

1. Pygmalion effect

- (a) The test scores of individual soldiers in the same platoon would probably not be independent. They are impacted by the fact that they go through training together. (Note it might be OK to use individual scores if an experiment was done inside the platoon; since this experiment uses men from multiple platoons we have to worry about this source of dependence.) Often you will hear people say that you should analyze data for the units to which randomization was applied – platoons were randomized to treatment so platoon average scores should be analyzed. This is essentially the same point; people say this because the treatment was applied to the whole platoon and this may induced dependence.
- (b) Note that this is basically a “paired” analysis even though there are three platoons per company. We compute a control average and then subtract from the Pygmalion average. This yields a single measurement per company. The numbers are 13.8, 11.6, -11.95, 10.0, 11.6, 8.0, 11.0, 1.0, -3.6. A one-sample t-test yields $t = \bar{d}/(s_d/\sqrt{n}) = 5.717/(8.712/\sqrt{9}) = 1.968$ which should be compared to the t_8 reference distribution. The one-sided p-value (we expect a positive mean) is .042 which means there is evidence to reject the null hypothesis. The study supports the Pygmalion effect.
- (c) Note that since we end up with a one-sample procedure the assumptions are that the companies be independent and that the company difference scores of part (b) are normally distributed. (Note no assumption of constant variance because there’s only one sample!)
- (d) The estimated variance is 44.47 and the standard deviation is 6.67. Note that this is the variation that we see in platoon scores for platoons within a single company that receive the same treatment.
- (e) The differences had standard deviation 8.712. This measures the variation in the Pygmalion effect across different companies. The main point here is that there are many variance calculations possible and it is important that you be able to describe what you have done. It is natural to see more variation in the difference scores – after all this includes a comparison of platoons that received different treatments. Note that one could also compute the variation of Pygmalion scores across companies (7.82) and the variation of control scores across companies (6.06).

2. SAS Randomized block

- (a) There seem to be several places where the curves cross suggesting a possible interaction between treatment and block. In particular there are a couple of varieties where the measurement for the treatment 4 is much higher than the measurement for treatment 2 while most varieties don’t have much difference. As usual it is hard to really judge though because the square root of the MSE is 1112 (so that points can be up to 2000 away from their “true” value just due to sampling error.
- (b) The normal probability plot shows that the residuals look roughly normal with perhaps one outlying observation. (This is likely the same observation that stands out in the plot for (a). It may be worth running the analysis with and without this variety to see if the conclusions change.) The plot shows fairly constant variance across treatments. There is some evidence of non-constant variance across varieties but this is hard (impossible) to judge with only 3 observations per block.
- (c) The F-test statistic for the treatment effect is 28.72 with $p < .0001$. There is strong evidence of treatment differences. Note that you have to seek out the treatment effect on the output; it is not related to the SS(Model).
- (d) We can use the Bonferroni procedure (or the Tukey procedure as they agree in this case) and detect that the control (mean = 4557.5) is worse than the two treatments (mean = 2275 for 4 weeks and mean = 1758 for 2 weeks). The difference between the two treatments (2 weeks or 4 weeks) does not appear to be significant. This can also be detected via the use of contrasts with weights (2, -1, -1) and (0, 1, -1). The first result is what we would expect - some herbicide application is better than none. Early application is a bit better than late application but the difference is not significant.
- (e) Blocks are not significantly different at the .05 level or even at the .10 level. This DOES NOT however mean that blocking was ineffective. This is an important point. Blocking is mildly effective in this study. The MSE would be higher had we not blocked. The quantity I mentioned in class estimates MSE for an unblocked study is 12% higher than the MSE in the blocked study. This means we’d need 12% more data (perhaps 18 measurements per treatment instead of the 16 blocks) in a completely randomized experiment to attain the same level of significance as we found in the blocked study.

