6 The 2^k Factorial Design

6.1 Introduction

Factorial Designs are widely used in Experiments involving several factors where it is necessary to study the joint effect of the factors on a response. The most important of these special cases is that of k factors, each at only two levels. These levels may be quantitative or qualitative. A complete replicate od such a design requires $2 \times 2 \times ... = 2^k$ observations and is called a 2^k factorial design.

6.2 The 2^2 design

The first design in the 2^k series is one with only two factors say A and B, each run at two levels. The is called a 2^2 factorial design. The levels of the factors may be arbitrarily called "low" and "high". As an example, consider an investigation into the effect of the concentration of the reactant and the amount of the catalyst on the conversion (yield) in a chemical process.

Factor		Treatment	Replicate			
A	B	Combination	I	II	III	Total
−	−	A low, B low	28	25	27	80
+	−	A high, B low	36	32	32	100
−	+	A low, B high	18	19	23	60
+	+	A high, B high	31	30	29	90

A = reactant concentration, B = catalyst amount,

y = recovery

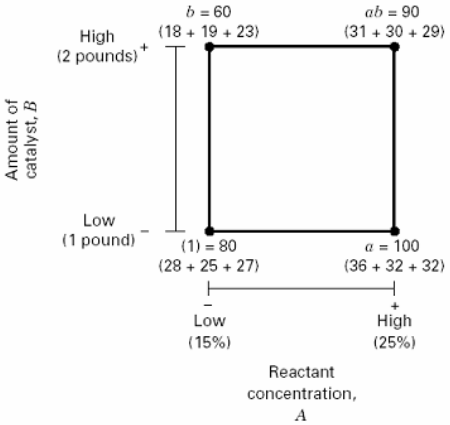


Figure 6-1 Treatment combinations in the 2^2 design.

Let a represent the treatment combination of A at the high level and B at the low level, b represent the treatment combination of A at the low level and B at the high level, ab represent the treatment combination of A at the high level and B at the high level. By convention, (1) is used to denote both factors at the low level.

Define the average effect of a factor as the change in response produced by a change in the level of that factor averaged over the levels of the other factor. Also, the symbols $(1), a, b, ab$ now represent the total of the response observation at all n replicates. The main effects of A and B is

$$A = \frac{1}{2n}[ab + a - b - (1)], \quad B = \frac{1}{2n}[ab + b - a - (1)]$$

The interaction effect AB is

$$AB = \frac{1}{2n}[ab + (1) - a - b]$$

Other expression:

$$A = \bar{y}_{A^+} - \bar{y}_{A^-}, \quad B = \bar{y}_{B^+} - \bar{y}_{B^-}$$

The Interaction effect AB can be expressed as the average of the right-to-left diagonal treatment combinations in the squares [ab and (1)] minus the average of the left-to-right diagonal treatment combinations (a and b), or

$$AB = \frac{ab + (1)}{2n} - \frac{a + b}{2n}$$

Using the experiment in Figure 6.1, We have that

$$A = \frac{1}{2(3)}(90 + 100 - 60 - 80) = \frac{25}{3}$$

Similarly, $B = -5.0, AB = \frac{5}{3}$.

Now we consider the sums of squares for A, B and AB. Note that a contrast is used in estimating A, namely,

$$\text{Contrast}_A = ab + a - b - (1)$$

This is the total contrast of A. Thus, we have

$$SS_A = \frac{[ab + a - b - (1)]^2}{4n}$$

Similarly,

$$SS_B = \frac{[ab - a + b - (1)]^2}{4n}, \quad SS_{AB} = \frac{[ab - a - b + (1)]^2}{4n}$$

Therefore, we have

$$SS_A = \frac{(50)}{4(3)} = \frac{625}{3}, \quad SS_B = \frac{(-30)^2}{4(3)} = 75.00, \quad SS_{AB} = \frac{(10)^2}{4(3)} = \frac{25}{3}$$

Further, we have

$$SS_T = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{3} y_{ijk}^2 - \frac{y_{\ldots}^2}{4(3)} = 9398.00 - 9075.00 = 323.00$$

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} = 323.00 - 625/3 - 75.00 - 25/3 = \frac{94}{3}$$

