Experiments with Several Crossed Factors +2 factors

Hao Zhang

- Introduction
- Statistical Models
- High Order oInteractions
- Analysis and SAS Code
- A Real Experiment-Popcorn-Microwave Experiment

Introduction

Experiments that involve more than two treatment factors are designed and analyzed using many of the same principles that were discussed previously for two-factor experiments. We continue to label the factors with uppercase Latin letters and their numbers of levels with the corresponding lowercase letters. An experiment that involves four factors, A, B, C, and D, having a, b, c, and d levels, respectively, for example, is known as an " $a \times b \times c \times d$ factorial experiment" (read " a by b by c by d ") and has a total of $v = abcd$ treatment combinations. A completely randomized design is conducted in the same way as an $a \times b$ factorial design.

However, for the analysis, the high order interactions may not exist or be ignored. Therefore a set of models that may be all appropriate for the design.

Statistical Models (3)

We will consider models involving three treatment factors A, B and C. The following cell-means model is just a rewrite of one-way model

$$
\begin{gathered}
Y_{ijkt} = \mu + \tau_{ijk} + \epsilon_{ijkt}, \\
\epsilon_{ijkt} \sim N\left(0, \sigma^2\right), \\
\epsilon_{ijkt}\text{'s mutually independent,} \\
t = 1, \ldots, r_{ijk}; \quad i = 1, \ldots, a;\ j = 1, \ldots, b;\ k = 1, \ldots, c.
\end{gathered}
$$

where τ_{ijk} are treatment effects and Y_{ijkt} represents the tth observation for the treatment ijk.

ABCD

We can use this cell-means model if our objective is to compare some specific treatment means.

The cell-means model does not reveal how each factor affects the treatment means or how the factors interact. To answer those questions, we use multi-way models. For example the three-way complete model is

Interaction of 2 Interaction all 3

(ABC)(ABC)

~~AA~~, AB, AC
~~BA~~, ~~BB~~, BC
~~CA~~, ~~CB~~, ~~CC~~

$$
\begin{gathered}
Y_{ijkt} = \mu + \overset{A}{\alpha_i} + \overset{B}{\beta_j} + \overset{C}{\gamma_k} + \overset{AB}{(\alpha\beta)_{ij}} + \overset{AC}{(\alpha\gamma)_{ik}} + \overset{BC}{(\beta\gamma)_{jk}} + \overset{ABC}{(\alpha\beta\gamma)_{ijk}} + \epsilon_{ijkt}, \\
\epsilon_{ijkt} \sim N\left(0, \sigma^2\right), \\
\epsilon_{ijkt}\text{ 's mutually independent ,} \\
t = 1, \ldots, r_{ijk}; \quad i = 1, \ldots, a;\ j = 1, \ldots, b;\ k = 1, \ldots, c.
\end{gathered}
$$

If prior to the experiment certain interaction effects are known to be negligible, these interaction terms shall be excluded in the model. For example, if the factors A and B are known not to interact in a three-factor experiment, then the $\cancel{A}B$ and $\cancel{A}BC$ interaction effects are negligible, so the terms $(\alpha\beta)_{ij}$ and $(\alpha\beta\gamma)_{ijk}$ are excluded from model.

However, when a model includes an interaction between a specific set of m factors, then all interaction terms involving subsets of those m factors should be included in the model. For example, a model that includes the effect of the three-factor interaction ABC would also include the effects of the AB, AC, and BC interactions as well as the main effects A, B, and C.

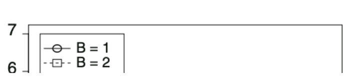
Use of a submodel or reduced model, when appropriate, is advantageous, because simpler models generally yield tighter confidence intervals and more powerful tests of hypotheses. However, if interaction terms are removed from the model when the factors do, in fact, interact, then the resulting analysis and conclusions may be totally incorrect.

A, B, C AB, AC, BC, ABC

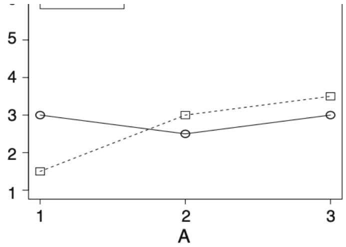
We use the three-way complete model if we are interested in investigating the main effects of all factors and all interaction effects. If for example, it is not our objectives to investigate the main effects of A and B, and the AB interaction effects, we can combine the two factors A and B into one. Then the three-way model then becomes a two-way model.

High Order oInteractions

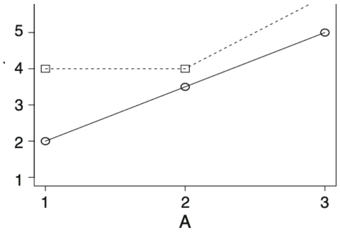
What does the three way ABC interaction mean? If the interaction effects of any pair (say, A and B) depends on the level of the third factor, then ABC interaction effects do exist.



