STAT 8200 — Design of Experiments for Research Workers
Lab 7 – Due: Monday, Oct. 21

This lab is intended to highlight several ideas in the analysis of two-way layouts:

1. For unbalanced data, there are three distinct types of sums of squares whose values are not the same. Type III Sums of Squares are typically the ones to base inference on.

2. For understanding interaction, and for determining whether or not interaction undermines the appropriateness of main effect comparisons, profile plots are useful. In addition, it is often helpful to produce profile plots in both of the possible forms:

   (i) plot separate profiles for each level of factor A, with the levels of factor B running along the horizontal axis; and

   (ii) plot separate profiles for each level of factor B, with the levels of factor A running along the horizontal axis.

   In particular, profile plots of type (i) can be helpful for understanding main effects of factor A, and whether or not such comparisons are meaningful; and profile plots of type (ii) can be helpful for understanding main effects of factor B, and whether or not such comparisons are meaningful.

3. Model diagnostics are just as important in the two-way layout (and other designs) as they were in the one-way layout. In particular, it may be necessary to transform the response.

4. Transformations can affect whether or not interactions are present. Sometimes (but, as it turns out, not in this example), the data interact on the original scale, but do not on the transformed scale (or vice versa).

5. Contrasts in the joint means are not always interaction contrasts. Sometimes, it is interesting to make such comparisons. In particular, in the presence of interaction, we may want to make comparisons across the levels of factor A (B) separately within each of the levels of factor B (A).

Example/Exercise:

An experiment was conducted to study the effects of bleach concentration (factor A) and type of stain (factor B) on the speed of stain removal from a piece of cloth. The bleach concentrations chosen for the experiment were 3, 5, and 7 teaspoonfuls of bleach per cup of water, and the types of stains were blue ink, jam, and tomato sauce. All 3×3 combinations of the levels of the two factors were observed, for a total of 9 treatments. The experimenter planned on a balanced design with $n = 5$ replicates per treatment, but there were some problems during data collection that led to the loss of data from some replicates in some treatments. The result was an unbalanced two-way layout, the data for which are displayed below:

Amount of Bleach	Stain Type	Time until Stain Removal (seconds)				
3	ink	3600	3920	3340	3173	.
3	jam	495	236	515	573	555
3	tomato	733	525	793	1026	510
5	ink	2029	2271	2156	2493	2805
5	jam	428	432	335	288	.
5	tomato	880	759	1138	780	1625
7	ink	3660	4105	4545	3569	3342
7	jam	410	225	437	350	140
7	tomato	347	584	781	.	.

Please transfer the file bleach.sas from the course web site to your USB drive, run it, and examine both the program and its output. Be sure to specify an appropriate path for the pdf output file (that is, change "mypath" on line 8 of bleach.sas to a path on your USB drive).

We consider points 1–5 listed above one at a time:

1. Unlike the weight gain in rats example, the design here is unbalanced. If we denote the number of replicates in the i, j^{th} treatment as n_{ij}, then here $n_{33} = 3$, $n_{11} = n_{22} = 4$ and $n_{12} = n_{13} = n_{21} = n_{23} = n_{31} = n_{32} = 5$. In unbalanced two-way layouts, the three types of sums of squares given by SAS (actually, there are four of them, but we're not going to talk about type IV SS's in this lab), are not all equal to one another. Each type of SS's can be requested on the MODEL statement in PROC GLM by the option / SSx; where $x = 1, 2$ or 3. In bleach.sas notice that I request all three, with the option /SS1 SS2 SS3; In bleach.pdf (p.2) we see that indeed, the three types of SS's are not all the same. The most important thing is to know which to use. For most applications, Type III SS's are most appropriate.

- Notice that the F tests corresponding to the type III SS's (p.2 of bleach.pdf) indicate that the interaction is significant ($p < .0001$) and both main effects are significant ($p = .0075$ for amtblch, and $p < .0001$ for staintyp). Because the interaction is significant, we need to examine the profile plot before making conclusions.

2. When ODS graphics are switched on, PROC GLM will automatically produce a profile plot when the fitted model is a two-way anova model. Although the two types of profile plots (i) and (ii) described above are possible, it only produces one of them, and chooses which one to produce based on the order of the factors on the CLASS statement. Specifically, it produces a plot where the levels of the first factor on the CLASS statement run along the horizontal axis, and a distinct profile (connected line segments) is plotted for each level of the second factor on the CLASS statement. Therefore, if you want to see both profile plots using ODS graphics, you need to refit the model, switching the order of the variables on the CLASS statement. In bleach.sas, I have done just that; the first call to PROC GLM yields a profile plot that shows the stain type profiles as we vary amount of bleach (p.4 of bleach.pdf) and the second call to PROC GLM produces the amount of bleach profiles as we vary stain type (p.8). Note that the only reason for the second call to PROC GLM is to produce the other profile plot, so in that call I have used the "ODS exclude ..." statement to suppress some redundant output.