Table 6-1 Analysis of Variance for the Experiment in Figure 6-1

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
A	208.33	1	208.33	53.15	0.0001
B	75.00	1	75.00	19.13	0.0024
AB	8.33	1	8.33	2.13	0.1826
Error	31.34	8	3.92		
Total	323.00	11			

Table 6.2 Algebraic Signs for Calculating Effects in the 2^2 Designs

Treatment Combination	Factorial Effect I	A	B	AB
(1)	+	-	-	+
a	+	+	-	-
b	+	-	+	-
ab	+	+	+	+

The Regression Model It is easy to express the results of the expression in terms of a regression model.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 + \epsilon$$

where x_1 is a coded variable that represents the reactant concentration, x_2 is a coded variable that represents the amount of catalyst, and β's are regression coefficients. The coded variable is

$$x_1 = \frac{\text{Conc} - (\text{Conc}_{low} + \text{Conc}_{high})/2}{(\text{Conc}_{low} - \text{Conc}_{high})/2}$$

$$x_1 = \frac{\text{Catalyst} - (\text{Catalyst}_{low} + \text{Catalyst}_{high})/2}{(\text{Catalyst}_{low} - \text{Catalyst}_{high})/2}$$

The fitted regression model is

$$\hat{y} = 27.5 + \left(\frac{25}{2(3)}\right) x_1 + \left(\frac{-5.00}{2}\right) x_2$$

Residuals and Model Adequacy The regression model can be used to obtain the predicted or fitted value of y at four points in the design. For example, at low level for two factors, we have

$$\hat{y} = 27.5 + \left(\frac{25}{2(3)}\right)(-1) + \left(\frac{-5.00}{2}\right)(-1) = 25.835$$

There are three observations at this treatment combination, and the residuals are

$$e_1 = 28 - 25.835 = 2.165, \quad e_2 = 25 - 25.835 = -0.835, \quad e_3 = 27 - 25.835 = 1.165$$

The response Surface

$$\hat{y} = 27.5 + \left(\frac{8.33}{2}\right)\left(\frac{\text{Conc} - 20}{5}\right) + \left(\frac{-5.00}{2}\right)\left(\frac{\text{Catalyst} - 1.5}{0.5}\right)$$

$$\hat{y} = 27.5 + 0.833\text{Conc} - 5.00\text{Catalyst}$$

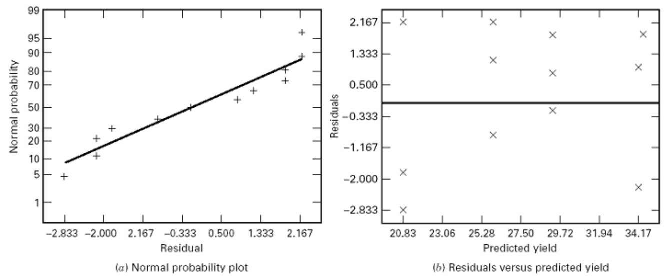


Figure 6-2 Residual plots for the chemical process experiment.

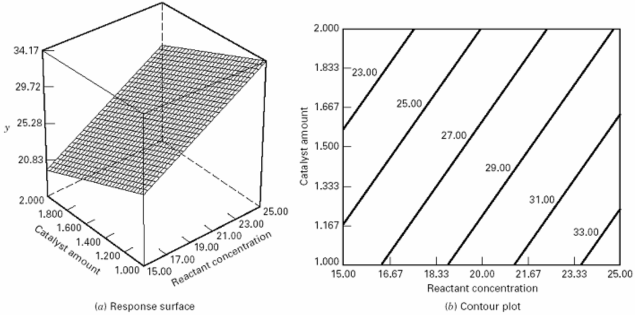
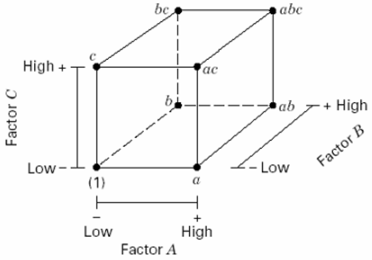


Figure 6-3 Response surface plot and contour plot of yield from the chemical process experiment.

6.3 The 2^3 design Suppose that three factors, A, B, and C, each at two levels, are of interest. The design is called a 2^3 factorial design.



