1/7

Examination 12-10-25 **Problem 1**

Single - factor experiment

Data and calculated mean values and effects:

Dosage (g)	Observations				$y_{i\bullet}$	$\bar{y}_{i\bullet}$	$\tau = \bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}$
20	24	28	37	30	119	29.75	-7.333
30	37	44	31	35	147	36.75	-0.333
40	42	47	52	38	179	44.75	7.667

A = dosage level **y = bioactivity**

a =3 **n=4** **N = a · n**

$\bar{y}_{\bullet\bullet}$ = 37.083

A.

Fixed effects model:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \qquad \begin{cases} i = 1, 2 \ldots\ldots a \\ j = 1, 2 \ldots\ldots n \end{cases}$$

$$N = a \cdot n$$

Sum of squares:

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{\cdot\cdot})^2 = \sum\sum y_{ij}^2 - \frac{y_{\cdot\cdot}^2}{N}$$

$$SS_A = n \cdot \sum_{i=1}^{a} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = \frac{1}{n} \sum y_{i\cdot}^2 - \frac{y_{\cdot\cdot}^2}{N}$$

$$SS_E = SS_T - SS_A$$

df: $N-a \quad N-1 \quad a-1$

Mean squares:

$$MS_A = \frac{SS_A}{a-1} \qquad MS_E = \frac{SS_E}{N-a}$$

Expected values:

$$E(MS_A) = \sigma^2 + \frac{n\sum \tau_i^2}{a-1}$$

$$E(MS_E) = \sigma^2$$

Hypothesis:

H_0: $\tau_1 = \tau_2 = \ldots \tau_a = 0$

H_a: $\tau_i \neq 0$ for at least one i

Test variable: $F_{obs} = MS_A / MS_E$

B. The significance of the difference between the pair of means,

$|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}|$, $j \neq i$

is examined with the LSD-method

$$LSD = t(1-\alpha/2, N-a) \cdot \sqrt{s^2(1/n + 1/n)}$$

Significance: $|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > LSD$

C. Residual analysis.

$$e_{ij} = y_{ij} - \hat{y}_{ij}, \quad i = 1 \dots 3, \quad j = 1 \dots 4$$

$$\hat{y}_{ij} = \hat{\mu} + \hat{\tau}_i$$

$$\hat{\mu} = \bar{y}_{\cdot\cdot}, \quad \hat{\tau}_i = \bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}$$

$$\hat{y}_{ij} = \bar{y}_{\cdot\cdot} + \bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot} = \bar{y}_{i\cdot}$$

$$e_{ij} = y_{ij} - \bar{y}_{i\cdot}$$

Calculations and conclusions:

see pages 4 – 7

Examination 12 10-25 Problem 1

A: ANOVA

Source	SS	df	MS	Fobs	sign	p
Dosage	450.67	2	225.33	7.0356	*	0.014
Residual	288.25	9	32.028			
Totcorr	738.92	11				

