Solution to additional exercise 10.5

The data stem from a greenhouse experiment to compare 2 varieties and 4 watering treatments with respect to growth of hibiscus plants. The outcome measured was the total root weight for a "parcel" of plants subject to the same conditions. The watering was applied to troughs, with 4 replicate troughs for each treatment. Within each trough, 2 parcels of different varieties were grown. Thus, the design is balanced with respect to both treatment factors.

The design has split-plot character because the watering treatments were applied to entire troughs, whereas the two varieties were grown within each trough. The troughs should be considered as "whole plots", and the "parcel" of plants of each variety within the troughs as "split plots" (or "sub plots"). There are no blocks in the design; on the contrary, the 4 troughs per watering conditions are replications. The corresponding statistical model may be written, denoting by y_i the total root weight in the ith parcel,

$$y_i = \mu + \alpha_{\text{water}(i)} + C_{\text{trough}(i)} + \beta_{\text{variety}(i)} + (\alpha\beta)_{\text{water}\times\text{variety}(i)} + \varepsilon_i, \quad i = 1, \ldots, 32,$$

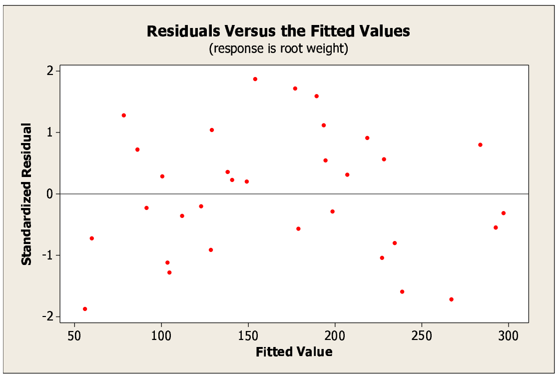
where the (split-plot) errors ε_i are assumed to be i.i.d. and $\sim$ N($0,\sigma^2$), and the (whole-plot) random effects C_{trough} are assumed to be i.i.d. as well and $\sim$ N($0,\sigma_C^2$). Here σ^2 is the subplot variance (variance within troughs), and σ_C^2 is the whole plot variance (variance between troughs).

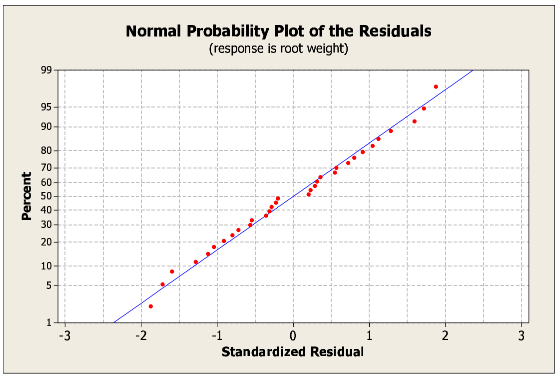
Alternatively, using a multi-index notation (ijk), where $i \sim$ watering type (whole-plot factor level), $j \sim$ variety (split-plot factor level), and $k \sim$ replication of trough (within watering type), the model can be written as

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + AC_{ik} + \varepsilon_{ijk},$$

where the (split-plot) errors ε_{ijk} are assumed to be i.i.d. and $\sim$ N($0,\sigma^2$), and the (whole-plot) random effects AC_{ik} are assumed to be i.i.d. as well and $\sim$ N($0,\sigma_{AC}^2$).

The data are balanced so the model may be analysed using the ANOVA-based method for mixed models. In Minitab we use the General Linear Model menu with random trough effects (nested within watering treatment). The ANOVA table is shown below, but we first carry out the model checking procedures. For split-plot residuals, a normal plot and a residual plot are directly available from the menu, and these are shown below.