SAS program:

```
options linesize=80;
filename weeds 'f:\Stat210\weedctrl.txt';
data weed;
    infile weeds firstobs=1;
    input variety $ treat loc biomass;
    if loc = 1;
**                                     **;
** plot score vs treatment (with block identified for each observation) **;
proc gplot;
```

```

symbol1 c=black i=join f=greek v=1;
symbol2 c=black i=join f=greek v=2;
symbol3 c=black i=join f=greek v=3;
symbol4 c=black i=join f=greek v=4;
symbol5 c=black i=join f=greek v=5;
symbol6 c=black i=join f=greek v=6;
symbol7 c=black i=join f=greek v=7;
symbol8 c=black i=join f=greek v=8;
symbol9 c=black i=join f=greek v=9;
symbol10 c=black i=join f=greek v=<;
symbol11 c=black i=join f=greek v=0;
symbol12 c=black i=join f=greek v=#;
symbol13 c=black i=join f=greek v=$;
symbol14 c=black i=join f=greek v=%;
symbol15 c=black i=join f=greek v=>;
symbol16 c=black i=join f=greek v=&;
plot biomass*treat=variety;
**
** fit anova model with multiple contrasts, save output **;
proc glm order=data;
  class variety treat;
  model biomass = variety treat;
  means treat / bon tukey;
  contrast 'lin' treat 2 -1 -1;
  contrast 'time' treat 0 1 -1;
  output out=resids p=yhat r=resid;
**
** create normal scores for residuals **;
proc rank normal=blom out=residnrm data=resids;
  var resid;
  ranks residnrm;
**
** residual plots for randomized block analysis **;
proc plot data=residnrm;
  symbol i=none;
  plot resid*residnrm resid*treat resid*variety;
run;

```

3. Random/fixed effects

- (a) The key point in the factor effects model is that the overall mean and the J treatment effects are too many parameters. All of these parameters can't be estimated without a constraint of some kind because the data (really the likelihood) can't distinguish between parameter values μ, τ_1, τ_2 and $\mu + c, \tau_1 - c, \tau_2 - c$ (for any c we want). Both sets of parameters yield identical means for all observations under the model. Many of you tried to prove that we must have $\sum_j \tau_j = 0$. This is not possible; we must have a restriction but it need not be that one. We like to choose that restriction because it gives the parameters a natural interpretation (μ is the grand mean and τ_j is the treatment j effect) but we don't have to choose that one.
- (b) Inference for treatment mean
- Fixed effects ($\sum_i \beta_i = 0$): Start by noting that \bar{Y}_j is a linear combination of normal random variables and hence will be normal. Then, $E(\bar{Y}_j) = \frac{1}{I}E(\sum_i(\mu + \beta_i + \tau_j + \epsilon_{ij})) = \frac{1}{I}E(I\mu + 0 + I\tau_j + \sum_i \epsilon_{ij}) = \mu + \tau_j$ and $Var(\bar{Y}_j) = \frac{1}{I^2} \sum_i Var(\mu + \beta_i + \tau_j + \epsilon_{ij}) = \sigma^2/I$. The problem asked how you'd use these results ... you can form a t confidence interval for $\mu + \tau_j$ from this mean and variance (after plugging in the MSE for σ^2).
 - Random effects ($\beta_i \sim N(0, \sigma_b^2)$): The initial part of the argument proceeds as above to find that \bar{Y}_j is normal. The mean calculation is similar to the one given above. In the variance calculation β_i and ϵ_{ij} are random and both contribute variance terms so that we end up with a mean of $\mu + \tau_j$ and a variance of $(\sigma_b^2 + \sigma^2)/I$. Here again you should interpret the result and not just point out the obvious fact that the variance has two terms now. The interpretation is that if we hope to draw inferences about treatment j that generalize to a population of blocks we should use a bigger variance to account for differences among blocks. If we only want to draw inferences about the blocks in the study (the fixed effects case) then we can get more precise inferences.
- (c) Here notice that $\bar{Y}_j - \bar{Y}_{j'}$ is $\frac{1}{I}(\sum_i Y_{ij} - Y_{ij'}) = \frac{1}{I}(\sum_i \tau_j - \tau_{j'} + \epsilon_{ij} - \epsilon_{ij'})$. Note that this difference in means doesn't depend on the block effects since they cancel out of the sum. The mean is $\tau_j - \tau_{j'}$ and the variance is $2\sigma^2/I$ (same as in the usual completely randomized ANOVA!). Here too the explanation should provide the intuition (about the block effects cancelling) rather than merely report that the calculations result in the same mean and variance.
- (d) The natural quantity to examine in this case is the sample contrast $\sum_j c_j \bar{Y}_j$. As in the previous part it helps to first simplify this sum as $\sum_j c_j(\mu + \beta_i + \tau_j + \bar{\epsilon}_j) = \sum_j c_j \tau_j + \sum_j c_j \bar{\epsilon}_j$ where we use the fact that the sum of contrast weights is zero and the first two terms do not depend on j . From this last expression we can easily see that the mean is $\sum_j c_j \tau_j$ and the variance is $\sum_j c_j^2 \sigma^2/I$. Here too the block effects cancel out in computing the contrast so it doesn't matter whether we think of them as fixed or random. (Another plus of contrasts!)