The dosage level has significant effect on bioactivity

B: Comparisons between the pairs of means with the LSD method

dosage 20-30: abs(y1m-y2m) = 7

20-40: abs(y1m-y3m) = 15

30-40: abs(y2m-y3m) = 8

LSD=t(0.975,N-a)*sqrt(MSE*(1/n+1/n)) = 9.05

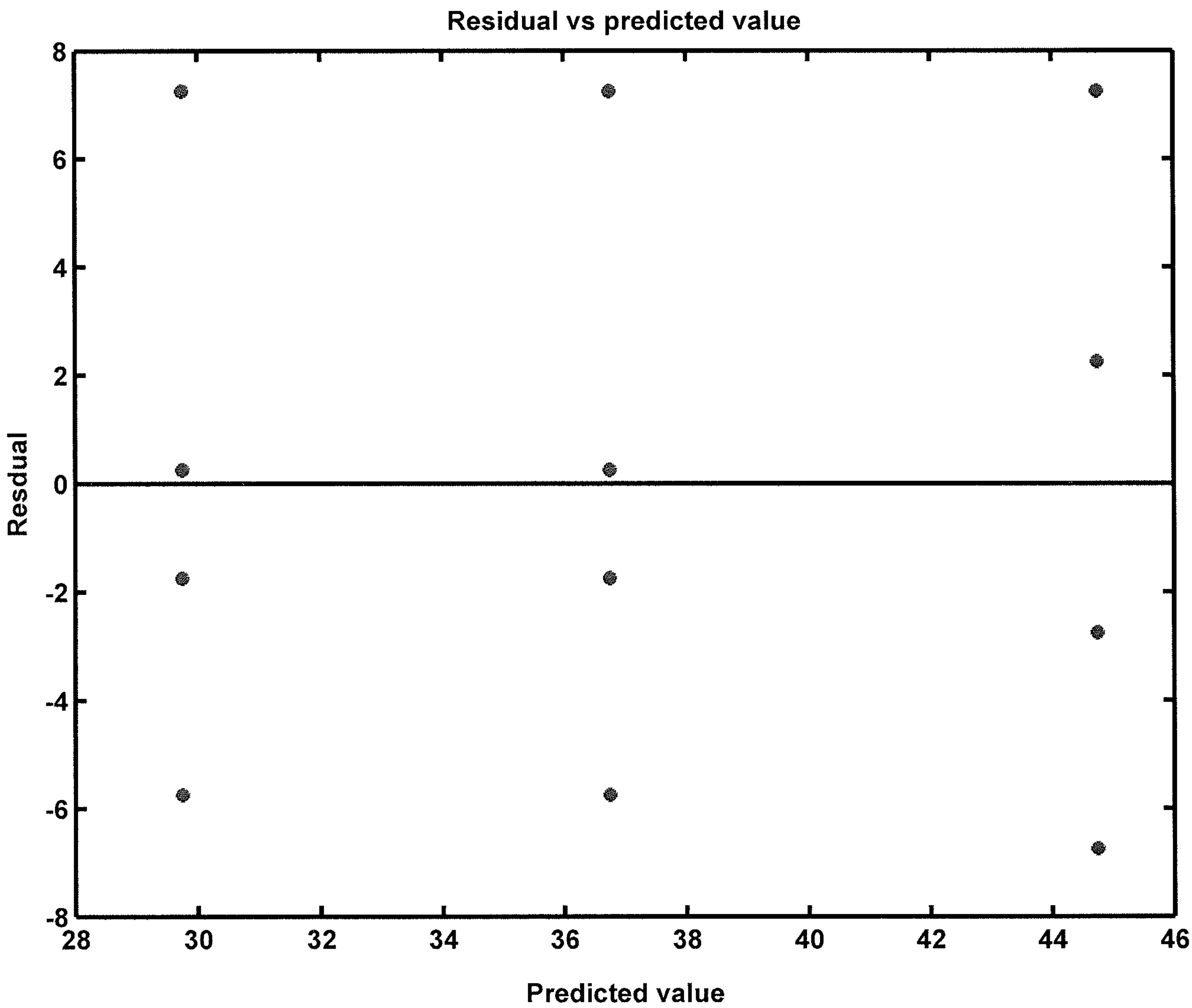
There is a differnce in the mean between the 20g and 40g dosages.

