Stat 602 The Design of Experiments

Yuqing Xu

Department of Statistics
University of Wisconsin
Madison, WI 53706, USA

April 21, 2016

Daniel's Method for Testing in 2^k Model

Some Facts by Normal Assumption

- Our original data are independent and normally distributed with constant variance.
- Effects contrasts in Table 10.3 gives us results that are also independent and normally distributed with constant variance.
- The expected value of any of these contrasts is zero if the corresponding null hypothesis is true.

Table 10.3: All contrasts for a 2^3 design.

	A	B	C	AB	AC	BC	ABC
(1)	–	–	–	+	+	+	–
a	+	–	–	–	–	+	+
b	–	+	–	–	+	–	+
ab	+	+	–	+	–	–	–
c	–	–	+	+	–	–	+
ac	+	–	+	–	+	–	–
bc	–	+	+	–	–	+	–
abc	+	+	+	+	+	+	+

Daniel's Method for Testing in 2^k Model

Technique of Daniel's Plot

- The main effect α can be estimated by:

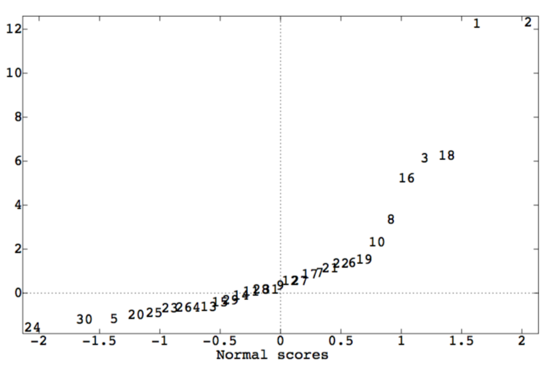
$$\hat{\alpha} = \frac{1}{8}[(y_a + y_{ab} + y_{ac} + y_{abc}) - (y_{(1)} + y_b + y_c + y_{bc})]$$

- Let $\hat{A} = 2\hat{\alpha}$.
- Contrasts corresponding to null effects should look like a sample from a normal distribution with mean zero and fixed variance.
- Contrasts corresponding to non-null effects will have different means and should look like outliers
- **Key Assumption:** We assume that we will have mostly null results, with a few non-null results that should look like **outliers**.

Note: These techniques will work poorly if there are many non-null effects.

Daniel's Method for Testing in 2^k Model

Example of Normal Plot



- Numbers indicates standard order
- Effect 1 and 2 the main effect of A, B, appears as a clear outlier, and the rest appear to follow a line.

Daniel's Method for Testing in 2^k Model

Cont. Normal Plot

- What about changing the positive $A, B, C \ldots$ to negative ones?
- It is supposed to be working too...
- The whole plot changes! No uniqueness.

Lenth's Method for Testing in 2^k Model

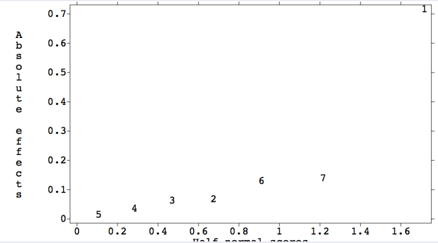
Lenth's Method

- Let $\hat{\theta}_1, \ldots, \hat{\theta}_I$ denotes estimated effects.
- First, let s_0 be 1.5 times the median of the absolute values of the contrast $\hat{\theta}$'s.
- Second, delete any contrasts results whose absolute values are greater than $2.5s_0$.
- Third, let the PSE be 1.5 times the median of the remaining absolute contrast results.
- Treat the PSE as a standard error for the contrasts with $(2^k - 1)/3$ degrees of freedom, and do t-tests.

Lenth's Method for Testing in 2^k Model

Example

- To use Lenth's method, we first need the median of the absolute factorial effects, 0.068 for these data.
- delete any absolute effects greater than $2.5 \times 0.068 = 0.17$; only the the main effect of A meets this cutoff.
- The median of the remaining absolute effects is 0.065
- The PSE is $1.5 \times 0.065 = 0.098$.



Extra Topic: Nested Model

Example

- Consider the vitamin A content of baby food carrots.
- We go to the grocery store and select **4 jars of carrots** at random **from each of the 3 brands** of baby food.
- We then take two samples from each jar and measure the vitamin A in every sample for a total of 24 responses.
- There is variation between the brands, variation be-tween individual jars for each brand, and variation between samples for every jar.

Why this example different?

- It does not make sense to consider jar main effects and brand by jar interaction.
- Jar 1 for brand A has absolutely nothing to do with jar 1 for brand B.

Extra Topic: Nested Model

Framework of Nested Model

- Factor B is nested in factor A if there is a completely different set of levels of B for every level of A.
- Thus the jars are nested in the brands and not crossed with the brands, because we have a completely new set of jars for every brand.
- We write nested models using parentheses in the subscripts to indicate the nesting.
- If brand is factor A and jar (nested in brand) is factor B, then the model is written

$$y_{jki} = \mu + \alpha_j + \beta_{k(j)} + \epsilon_{i(jk)}$$

- The $k(j)$ indicates that the factor corresponding to $k = 1, 2, 3, 4$ (factor B) is nested in the level j of factor A: $j = 1, 2, 3$.
- There is a different β_k for each level j of A.

Extra Topic: Nested Model

Skeleton of ANOVA

$$y_{jki} = \mu + \alpha_j + \beta_{k(j)} + \epsilon_{i(jk)}$$

Source	df	EMS
A	$a-1$	$\sigma^2 + n\sigma_b^2 + nb\sigma_a^2$
B(A)	a(b-1)	$\sigma^2 + n\sigma_b^2$
Error	ab(n-1)	σ^2

- The sum of squares and degrees of freedom for B(A) are the totals of those quantities for B and AB from the factorial.
- Similarly, the estimated effects are found by addition:

$$\hat{\beta_{k(j)}} = \hat{\beta}_k + (\hat{\alpha\beta})_{jk}$$

Extra Topic: Nested Model

Why Nested Model

- Subsampling produces small units by one or more layers of selection from larger batchs of units.(baby carrots example)
- In other experiments, nesting is the only available method.
- e.g. consider a genetics experiment with females nested in males. We need to be able to identify the father of the female offsprings ...
- Hence we do not choose to use a nested model for an experiment. We use a nested model because the treatment structure of the experiment was nested.