Solutions for the following questions should be returned to Moodle's quiz platform. Platform will be opened soon. The last return time is 26 March at 12.00.

1. Consider the dataset cmort.txt:

```
> data<-read.table("cmort.txt", sep="\t", dec=".", head=TRUE)
> head(data)
  t      y
1 1  97.85
2 2 104.64
3 3  94.36
4 4  98.05
5 5  95.85
6 6  95.98
> tail(data)
      t     y
503 503 73.46
504 504 79.03
505 505 76.56
506 506 78.52
507 507 89.43
508 508 85.49
```

```
Description

Average weekly cardiovascular mortality in Los Angeles County;
508 six-day smoothed averages obtained by filtering daily values
over the 10 year period 1970-1979.
```

Consider modeling the time series y_t by the linear model

$$\mathcal{M}: \; y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 \cos(2\pi t/52) + \beta_5 \sin(2\pi t/52) + \varepsilon_t,$$

where random error term ε_t is assumed to follow first order auto-regressive AR(1) process

$$\varepsilon_t = \phi \varepsilon_{t-1} + u_t,$$

with white noise u_t being independent normally distributed $u_t \sim N(0, \sigma^2)$.

(a) Calculate the (restricted) maximum likelihood estimate for the parameter ϕ.

(1 points)

(b) Consider the alternative model

$$\mathcal{M}_0: \; y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \varepsilon_t,$$

with

$$\varepsilon_t = \phi\varepsilon_{t-1} + u_t, \qquad u_t \sim N(0, \sigma^2).$$

Test the hypotheses

$$H_0: \text{Model } \mathcal{M}_0 \text{ is the true model},$$
$$H_1: \text{Model } \mathcal{M} \text{ is the true model}.$$

Select appropriate test statistic to test the above hypotheses.

(1 points)

(c) Consider predicting the value of the new observation y_f at the next time point $t_* = 509$, i.e., $y_f = y_{t_*} = y_{509}$ by the model $\mathcal{M}$. Calculate the (empirical) best linear unbiased prediction for the y_f.

(2 points)

(d) Consider predicting the value of the new observation y_f at the next time point $t_* = 509$, i.e., $y_f = y_{t_*} = y_{509}$ by the model $\mathcal{M}$. Construct 80% prediction interval for the y_f.

(2 points)

2. Consider the dataset stageforest.txt:

```
> data<-read.table("stageforest.txt", sep="\t", dec=".", header=TRUE)
> head(data)
  Tree.ID Age  Forest.ID dbhib.cm height.m
1       1  55 Clearwater   37.084 21.76272
2       1  45 Clearwater   31.496 18.71472
3       1  35 Clearwater   22.352 12.22248
4       1  25 Clearwater   17.780  8.71728
5       1  15 Clearwater   10.160  5.97408
6       2 107 Clearwater   50.800 31.51632
> tail(data)
    Tree.ID Age Forest.ID dbhib.cm height.m
537      85  68   Wallowa   36.322 29.77896
538      85  58   Wallowa   31.750 28.98648
539      85  48   Wallowa   24.892 25.72512
540      85  38   Wallowa   17.526 18.31848
541      85  28   Wallowa   11.684 11.24712
542      85  18   Wallowa    6.604  6.18744

Description

The data are stem measures from 66 trees.
The trees were selected as having been dominant throughout their
lives with no visible evidence of damage or forks.
The trees came from stands throughout the inland range of the species.

A data frame with 542 observations on the following variables.
Tree.ID - A factor uniquely identifying the tree.
Age - Age of tree
Forest.ID - The national forest in which the tree was.
dbhib.cm - Diameter (cm.) at 1.37 m (4’6?).
height.m - Height of tree (m)

The national forests are: Kaniksu, Coeur d’Alene, St. Joe, Clearwater, Nez Perce,
Clark Fork, Umatilla, Wallowa, and Payette.

Note that values are strangely in wrong order respect the age variable.
```

Denote the variables as following

$$Y = \text{height.m}, \quad X = \text{dbhib.cm}, \quad Z = \text{Forest.ID}, \quad T = \text{Age}.$$

(a) Consider the linear mixed effects model

$$\mathcal{M}_1 : \quad Y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 x_{it}^2 + b_h + \varepsilon_{it},$$

where for the tree i, the index t is related to variable $T = \mathsf{Age}$, and the index h denotes values of the variable $Z = \mathsf{Forest.ID}$. The random effects b_h, and the random error term ε_{it} are assumed to follow normal distributions $b_h \sim N(0, \sigma_z^2), \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$, respectively.
Calculate the (restricted) maximum likelihood estimate for the variance parameter σ_z^2.

(2 points)

(b) Which forest area has the largest predicted random effect b_h in the model $\mathcal{M}_1$?

i. Clark Fork
ii. Clearwater
iii. Coeur dAlene
iv. Kaniksu
v. Nez Perce
vi. Payette
vii. St. Joe
viii. Umatilla
ix. Wallowa

(1 point)

(c) Consider then the linear mixed effects model

$$\mathcal{M}_2 : \quad Y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 x_{it}^2 + b_{i0} + b_{i1} x_{it} + b_h + \varepsilon_{it},$$

where the random effects b_{i0}, b_{i1}, b_h, and the random error term ε_{it} are assumed to follow normal distributions $b_{i0} \sim N(0, \sigma_{b_0}^2), b_{i1} \sim N(0, \sigma_{b_1}^2), b_h \sim N(0, \sigma_z^2), \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$, respectively. Calculate the maximum likelihood prediction $\tilde{\mu}_{it_*}$ for the expected value μ_{it_*} when

`Tree.ID=67, Forest.ID=Wallowa, dbhib.cm=35.5.`

Note that we have to assume here that the value dbhib.cm=35.5 is recorded at known age t_*.