C: Residuals

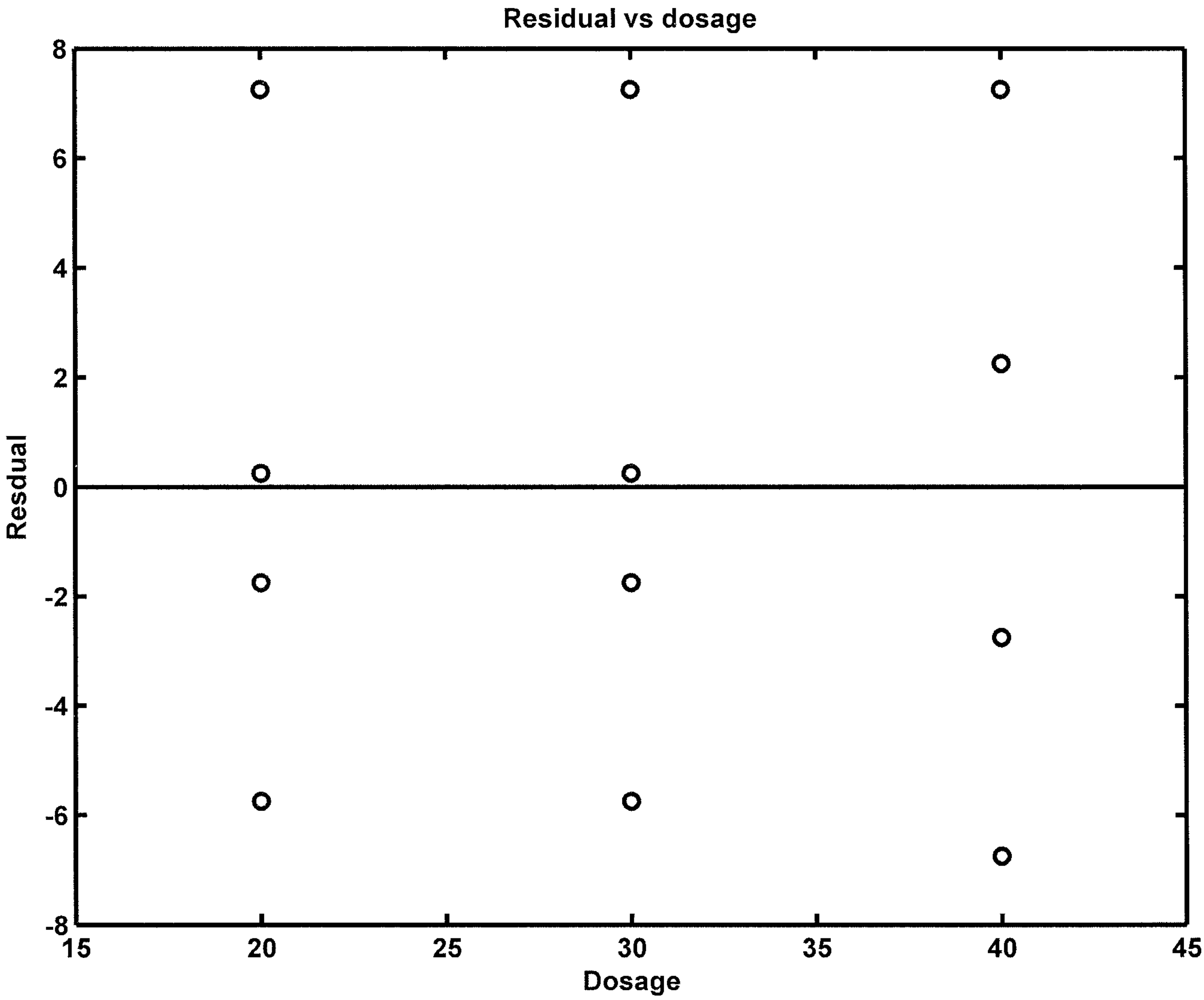
-5.75	-1.75	7.25	0.25
0.25	7.25	-5.75	-1.75
-2.75	2.25	7.25	-6.75

The residuals are plotted vs predicted value,and dosage, and also as a "normal plot".

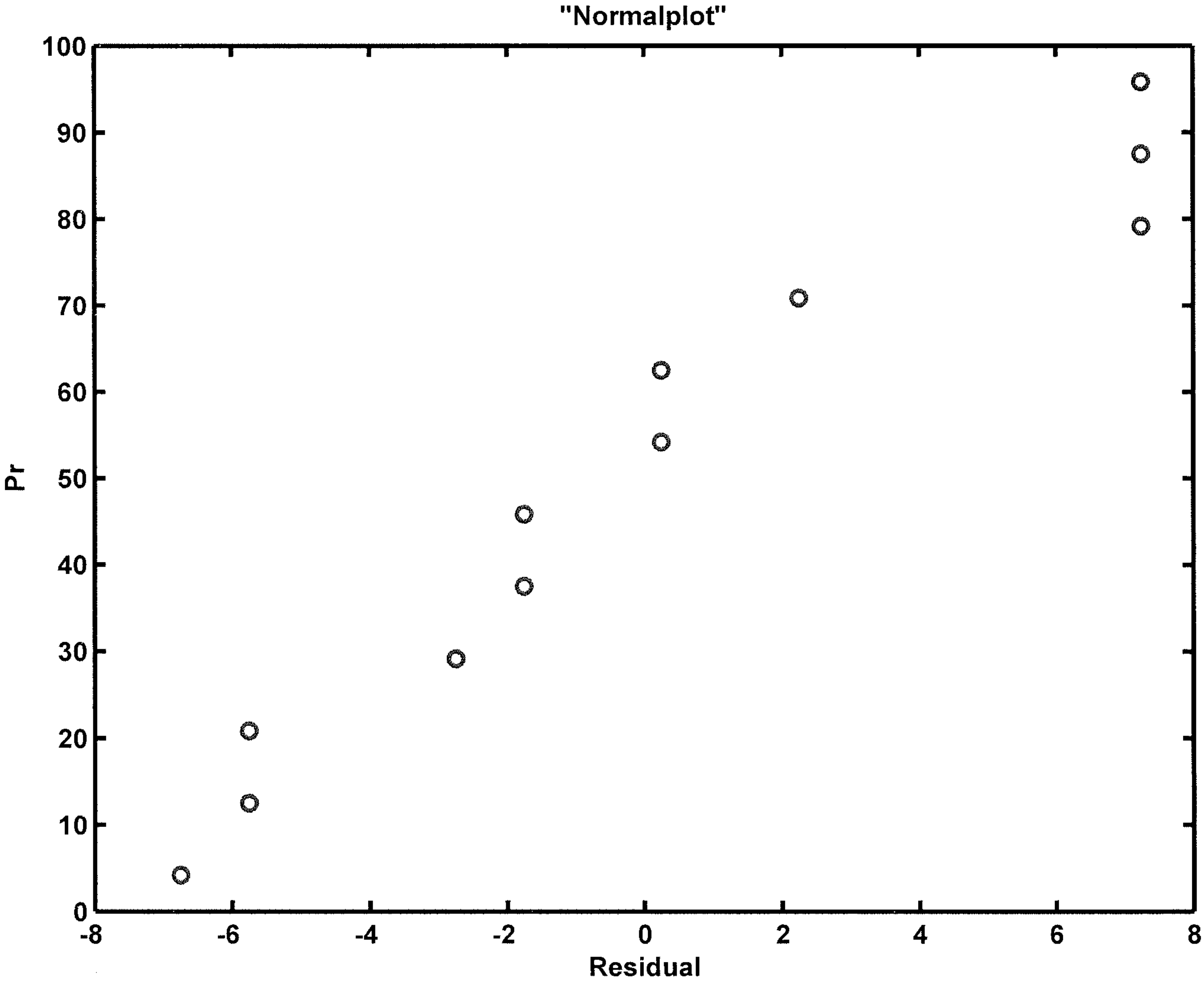
There is nothing too unusual about the residual plots.