4. Expected mean squares

- (a) Start with the definition $SSTR = \sum_{j=1}^r n_b(\bar{Y}_{.j} - \bar{Y}_{..})^2$. I'm going to assume fixed block effects. As in the previous problem (but with new notation - sorry about that), we get $\bar{Y}_{.j} = \mu + \tau_j + \bar{\epsilon}_{.j}$ and $\bar{Y}_{..} = \mu + \bar{\epsilon}_{..}$. The difference

squared is $\tau_j^2 + 2\tau_j(\bar{\epsilon}_{.j} - \bar{\epsilon}_{..}) + (\bar{\epsilon}_{.j} - \bar{\epsilon}_{..})^2$. Taking the sum (after multiplying by n_b) and the expected value gives $\sum_j n_b \tau_j^2 + \sum_j n_b E(\bar{\epsilon}_{.j} - \bar{\epsilon}_{..})^2 = \sum_j n_b \tau_j^2 + r n_b \text{Var}(\bar{\epsilon}_{.j} - \bar{\epsilon}_{..})$ where the expected value of the cross term is zero and the final term is recognized as a variance that is the same for each j . This last variance is a messy one; the short way to calculate it is to recognize that $\text{Var}(\bar{\epsilon}_{.j} - \bar{\epsilon}_{..}) = \text{Var}(\bar{\epsilon}_{.j}) + \text{Var}(\bar{\epsilon}_{..}) - 2\text{Cov}(\bar{\epsilon}_{.j}, \bar{\epsilon}_{..}) = \sigma^2/n_b + \sigma^2/rn_b - 2\sigma^2/rn_b = (r-1)\sigma^2/n_b r$. Putting this altogether gives $SSTR = (r-1)\sigma^2 + \sum_j n_b \tau_j^2$ and the $MSTR$ is obtained when we divide by the d.f. which is $r-1$.

- (b) You did not prove (mercifully) that $E(MSE) = \sigma^2$ but it does. If so we can note that under H_o (no treatment effect) the two expected mean squares are both equal to σ^2 whereas if H_o is false then the $E(MSTR)$ is larger than $E(MSE)$. This means looking at their ration $MSTR/MSE$ is a way to test if there are treatment effects (which of course leads to the F-test).
- (c) This question was a bit unclear. A treatment-block interaction is a situation in which the treatment effect varies from block-to-block. Remember that we explicitly rule that out in our model! If there are such interactions, then they will be absorbed into the MSE . Thus the MSE will include individual variation σ^2 plus any "error" that is due to interactions. The MSE is greater than it should be and that's what makes the test conservative.