The Theory of Designed Experiments

4. Treatments: Structures, Contrasts and Models

UNESP-Botucatu, August 2010

Treatment Structures

The precise form of treatments depends on the experimental objectives. Some distinctions:

- Treatments which are combinations of levels of several factors, or treatments which are levels of a single factor.
- Qualitative or quantitative factor levels.
- Factor levels which are of specific interest (fixed effects), or qualitative factor levels which are a sample from a population (random effects).

Treatment Structures

The initial form of analysis of variance is the same in each case, i.e. if we do not break down Treatment sum of squares into different components, and is defined by the unit structure.

Often we are interested in specific treatment *contrasts*, e.g. main effects and interactions when the treatments have a factorial structure. These correspond to a breakdown of the Treatment sum of squares into single degree of freedom components.

We might fit particular models of the treatment effect, especially for treatment factors with continuous levels. These might require fewer degrees of freedom than the Treatment sum of squares.

Factorial Structures

If the objectives go beyond discovering which effects are non-zero, an important idea is that of *hidden replication* - the replication needed is in the levels of individual factors and combinations of pairs of factors.

Consider r replicates of a 2^3 factorial experiment. Write the model for treatment effects as

$$t_{rst} = p_r + q_s + r_t + (pq)_{rs} + (pr)_{rt} + (qr)_{st} + (pqr)_{rst},$$

with $p_0 + p_1 = 0$, $(pq)_{00} + (pq)_{01} = 0$ and $(pq)_{00} + (pq)_{10} = 0$, etc.

Factorial Structures

The effects are estimated as follows:

$$\begin{aligned}
\hat{p}_1 &= \frac{1}{2}(\bar{y}_{1\cdot\cdot} - \bar{y}_{0\cdot\cdot}) \\
\widehat{(pq)}_{11} &= \frac{1}{4}(\bar{y}_{00\cdot} + \bar{y}_{11\cdot} - \bar{y}_{01\cdot} - \bar{y}_{10\cdot}) \\
\widehat{(pqr)}_{111} &= \frac{1}{8}(\bar{y}_{001} + \bar{y}_{010} + \bar{y}_{100} + \bar{y}_{111} \\
&\quad -\bar{y}_{000} - \bar{y}_{011} - \bar{y}_{101} - \bar{y}_{011})
\end{aligned}$$

Each of these has variance $\frac{\sigma^2}{8r}$.

Factorial Structures

The effect of P with the other factors at their high levels is

$$\begin{aligned} t_{111} - t_{011} &= p_1 - p_0 + (pq)_{11} - (pq)_{01} \\ &\quad + (pr)_{11} - (pr)_{01} + (pqr)_{111} - (pqr)_{011}, \end{aligned}$$

estimated with variance $\frac{\sigma^2}{r}$.

If a model without the three-factor interaction is used, the variance becomes $\frac{3\sigma^2}{4r}$.

If a main effects only model is used, the variance becomes $\frac{\sigma^2}{4r}$.

Fitting reduced models without some of the factorial effects has the following advantages:

- Simplicity of interpretation and improved understanding.
- Lower variances of estimated comparisons and predicted responses.

Factorial Structures

Thus using factorial treatments, rather than doing separate experiments with each factor has three advantages:

- interactions can be discovered;
- main effects are estimated more efficiently if there are no interactions;
- conclusions drawn about main effects are more general, as they apply to all levels of the other factors.

Nearly Saturated Structures

To investigate the effects of temperature (coded as X_1) and pressure (coded as X_2) on the yield of a reaction, the first experiment might use the design:

Treat	X_1	X_2
1	-1	-1
2	-1	1
3	1	-1
4	1	1
5	0	0
5	0	0
5	0	0
5	0	0

Nearly Saturated Structures

We obtain the following analysis of variance:

Source	df
Temperature_L	1
Pressure_L	1
$\text{Temp}_L \times \text{Press}_L$	1
Residual:	4:
Lack of fit	1
Pure error	3
Total	7

Nearly Saturated Structures

From the randomization viewpoint, we obtain 4 orthogonal treatment contrasts:

Source	df
Treatments:	4:
Temperature_L	1
Pressure_L	1
$\text{Temp}_L \times \text{Press}_L$	1
Lack of fit	1
Residual	3
Total	7

Nearly Saturated Structures

This makes clearer the meaning of "pure error" - it is just the usual unbiased estimator of σ^2.

If lack of fit, or the interaction, are close to zero, should we replace the unbiased estimator of σ^2 with a biased one including these terms in the residual?

No, unless there is some reason why we must get as good an estimate from this small experiment as possible. This is rarely the case.