A:3
B:2
C:2



(a) Level 1 of C



(b) Level 2 of C

* A and B does interact regardless of C, but they interact differently across different levels of C. (lines are not the same)

Hence: How A and B interacts is depended on C

* three-way interaction *

The above plots show the three way ABC-interaction does exist because the way A and B interacts depends on the level of C!

Analysis and SAS Code

The contrast is defined the same way as before: A linear combination of the treatment means with the contrast coefficients adding up to be 0. We may also define specific contrasts such as main effects contracts for each factor, two-way interaction effects, three-way interaction effects, etc.

The least squares estimate of the treatment means are the sampling means. Therefore, a contrast can be estimated by a corresponding linear combination of sample means, whose variance and standard derivation can be easily obtained. The error variance σ^2 is estimated by the mean squares for error (MSE). Therefore the confidence interval for each individual contrast can be constructed.

The simultaneous confidence intervals can be obtained by applying Bonferroni, Scheffé, Tukey, and Dunnett methods similar to in the two-way models.

The test of hypothesis about the mean effects or interaction effects is conducted similarly by applying an F-test.

The following ANOVA table is for the three-way complete model.

Source of variation	Degrees of freedom	Sum of squares	Mean square	F-Ratio	p-value
A	a-1	SSA	MSA	MSA/MSE	
B	b-1	SSB	MSB	MSB/MSE	
C	c-1	SSC	MSC	MSC/MSE	
AB	(a-1)(b-1)	SSAB	MSAB	MSAB/MSE	
AC	(a-1)(c-1)	SSAC	MSAC	MSAC/MSE	
BC	(b-1)(c-1)	SSBC	MSBC	MSBC/MSE	
ABC	(a-1)(b-1)(c-1)	SSABC	MSABC	MSABC/MSE	
Error	n-abc	SSE	MSE		
Total	n-1				

A Real Experiment-Popcorn-Microwave Experiment

The experiment described section 7.4 was to compare brands of microwave popcorn. The objective of the experiment was to find out which brand gives rise to the best popcorn in terms of the proportion of popped kernels. The experiment was restricted to popcorn produced in a microwave oven.

A: 3 The first treatment factor was "brand."' Three levels were selected, including two national brands (levels 1 and 2) and one local brand (level 3). All three brands are packaged for household consumers in boxes of 3.5 ounce packages, and a random selection of packages was used in this experiment.

B: 2 Power of the microwave oven was identified as a possible major source of variation and was included as a second treatment factor. Two available microwave ovens had power ratings of 500W and 600W.

C: 3 Popping time was taken as a third treatment factor. The usual instructions provided with microwave popcorn are to microwave it until rapid popping slows to 2 to 3 seconds between pops. Five preliminary trials using brand 3, a 600W microwave oven, and times equally spaced from 3 to $5\ \mathrm{min}$ suggested that the best time was between 4 and $5\ \mathrm{min}$. Hence, time levels of $4, 4.5$, and $5\ \mathrm{min}$ were selected for the experiment and coded 1-3, respectively.

1 2 3

Read Section 7.4 for detail about the measurements to be taken and how the sample size was determined. The data is provided in the following table

Brand (i)	Power (j)	Time (k)		
		1	2	3
1	1	73.8, 65.5	70.3, 91.0	72.7, 81.9
1	2	70.8, 75.3	78.7, 88.7	74.1, 72.1
2	1	73.7, 65.8	93.4, 76.3	45.3, 47.6
2	2	79.3, 86.5	92.2, 84.7	66.3, 45.7
3	1	62.5, 65.0	50.1, 81.5	51.4, 67.7
3	2	82.1, 74.5	71.5, 80.0	64.0, 77.0

We use the three-way complete model to analyze the data.

SAS code:

```
data popmic;A  B
  input brand power @@;
C do time=1 to 3;
    do rep=1 to 2; drop rep; 2 Replicates for each level of time
      input y @@;
      output;
  end; end;
  lines;  time1      time2      time3
  1 1  73.8 65.5  70.3 91.0  72.7 81.9
  1 2  70.8 75.3  78.7 88.7  74.1 72.1
  2 1  73.7 65.8  93.4 76.3  45.3 47.6
  2 2  79.3 86.5  92.2 84.7  66.3 45.7
  3 1  62.5 65.0  50.1 81.5  51.4 67.7
  3 2  82.1 74.5  71.5 80.0  64.0 77.0
;
run;

proc print data=popmic;
run;

proc glm data=popmic;
class brand power time;
model y=brand power time brand*power brand*time power*time brand*power*time;
run;     A     B   C     AB        AC        BC        ABC
         1     2   3     4         5         6         7
```

