Exam 13-01-17

Problem 1

Solution 1: $\bar{y}_1 = 9.950$ $s_1^2 = 0.20286$

$s_1 = 0.4504$ $n_1 = 8$

Solution 2: $\bar{y}_2 = 10.363$ $s_2^2 = 0.05411$

$s_2 = 0.23261$ $n_2 = 8$

A: Test if the variances are equal.

$H_0: \sigma_1^2 = \sigma_2^2$

$H_a: \sigma_1^2 \neq \sigma_2^2$

Test variable: $F_{obs} = \dfrac{s_1^2}{s_2^2}$

Reject H_0 if

$F_{obs} > F(0.975, n_1-1, n_2-1) = F_1$

or

$F_{obs} < F(0.025, n_1-1, n_2-1) = F_2$

$F(0.025, n_1-1, n_2-1) = 1/F(0.975, n_2-1, n_1-1)$

$F_{obs} = 3.75$

$F_1 = 4.99$ $F_2 = 0.20$

don't reject H_0

B: Are the mean values equal?

Hypotheses: $H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \neq \mu_2$

Test variable: $t_{obs} = \dfrac{|\bar{y}_1 - \bar{y}_2|}{s \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

Pooling of variances:

$$s^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 - 1 + n_2 - 1}$$

$$s^2 = \frac{7(s_1^2 + s_2^2)}{14} = 0.5(s_1^2 + s_2^2)$$

$s^2 = 0.1285$ $s = 0.3584$

Degrees of freedom: $\nu = 14$

$$t_{obs} = \frac{|9.950 - 10.3631|}{\sqrt{0.1285 \cdot \frac{2}{8}}} = 2.3016$$

P-value: 0.037

$t_{stat} = t(0.975, 14) = 2.14$

Conclusion: The solutions have **not** the same etch rate.

Examination 13-01-17 **Problem 2**

Nested design

Machine (i) a=3	I			II			III		
Station (j) b=3	1	2	3	1	2	3	1	2	3
Yield (k) n=2	35	34	37	31	33	32	31	33	32
	32	36	38	34	34	35	32	31	30
$y_{ij\cdot}$	67	70	75	65	67	67	63	64	62
$y_{i\cdot\cdot}$	212			199			189		
$y_{\cdot\cdot\cdot}$	600								

1/7

Examination 13-01-17 **Problem 2**

Nested design

Machine (i) a=3	**I**			**II**			**III**		
Station (j) b=3	1	2	3	1	2	3	1	2	3
Yield (k) n=2	35	34	37	31	33	32	31	33	32
	32	36	38	34	34	35	32	31	30
mean($y_{ij\cdot}$)	33.5	35	37.5	32.5	33.5	33.5	31.5	32	31
mean($y_{i\cdot\cdot}$)	35.333			33.167			31.5		
mean($y_{\cdot\cdot\cdot}$)	33.333								
mean($y_{ij\cdot}$) - mean($y_{i\cdot\cdot}$) = $\beta_{j(i)}$	-1.833	-0.333	2.167	-0.667	0.333	0.333	0	0.5	-0.5
mean($y_{i\cdot\cdot}$) - mean($y_{\cdot\cdot\cdot}$) = τ_i	2.0			-0.167			-1.833		
residual	1.5	-1	-0.5	-1.5	-0.5	-1.5	-0.5	1	1
y_{ijk} - mean($y_{ij\cdot}$)	-1.5	1	0.5	1.5	0.5	1.5	0.5	-1	-1

2/7

A: sum of squares

$$SS_A = \frac{1}{bn}\sum^{a} y_{i..}^2 - \frac{y^2_{...}}{abn}$$ A = machine, B = Station

$$SS_{B(A)} = \frac{1}{n}\sum^{a}\sum^{b} y_{ij.}^2 - \frac{1}{bn}\sum^{a} y_{i..}^2$$

$$SS_T = \sum^{a}\sum^{b}\sum^{n} y^2_{ijk} - \frac{y^2_{...}}{abn}$$

$$SS_E = \sum^{a}\sum^{b}\sum^{n} y^2_{ijk} - \frac{1}{n}\sum^{a}\sum^{b} y_{ij.}^2$$