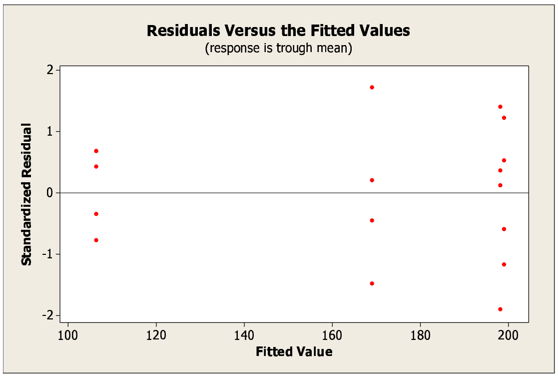


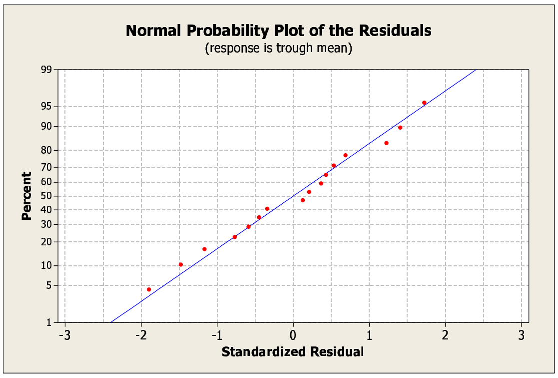


These plots look very nice. The residual plot shows an apparently random scatter of points, and there are no extreme standardized residuals. The normal plot looks almost completely straight; the

Anderson-Darling test for normality gives a P-value of 0.95.

The whole-plot residuals are checked in an analysis of trough means (because the random effects are for troughs). The only factor entering the model for trough means is watering type, so the analysis is a one-way ANOVA. We show only the residual plots from that analysis (the P-value for effect of watering type is the same as in the full split-plot model).





These plots also look nice. The normal plot is almost straight; the Anderson-Darling test for normality has $P = 0.96$. In the residual plot all standardized residuals are numerically quite small (less than 2), and the only obvious model deviation is that one of the watering treatments (I) has somewhat less variation than the other treatments. The lowest standard deviation (among trough means) is about half of the other values, but with only four replications in each group this is far from statistical significance, and therefore nothing to worry about.

The ANOVA table (below) shows the interaction between watering and varieties to be non-significant (P=0.12) but maybe yet of a magnitude to be of some interest. We explore the interaction below. The main effect of varieties is certainly of greater magnitude than the interaction, and clearly significant. The main effect of watering treatments is just above the significance level (borderline significant), but it is still of interest to further describe the weak different treatments effects, both overall and within the interaction.

Factor	Type	Levels	Values
water	fixed	4	I, II, III, IV
variety	fixed	2	H, R
trough(water)	random	16	1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4, 1, 2, 3, 4

Analysis of Variance for root weight, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
water	3	45336.8	45336.8	15112.3	3.03	0.071
variety	1	34716.1	34716.1	34716.1	35.29	0.000
water*variety	3	6921.1	6921.1	2307.0	2.35	0.124
trough(water)	12	59873.8	59873.8	4989.5	5.07	0.004
Error	12	11805.8	11805.8	983.8		
Total	31	158653.5				

S = 31.3658   R-Sq = 92.56%   R-Sq(adj) = 80.78%

There is a substantial (and clearly significant) variation between troughs, and the estimated variance components are

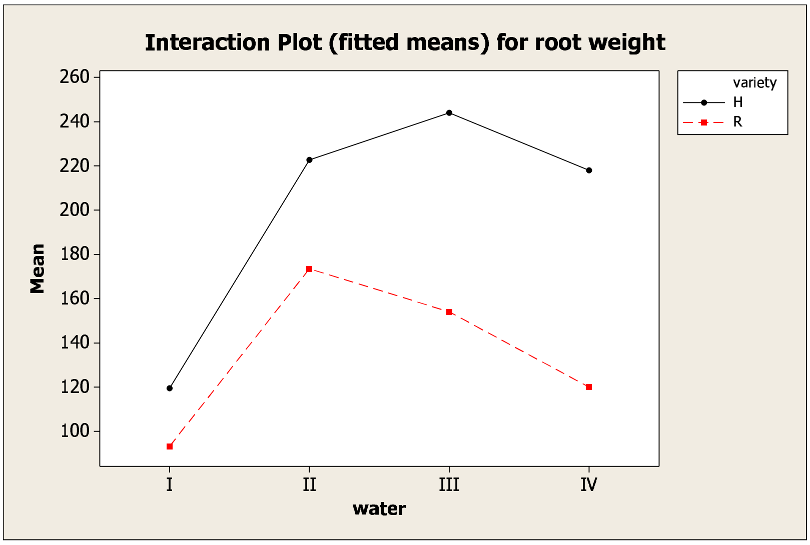
$$\begin{aligned} \hat{\sigma}^2 &= 983.8 \quad \text{(subplot variance)}, \\ \hat{\sigma}_C^2 &= 2002.8 \quad \text{(whole plot variance)}, \end{aligned}$$

where the whole plot variance estimate is computed from the table as (MS(Trough)–MSE)/2. Thus, the total unexplained variance is estimated at

$$\hat{\sigma}_{\text{tot}}^2 = \hat{\sigma}^2 + \hat{\sigma}_C^2 = 2986.8.$$

The proportion of variance at the trough level is $2002.8/2986.8 = 0.67$ (or 67%), and this is also the (intra-class) correlation coefficient for two observations within the same trough.

