Lec 8: 2^k Factorial Design

Ying Li

December 1, 2011

2^k factorial design

- k factors;
- Each at two levels;
- A complete replicate requires $2 \times 2 \times \cdots \times 2 = 2^k$

Field of Application

- k factors
- Which of them are important?
- Investigate 2 levels per factor in n replicate.
- Complete factorial design: $2^k n$ observations.

Assumption through this lecture

- The factors are fixed.
- The designs are completely randomized.
- The usual normality assumption are satisfied.

2^2 Design

Replicate	Temperature	Dough Liquid	Fermentation Time
I	35	Milk	76
I	35	Water	75
I	39	Milk	69
I	39	Water	65
II	35	Milk	76
II	35	Water	79
II	39	Milk	64
II	39	Water	61

2^2 Design

Replication	T	D	Y
I	1	1	76
I	1	-1	75
I	-1	1	69
I	-1	-1	65
II	1	1	76
II	1	-1	79
II	-1	1	64
II	-1	-1	61

Treatment combinations in the 2^2 design

	D	
T	-1	1
-1	126	133
1	154	152

Sums of squares

SS are easily calculated in 2^2 experiments:

$$SS = \frac{Contrast^2}{4n}$$

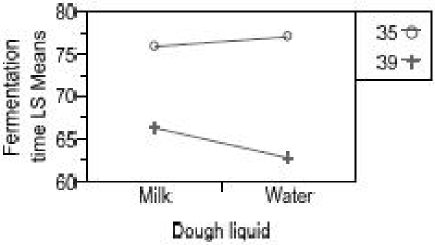
Example

$$SS_T = \frac{(TD + T - D - (1))^2}{4n}$$

ANOVA table for the example

	DF	SS	MS	F	P
T	1	276,125	276,125	38,75	0,003
D	1	3,125	3,125	0,44	0,544
T*D	1	10,125	10,125	1,42	0,299
Error	4	28,500	7,125		
Total	7	317,875			

Interaction Plot



The regression model estimates

Term	Estimate	Std Error	t	P
Intercept	70,625	0,944	78,84	0,000
T	5,875	0,944	6,23	0,003
D	0,625	0,944	0,66	0,544
T*D	-1,125	0,944	-1,19	0,299

2^k are optimal designs

The 2^k design for fitting the first-order model or the first order model with interaction is:

- **D-optimal design** for minimizing the variance of the model regression coefficients.
- **G-optimal design** for minimizing the maximum prediction variance.
- **I-optimal design** for its smallest possible value of the average prediction variance.

The general 2^k design

2^k factorial design

- A design with k factors each at 2 levels.
- The statistical model for a 2^k design would include:
  - k main effect
  - C_k^2 two-factor interactions
  - C_k^3 three-factor interactions
  - ...
  - k-factor interaction.

Calculation of effects

Effects are easily calculated in 2^k experiments

$$Effect = \frac{Contrast}{2^{k-1}n}$$

Sums of squares

SS are easily calculated in 2^k experiments

$$SS = \frac{Contrast^2}{2^k n}$$

2^4 experiments

Temp	Dough Liquid	Sugar	Fat	Repl. 1	Repl. 2	Sum
35	Milk	With	Oil	76	76	152
35	Milk	With	Butter	70	74	144
35	Milk	Utan	Oil	67	75	142
35	Milk	Utan	Butter	71	79	150
35	Water	With	Oil	75	79	154
35	Water	With	Butter	75	72	147
35	Water	Utan	Oil	66	64	130
35	Water	Utan	Butter	67	64	131
39	Milk	With	Oil	69	64	133
39	Milk	With	Butter	64	58	122
39	Milk	Utan	Oil	63	61	124
39	Milk	Utan	Butter	57	59	116
39	Water	With	Oil	65	61	126
39	Water	With	Butter	63	59	122
39	Water	Utan	Oil	55	52	107
39	Water	Utan	Butter	58	53	111

2^4 experiments

T	D	S	F	Repl 1	Repl 2	Sum
1	1	1	1	76	76	tdsf = 152
1	1	1	-1	70	74	tds = 144
1	1	-1	1	67	75	tdf = 142
1	1	-1	-1	71	79	td = 150
1	-1	1	1	75	79	tsf = 154
1	-1	1	-1	75	72	ts = 147
1	-1	-1	1	66	64	tf = 130
1	-1	-1	-1	67	64	t = 131
-1	1	1	1	69	64	dsf = 133
-1	1	1	-1	64	58	ds = 122
-1	1	-1	1	63	61	df = 124
-1	1	-1	-1	57	59	d = 116
-1	-1	1	1	65	61	sf = 126
-1	-1	1	-1	63	59	s = 122
-1	-1	-1	1	55	52	f = 107
-1	-1	-1	-1	58	53	(1) = 111

$$T = \frac{76+70+\ldots+64+64-69-64-\ldots-52-53}{2^{4-1}\cdot 2} = 11{,}8125$$

$$SS_{Temp} = \frac{(76+70+\ldots+64+64-69-64-\ldots-52-53)^2}{4^4\cdot 2} = 1116$$