Mean squares:

$$MS_A = \frac{SS_A}{a-1} \qquad MS_{B(A)} = \frac{SS_{B(A)}}{a(b-1)}$$

$$MS_E = \frac{SS_E}{ab(n-1)}$$

Expected mean squares:

A and B fixed

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

$$E(MS_{B(A)}) = \sigma^2 + \frac{n\sum\sum \beta_{j(i)}^2}{a(b-1)}$$

H_0: $\tau_i = 0$ is tested by $F_{obs} = \frac{MS_A}{MSE}$

no difference between machines

H_0: $\beta_{j(i)} = 0$ is tested by $F_{obs} = \frac{MS_{B(A)}}{MSE}$

no difference between stations.

ANOVA

Source	SS	df	MS	F_{obs}
A (machine)	44.33	2	22.17	9.50 ×
B (station)	18.67	6	3.11	1.33
Error	21.00	9	2.33	
Total (corr)	84	17		

$F(0.95, 2, 9) = 4.26$ $\alpha = 0.05$ P-value: 0.0061

$F(0.95, 6, 9) = 3.37$ P-value: 0.335

The factor Machine has a significant effect on the yield.

The choice of machine is important.

B. Significant difference between the machines is examined by the LSD-method or by t-test.

$$LSD = t_{stat} \cdot \sqrt{s^2\left(\frac{1}{bn} + \frac{1}{bn}\right)}$$

$$t_{stat} = t(0.975, ab(n-1)) =$$

$$= t(0.975, 9) = 2.262$$

$$LSD = 2.262 \cdot \sqrt{2.333\left(\frac{1}{6} + \frac{1}{6}\right)} = 1.995$$

$$\Delta \bar{y}_{i,i+1} = \left|\bar{y}_{i\cdot\cdot} - \bar{y}_{i+1\cdot\cdot}\right| = \left|\tau_i - \tau_{i+1}\right|$$

Significant difference if

$$\Delta \bar{y}_{i,i+1} > LSD$$

Alt: t-test

$$t_{obs} = \frac{\left|\bar{y}_{i\cdot\cdot} - \bar{y}_{i+1\cdot\cdot}\right|}{\sqrt{s^2\left(\frac{1}{bn} + \frac{1}{bn}\right)}}$$

Significant difference if $t_{obs} > t_{stat}$

machine	mean. diff	t_{obs}
I – II	2.167	2.46
I – III	3.833	4.35
II – III	1.666	1.89

Machine I is preferable; best combined with station 3

Diagnostic checking

Residual analysis:

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \varepsilon_{(ij)k}$$

$$e_{ijk} = y_{ijk} - \hat{y}_{ijk}$$

$$\hat{y}_{ijk} = \hat{\mu} + \hat{\tau}_i + \hat{\beta}_{j(i)}$$

$$\Sigma \hat{\tau}_i = 0 \quad \Sigma \beta_{j(i)} = 0 \qquad i = 1, 2 \ldots a$$

$$\hat{\mu} = \bar{y}_{...} \qquad \hat{\tau}_i = \bar{y}_{i..} - \bar{y}_{...}$$

$$\hat{\beta}_{j(i)} = \bar{y}_{ij.} - \bar{y}_{i..}$$

$$\hat{y}_{ijk} = \bar{y}_{...} + (\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{ij.} - \bar{y}_{i..}) =$$