Partial outcome is repreduced below.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	17	4065.728889	239.160523	2.73	0.0206
Error	18	1577.870000	87.659444		
Corrected Total	35	5643.598889			

		Source	DF	Type III SS	Mean Square	F Value	Pr > F	
	A	brand	3-1 = 2	331.100556	165.550278	1.89	0.1801	
√	B	power	2-1 = 1	455.111111	455.111111	5.19	0.0351	
	C	time	3-1 = 2	1554.575556	777.287778	8.87	0.0021	
X	AB	brand*power	2x1 = 2	196.040556	98.020278	1.12	0.3485	no AB
√	AC	brand*time	2x2 = 4	1433.857778	358.464444	4.09	0.0157	
X	BC	power*time	1x2 = 2	47.708889	23.854444	0.27	0.7648	no BC
X	ABC	brand*power*time	2x1x2 = 4	47.334444	11.833611	0.13	0.9673	no ABC

no 3 way interaction

We see that power is significant and does not interact with brand and time. Therefore it makes sense to compare the main effects of power. We also like to know how the significant interaction effects of brand and time look like.

confidence limits

```
B   lsmeans power/cl pdiff; α=0.01
AC  lsmeans brand*time;
```

power	y LSMEAN	95% Confidence Limits	
1 500W	68.638889	64.002573	73.275205
2 600W	75.750000	71.113684	80.386316

B ☺ use 600w for more popcorn kernel

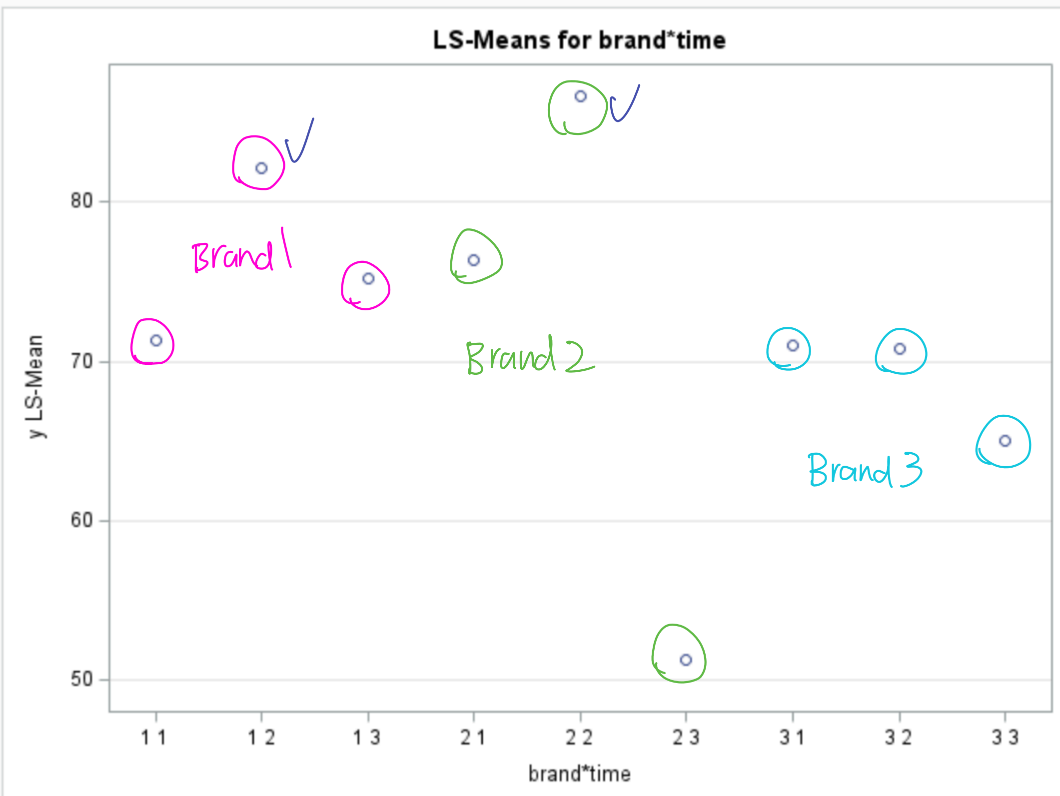
Least Squares Means for Effect power			
i	j	Difference Between Means	95% Confidence Limits for LSMean(i)-LSMean(j)

1	2	-7.111111	-13.667852	-0.554370

The GLM Procedure
Least Squares Means

AC

brand	time	y LSMEAN
1	1	71.3500000
1	2	82.1750000
1	3	75.2000000
2	1	76.3250000
2	2	86.6500000
2	3	51.2250000
3	1	71.0250000
3	2	70.7750000
3	3	65.0250000



We see Time level 2 has the highest mean for Brand 1 and 2 but not for Brand 3.

Activity Details

Task: View this topic