model statement (**) for use of these terms.) Explain using appropriate values from the printout.

d) What are approximate BLUPs for
   • \( \mu + \alpha_i + \beta_i + \alpha \beta_{1i} \) (a long-run average of measurements of part 1 by operator 1) (give a numerical value)

   • \( \mu + \alpha_i + \beta_i + \alpha \beta_{51} + \epsilon_{51} \) (a measurement on a new punch by operator #1) Here, give both the BLUP AND an appropriate standard error (give numerical values).
3. Suppose that for \( i, j = 1, 2 \), \( y_{ij} = \mu + \alpha_i + \epsilon_{ij} \) for independent variables

\[ \alpha_i \sim \text{iid } \mathcal{N}(0, \sigma^2) \text{ and } \epsilon_{ij} \sim \text{iid } \mathcal{N}(0, \sigma^2) . \]

Take \( W = BY \) for \( Y = (y_{11}, y_{12}, y_{21}, y_{22})^\top \) and

\[
B = \begin{pmatrix}
1 & 1 & -1 & -1 \\
1 & -1 & 0 & 0 \\
0 & 0 & 1 & -1
\end{pmatrix}.
\]

a) Argue that REML estimation of \( \sigma^2_a \) and \( \sigma^2 \) can be based on \( W \) and write out explicitly the function of \( w_1, w_2, w_3, \sigma^2_a, \) and \( \sigma^2 \) that can be maximized as a function of \( \sigma^2_a \) and \( \sigma^2 \) to produce REML estimates.

b) The restricted loglikelihood from a) can be written as a function of log-variances \( \gamma_1 = \log \sigma^2 \) and \( \gamma_2 = \log \sigma^2_a \). For a particular \( W \), this is maximized at \( \hat{\gamma}_1 = .6931 \) and \( \hat{\gamma}_2 = 1.3863 \). The Hessian (matrix of second partial derivatives) of this function at its maximizer is

\[
\begin{pmatrix}
-1.5 & -1 \\
-1 & -2
\end{pmatrix}.
\]

Find approximate 95% confidence limits for \( \sigma_a \) based on this information. (Plug in and evaluate.)
4. Suppose that for $i = 1, 2, 3$ independent observations $y_{ij} \sim \text{iid } N\left(\mu_i, \sigma_i^2\right)$ for $j = 1, \ldots, n_i$ (that is, we have independent samples of sizes $n_i$ from three different normal distributions). For constants $c_1, c_2, \text{ and } c_3$, the random variable $c_1\bar{y}_1 + c_2\bar{y}_2 + c_3\bar{y}_3$ has variance

$$V = \frac{c_1^2}{n_1}\sigma_1^2 + \frac{c_2^2}{n_2}\sigma_2^2 + \frac{c_3^2}{n_3}\sigma_3^2$$

Based on variances for the 3 samples ($s_1^2, s_2^2, \text{ and } s_3^2$) use the Cochran-Satterthwaite approximation to identify approximate 95% confidence limits for $V$. (Say explicitly what percentage point(s) of exactly what distribution will be needed.)
For Problem #1

\begin{verbatim}
> pressure<-c(2.74,6.10,9.76,14.45,18.92,26.74,33.28,42.23)
> conc<-c(36.6,51.4,64.3,78.7,91.5,110.3,122.9,143.7)

> nlrfit<-nls(formula=conc~H*pressure+LC*((LA*pressure)/(1+LA*pressure)),start=c(H=2,LC=50,LA=.5),trace=T)
446.4045 :   2.0 50.0  0.5
11.02893 :   2.1865348 54.6735257  0.4116493
10.34221 :   2.1785177 55.0649938  0.4177164
10.34211 :   2.1789683 55.0447196  0.4182403
10.34211 :   2.1790082 55.0429003  0.4182864
10.34211 :   2.1790117 55.0427404  0.4182904

> summary(nlrfit)
Formula: conc ~ H * pressure + ((LC * LA)/(1 + LA * pressure)) * pressure

Parameters:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| H        | 2.17901    | 0.08574 | 25.414   | 1.76e-06 *** |
| LC       | 55.04274   | 3.49365 | 15.755   | 1.87e-05 *** |
| LA       | 0.41829    | 0.08051 | 5.195    | 0.00348 **   |

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.438 on 5 degrees of freedom