Residual vs predicted value
Resdual
Predicted value
8
6
4
2
0
-2
-4
-6
-8
28
30
32
34
36
38
40
42
44
46

6/7



Residual vs dosage
Resdual
Dosage
8
6
4
2
0
-2
-4
-6
-8
15
20
25
30
35
40
45


"Normalplot"
Pr
Residual
100
90
80
70
60
50
40
30
20
10
0
-8
-6
-4
-2
0
2
4
6
8

Problem 2.

Factors: solution (row)
days (column)

Response (y): bacteria growth

Model: $y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}$

τ_i = row effect (solution) $i = 1, 2 .. a$
$a = 3$

β_j = column effect (days), $j = 1, 2 .. b$
$b = 4$

μ = overall mean
ε_{ij} = random error $NID(0, \sigma^2)$

Experimental data, mean values and effects, see page 3

Sum of squares:

$SS_T = \sum\sum(y_{ij} - \bar{y}_{..})^2 = \sum\sum y_{ij}^2 - y_{..}^2/N$

$SS_T = 1862.3$

$SS\,(\text{solution}) = b \cdot \sum(\bar{y}_{i.} - \bar{y}_{..})^2 = \frac{1}{b}\sum y_{i.}^2 - y_{..}^2/N$

$SS\,(\text{solution}) = 703.5$

$SS\,(\text{days}) = a \cdot \sum(\bar{y}_{.j} - y_{..})^2 = \frac{1}{a}\sum y_{.j}^2 - y_{..}^2/N$

$SS\,(\text{days}) = 1862.3$

$SS_E = SS_T - SS(\text{solution}) - SS(\text{days})$

$SS_E = 51.83$

ANOVA - table and comments,
see page 4.

Examination 12_10_25 Problem 2

Randomized block design

Data and calculated mean values and parameters

Solution	Days						
	1	2	3	4	$y_{i\bullet}$	$\bar{y}_{i\bullet}$	$\tau = \bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}$
I	13	22	18	39	92	23	4.25
II	16	24	17	44	101	25.25	6.5
III	5	4	1	22	32	8	-10.75
$y_{\bullet j}$	34	50	36	105			
$\bar{y}_{\bullet j}$	11.33	16.67	12	35			
$\beta_j = \bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet}$	-7.42	-2.08	-6.75	16.25			

A = solution B = days

$a = 3$ $b = 4$ $\bar{y}_{\bullet\bullet} = 18.75$

Model parameters : $\mu = 18.75$ $\tau = [4.25 \;\; 6.5 \;\; -10.75]$ $\beta = [-7.42 \;\; -2.08 \;\; -6.75 \;\; 16.25]$

3/4

4/4

Examination 12-10-25 Problem 2

A: ANOVA

Source	SS	df	MS	Fobs	sign
Solution	703.5	2	351.75	40.72	*
Day	1106.9	3	368.97	42.71	*
Residual	51.83	6	8.64		
Totcorr	1862.3	11			

Fstatsolution=F(1-alfa,a-1,(a-1)*(b-1)) = 5.14

Fstatday=F(1-alfa,b-1,(a-1)*(b-1)) = 4.75

Solutions and days have a significant effect on retarding bacteria growth.

