IE 361 Module 49

Design and Analysis of Experiments Part 9 (Inference for p-Way Factorial Studies)

Reading: Section 5.3 *Statistical Methods for Quality Assurance*

ISU and Analytics Iowa LLC

Confidence Intervals for Factorial Effects

Having defined and computed fitted effects for a 2^p factorial, one needs ways of judging whether what has been computed is really anything more than background noise/experimental variation. The most reliable method of doing this is based on the same kind of inference we employed in the two-factor case in Module 45. That is, each fitted effect produced by the Yates algorithm is a linear combination of sample means, an $\hat{L}$ in the notation of Module 42. Thus one can attach "margins of error" to such fitted effects using the basic method for estimating a linear combination of means (L's) presented in Module 42. And the general formula from Module 42 takes a particularly simple form in the case of fitted effects from a 2^p study.

Confidence Intervals for Factorial Effects

Corresponding to each 2^p factorial fitted effect, $\widehat{E}$ (an $\widehat{L}$), is a theoretical/population effect (a corresponding L). Confidence limits for the theoretical effect are then

$$\widehat{E} \pm t s_{\text{pooled}} \frac{1}{2^p} \sqrt{\frac{1}{n_{(1)}} + \frac{1}{n_{\text{a}}} + \frac{1}{n_{\text{b}}} + \frac{1}{n_{\text{ab}}} + \cdots}$$

In the case that the data are balanced (all samples are of size m) this formula reduces to

$$\widehat{E} \pm t s_{\text{pooled}} \sqrt{\frac{1}{m 2^p}}$$

These formulas provide the margins of error necessary to judge whether fitted effects clearly represent some non-zero real effects of the experimental factors.

In the pilot plant study, the $2^3 = 8$ sample variances can be pooled (since each of the 8 sample sizes is 2) to make

$$\begin{aligned} s_{\text{pooled}} &= \sqrt{\frac{2+8+32+2+8+8+2+2}{1+1+1+1+1+1+1+1}} \\ &= 2.83\,\text{g} \end{aligned}$$

with degrees of freedom $n - r = 16 - 8 = 8$. So (since the sample sizes are all $m = 2$) 95% two-sided confidence limits for each effect E are

$$\widehat{E} \pm 2.306\,(2.83)\sqrt{\frac{1}{2 \cdot 2^3}} \quad \text{i.e.} \quad \widehat{E} \pm 1.63\,\text{g}$$

Looking again at the result of applying the Yates algorithm, we see that only the A main effects, AC two factor interactions, and the B main effects are clearly detectable above the background variation in the pilot plant study. (Only these effects are larger in magnitude than the 1.63 g margin of error.)

The Importance of SOME Replication

The use of confidence limits for effects is the best available method for judging the statistical detectability of 2^p factorial effects. But it requires that there be some replication somewhere in a 2^p study, so that s_{pooled} can be calculated. As long as someone knowledgeable is in charge of an experiment, this is not typically an onerous requirement. Getting repeat runs at a few (if not all) sets of experimental conditions is typically not as problematic as potentially leaving oneself with ambiguous results. But unfortunately there are times when a less knowledgeable person is in charge, and one must analyze data from an *unreplicated* 2^p study. This is a far from ideal situation and the best available analysis method is not nearly as reliable as what can be done on the basis of some replication.

"Effect Sparsity" and Normal Plotting of Fitted Effects

All one can do when there is no replication in a 2^p factorial is to rely on the likelihood of **effect sparsity** and try to identify those effects that are clearly "bigger than noise" using normal plotting. That is,

- a kind of "Pareto principle" of effects says that in many situations there are really relatively few (and typically simple/low order) effects that really drive experimental results, and relative to the "big" effects, "most" others are small, almost looking like "noise" in comparison, and
- when effect sparsity holds, one can often identify the "few big actors" in a 2^p study by normal plotting the output of the Yates algorithm (except the first or overall mean value produced by Yates), looking for those few fitted effects that "plot off the line established by the majority of the fitted effects."

Below are some fictitious unreplicated 2^4 factorial data and corresponding fitted effects (the reader should apply the 4-cycle Yates algorithm and verify that he or she can compute these fitted effects!!!).

comb	y	comb	y	comb	y	comb	y
(1)	5.2	c	2.8	d	9.6	cd	9.2
a	13.0	ac	10.6	ad	13.8	acd	14.2
b	4.0	bc	4.0	bd	11.2	bcd	10.0
ab	12.6	abc	11.8	abd	13.8	abcd	14.2

$$a_2 = 3.0, b_2 = .2, ab_{22} = -.1, c_2 = -.4, ac_{22} = .1, bc_{22} = .2,$$
$$abc_{222} = 0, d_2 = 2.0, ad_{22} = -1.0, bd_{22} = .1, abd_{222} = -.2,$$
$$cd_{22} = .3, acd_{222} = .2, bcd_{222} = -.3, \quad \text{and} \quad abcd_{2222} = .1$$

A normal plot of these $2^4 - 1 = 15$ fitted effects (the overall mean is typically not plotted) is on panel 8.

The plot below suggests that even if the majority of fitted effects reflect only experimental noise, the fitted main effects of A and D (and probably the AD 2 factor interaction) are of a size that they "do not fit with the others" and are thus "clearly visible above the background variation in this study."

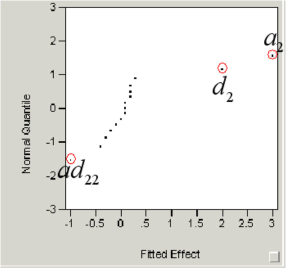

Figure: Normal Plot of 15 Fitted Effects From a Fictitious Unreplicated 2^4 Factorial Study