   The plot on p.4 shows that for each amount of bleach, the ordering of the three stain types is the same: jam is the easiest stain to remove, followed by tomato sauce and then ink, which is much more difficult than the others. The difference in the mean time to removal between stain types (e.g., between ink and tomato sauce) depends upon amount of bleach (which is, of course, the nature of interaction), but because the ordering of the three stain types remains constant it would be a simplification of what is going on to describe main effects in stain type or to make comparisons of the marginal means for stain type, but not a gross misrepresentation of the effects of this factor. In contrast, the plot on p.8 exhibits crossing profiles. E.g., 5 units of bleach produced the lowest average time until stain removal for ink stains, but the highest average removal time for tomato sauce stains. For crossing profiles like this, it definitely is not appropriate to make marginal comparisons (i.e., main effect comparisons) across the levels of factor A, amount of bleach.

3. Before we get too carried away with our conclusions, we probably should do some model diagnostics to determine whether our two-way anova model was appropriate. We should be most concerned about the assumption of constant variance. To check this, we examine a residuals versus fitted plot. This plot (obtained with ODS graphics on p.3 of bleach.pdf) shows a problem: it appears that the variance increases with the mean. To rectify this, we consider transforming the data. To choose the appropriate transformation, we can proceed just as we did in the one-way layout: perform a regression of the log treatment standard deviations on the log treatment means, and use the slope of this regression to estimate the

power transformation parameter.* In the two-way layout, the treatments are the combinations of the levels of factors A and B, so we will regress $log(s_{ij})$ on $log(\bar{y}_{ij\cdot})$ where s_{ij} represents the sample standard deviations of those observations in the i, j^{th} treatment. We compute the necessary quantities using PROC MEANS and run the regression using PROC REG in bleach.sas. The results appear on p.9 of bleach.pdf. The estimated slope is .53, so we take $\hat{\delta} = 0.5 \approx 0.53$, or $\hat{\lambda} = 1 - .5 = .5$ and choose the $y^{.5} = \sqrt{y}$ transformation to fix the non-constant variance problem.

After transforming the data (taking square roots) we re-fit the two-way layout model and re-do the steps we took before (produce profile plots and a plot of residuals versus fitteds). The ANOVA table for the transformed data appears on p.11 of bleach.pdf, profile plots on p.13 and p.18, and the residual plot on p.12. Notice from the residual plot that we seem to have fixed our non-constant variance problem.

4. Often interactions are present on one scale, but disappear after re-scaling (transforming) the data. This is not too hard to see in the case of the log transformation. Sometimes two factors interact because they have a multiplicative effect on the response. Since the log function has the property that $\log(XY) = \log(X) + \log(Y)$, it changes multiplicative effects into additive ones, thereby removing multiplicative interactions. However, it is not always the case that transformations can be found that remove interactions. In addition, while lack of interaction is typically easier to interpret than interaction, this gain in interpretability can be offset if the transformed scale is less meaningful to the researcher than the original one. Furthermore, a transformation to remove interaction may have ill side effects; e.g., if the data had constant variance on the original scale, a transformation to remove interaction may give the data nonconstant variance on the transformed scale.

   In summary, I don't think that it is, in general, wise to always try to transform away interactions in the data. It is more appropriate to transform to alleviate a problem like non-constant variance. If such a transformation happens to remove interaction as well, then that's just a happy by-product. In the bleach data, the square root transformation turns out to be quite helpful in producing constant variance, but does not transform away the interaction between amount of bleach and stain type. In fact, the profile plots after transformation look very similar to those before transformation. Also, the F tests for interaction and for stain type are still significant. The F test for amount of bleach is now nonsignificant. The appropriate conclusions are the same as we made earlier before transforming the data. In particular, while there is a significant interaction, marginal comparisons across the stain types are meaningful, but marginal mean comparisons across amounts of bleach are not.

5. Lastly, I want to illustrate that sometimes we may want to make comparisons

* Alternatively, we could use the Box-Cox approach (implemented in PROC TRANSREG, for example) to estimate the transformation parameter, but I will stick with the regression method in this example.

among the joint means that are not interaction contrasts. For example, the presence of an interaction in these data compromise marginal comparisons across different amounts of bleach. If we want to make comparisons across different amounts of bleach, it is important that we make them separately, within each stain type. Since amount of bleach is a quantitative factor with evenly spaced levels, a natural set of questions to address is

a. For ink stains, does the mean response change linearly with amount of bleach? (In particular, we might think that the more bleach we add, the quicker the stain will go away, and we might wonder if that relationship might be linear.)

b. For jam stains, does the mean response change linearly with amount of bleach?

c. For tomato sauce stains, does the mean response change linearly with amount of bleach?

How would we form contrasts to address these questions?

Here are the joint means

Amount of Bleach	Stain Type	Treatment (joint) mean
3	ink	μ_{11}
3	jam	μ_{12}
3	tomato	μ_{13}
5	ink	μ_{21}
5	jam	μ_{22}
5	tomato	μ_{23}
7	ink	μ_{31}
7	jam	μ_{32}
7	tomato	μ_{33}

The joint means involving ink only are μ_{11}, μ_{21}, and μ_{31}. Essentially what we want to do for question (a), is to test an "at least linear" contrast and a "nonlinear" or "lack of fit from linearity" contrast where those contrasts just involve μ_{11}, μ_{21}, and μ_{31}. Therefore, we need the linear and quadratic contrast coefficients for 3 means from the back of the book. Since there are only 3 means involved here, the quadratic contrast quantifies all of the nonlinearity, so that the lack of fit from linearity contrast would just equal the quadratic contrast.