(2 points)

(d) Construct 80% prediction interval for the random variable Y_{it}, when

`Tree.ID=67, Forest.ID=Wallowa, dbhib.cm=35.5.`

(1 point)

3. (a) Let us consider the simple time series model

$$y_t = \beta_0 + \varepsilon_t,$$

where the error term ε_t follows the first order auto-regressive process

$$\varepsilon_t = \phi\varepsilon_{t-1} + u_t, \qquad u_t \sim N(0, \sigma^2),$$

where ϕ is assumed to be known. With the observable data $\mathbf{y} = (y_1, y_2, \ldots, y_t)'$, let us consider prediction of next observation y_{t+1}. The joint distribution of $\mathbf{y}$ and y_{t+1} is

$$\begin{pmatrix} \mathbf{y}_t \\ y_{t+1} \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{1}\beta_0 \\ \beta_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \mathbf{V} & \mathbf{w} \\ \mathbf{w}' & \frac{1}{1-\phi^2} \end{pmatrix} \right),$$

where

$$\mathbf{V} = \frac{1}{1-\phi^2} \begin{pmatrix} 1 & \phi & \phi^2 & \ldots & \phi^{t-2} & \phi^{t-1} \\ \phi & 1 & \phi & \ldots & \phi^{t-3} & \phi^{t-2} \\ & \vdots & & & & \\ \phi^{t-2} & \phi^{t-3} & \phi^{t-4} & \ldots & 1 & \phi \\ \phi^{t-1} & \phi^{t-2} & \phi^{t-3} & \ldots & \phi & 1 \end{pmatrix}$$

and

$$\mathbf{w} = \frac{1}{1-\phi^2} \begin{pmatrix} \phi^t \\ \phi^{t-1} \\ \vdots \\ \phi^2 \\ \phi \end{pmatrix}.$$

i. Show that

$$\phi\mathbf{V}\mathbf{i} = \mathbf{w}, \text{ where } \mathbf{i} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

ii. Prove that the BLUP of y_{t+1} is

$$\hat{y}_{t+1} = (1-\phi)\hat{\beta}_0 + \phi y_t,$$

where $\hat{\beta}_0 = (\mathbf{1}'\mathbf{V}^{-1}\mathbf{1})^{-1}\mathbf{1}'\mathbf{V}{-}1\mathbf{y}$.

(2 points)

(b) In question 2 c), we consider the model

$$\mathcal{M}_2: \quad Y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 x_{it}^2 + b_{i0} + b_{i1} x_{it} + b_h + \varepsilon_{it}.$$

Write in details the model $\mathcal{M}_2$ in matrix form

$$\begin{aligned} \mathbf{y}|\mathbf{b} &\sim N(\boldsymbol{\mu}_{\mathbf{b}}, \sigma^2 \mathbf{I}_n), \\ \boldsymbol{\mu}_{\mathbf{b}} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}, \\ \mathbf{b} &\sim N(\mathbf{0}, \mathbf{G}). \end{aligned}$$

That is, explain what kind of partitioned forms the matrices and vectors $\mathbf{X}, \mathbf{Z}, \mathbf{G}, \mathbf{V}, \boldsymbol{\beta}, \mathbf{b}$ are having when the model is $\mathcal{M}_2$ and the dataset is `stageforest.txt`.

(2 points)

(c) Consider the prediction of y_f in the simple linear model

$$y_i = \beta_0 + \varepsilon_i,$$

with the joint distribution being

$$\begin{pmatrix} \mathbf{y} \\ y_f \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{1}\beta_0 \\ \beta_0 \end{pmatrix}, \sigma^2 \begin{pmatrix} \mathbf{V} & \mathbf{w} \\ \mathbf{w}' & v_f \end{pmatrix} \right).$$

Let the covariance matrix of $\left(\begin{smallmatrix} \mathbf{y} \\ y_f \end{smallmatrix} \right)$ has an intraclass correlation structure

$$\sigma^2 \begin{pmatrix} \mathbf{V} & \mathbf{w} \\ \mathbf{w}' & v_f \end{pmatrix} = \sigma^2 \begin{pmatrix} (1-\rho)\mathbf{I} + \rho\mathbf{1}\mathbf{1}' & \rho\mathbf{1} \\ \rho\mathbf{1}' & 1 \end{pmatrix},$$

where $-1 < \rho < 1$ is assumed to known. Show, with help of the fundamental BLUP equations, that the sample mean

$$\hat{\beta}_0 = (\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}'\mathbf{y} = \frac{1}{n}\mathbf{1}'\mathbf{y} = \frac{y_1 + y_2 + \cdots + y_n}{n} = \bar{y}$$

is the BLUP of y_f.

(2 points)