

1.(a) A model for the overall quality score data is as the following:

$$Y_{ijk} = \mu + \alpha_i + u_{ij} + e_{ijk},$$

where

(1) Y_{ijk} is the overall quality score for the k^{th} cut of the j^{th} cooler at the i^{th} temperature level;

(i.e., $i=1,2,3$; $j=1,2,3,4$; $k=1,2$.)

(2) α_i is the fixed effect of i^{th} temperature level;

(3) $u_{ij} \stackrel{iid}{\sim} N(0, \sigma_u^2)$ is the random effect of the j^{th} cooler at the i^{th} temperature level;

(4) $e_{ijk} \stackrel{iid}{\sim} N(0, \sigma_e^2)$ is the random error.

1(b)	<u>Source</u>	<u>D.f.</u>	<u>Sum of Squares</u>
	Temperature	$3-1=2$	$\sum_{i=1}^3 \sum_{j=1}^4 \sum_{k=1}^2 (\bar{y}_{i..} - \bar{y}_{...})^2$
	Cooler (Temp.)	$3 \cdot (4-1) = 9$	$\sum_{i=1}^3 \sum_{j=1}^4 \sum_{k=1}^2 (\bar{y}_{ij.} - \bar{y}_{i..})^2$
	Error	$23 - 2 - 9 = 12$	$\sum_{i=1}^3 \sum_{j=1}^4 \sum_{k=1}^2 (y_{ijk} - \bar{y}_{ij.})^2$
	<u>Corrected Total</u>	$3 \times 4 \times 2 - 1 = 23$	$\sum_{i=1}^3 \sum_{j=1}^4 \sum_{k=1}^2 (y_{ijk} - \bar{y}_{...})^2$

1.(c) To test for the temperature main effect, i.e.,
to test $H_0: \alpha_1 = \alpha_2 = \alpha_3$, I would compute
the F-statistic:

$$F = \frac{MS_{\text{Temperature}}}{MS_{\text{Cooler(Temp)}}$$

which has a central F dist'n with (2, 9) d.f.
under H_0 .

2.(a) Recall the model from Problem 2 of HW7:

④

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + u_{ik} + e_{ijk},$$

where $i=1,2$; $j=1,2$; $k=1,2,3,4$;

u_{ik} is the random effect of litters, $u_{ik} \text{ iid } N(0, \sigma_u^2)$;

$e_{ijk} \text{ iid } N(0, \sigma_e^2)$

According to the simulation code in R, we have

$$\mu = 4; \alpha_2 = 1; \beta_2 = -1.5; \gamma_{22} = 0.5.$$

Then, the true means for each combination of diet and drug are calculated in the following table:

	drug 1	drug 2
diet 1	$\mu = 4$	$\mu + \beta_2 = 4 - 1.5 = 2.5$
diet 2	$\mu + \alpha_2 = 5$	$\mu + \alpha_2 + \beta_2 + \gamma_{22} = 4 + 1 - 1.5 + 0.5 = 4$

Alternatively, we could examine `X%*%beta` directly to get the above results..

2.(b) The true values of the variance components are:

$$\sigma_u^2 = 0.8^2 = 0.64$$

$$\sigma_e^2 = 0.3^2 = 0.09$$

⊕

2.(c) An ANOVA table for the simulated data is

(6)

<u>Source</u>	<u>D.f.</u>	<u>Sum Sq.</u>	<u>Mean Sq.</u>
Diet	1	7.3577	7.3577
Litter(Diet)	6	10.5457	1.7576
Drug	1	4.7198	4.7198
Diet x Drug	1	0.2678	0.2678
Error	6	0.4148	0.0691

i. By Problem 2 of HW7,

$$E(MS_{Error}) = \sigma_e^2$$

$$E(MS_{Litter(Diet)}) = M\sigma_u^2 + \sigma_e^2 = 2\sigma_u^2 + \sigma_e^2 \quad \text{because } M=2.$$

\therefore Unbiased estimators of σ_e^2 and σ_u^2 are:

$$\hat{\sigma}_e^2 = MS_{Error} = 0.0691$$

$$\hat{\sigma}_u^2 = \frac{MS_{Litter(Diet)} - MS_{Error}}{2} = \frac{1.7576 - 0.0691}{2} = 0.84425$$

