Additional Exercise 9.1

Data: Measurements of peak expiratory flow (PEF) of 13 children 8 hours after treatment with two different inhalation medicines. We may either use a single-index or multi-index notation:

$$\begin{aligned} y_i &= \text{PEF for treatment episode } i,\ i = 1, \ldots, 26, \quad \text{or} \\ y_{ijk} &= \text{PEF for child } j \text{ after treatment } i \text{ in period } k,\ i = \text{F,S};\ j = 1, \ldots, 13;\ k = 1, 2. \end{aligned}$$

(a) Fixed effects modelling.

The design is a AB/BA cross-over design with 7 children for treatment sequence FS and 6 children for sequence SF. If no carry-over effect is assumed, the statistical model should include only treatment, period and subject effects:

$$\begin{aligned} y_i &= \mu + \alpha_{\text{tx}(i)} + \beta_{\text{period}(i)} + \gamma_{\text{child}(i)} + \varepsilon_i, \quad \text{or} \\ y_{ijk} &= \mu + \alpha_i + \beta_k + \gamma_j + \varepsilon_{ijk}, \end{aligned}$$

depending on the chosen notation. The errors are assumed independent and distributed as $\sim \text{N}(0, \sigma^2)$. We first show the results from direct analysis of this model.

General Linear Model: pef versus id, period, tx

Method

Factor coding (1, 0)

Factor Information

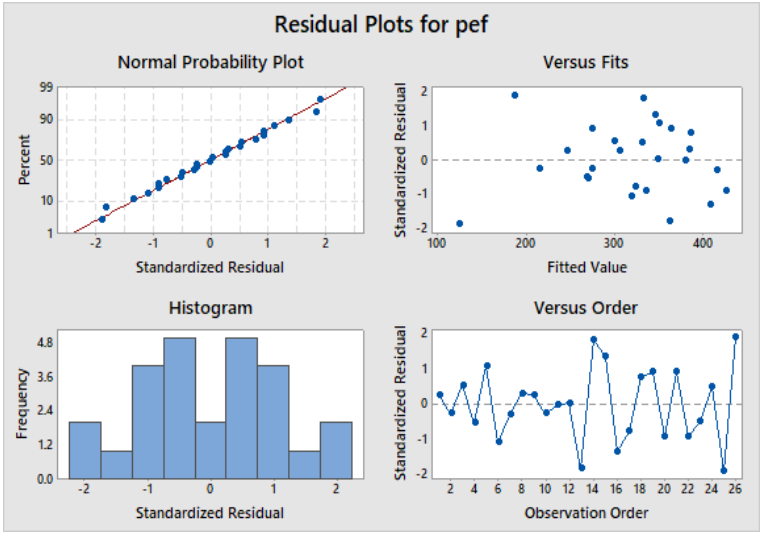
Factor	Type	Levels	Values
id	Fixed	13	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14
period	Fixed	2	1, 2
tx	Fixed	2	F, S

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
id	12	115213	83.19%	115213	9601.1	12.79	0.000
period	1	985	0.71%	1632	1632.1	2.17	0.168
tx	1	14036	10.14%	14036	14035.9	18.70	0.001
Error	11	8254	5.96%	8254	750.4		
Total	25	138488	100.00%				

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)
27.3935	94.04%	86.45%	46516.8	66.41%



The normal plot looks fine, but the residual plot shows some tendency of "inverse cone"-shape (decreasing variance with increasing values). This is essentially due to one pair of large residuals (note that all residuals come in pairs with positive and negative signs but the same numerical value) at the lower end of the predicted values. No residuals are suspiciously large. Box-Cox analysis in Minitab (not shown) or Stata gives an optimal power of 1.82 and confidence intervals that do not include 1, and thus provides evidence in favour of a transformation. However, the tests of homoscedasticity (in Stata) are non-significant. Because power transformations with powers > 1 are somewhat unusual and its justification in the present case is not intuitively obvious, we decide to work with the untransformed data.

The ANOVA table shows a clear subject effect, no period effect and a clearly significant treatment effect ($P < 0.0005$). Note that sequential and adjusted sum of squares are not the same, because the design is not balanced in the sequences (due to child no. 8 dropping out). The (least squares) means and pairwise comparisons below show that the new drug, Formoterol, on the average increases the PEF by 46.6 units with the corresponding confidence interval well away from zero.

Means

Term	Fitted Mean	SE Mean
period		
1	310.52	7.61
2	326.41	7.61
tx		
F	341.77	7.61
S	295.16	7.61

Fisher Pairwise Comparisons: tx

Grouping Information Using Fisher LSD Method and 95% Confidence

tx	N	Mean	Grouping
F	13	341.765	A
S	13	295.158	B

Means that do not share a letter are significantly different.