ANOVA table

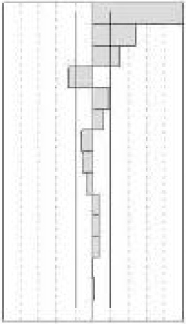
	DF	SS	MS	F	P
T	1	1116,28	1116,28	112,685	<,0001*
D	1	94,53	94,53	9,543	0,0070*
T*D	1	0,28	0,28	0,028	0,8683
S	1	247,53	247,53	24,987	0,0001*
T*S	1	0,03	0,03	0,003	0,9559
D*S	1	81,28	81,28	8,205	0,0112*
T*D*S	1	13,78	13,78	1,391	0,2554
F	1	19,53	19,53	1,972	0,1794
T*F	1	5,28	5,28	0,533	0,4759
D*F	1	5,28	5,28	0,533	0,4759
T*D*F	1	19,53	19,53	1,972	0,1794
S*F	1	38,28	38,28	3,864	0,0669
T*S*F	1	5,28	5,28	0,533	0,4759
V*S*F	1	0,28	0,28	0,028	0,8683
T*D*S*F	1	5,28	5,28	0,533	0,4759
Error	16	158,5	9,91		
Total	31				

Regression coefficients

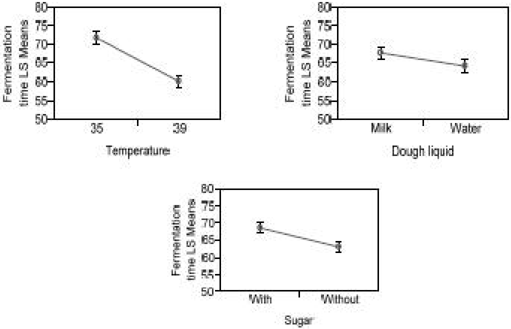
Response Fermentation time

Sorted Parameter Estimates

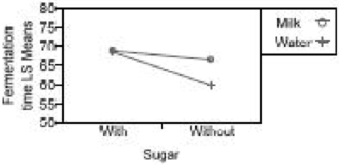
Term	Estimate	Std Error	t Ratio	Prob>|t|
T	5,90625	0,55639	10,62	<,0001*
S	2,78125	0,55639	5,00	0,0001*
D	1,71875	0,55639	3,09	0,0070*
D*S	-1,59375	0,55639	-2,86	0,0112*
S*F	1,09375	0,55639	1,97	0,0669
F	0,78125	0,55639	1,40	0,1794
T*D*F	-0,78125	0,55639	-1,40	0,1794
T*D*S	-0,65625	0,55639	-1,18	0,2554
T*F	-0,40625	0,55639	-0,73	0,4759
D*F	0,40625	0,55639	0,73	0,4759
T*S*F	0,40625	0,55639	0,73	0,4759
T*D*S*F	0,40625	0,55639	0,73	0,4759
T*D	-0,09375	0,55639	-0,17	0,8683
D*S*F	0,09375	0,55639	0,17	0,8683
T*S	-0,03125	0,55639	-0,06	0,9559

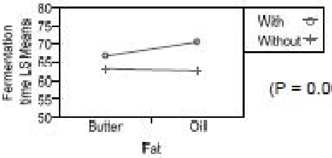


Significant effects



Significant interactions





(P = 0.069)

No replicates

- Large number of combinations

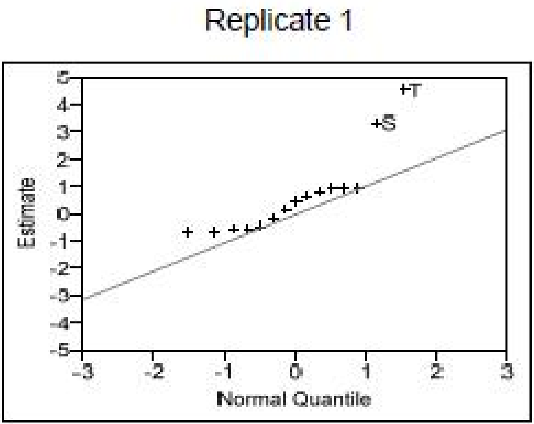
No replicates

- Large number of combinations
- Without replicates:
  - All effects can be estimated.
  - The variance cannot be estimated.
- Could the effects be evaluated without an estimates of the variance?

No replicates

- Large number of combinations
- Without replicates:
  - All effects can be estimated.
  - The variance cannot be estimated.
- Could the effects be evaluated without an estimates of the variance?
  - Normal probability plot.
  - Half normal probability plot.

Normal probability plot



Half normal plot