So, here are our contrasts in terms of the μ_{ij}'s for question (a):

$$\begin{aligned}
(\text{“at least” linear}):\quad & \psi_1 = (-1)\mu_{11} + (0)\mu_{21} + (1)\mu_{31} = -\mu_{11} + \mu_{31} \\
(\text{lack of fit from linearity}):\quad & \psi_2 = (1)\mu_{11} + (-2)\mu_{21} + (1)\mu_{31} = \mu_{11} - 2\mu_{21} + \mu_{31}
\end{aligned}$$

Note though, that the model that we fit uses the effects parameterization, so we have to translate these contrasts into coefficients on the α_i's, β_j's, and $(\alpha\beta)_{ij}$'s. Recall $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$. Therefore,

$$\begin{aligned}
\psi_1 &= -\mu_{11} + \mu_{31} \\
&= -(\mu + \alpha_1 + \beta_1 + (\alpha\beta)_{11}) + (\mu + \alpha_3 + \beta_1 + (\alpha\beta)_{31}) \\
&= -\alpha_1 + \alpha_3 - (\alpha\beta)_{11} + (\alpha\beta)_{31} \\
&= (-1)\alpha_1 + (0)\alpha_2 + (1)\alpha_3 + (-1)(\alpha\beta)_{11} + (0)(\alpha\beta)_{12} + (0)(\alpha\beta)_{13} \\
&\quad + (0)(\alpha\beta)_{21} + (0)(\alpha\beta)_{22} + (0)(\alpha\beta)_{23} + (1)(\alpha\beta)_{31} + (0)(\alpha\beta)_{32} + (0)(\alpha\beta)_{33}
\end{aligned}$$

Similarly,

$$\begin{aligned}
\psi_2 &= \mu_{11} - 2\mu_{21} + \mu_{31} \\
&= 1(\mu + \alpha_1 + \beta_1 + (\alpha\beta)_{11}) - 2(\mu + \alpha_2 + \beta_1 + (\alpha\beta)_{21}) + 1(\mu + \alpha_3 + \beta_1 + (\alpha\beta)_{13}) \\
&= \alpha_1 - 2\alpha_2 + \alpha_3 + (\alpha\beta)_{11} - 2(\alpha\beta)_{21} + (\alpha\beta)_{31} \\
&= (1)\alpha_1 + (-2)\alpha_2 + (1)\alpha_3 + (1)(\alpha\beta)_{11} + (0)(\alpha\beta)_{12} + (0)(\alpha\beta)_{13} \\
&\quad + (-2)(\alpha\beta)_{21} + (0)(\alpha\beta)_{22} + (0)(\alpha\beta)_{23} + (1)(\alpha\beta)_{31} + (0)(\alpha\beta)_{32} + (0)(\alpha\beta)_{33}
\end{aligned}$$

These two contrasts are implemented in the third call to PROC GLM, as well as two more that address question (b). Notice that in the output (p.17) we conclude that, for ink stains, the mean response does not exhibit an overall trend with amount of bleach (linear contrast gives $F_{1,32} = 1.15$, $p = .2914$), and whatever pattern of change the mean exhibits as amount of bleach increases is highly non-linear ($F_{1,32} = 32.91$, $p < .0001$). And for jam stains, the pattern of change shows little nonlinearity ($F_{1,32} = 0.01$, $p = .9073$), but the trend is not significantly different from zero (zero meaning a flat trend) ($F_{1,32} = 3.11$, $p = .0874$).

STAT 8200 — Lab 7

Name:____

Exercise: Uncomment and complete the fifth and sixth contrast statements in the third call to PROC GLM in bleach.sas so that they answer question (c) above. That is, for tomato sauce stains, is the relationship between the mean response and the amount of bleach linear, or is it nonlinear? Support your answer with evidence from the two contrasts you performed.

Finally, note that if we specify the model in the cell means parameterization, specifying joint mean contrasts is a bit easier since we don't have to translate our contrasts in the μ_{ij}s into contrasts in terms of the effects parameters. This can be done by changing our MODEL statement from **MODEL Y=A B A*B;** to **MODEL Y=A*B/noint;**. With this change, the A*B terms become, in effect, the μ_{ij}s.

The final call to PROC GLM in bleach.sas refits our two-way anova model for the square root of time using a cell-means parameterization and conducts linear and quadratic tests of the effect of amount of bleach separately for each stain type. As you'll see, the specification of the contrasts is a bit more straight-forward than in the effects model and, of course, gives the same answers (cf. the results on pp. 17 & 19).

Please hand in p.7, including your answers. Remember to write your name at the top. You may keep pages 1–6 for your notes.