Fisher Individual Tests for Differences of Means

Difference of tx Levels	Difference of Means	SE of Difference	Individual 95% CI	T-Value	P-Value
S - F	-46.6	10.8	(-70.3, -22.9)	-4.32	0.001

Simultaneous confidence level = 95.00%

We next show how to obtain the same inference for the treatment from a two-sample analysis of differences. The Minitab commands to analyse the period differences 1-2 are given below.

```
MTB > Unstack ('id' 'order' 'pef');
SUBC>   Subscripts 'period';
SUBC>   After;
SUBC>   VarNames.
MTB > Name C12 'diffper'
MTB > Let 'diffper' = 'pef_1'-'pef_2'
MTB > TwoT 'diffper' 'order_1';
SUBC>   Confidence 95.0;
SUBC>   Test 0.0;
SUBC>   Alternative 0;
SUBC>   Pooled.
MTB >
```

Two-Sample T-Test and CI: diffper, order_1

Method

μ_1: mean of diffper when order_1 = FS
μ_2: mean of diffper when order_1 = SF
Difference: $\mu_1 - \mu_2$

Equal variances are assumed for this analysis.

Descriptive Statistics: diffper

order_1	N	Mean	StDev	SE Mean
FS	7	30.7	33.0	12
SF	6	-62.5	44.7	18

Estimation for Difference

Difference	Pooled StDev	95% CI for Difference
93.2	38.7	(45.8, 140.7)

Test

Null hypothesis H_0: $\mu_1 - \mu_2 = 0$
Alternative hypothesis H_1: $\mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
4.32	11	0.001

The two-sample t-test with equal variances does indeed reproduce the strong significance from above; note that $t^2 = 4.32^2 = 18.7 = F$ (for treatment). It is also seen that the estimated difference between the sequences is twice the estimated difference between the treatments.

(b) Random effects modelling.

The model with subject random effects can be run in the `General Linear Model` menu in Minitab (by selecting random effects for the factor `id`) as long as the interaction period∗tx is not included (it will simply be dropped if we try to include it). We show the added output from the random effects model; note that both the ANOVA table and the residuals are completely unchanged.

Expected Mean Squares, using Adjusted SS

	Source	Expected Mean Square for Each Term
1	id	(4) + 2.0000 (1)
2	period	(4) + Q[2]
3	tx	(4) + Q[3]
4	Error	(4)

Variance Components, using Adjusted SS

Source	Variance	% of Total	StDev	% of Total
id	4425.36	85.50%	66.5234	92.47%
Error	750.406	14.50%	27.3935	38.08%
Total	5175.76		71.9428	

The estimated variance component for subjects is very large: 4425 compared to 750 for the within-subject variation, corresponding to 85% of the unexplained variation. This shows that the cross-over design is much more efficient than an ordinary completely randomised design would have been.

We finally show output from running both this model and the extended model with the period∗tx interaction in the `Mixed Effects Model` menu.

Mixed Effects Model: pef versus id, period, tx

Method

Variance estimation	Restricted maximum likelihood
DF for fixed effects	Kenward-Roger

Factor Information

Factor	Type	Levels	Values
id	Random	13	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14
period	Fixed	2	1, 2
tx	Fixed	2	F, S

Variance Components

Source	Var	% of Total	SE Var	Z-Value	P-Value
id	4425.357975	85.50%	1966.340090	2.250556	0.012
Error	750.405844	14.50%	319.974127	2.345208	0.010
Total	5175.763820				

-2 Log likelihood = 257.901906

Tests of Fixed Effects

Term	DF Num	DF Den	F-Value	P-Value
period	1.00	11.00	2.17	0.168
tx	1.00	11.00	18.70	0.001

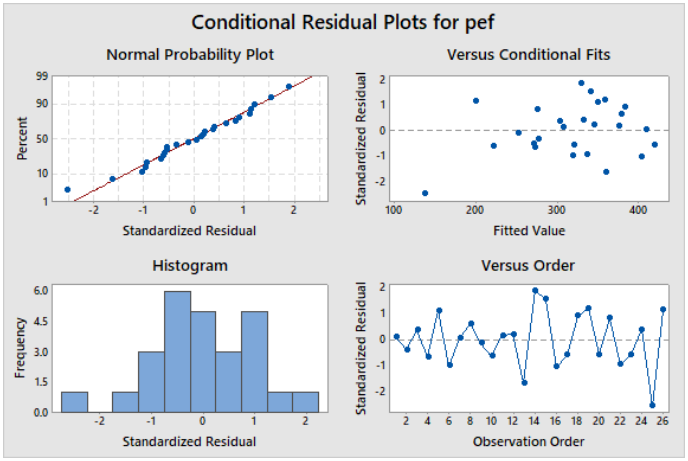
Model Summary

S	R-sq	R-sq(adj)
27.3935	93.53%	92.97%

Coefficients

Term	Coef	SE Coef	DF	T-Value	P-Value
Constant	318.461538	19.216501	12.00	16.572296	0.000
period					
1	-7.946429	5.388280	11.00	-1.474762	0.168
tx					
F	23.303571	5.388280	11.00	4.324863	0.001