Comparisons between the pairs of solution means with the LSD method

solution I-II: abs(y1m-y2m) = 2.25

I-III: abs(y1m-y3m) = 15

II-III: abs(y2m-y3m) = 17.25

LSD=t(0.975,(a-1)*(b-1))*sqrt(MSE*2/b)´= 5.09

The LSD procedure indicates that solution III is significantly different from the other two.

Problem 3

Model: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 +$
$+ \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \varepsilon$

A. Model fitting

$\underline{b} = (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \cdot \underline{\underline{X}}^T \underline{y}$

$\underline{\underline{X}}$ = calculation matrix

$\underline{y}$ = response vector

Significance:

I. t - test. $t_{obs_j} > t_{stat}$

$$t_{obs_j} = \frac{b_j}{se(b_j)}$$

$$t_{stat} = t(0.975, N-P)$$

II. F - test $F_{obs_j} > F_{stat}$

$$F_{obs_j} = \frac{SS(b_j)/1}{SS_E/(N-P)}$$

$$F_{stat} = F(0.95, 1, N-P)$$

$t^2_{obs_j} = F_{obs_j}$ $\qquad t^2_{stat} = F_{stat}$

B. Lack of fit,

$$SS_E = SS_{PE} + SS_{LOF}$$

Replicates in the center point give SS_{PE}:

$$SS_{PE} = \sum_{i=1}^{n_c} (y_{ci} - \bar{y}_c)^2$$

n_c = number of replicates.

$$df_{PE} = n_c - 1$$

$$SS_{LOF} = SS_E - SS_{PE}$$

$$df_{LOF} = N - p - (n_c - 1)$$

Test:

$$F_{LOF} = \frac{SS_{LOF} / (N - p - n_c + 1)}{SS_{PE} / (n_c - 1)}$$

$$F_{stat} = F(0.95, N - p - n_c + 1, n_c - 1)$$

No lack of fit if $F_{LOF} < F_{stat}$

C: Model with pure quadratic terms.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \underbrace{\beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2} + \varepsilon$$

With the given data only the sum of the parameters for the pure quadratic terms can be estimated

$$b_{11} + b_{22} + b_{33} = b_{quad} = \bar{y}_f - \bar{y}_c$$

$\bar{y}_f$ = mean value of the responses in the factorial points

n_f = number of factorial points

$\bar{y}_c$ = mean value of responses in the center point.

$$SS_{purequad} = \frac{n_f \cdot n_c (\bar{y}_f - \bar{y}_c)^2}{n_f + n_c}$$

with $df = 1$

$$F_{obs} = \frac{SS_{purequad}/1}{SS_E/(N-p)}$$

$$F_{stat} = F(0.95, 1, N-p)$$

Significance if $F_{obs} > F_{stat}$.

Summary of all calculations, (A, B, C) with conclusions; see pages 5 and 6

D. Fitting a second-order model.

Complementary addition of 2 k axial points to a 2^k-design to form a central composite design (CCD)

Add here 2 × 3 axial points to the 2^3-design at a distance α from the design center:

The design is made rotatable by setting $\alpha = (n_f)^{1/4} = 8^{1/4} = 1.68$.

To get a low and almost constant prediction variance for the CCD the best choice is
$\alpha = \sqrt{3} = 1.73$, which gives a spherical CCD.

The axial points are added as:

$$
\begin{array}{ccc}
x_1 & x_2 & x_3 \\
\end{array}
$$
$$
\begin{bmatrix}
-\alpha & 0 & 0 \\
\alpha & 0 & 0 \\
0 & -\alpha & 0 \\
0 & \alpha & 0 \\
0 & 0 & -\alpha \\
0 & 0 & \alpha
\end{bmatrix}
$$

Examination 12-10 25 Poblem 3

Calculation matrix and observations

		x0	x1	x2	x3	x1x2	x1x3	x2x3
exp	obs	M	A	B	C	AB	AC	BC
(1)	10	1	-1	-1	-1	1	1	1
a	12	1	1	-1	-1	-1	-1	1
b	4	1	-1	1	-1	-1	1	-1
ab	4	1	1	1	-1	1	-1	-1
c	9	1	-1	-1	1	1	-1	-1
ac	10	1	1	-1	1	-1	1	-1
bc	5	1	-1	1	1	-1	-1	1
abc	4	1	1	1	1	1	1	1
center	6	1	0	0	0	0	0	0
center	7	1	0	0	0	0	0	0
center	8	1	0	0	0	0	0	0
center	7	1	0	0	0	0	0	0