(a) Geometric view

Run	Factor A	Factor B	Factor C
1	−	−	−
2	+	−	−
3	−	+	−
4	+	+	−
5	−	−	+
6	+	−	+
7	−	+	+
8	+	+	+

(b) The design matrix

Figure 6-4 The 2^3 factorial design.

Main Effects for the factors:

$$A = \bar{y}_{A^+} - \bar{y}_{A^-} = \frac{1}{4n}[abc + ab + ac + a - bc - b - c - (1)]$$

$$SS_A = \frac{(abc + ab + ac + a - bc - b - c - (1))^2}{8n}$$

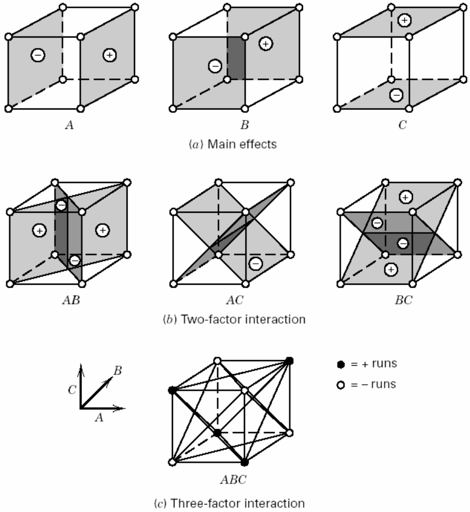


Figure 6-5 Geometric presentation of contrasts corresponding to the main effects and interactions in the 2^3 design.

Table 6-3 Algebraic Signs for Calculating Effects in the 2^3 Design

Treatment Combination	Factorial Effect							
	I	A	B	AB	C	AC	BC	ABC
(1)	+	−	−	+	−	+	+	−
a	+	+	−	−	−	−	+	+
b	+	−	+	−	−	+	−	+
ab	+	+	+	+	−	−	−	−
c	+	−	−	+	+	−	−	+
ac	+	+	−	−	+	+	−	−
bc	+	−	+	−	+	−	+	−
abc	+	+	+	+	+	+	+	+

Example 6.1 A 2^3 design was used to develop a nitride etch process on a single-wafer plasma etching tool. The design factors are the gap between the electrodes, the gas flow (C_2F_6 is used as the reactant gas), and the RF power applied to the cathode.

Table 6-4 The Plasma Etch Experiment, Example 6-1

	Coded Factors			Etch Rate			Factor Levels		
Run	A	B	C	Replicate 1	Replicate 2	Total		Low (-1)	High $(+1)$
1	-1	-1	-1	550	604	$(1) = 1154$	A (Gap, cm)	0.80	1.20
2	1	-1	-1	669	650	$a = 1319$	B (C_2F_6 flow, SCCM)	125	200
3	-1	1	-1	633	601	$b = 1234$	C (Power, W)	275	325
4	1	1	-1	642	635	$ab = 1277$			
5	-1	-1	1	1037	1052	$c = 2089$			
6	1	-1	1	749	868	$ac = 1617$			
7	-1	1	1	1075	1063	$bc = 2178$			
8	1	1	1	729	860	$abc = 1589$			

Table 6-5 Effect Estimate Summary for Example 6-1

Factor	Effect Estimate	Sum of Squares	Percent Contribution
A	-101.625	41,310.5625	7.7736
B	7.375	217.5625	0.0409
C	306.125	374,850.0625	70.5373
AB	-24.875	2475.0625	0.4657
AC	-153.625	94,402.5625	17.7642
BC	-2.125	18.0625	0.0034
ABC	5.625	126.5625	0.0238

Table 6-6 Analysis of Variance for the Plasma Etching Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
Gap (A)	41,310.5625	1	41,310.5625	18.34	0.0027
Gas flow (B)	217.5625	1	217.5625	0.10	0.7639
Power (C)	374,850.0625	1	374,850.0625	166.41	0.0001
AB	2475.0625	1	2475.0625	1.10	0.3252
AC	94,402.5625	1	94,402.5625	41.91	0.0002
BC	18.0625	1	18.0625	0.01	0.9308
ABC	126.5625	1	126.5625	0.06	0.8186
Error	18,020.5000	8	2252.5625		
Total	531,420.9375	15			

The Regression Model and response Surface. The regression model for predicting etch rate is