Conditional Residual Plots for pef



Mixed Effects Model: pef versus id, period, tx

Method

Variance estimation	Restricted maximum likelihood
DF for fixed effects	Kenward-Roger

Factor Information

Factor	Type	Levels	Values
id	Random	13	1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14
period	Fixed	2	1, 2
tx	Fixed	2	F, S

Variance Components

Source	Var	% of Total	SE Var	Z-Value	P-Value
id	4846.536797	86.59%	2232.297753	2.171098	0.015
Error	750.405844	13.41%	319.974127	2.345208	0.010
Total	5596.942641				

-2 Log likelihood = 250.071632

Tests of Fixed Effects

Term	DF Num	DF Den	F-Value	P-Value
period	1.00	11.00	2.17	0.168
tx	1.00	11.00	18.70	0.001
period*tx	1.00	11.00	0.03	0.861

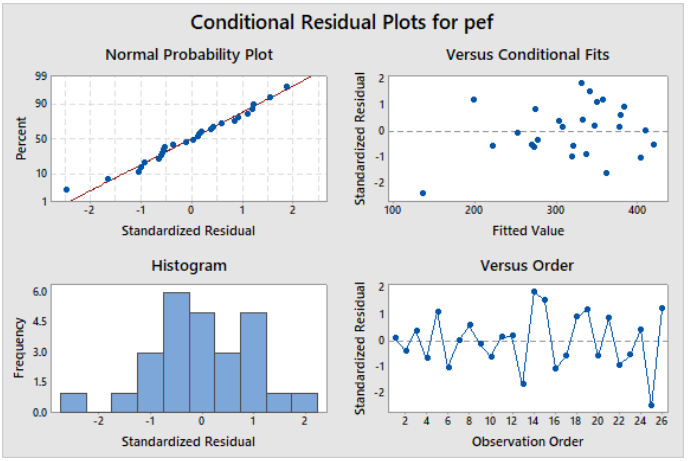
Model Summary

S	R-sq	R-sq(adj)
27.3935	93.61%	92.74%

Coefficients

Term	Coef	SE Coef	DF	T-Value	P-Value
Constant	318.184524	20.101323	11.00	15.829034	0.000
period					
1	-7.946429	5.388280	11.00	-1.474762	0.168
tx					
F	23.303571	5.388280	11.00	4.324863	0.001
period*tx					
1 F	3.601190	20.101323	11.00	0.179152	0.861

Conditional Residual Plots for pef



When fitted with likelihood-based (REML) estimation, the random effects model without the interaction gives similar but not entirely identical results to the ANOVA-based analysis. The variance components are identical (as is true for all "nice" designs, when the likelihood-based analysis uses

REML estimation). The tests for the fixed effects are also identifical, due to the use of the Kenward-Roger estimation method (the Stata default z-test and χ^2-tests give somewhat smaller P-values). However, the residuals are not quite the same; the Minitab term "Conditional Residual" means that the predicted random effects are included in the fitted values (this is probably most natural, and corresponds most closely to the ANOVA-method). The display of the estimates shows the estimated treatment difference between S and F to be the same (in the parametrization used, the omitted coefficient is not zero but defined by the sum of parameters being equal to zero).

The treatment by period interaction shows absolutely no significance. It is somewhat counterintuitive that the variance component for the subjects *increases* in the larger model; this can happen when added effects are non-significant. In the fixed effects parametrization in Minitab, the estimates and inference for the main effects are essentially unchanged, and they will in any case still be valid in the presence of the interaction; note that this is *not* the case in parametrizations with a reference category! The standardized residuals are not entirely identical but very similar.

If there had been a substantial treatment by period interaction, it would mean that the difference between treatments depends on the period. Such an effect would most likely be a carry-over effect (although there could also be genuinely different treatment effects in periods). For example, treatment A differs in periods 1 and 2 by being either the first treatment applied (AB) or by being the follow-up treatment after treatment B was applied (BA). In the simple AB/BA-design we cannot from the data alone distinguish between carry-over effects and a genuine treatment by period interaction. With a totally non-significant interaction, we can however conclude that no carry-over effects seem to be present in these data.