Number of iterations to convergence: 5
Achieved convergence tolerance: 1.959e-06

> confint(nlrfit)
Waiting for profiling to be done...
51.67906 :  55.0427404  0.4182904
13.45704 :  58.8543577  0.3425121

:
:
:

47.62486 :   2.455613 43.357777
64.20141 :   2.523317 40.680263
61.44185 :   2.497246 41.751108
79.19115 :   2.562831 39.220090
76.03984 :   2.53446 40.34596
94.31986 :   2.598318 37.934771
90.55775 :   2.566830 39.146689

2.5%     97.5%
H  1.9317119  2.3913757
LC 46.7938116 65.8979832
LA 0.2616633  0.7487824

> vcov(nlrfit)

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>LC</th>
<th>LA</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>0.007351596</td>
<td>-0.2862178</td>
<td>0.005690264</td>
</tr>
<tr>
<td>LC</td>
<td>-0.286217844</td>
<td>12.2056221</td>
<td>-0.259450579</td>
</tr>
<tr>
<td>LA</td>
<td>0.005690264</td>
<td>-0.2594506</td>
<td>0.006482498</td>
</tr>
</tbody>
</table>
\end{verbatim}

For Problem #2
> height
  495 497 498 498 496 496 498 497
> Part
  [1] 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4 4 4 4
Levels: 1 2 3 4
> Operator
  [1] 1 1 2 2 3 3 1 1 2 2 3 3 1 1 2 2 3 3 1 1 2 2 3 3
Levels: 1 2 3
> PartbyOperator
  [1] 1 1 2 2 3 3 1 1 2 2 3 3 1 1 2 2 3 3 1 1 2 2 3 3
Levels: 1 2 3

> lmeRandR<-lme(height~1+Operator,random=-1|Part/PartbyOperator)

> summary(lmeRandR)
Linear mixed-effects model fit by REML
Data: NULL
  AIC     BIC    logLik
  85.73936 92.0065 -36.86968
Random effects:
  Formula: ~1 | Part
    (Intercept)
    StdDev:    1.596437
  Formula: ~1 | PartbyOperator %in% Part
    (Intercept)  Residual
    StdDev:   0.4930067 0.9128709
Fixed effects: height ~ 1 + Operator
  Value Std.Error DF t-value p-value
  (Intercept) 498.625 0.8955912 12 556.7552  0.0000
  Operator2    -1.000 0.5743354  6  -1.7411  0.1323
  Operator3    -0.625 0.5743354  6  -1.0882  0.3183
Correlation:
  (Intr) Oprtr2
  Operator2 -0.321
  Operator3 -0.321  0.500
Standardized Within-Group Residuals:
       Min        Q1       Med        Q3       Max
-1.68303446 -0.54044333 -0.07409349  0.57924966  1.89128663
Number of Observations: 24
Number of Groups:
  Part PartbyOperator %in% Part
      4          12

> fixed.effects(lmeRandR)
(Intercept)   Operator2   Operator3
  498.625     -1.000     -0.625

> vcov(lmeRandR)
   (Intercept)    Operator2   Operator3
(Intercept)   0.8020835   -0.1649306  -0.1649306
Operator2    -0.1649306   0.3298611   0.1649306
Operator3    -0.1649306   0.1649306   0.3298611

> intervals(lmeRandR)
Approximate 95% confidence intervals

Fixed effects:

<table>
<thead>
<tr>
<th></th>
<th>lower</th>
<th>est.</th>
<th>upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>496.673675</td>
<td>498.625</td>
<td>500.576325</td>
</tr>
<tr>
<td>Operator2</td>
<td>-2.405348</td>
<td>-1.000</td>
<td>0.405348</td>
</tr>
<tr>
<td>Operator3</td>
<td>-2.030348</td>
<td>-0.625</td>
<td>0.780348</td>
</tr>
</tbody>
</table>

Random Effects:

Level: Part

<table>
<thead>
<tr>
<th></th>
<th>lower</th>
<th>est.</th>
<th>upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>sd((Intercept))</td>
<td>0.6722698</td>
<td>1.596437</td>
<td>3.791055</td>
</tr>
</tbody>
</table>