A: ANOVA

Source	Param	SS	df	MS	Fobs	sign
Model		77.5	6	12.92	29.81	*
b0	7.167	616.3	1	616.3	1422.3	*
b1	0.25	0.5	1	0.5	1.154	
b2	-3	72	1	72	166.2	*
b3	-0.25	0.5	1	0.5	1.154	
b1b2	-0.5	2	1	2	4.615	
b1b3	-0.25	0.5	1	0.5	1.154	
b2b3	0.5	2	1	2	4.615	
Residual		2.167	5	0.4333		
Stotcorr		79.67	11			

Fstatmod = F(0.95,p-1,N-p) = 4.95 Fstatparam = F(0.95,1,N-p) = 6.61

R2 = 0.9728 R2adj = 0.9402

B: Lack of fit

SSPE = 2   SSLOF = 0.1667   FLOF = 0.125

FstatLOF = F(0.95,N-p-nc+1,nc-1) = 9.55   No lack of fit

C: Pure quadratic terms

ycm = 7   yfm = 7.25   nc = 4   nf = 8

bquad=0.25   SSquad = 0.167   Fquad = 0.385

Fstatquad = F(0.95,1,N-p) = 6.61   No significance

Problem 4

A. A simplex is a regular sided figure.

Here with 3 variables the design must be a regular tetrahedron.

Proof: Equal distances between the vertices.

Coordinates for the vertices:

$z_1 = [0 \quad \sqrt{2} \quad -1]$ $z_3 = [0 \quad -\sqrt{2} \quad -1]$

$z_2 = [-\sqrt{2} \quad 0 \quad 1]$ $z_4 = [\sqrt{2} \quad 0 \quad 1]$

Distances between two vertices:

$d_{ij} = || z_i - z_j ||$, $i = 1, 2, 3$, $j = 2, 3, 4$
$i \neq j$

$d_{ij} = 2\sqrt{2}$ for all i, j

Alternative:

Equal distances from centre point to vertices.

$d_{ic} = || z_i ||$ $i = 1, 2 \ldots 4$

$d_i = \sqrt{3}$ for all i;

B. Fitting first order model.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + b_3 x_3 + \varepsilon$$

$$\underline{\underline{X}} = \begin{bmatrix} 1 & 6 & \sqrt{2} & -1 \\ 1 & -\sqrt{2} & 0 & 1 \\ 1 & 0 & -\sqrt{2} & -1 \\ 1 & \sqrt{2} & 0 & 1 \end{bmatrix} \qquad \underline{y} = \begin{bmatrix} 18.5 \\ 19.8 \\ 17.4 \\ 22.5 \end{bmatrix}$$

$$\underline{b} = (\underline{\underline{X}}^T \underline{\underline{X}})^{-1} \underline{\underline{X}}^T \underline{y}$$

$$\underline{b} = \begin{bmatrix} 19.55 \\ 0.9546 \\ 0.3889 \\ 1.6 \end{bmatrix}$$

C. Path of steepest ascent

Path of steepest ascent = the gradient direction.

$$\underline{g} = \nabla y = \begin{bmatrix} \frac{\partial y}{\partial x_1} & \frac{\partial y}{\partial x_2} & \frac{\partial y}{\partial x_3} \end{bmatrix}^T = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

The steps along the path are proportional to the parameters $\underline{b}$

$$\underline{\Delta x} = \text{stepfactor} \cdot k \cdot \underline{g}$$

$$\underline{x}_{new} = \underline{x}_0 + \underline{\Delta x}$$

x_3 = konstant = 1 $\qquad x_0 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$\underline{g} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0.9546 \\ 0.3889 \end{bmatrix}$$

$$k = 1/\max[\mathrm{abs}(g)] = 1/\mathrm{abs}(g_1) = \frac{1}{0.9546}$$

$$\underline{\Delta x} = \text{stepfactor} \cdot \begin{bmatrix} 1 \\ 0.4074 \end{bmatrix}$$

$$\underline{x}_{new} = \underline{x}_0 + \text{stepfactor} \cdot \begin{bmatrix} 1 \\ 0.4074 \end{bmatrix}$$

An overall step is taken to the final point.

$x_{1max} = \sqrt{2} \implies$ stepfactor $= \sqrt{2}$

$$\underline{x}_{final} = \underline{x}_0 + \sqrt{2} \cdot \begin{bmatrix} 1 \\ 0.4074 \end{bmatrix} = \begin{bmatrix} 1{,}414 \\ 0.576 \end{bmatrix}$$

$x_2 = 0.576$

$y_0 = 21.15 \qquad y_{final} = 22.72$

Remarks

(not included in the problem statement)

Steps, which is taken in the gradient direction from the starting point $\underline{x}_0 = [0\ 0\ 1]^T$

at once result in process variables outside the simplex region.

