Lab week 11

2^k designs

An unreplicated 2^k factorial design is also sometimes called a "single replicate" of the 2^k experiment. You would find these types of designs used where the process for instance is very expensive, or it takes a long time to run an experiment. In these cases, for the purpose of saving time or money, we want to run a screening experiment with as few observations as possible. As a matter of fact, the general rule of thumb is that you would have at least two replicates. This would be a minimum in order to get an estimate of variation - but when we are in a tight situation, we might not be able to afford this due to time or expense. We will look at an example with one observation per cell, no replications, and what we can do in this case.

Exercise 1. In an experiment, the goal is to determine how the yield of an adhesive application process can be improved by adjusting three process parameters: mixture ratio, curing temperature, and curing time. For each of these input parameters, two levels will be defined for use in this 2-level experiment. For the mixture ratio, the high level is set at 55%, while the low level is set at 45%. For the curing temperature, the high level is set at 150 deg C while the low level is set at 100 deg C. For the curing time, the high level is set at 90 minutes, while the low level is set at 30 minutes. The output response monitored is process yield. Assume further that the data were gathered by performing just a single replicate (n=1) per combination treatment.

Run	Combination	Mix Ratio (a)	Temp(b)	Time(c)	Response
1	(1)	45% (-)	100C (-)	30m (-)	8
2	a	55% (+)	100C (-)	30m (-)	9
3	b	45% (-)	150C (+)	30m (-)	34
4	ab	55% (+)	150C (+)	30m (-)	52
5	c	45% (-)	100C (-)	90m (+)	16
6	ac	55% (+)	100C (-)	90m (+)	22
7	bc	45% (-)	150C (+)	90m (+)	45
8	abc	55% (+)	150C (+)	90m (+)	56

Effect-the average of observations when the factor is **high** minus the average when it is **low**.

In general for 2^k factorials the effect of each factor and interaction is:

Effect = $\frac{1}{2^{k-1}n} \times$ [contrast of the totals].

The following calculations for the main and interaction effects of these 3 factors are obtained:

A = 1/(4n) x [-(1)+a-b+ab-c+ac-bc+abc] = [-8+9-34+52-16+22-45+56] = 1/4 x 36 = 9

B = 1/4 x [-8-9+34+52-16-22+45+56] = 1/4 x 132 = 33

AB = 1/4 x [+8-9-34+52+16-22-45+56] = 1/4 x 22 = 5.5

C = 1/4 x [-8-9-34-52+16+22+45+56] = 1/4 x 36 = 9

AC = 1/4 x [+8-9+34-52-16+22-45+56] = 1/4 x -2 = -0.5

BC = 1/4 x [+8+9-34-52-16-22+45+56] = 1/4 x -6 = -1.5

ABC = 1/4 x [-8+9+34-52+16-22-45+56] = 1/4 x -12 = -3

If we want to test an effect, for instance, say A = 0, then we can construct an F-test based on the sum of squares. The equation for the sum of squares due to an effect is defined as:

$$SS(Effect) = \frac{N \times (Effect)^2}{4}$$

where N is the total number of measurements, 8 in this case.

SS(A $= 9^2/4$=162

SS(B) $= 33^2/4 = 2178$

SS(AB) $= 5.5^2/4 = 60.5$

SS(C) $= 9^2/4 = 162$

SS(AC) $= -0.5^2/4 = 0.5$

SS(BC) $= -1.5^2/4 = 4.5$

SS(ABC) $= -3^2/4 = 18$

Each sum of squares for effects has one degree of freedom.

If all interaction are considered and no replicates are made, there is not enough degrees of freedom to test for significance of main effects and interactions. If such tests are desired the highest order interaction should be dropped from considerations.

The analysis in R is done as follows:

yield <- c(8,9,34,52,16,22,45,56)

Mix <- c(-1,1,-1,1,-1,1,-1,1)

Temp<-c(-1,-1,1,1,-1,-1,1,1)

Time<-c(-1,-1,-1,-1,1,1,1,1)

$model < -lm(yield \sim Mix + Temp + Time + Mix * Temp + Mix * Time + Temp * Time + Mix * Time * Temp)$

anova(model)

Analysis of Variance Table

Response: yield

	Df	Sum Sq	Mean Sq	F value	Pr(¿F)
Mix	1	162.0	162.0		
Temp	1	2178.0	2178.0		
Time	1	162.0	162.0		
Mix:Temp	1	60.5	60.5		
Mix:Time	1	0.5	0.5		
Temp:Time	1	4.5	4.5		
Mix:Temp:Time	1	18.0	18.0		
Residuals	0	0.0			

Warning message:

In anova.lm(model) :

ANOVA F-tests on an essentially perfect fit are unreliable

In these experiment one really cannot model the "noise" or variability very well. These experiments cannot really test whether or not the assumptions are being met - again this is another shortcoming, or the price of the efficiency of these experiment designs. With no replication, fitting the full model results in zero degrees of freedom for error. Potential solutions to this problem might be: pooling high-order interactions to estimate error or dropping entire factors from the model.

Exercise 2. In a liquid chromatography experiment, the dependence of the retention parameter, k', on three factors was investigated. The factors are: pH (P), the concentration of a counter-ion (T), and the concentration of the organic solvent in the mobile phase (C). Two levels were used for each factor an two replicate measurements made or each combination. The table below gives the results.

Run	Combination	P	T	C	Response
1	(1)	(-)	(-)	(-)	4.6 , 4.8
2	p	(+)	(-)	(-)	9.8, 10
3	t	(-)	(+)	(-)	6.5, 7.5
4	pt	(+)	(+)	(-)	14.5, 15.5
5	c	(-)	(-)	(+)	2.6, 2.8
6	pc	(+)	(-)	(+)	5.1 , 5.5
7	tc	(-)	(+)	(+)	3.1, 3.3
8	ptc	(+)	(+)	(+)	5.6, 6. 4

Use R to test for significance the main effects, the first and second order interactions.

y <- c(4.6 , 4.8, 9.8, 10, 6.5, 7.5, 14.5, 15.5, 2.6, 2.8, 5.1 , 5.5, 3.1, 3.3, 5.6, 6.4)

P <- rep(c(-1,1,-1,1,-1,1,-1,1), each=2)

T<-rep(c(-1,-1,1,1,-1,-1,1,1),each=2)

C<-rep(c(-1,-1,-1,-1,1,1,1,1),each=2)

model <- lm(y $\sim T + C + P + T * C + T * P + C * P + T * C * P$)

anova(model)

Analysis of Variance Table

Response: y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
T	1	18.49	18.490	99.9459	8.505e-06
C	1	94.09	94.090	508.5946	1.582e-08
P	1	86.49	86.490	467.5135	2.206e-08
$T:C$	1	9.61	9.610	51.9459	9.179e-05
$T:P$	1	2.25	2.250	12.1622	0.00823
$C:P$	1	15.21	15.210	82.2162	1.754e-05
$T:C:P$	1	1.69	1.690	9.1351	0.01650
Residuals	8	1.48	0.185		