Level: PartbyOperator

<table>
<thead>
<tr>
<th></th>
<th>lower</th>
<th>est.</th>
<th>upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>sd((Intercept))</td>
<td>0.09020527</td>
<td>0.4930067</td>
<td>2.694472</td>
</tr>
</tbody>
</table>

Within-group standard error:

<table>
<thead>
<tr>
<th></th>
<th>lower</th>
<th>est.</th>
<th>upper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.5976521</td>
<td>0.9128709</td>
<td>1.3943452</td>
</tr>
</tbody>
</table>

> random.effects(lmeRandR)
Level: Part

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>2.2247074</td>
<td>-0.2301421</td>
<td>-1.1507107</td>
<td>-0.8438545</td>
</tr>
</tbody>
</table>

Level: PartbyOperator %in% Part

<table>
<thead>
<tr>
<th></th>
<th>1/1</th>
<th>1/2</th>
<th>1/3</th>
<th>2/1</th>
<th>2/2</th>
<th>2/3</th>
<th>3/1</th>
<th>3/2</th>
<th>3/3</th>
<th>4/1</th>
<th>4/2</th>
<th>4/3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.313050149</td>
<td>0.423792081</td>
<td>0.101423605</td>
<td>0.038736586</td>
<td>-0.145473971</td>
<td>0.084789226</td>
<td>0.193682932</td>
<td>0.009472374</td>
<td>-0.312896102</td>
<td>0.080630631</td>
<td>-0.287790484</td>
<td>0.126683270</td>
</tr>
</tbody>
</table>

> predict(lmeRandR,level=0:2)

<table>
<thead>
<tr>
<th></th>
<th>Part</th>
<th>PartbyOperator</th>
<th>predict.fixed</th>
<th>predict.Part</th>
<th>predict.PartbyOperator</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1/1</td>
<td>498.625</td>
<td>500.8497</td>
<td>500.5367</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1/1</td>
<td>498.625</td>
<td>500.8497</td>
<td>500.5367</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1/2</td>
<td>497.625</td>
<td>499.8497</td>
<td>500.2735</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1/2</td>
<td>497.625</td>
<td>499.8497</td>
<td>500.2735</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1/3</td>
<td>498.000</td>
<td>500.2247</td>
<td>500.3261</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1/3</td>
<td>498.000</td>
<td>500.2247</td>
<td>500.3261</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>2/1</td>
<td>498.625</td>
<td>498.3949</td>
<td>498.4336</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>2/1</td>
<td>498.625</td>
<td>498.3949</td>
<td>498.4336</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>2/2</td>
<td>497.625</td>
<td>497.3949</td>
<td>497.2494</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>2/2</td>
<td>497.625</td>
<td>497.3949</td>
<td>497.2494</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>2/3</td>
<td>498.000</td>
<td>497.7699</td>
<td>497.8546</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>2/3</td>
<td>498.000</td>
<td>497.7699</td>
<td>497.8546</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>3/1</td>
<td>498.625</td>
<td>497.4743</td>
<td>497.6680</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>3/1</td>
<td>498.625</td>
<td>497.4743</td>
<td>497.6680</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>3/2</td>
<td>497.625</td>
<td>496.4743</td>
<td>496.4838</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>----</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>16</td>
<td>3</td>
<td>3/2</td>
<td>497.625</td>
<td>496.4743</td>
<td>496.4838</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>3/3</td>
<td>498.000</td>
<td>496.8493</td>
<td>496.5364</td>
</tr>
<tr>
<td>18</td>
<td>3</td>
<td>3/3</td>
<td>498.000</td>
<td>496.8493</td>
<td>496.5364</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>4/1</td>
<td>498.625</td>
<td>497.7811</td>
<td>497.8618</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>4/1</td>
<td>498.625</td>
<td>497.7811</td>
<td>497.8618</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>4/2</td>
<td>497.625</td>
<td>496.7811</td>
<td>496.4934</td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>4/2</td>
<td>497.625</td>
<td>496.7811</td>
<td>496.4934</td>
</tr>
<tr>
<td>23</td>
<td>4</td>
<td>4/3</td>
<td>498.000</td>
<td>497.1561</td>
<td>497.2828</td>
</tr>
<tr>
<td>24</td>
<td>4</td>
<td>4/3</td>
<td>498.000</td>
<td>497.1561</td>
<td>497.2828</td>
</tr>
</tbody>
</table>