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_{13} x_1 x_3 = 776.0625 + \left(\frac{-101.625}{2}\right) x_1 + \left(\frac{306.125}{2}\right) x_3 + \left(\frac{-153.625}{2}\right) x_1 x_3$$

Computer Solution

$$SS_{model} = SS_A + SS_B + SS_C + SS_{AB} + SS_{AC} + SS_{BC} + SS_{ABC} = 5.134 \times 10^5$$

Thus, the statistic

$$F_0 = \frac{MS_{Model}}{MS_E} = \frac{73342.92}{2252.56} = 32.56$$

is testing the hypothesis

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_{12} = \beta_{13} = \beta 23 = \beta 123 = 0, \quad v.s. \quad H_1 : \text{ at least one} \beta = 0$$

$$R^2 = \frac{SS_{Model}}{SS_{Total}} = \frac{5.134 \times 10^5}{5.314 \times 10^5} = 0.9661$$

Model Coefficients – Full Model

Factor	Coefficient Estimated	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	776.06	1	11.87	748.70	803.42	
A-Gap	−50.81	1	11.87	−78.17	−23.45	1.00
B-Gas flow	3.69	1	11.87	−23.67	31.05	1.00
C-Power	153.06	1	11.87	125.70	180.42	1.00
AB	−12.44	1	11.87	−39.80	14.92	1.00
AC	−76.81	1	11.87	−104.17	−49.45	1.00
BC	−1.06	1	11.87	−28.42	26.30	1.00
ABC	2.81	1	11.87	−24.55	30.17	1.00

Model Coefficients – Reduced Model

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	776.06	1	10.42	753.35	798.77	
A-Gap	−50.81	1	10.42	−73.52	28.10	1.00
C-Power	153.06	1	10.42	130.35	175.77	1.00
AC	−76.81	1	10.42	−99.52	−54.10	1.00

Refine Model C Remove Nonsignificant Factors

Table 6-7 (continued)

Response: Etch rate
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	5.106E+005	3	1.702E+005	97.91	<0.0001
A	*41310.56*	*1*	*41310.56*	*23.77*	*0.0004*
C	*3.749E+005*	*1*	*3.749E+005*	*215.66*	*<0.0001*
AC	*94402.56*	*1*	*94402.56*	*54.31*	*<0.0001*
Residual	20857.75	12	1738.15		
Lack of Fit	*2837.25*	*4*	*709.31*	*0.31*	*0.8604*
Pure Error	*18020.50*	*8*	*2252.56*		
Cor Total	5.314E+005	15			

Std. Dev.	41.69	R-Squared	0.9608
Mean	776.06	Adj R-Squared	0.9509
C.V.	5.37	Pred R-Squared	0.9302
PRESS	37080.44	Adeq Precision	22.055

The Regression Model

Final Equation in Terms of Coded Factors:

Etch rate	=
+776.06	
−50.81	* A
+153.06	* C
−76.81	* A * C

Final Equation in Terms of Actual Factors:

Etch rate	=
−5415.37500	
+4354.68750	* Gap
+21.48500	* Power
−15.36250	* Gap * Power

Dispersion Effects

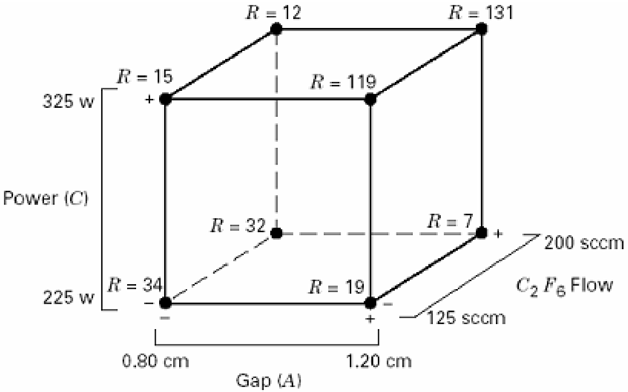


Figure 6-8 Ranges of etch rates for Example 6-1.

6.4 The General 2^k Design

The methods of analysis that we have presented thus far may be generalized to the case of a 2^k **factorial design**, that is, a design with k factors each at two levels. The statistical model for a 2^k design would include k main effects, $\binom{k}{2}$ two-factor interactions, $\binom{k}{3}$ three-factor interactions, ..., one k-factor interaction. The treatment combinations may be written in standard order by introducing the factors one at a time, with each new factor being successively combined with those that precede it. The standard order for a 2^4 design is (1), a, b, ab, c, ac, bc, abc, d, ad, bd, abd, cd, acd, bcd, and $abcd$.