2.(c) ii. By part (h) of HW7, a 95% confidence interval for the difference between diet 1 drug 1 and diet 2 drug 1 is,

$$(\bar{y}_{11.} - \bar{y}_{21.} - t_{d, 0.975} \sqrt{\frac{2(\hat{\sigma}_u^2 + \hat{\sigma}_e^2)}{L}}, \quad \bar{y}_{11.} - \bar{y}_{21.} + t_{d, 0.975} \sqrt{\frac{2(\hat{\sigma}_u^2 + \hat{\sigma}_e^2)}{L}}),$$

where $L=4$; $\hat{\sigma}_u^2$ and $\hat{\sigma}_e^2$ are calculated in 2(c) i. ;

$$d = \frac{\left\{ \frac{2}{L} \left[\frac{1}{M} (MS_{\text{Litter(Diet)}} - MS_{\text{Error}}) + MS_{\text{Error}} \right] \right\}^2}{\frac{\left[\frac{2}{L} \cdot \frac{1}{M} (MS_{\text{Litter(Diet)}}) \right]^2}{D(L-1)} + \frac{\left[\frac{2}{L} \left(1 - \frac{1}{M}\right) MS_{\text{Error}} \right]^2}{D(M-1)(L-1)}}$$

where $D=M=2$; $L=4$.

$$\therefore d \approx 6.471$$

Thus, a 95% C.I. for the difference between diet 1 drug 1 mean and the diet 2 drug 1 mean is

$$(-2.722, 0.527)$$

2.(c) iii. By Part (d) of Problem 2 of HW7, to test for the diet main effects, we use

$$F = \frac{MS_{\text{Diet}}}{MS_{\text{Litter(Diet)}}} = 4.18616.$$

Under H_0 , $F \sim F(D-1, D(L-1))$, which is $F(1, 6)$.

Thus, the p -value is $\Pr(F(1,6) \geq 4.18616) \doteq 0.08672$.

Thus, we do not reject the null hypothesis, i.e., there is no significant evidence that there exist diet main effects.

2.(d). Using the following code, we can the standard deviations of the random effects :

```
> lme (y ~ diet + drug + diet; drug, method = "ML", random = ~1 | Litter)
```

Then, we get

$$\hat{\sigma}_e = 0.2277025$$

$$\hat{\sigma}_u = 0.7957268$$

$$\therefore \hat{\sigma}_{e,MLE}^2 = \hat{\sigma}_e^2 = 0.2277025^2$$

$$\hat{\sigma}_{u,MLE}^2 = \hat{\sigma}_u^2 = 0.7957268^2$$

$$\approx 0.05185$$

$$\approx 0.63318$$

2(e). The default method of lme function is "REML".

Therefore, the following code in the R code file provided with this problem can give us REML estimates of Std. Dev. of the variance components.

```
> O = lme(y ~ diet + drug + diet:drug, random = ~|Litter)
> O
```

Then, $\hat{\sigma}_{u,REML}^2 = 0.9188262^2 \doteq 0.8442$

$$\hat{\sigma}_{e,REML}^2 = 0.2629282^2 \doteq 0.0691,$$

which are the same as the results in 2.(c) i.

2.(f) The results are exactly the same with those in 2.(c) iii).

②

3. Write down the model in the form of $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$,

$$\text{where } X = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix}, \underline{\beta} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \underline{\varepsilon} \sim N(\underline{0}, \Sigma),$$

$$\text{where } \Sigma = \sigma^2 \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & \frac{1}{2} & 1 \end{bmatrix}.$$

Clearly, $\text{rank}(X) = 2$, $n = 3$. Thus, $n - \text{rank}(X) = 3 - 2 = 1$.

Thus, we only need to find one \underline{a} such that $\underline{a}'X = \underline{0}'$.

It is easy to see that $\underline{a} = [1, -1, 0]'$ satisfies $\underline{a}'X = \underline{0}'$.

Then, $W = \underline{a}'\underline{Y} = Y_1 - Y_2$. Use W as data to find MLE of σ^2

based on W . In this way, we will get the REML of σ^2 .

Note that $W = Y_1 - Y_2 \sim N(0, \sigma^2)$ because $E(W) = EY_1 - EY_2 = \mu_1 - \mu_1 = 0$

$$\text{and } \text{Var}(Y_1 - Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) - 2\text{Cov}(Y_1, Y_2) = \sigma^2 + \sigma^2 - 2 \cdot \frac{\sigma^2}{2} = \sigma^2.$$

Then, the log likelihood function is

$$L(\sigma^2 | W) = -\frac{1}{2} \log |\sigma^2| - \frac{1}{2} \frac{1}{\sigma^2} (Y_1 - Y_2)^2 - \frac{1}{2} \log(2\pi)$$

$$\text{then, } \left. \frac{dL}{d\sigma^2} \right|_{\hat{\sigma}^2} = -\frac{1}{2} \frac{1}{\sigma^2} + \frac{1}{2} (Y_1 - Y_2)^2 \cdot \frac{1}{\sigma^4} - 0 \Big|_{\hat{\sigma}^2} = 0$$

$$\Rightarrow \hat{\sigma}^2 = (Y_1 - Y_2)^2,$$

i.e., the REML estimate of σ^2 is $(Y_1 - Y_2)^2$.