First, we describe how the statistical inference would proceed if it was decided to take into account the interaction between varieties and watering types. The interaction plot (below) shows curves for the two varieties that are not parallel, and where differences between varieties differ considerably across watering types. In detail, the curves are almost parallel for watering types I and II, and again for watering types III and IV, but not across all watering types. As the design is balanced, all inferences would be based on the (least squares) means displayed in the table below the plot.



Mean / Variety	I	II	III	IV	Total
	Watering type				
H	119.50	222.75	244.00	218.00	201.06
R	93.25	173.50	154.00	120.00	135.19
Total	106.38	198.13	199.00	169.00	168.13

The standard error of the variety × watering means in the table is estimated as $\sqrt{\hat{\sigma}_{\text{tot}}^2/4} = 27.3$, where the denominator (4) follows from the fact that every mean is over 4 observations. As detailed in Note 5.1 in the notes on linear mixed models, there is no exact χ^2-distribution for this estimate, but an approximation may be worked out from the formulae,

$$\begin{aligned} k &= \text{MSAC}/(\text{MSAC} + (2-1)\text{MSE}) = 4989.5/(4989.5 + 983.8) = 0.8353, \\ 1/\text{df} &= k^2/\text{DFAC} + (1-k)^2/\text{DFE} = 0.8353^2/12 + (1-0.8353)^2/12 = 0.0604, \\ \text{df} &= 1/0.604 = 16.6 \approx 17. \end{aligned}$$

As $t(17, .975) = 2.11$, we could compute (approximate) 95% confidence intervals for the variety $\times$ watering combinations by the formula: estimate $\pm\, 2.11 \cdot 27.3 = 57.6$. The most interesting comparisons in the plot/table are probably within varieties, in order to determine the optimal watering type for each variety. These comparisons are *across* whole-plots so they also involve $\hat{\sigma}^2_{\text{tot}}$; an (approximate) LSD-value is $2.11\sqrt{2\hat{\sigma}^2_{\text{tot}}/4} = 81.5$. LSD comparisons therefore lead us to conclude that there no significant differences between watering types for variety R (although I and II are close), whereas for variety H only watering type I differs from the three other types. Comparisons within whole-plot factor levels, i.e. between varieties for each watering type, use the subplot variation: $\text{LSD}_{.95} = t(12, .975)\sqrt{2\,\text{MSE}/4} = 48.3$. Here, the two varieties cannot be considered significantly different under watering treatment I and only barely under treatment II. Recall that the LSD method does not adjust for multiple comparisons, and corresponds to individual type I errors (of 5%), and that the simultaneous error is larger.

Finally, we outline the analysis after the ANOVA table under the assumption that the interaction is considered as non-significant and ignored. Comparisons between varieties would use the subplot error, but in our case there are only two varieties, and no further discrimination is necessary. The means show that variety H has higher root weights than variety R. Standard errors for these means would use $\hat{\sigma}^2_{\text{tot}}$, and are therefore computed as $\hat{\sigma}_{\text{tot}}/\sqrt{16} = 13.7$ (because each mean involves 16 observations). For the whole-plot factor, the entire analysis can be carried out at the whole-plot level. Therefore, it is easy to get the correct standard errors and comparisons (both in Stata and Minitab). If we want to work directly from the full ANOVA table, we use the whole-plot mean square (*not* the variance component) for inference. For example, the standard errors of the means is $\sqrt{\text{MSAC}/8} = 25.0$ (again, each mean involves 8 observations), and $\text{LSD}_{.95} = t(12, .975)\sqrt{2\,\text{MSAC}/8} = 77.0$. The four water treatment means are 106.4, 198.1, 199.0 and 169.0, so that only treatment I would seem to differ (weakly) from treatments II and III. Note how substantially different conclusions result from ignoring the interaction compared to taking it into account…