Nearly Saturated Structures

Block I			Block II		
Treat	X_1	X_2	Treat	X_1	X_2
1	-1	-1	1	-1	-1
2	-1	1	2	-1	1
3	1	-1	3	1	-1
4	1	1	4	1	1
5	0	0	5	0	0
5	0	0	5	0	0

How many degrees of freedom for pure error are there?

Nearly Saturated Structures

Only the randomization analysis gives a sensible answer.

Source	df
Blocks	1
Treatments:	4:
Temperature_L	1
Pressure_L	1
$\text{Temp}_L \times \text{Press}_L$	1
Lack of fit	1
Residual:	6:
"Non-additivity"	4
"Pure error"	2
Total	11

Saturated Structures

To study the effects of catalyst (two types), amount of chemical A (low/high), amount of chemical B (low/high), stirring (yes/no) and shaking (yes/no) on a chemical system, a single replicate of 2^5 factorial treatment combinations was run in a completely randomized design.

Source	df
Catalyst	1
Chemical A	1
Chemical B	1
Stirring	1
Shaking	1
2-factor interactions	10
3-factor interactions	10
4-factor interactions	5
5-factor interaction	1
Residual	0
Total	31

Saturated Structures

We get best linear unbiased estimators (BLUEs) of all factorial effects, but no estimator of σ^2 with which to carry out inference.

Possible solutions:

1. Do nothing! Exploratory analysis is all we need, e.g. Normal plot of estimated effects.
2. Estimate σ^2 from the small effects in the Normal plot. This is a biased estimator.
3. Assume *a priori* that high order interactions will be zero. This is a strong assumption and several small, but non-zero, effects can cause unquantifiable bias. Given this assumption, 2 replicates of a half-fraction would be better.

Saturated Structures

4. Use a prior estimate of σ^2. This requires a strong assumption.
5. Use a prior distribution for σ^2. The experiment provides no information about σ^2, so this still requires a fairly strong assumption.
6. Perform a fully Bayesian analysis, with priors on each effect and updating all priors using Bayes' Theorem.

I recommend 1 or 6, depending on what conclusions we want to draw.

Supersaturated Structures

These allow us to study more factors than there are experimental units, e.g. crash testing cars.

No sensible randomization analysis is possible and perhaps a fully Bayesian analysis is the only sensible one.

Quantitative Treatment Structures

Consider an experiment in enzyme kinetics with 3 replicates at each of the substrate concentrations 10, 20, 40, 80, 160, completely randomized.

Biochemical theory implies

$$E(Y_{ij}) = \frac{\theta_0 x_i}{\theta_1 + x_i}.$$

It is natural to fit this by nonlinear least squares (NLLS) - this gives estimates of θ_0, θ_1 and σ^2.

Quantitative Treatment Structures

But this is not an unbiased estimator of σ^2. We get this from the randomization analysis. The difference represents lack of fit.

Source	df
Treatments:	4:
Michaelis-Menten model	1
Lack of fit	3
Residual	10
Total	14

Quantitative Treatment Structures

The Michaelis-Menten component in the analysis of variance does not correspond to a *linear* contrast, but the interpretation is the same.

If we use a randomized complete block design, we simply add a random effect for blocks as usual.

Source	df
Blocks	2
Treatments:	4:
Michaelis-Menten model	1
Lack of fit	3
Residual	8
Total	14

Quantitative Treatment Structures

The error structure of NLLS has been questioned and instead a transform-both-sides model suggested:

$$Y_{ij}^{(\lambda)} = \left(\frac{\theta_0 x_i}{\theta_1 + x_i} \right)^{(\lambda)} + \epsilon_{ij},$$

where

$$Y^{(\lambda)} = \begin{cases} \frac{Y^\lambda - 1}{\lambda}, & \lambda \neq 0; \\ \log Y, & \lambda = 0. \end{cases}$$

This is a good idea, but can be difficult to fit in practice.

Randomization theory allows a two-stage analysis which *greatly* simplifies the computation.

Quantitative Treatment Structures

The assumption now is that unit and treatment effects are additive *on some transformed scale*, i.e.

$$y_{i(r)}^{(\lambda)} = \mu + t_r + e_i.$$

We fit the model

$$Y_{i(r)}^{(\lambda)} = \mu + t_r + \epsilon_i,$$

i.e. estimate a Box-Cox transformation for a completely randomized design model.

Now fixing $\hat{\lambda}$, use NLLS to fit

$$Y_{ij}^{(\hat{\lambda})} = \left(\frac{\theta_0 x_i}{\theta_1 + x_i} \right)^{(\hat{\lambda})} + \epsilon_{ij},$$

adjusting the residual term by one df.