We determine the contrast for effect $AB \cdots K$ by expanding the right-hand side of

$$\text{Contrast}_{AB\cdots K} = (a \pm 1)(b \pm 1) \cdots (k \pm 1)$$

The Effects and sums of squares can be computed as

$$AB \cdots K = \frac{2}{n2^k}(\text{Contrast}_{AB\cdots K})$$

and

$$SS_{AB\cdots K} = \frac{1}{n2^k}(\text{Contrast}_{AB\cdots K})^2.$$

6.5 A Single Replicate of the 2^k Design For even a moderate number of factors, the total number of treatment combinations in a 2^k factorial design is large. Because resources are usually limited, the number of replicates that the experimenter can em-ploy may be restricted. Frequently, available resources only allow a single replicate of the design to be run.

An obvious risk when conducting an experiment that has only one run at each test combination is that we may be fitting a model to noise.

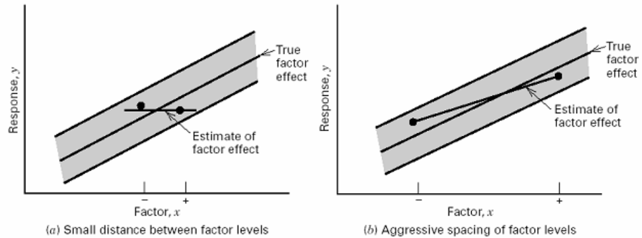


Figure 6-9 The impact of the choice of factor levels in an unreplicated design.

The single-replicate of a 2^k design is sometimes called an unreplicated factorial.

Example 6.2 A Single Replicate of the 2^4 Design

Table 6-10 Pilot Plant Filtration Rate Experiment

Run Number	Factor A	B	C	D	Run Label	Filtration Rate (gal/h)
1	−	−	−	−	(1)	45
2	+	−	−	−	*a*	71
3	−	+	−	−	*b*	48
4	+	+	−	−	*ab*	65
5	−	−	+	−	*c*	68
6	+	−	+	−	*ac*	60
7	−	+	+	−	*bc*	80
8	+	+	+	−	*abc*	65
9	−	−	−	+	*d*	43
10	+	−	−	+	*ad*	100
11	−	+	−	+	*bd*	45
12	+	+	−	+	*abd*	104
13	−	−	+	+	*cd*	75
14	+	−	+	+	*acd*	86
15	−	+	+	+	*bcd*	70
16	+	+	+	+	*abcd*	96

Example 6.2 A Single Replicate of the 2^4 Design

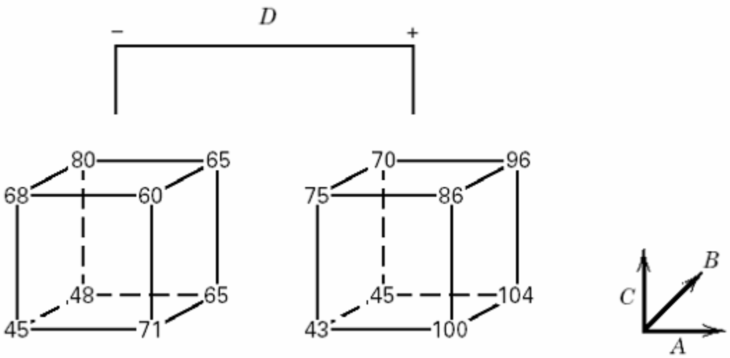


Figure 6-10 Data from the pilot plant filtration rate experiment for Example 6-2.