To avoid this, variable x_2 must be kept constant, $x_2 = 0$

$x_{limit} = [\sqrt{2}\ \ 0\ \ 1]^T$

$y_{limit} = 22.5$

The gradient and limit paths are shown in fig. 1 and 2.

5/5

Problem 4

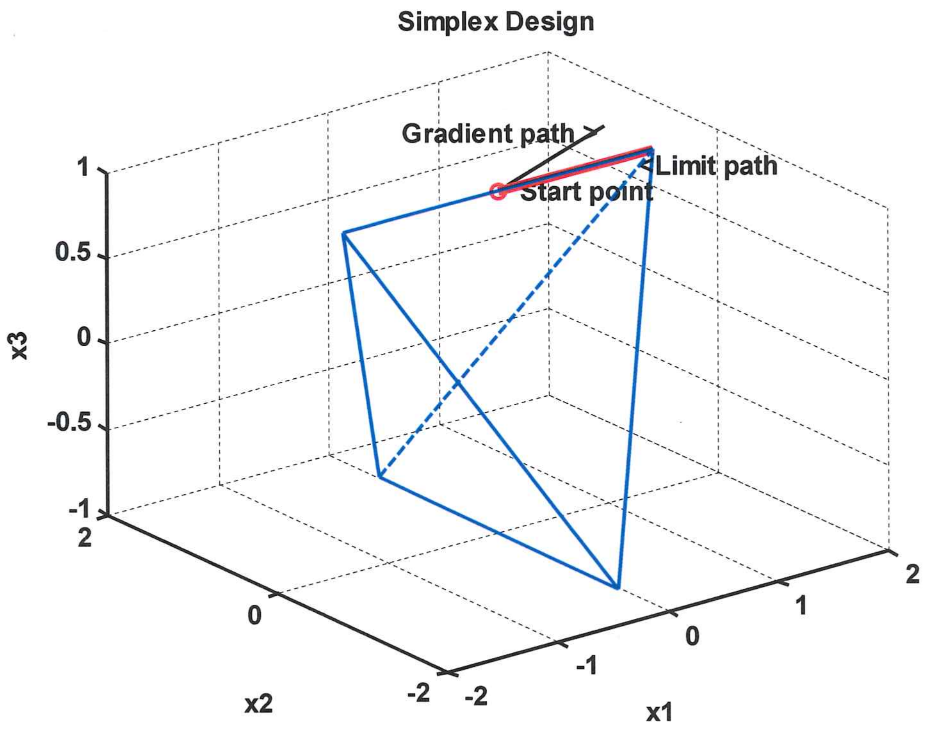


Fig.1 Regular tetrahedron

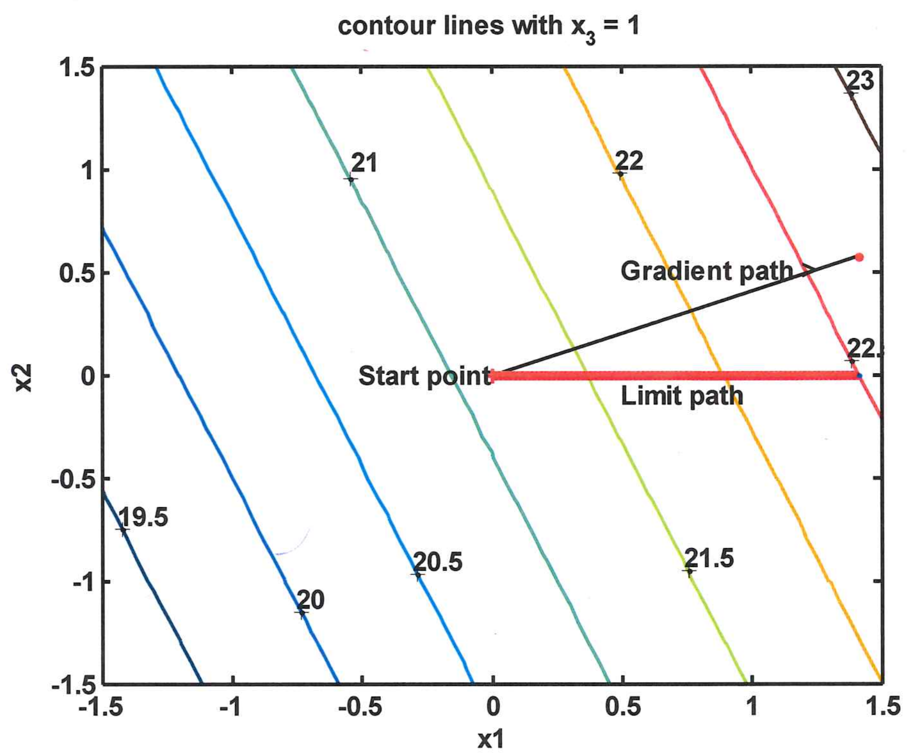


Fig.2 Contour plot

Problem 5

A: Alias structure

2^{3-1} fractional factorial design.

