Solutions to Ex. 2

Have you tried the questions yet? If not, I recommend you make a serious attempt at the questions before looking at these solutions. You can learn a lot by trying and failing in maths. If you read the solution first, you lose that experience.

i) $y_{ijk} = \beta_j + b_i + b_{ij} + \epsilon_{ijk}$

(1) i, j, k notation
(1) β_j
(1) b_i & b_{ij}

i = worker
j = machine
k = measurement (1)

(1) β_j is the fixed effect for machine j
(1) b_i is the random effect for worker i
b_{ij} is the random effect for machine j for worker i

y_{ijk} is the productivity score
ϵ_{ijk} is the random error

$b_i \sim N(0, \sigma_b^2)$
$b_{ij} \sim N(0, \sigma_c^2)$ (1)
$\epsilon_{ijk} \sim N(0, \sigma^2)$

/7

ii) The workers are randomly chosen so are not of direct interest (1). Interest is really in the productivity of the machines in the population of all workers (1), so machines are the fixed effect of interest (1)

/3

iii) For b_i, b_{ij} and ϵ_{ijk}, residuals relating to these quantities are level 0, 1 & 2 residuals respectively. (2)
Should check qqplots (1) for all three sets of residuals

/3

1.

iv) The null is that the nested random effects (1) for machine in each worker are zero. (1) This null is clearly rejected, there terms look like being needed. (1)

/3

v) $$\text{Var}\begin{pmatrix} b_1 \\ b_3 \\ b_5 \end{pmatrix} = \begin{pmatrix} \psi_1^2 & \psi_2^2 & \psi_2^2 \\ \psi_2^2 & \psi_1^2 & \psi_2^2 \\ \psi_2^2 & \psi_2^2 & \psi_1^2 \end{pmatrix}$$

(1) identical diagonals & identical off-diagonals

(1) 3×3 matrix

For mach2.lme

$$\text{Var}\begin{pmatrix} b_1 \\ b_3 \\ b_5 \end{pmatrix} = \begin{pmatrix} \psi_1^2 & 0 & 0 \\ 0 & \psi_1^2 & 0 \\ 0 & 0 & \psi_1^2 \end{pmatrix}$$

(1) identical diags

(1) off diagonals are 0.

ie zero covariance between worker random effects.

/4

2. Mechanism (i) is MAR and ignorable, (ii) is MCAR and ignorable, (iii) is NMAR and non-ignorable.

3. We could try to detect between MAR and MCAR mechanisms by building a regression model to predict missingness from the observed data. For example, if we have two covariates y_1 and y_2, with some missing values in y_2, we could fit a logistic regression model of the form

$$\text{logit}(\mathbb{P}(M_{i2} = 1|y_{i1}, \beta)) = \beta_0 + \beta_1 y_{i1} + e_i$$

   This is a MAR mechanism if $\beta_1 \neq 0$ and a MCAR mechanism if $\beta_1 = 0$. Thus we would test the hypothesis

$$H_0 : \beta_1 = 0 \text{ vs } \beta_1 \neq 0$$

   to detect between MCAR and MAR.

   It is generally impossible to detect whether a missing data mechanism is MAR or NMAR as the information we would need to determine this is generally missing.