Example 6.2 A Single Replicate of the 2^4 Design

Table 6-12 Factor Effect Estimates and Sums of Squares for the 2^4 Factorial in Example 6-2

Model Term	Effect Estimate	Sum of Squares	Percent Contribution
A	21.625	1870.56	32.6397
B	3.125	39.0625	0.681608
C	9.875	390.062	6.80626
D	14.625	855.563	14.9288
AB	0.125	0.0625	0.00109057
AC	−18.125	1314.06	22.9293
AD	16.625	1105.56	19.2911
BC	2.375	22.5625	0.393696
BD	−0.375	0.5625	0.00981515
CD	−1.125	5.0625	0.0883363
ABC	1.875	14.0625	0.245379
ABD	4.125	68.0625	1.18763
ACD	−1.625	10.5625	0.184307
BCD	−2.625	27.5625	0.480942
$ABCD$	1.375	7.5625	0.131959

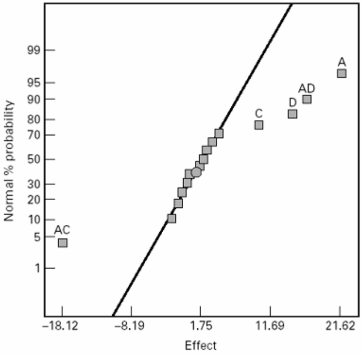


Figure 6-11 Normal probability plot of the effects for the 2^4 factorial in Example 6-2.

Example 6.2 A Single Replicate of the 2^4 Design

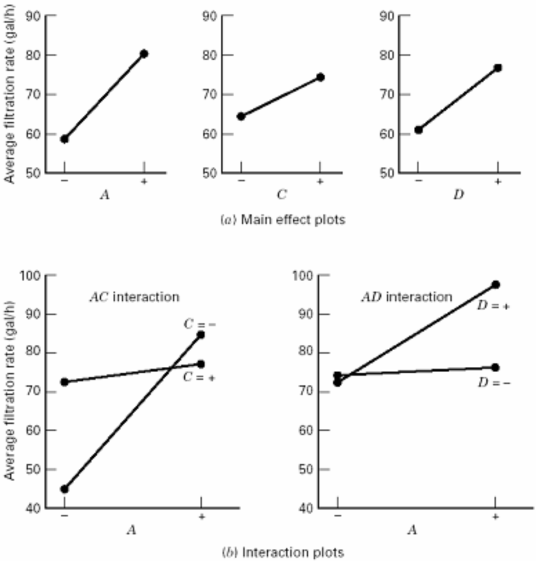


Figure 6-12 Main effect and interaction plots for Example 6-2.

Example 6.2 A Single Replicate of the 2^4 Design

Table 6-13 Analysis of Variance for the Pilot Plant Filtration Rate Experiment in A, C, and D

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
A	1870.56	1	1870.56	83.36	<0.0001
C	390.06	1	390.06	17.38	<0.0001
D	855.56	1	855.56	38.13	<0.0001
AC	1314.06	1	1314.06	58.56	<0.0001
AD	1105.56	1	1105.56	49.27	<0.0001
CD	5.06	1	5.06	<1	
ACD	10.56	1	10.56	<1	
Error	179.52	8	22.44		
Total	5730.94	15			

The Response Surface

$$\hat{y} = 70.06 + \left(\frac{21.625}{2}\right)x_1 + \left(\frac{9.875}{2}\right)x_3 + \left(\frac{14.625}{2}\right)x_4 - \left(\frac{18.125}{2}\right)x_1x_3 + \left(\frac{16.625}{2}\right)x_1x_4$$

Example 6.2 A Single Replicate of the 2^4 Design

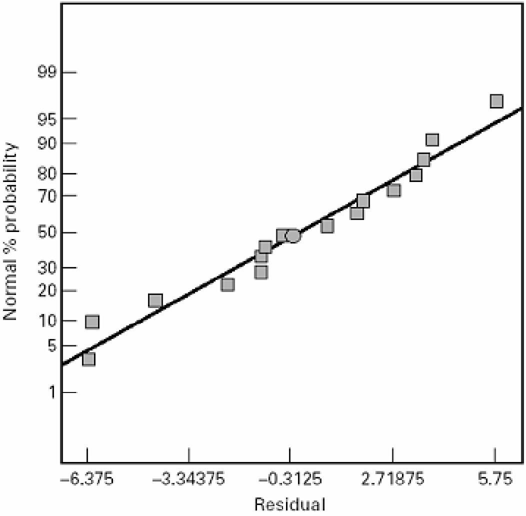


Figure 6-13 Normal probability plot of residuals for Example 6-2.

6.6 Additional Examples of Unreplicated 2^k Design

6.7 2^k Design are Optimal Designs

6.8 The Addition of Center Points to the 2^k Design