C is confounded with AB, C = AB

Generator: ABC

Defining relation for the reduced plan:

I = ABC ; resolution: III

Alias:

[A] → A + BC

[B] → B + AC

[C] → C + AB

B and C tasks:

ANOVA - table and conclusions, see page 3

D. Final model in terms of coded factors:

$y = b_0 + b_1 x_A + b_2 x_B + b_3 x_c$

$y = 778.88 - 51.88\, x_A - 73.13\, x_B + 140.63\, x_3$

Coding:

$x_A = \dfrac{A - 1.00}{0.20}$ A: 0.80 - 1.20

$x_A = 5A - 5$

$$x_B = \frac{B - 162.5}{37.5}$$ B: 125–200

$$x_B = \frac{B}{37.5} - \frac{162.5}{37.5}$$

$$x_C = \frac{C - 262.5}{62.5}$$ C: 200–325

$$x_C = \frac{C}{62.5} - \frac{262.5}{62.5}$$

The expressions for x_A, x_B and x_C are inserted in the model equation above and give the etch rate in terms of the actual factors.

$$y = 764.51 - 259.38\,A - 1.95\,B + 2.25\,C$$

Examination 12_10_25 Problem 5

Design for two replicates

exp	obs	M	A	B	C=AB
c	1037	1	-1	-1	1
a	669	1	1	-1	-1
b	633	1	-1	1	-1
abc	729	1	1	1	1
c	1052	1	-1	-1	1
a	650	1	1	-1	-1
b	601	1	-1	1	-1
abc	860	1	1	1	1

ANOVA

Source	Param	SS	df	MS	Fobs	sign
Model		222509.4	3	74169.79	31.61	*
A	-51.875	21528.13	1	21528.13	9.18	*
B	-73.125	42778.13	1	42778.13	18.23	*
C=AB	140.625	158203.1	1	158203.1	67.42	*
Residual		9385.5	4	2 346.37		
Totcorr		231894.9	7			

intercept: b0 = 778.875

SSReg=SS(A) + SS(B) + SS(C) = 222509.4 df = 3

SSE = SST(corr) - SSReg = 9385.5 SSE = SSPE

parameter standard error: se = 17.13

Fstatmodel = F(0.95,3,4) = 6.59

Fstatparam = F(0.95,1,4) = 7.71

R2 = 0.95953 R2adj = 0.9291

Model in terms of coded variables: y = b0 + b1xA + b2xB + b3xC

It is not unexpected that factor C is significant, since it is confounded with AB.

25/10 6 $E = ABD$ ABDE = GENERATOR

A	B	C	D	E = ABD	Exp
−	−	−	−	−	(1)
+	−	−	−	+	ae
−	+	−	−	+	be
+	+	−	−	−	ab
−	−	+	−	−	c
+	−	+	−	+	ace
−	+	+	−	+	bce
+	+	+	−	−	abc
−	−	−	+	+	de
+	−	−	+	−	ad
−	+	−	+	−	bd
+	+	−	+	+	abde
−	−	+	+	+	cde
+	−	+	+	−	acd
−	+	+	+	−	bcd
+	+	+	+	+	abcde

I = ABDE

A − BDE
B − ADE
C − ABDEC
D − ABE
E − ABD
AB − DE ⟷ DE − AB
AC − BCDE
AD − BE
AE − BD
BC − ACDE
CD = ABCE
CE = ABCD

DISADVANTAGE 2 ORDER ALIAS - 2 ORDER

RESOLUTION $\underline{IV}$

FOLDOVER   BYT TECKE PÅ MINSTA TYDLIGA EFFEKT

SEPARATE AB - DE, BD - AE   BYT TECKEN PÅ AD (E)

OR NEW 2^{5-1} CORRECT GENERATOR ABCDE

SEPARATE 2 ORDER EFFECTS

CHALMERS OF COURSE

CALTECH
MIT
ETH
IMPERIAL
UCLA

⇒ WILL ALSO DO