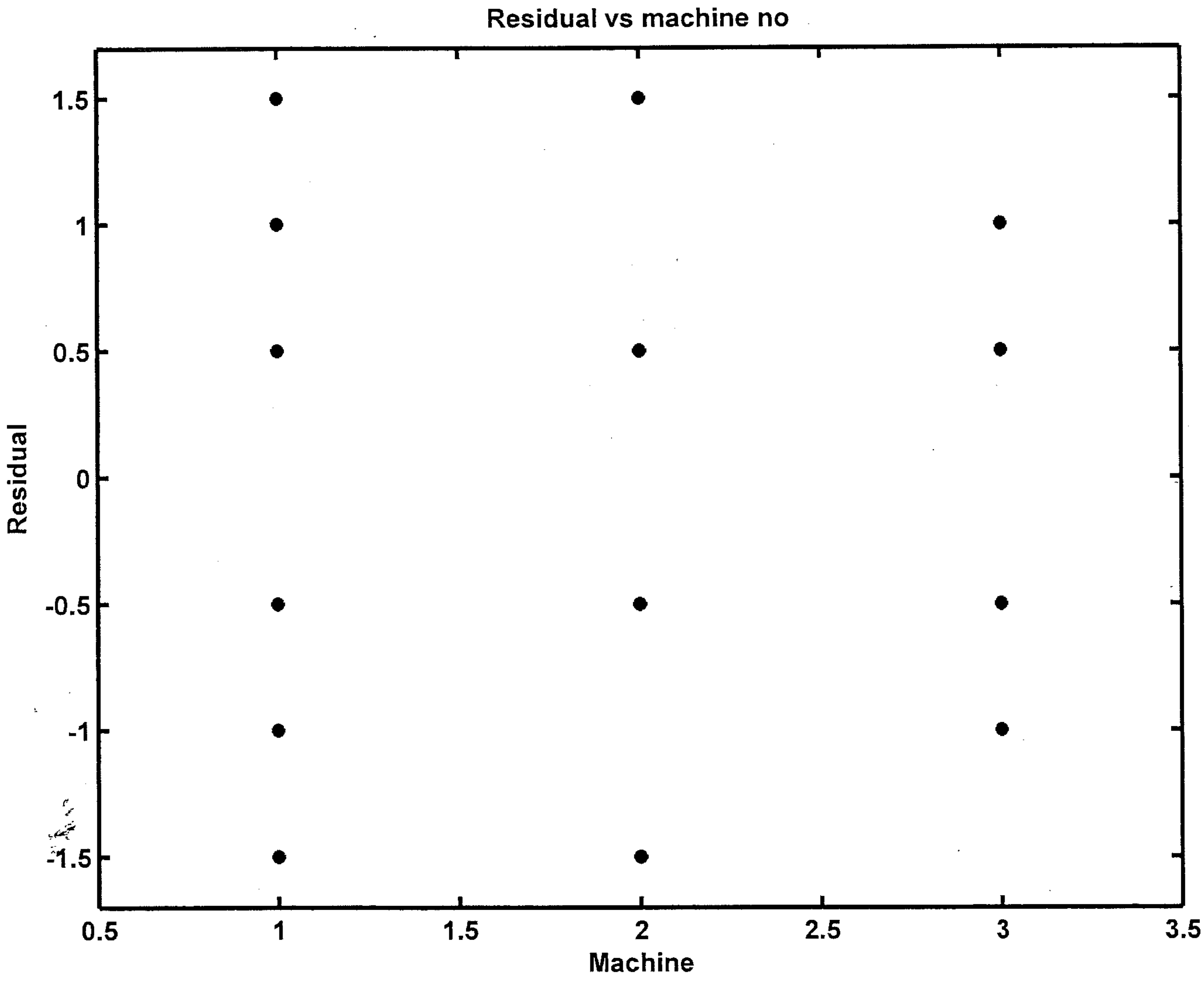
$$= \bar{y}_{ij.}$$

$\bar{y}_{ij.}$ = individual station averages

$$e_{ijk} = y_{ijk} - \bar{y}_{ij.}$$

See graf. e_{ijk} vs station number.

Smallest spread in residuals for station 3, but large negative effect on yield.


Residual vs machine no
Residual
1.5
1
0.5
0
-0.5
-1
-1.5
0.5
1
1.5
2
2.5
3
3.5
Machine

Problem 3

A fractional factorial design has been used to study the effect of five factors on the free height of leaf springs with 16 runs.

Inspection of the design matrix gives that D is confounded with ABC, $D \rightarrow ABC$

Generator: ABCD

Defining relation: $I = ABCD$

Resolution: IV

Alias structures:

$[A] = A + BCD$

$[B] = B + ACD$

$[C] = C + ABD$

$[D] = D + ABC$

$[E] = E + ABCDE$

$[AB] = AB + CD$

$[AC] = AC + BD$

$[AD] = AD + BC$

$[AE] = AE + BCDE$

$[BC] = BC + AD$

$[BD] = BD + AC$

$[BE] = BE + ACDE$

$[CD] = CD + AB$

$[CE] = CE + ABDE$

$[DE] = DE + ABCE$

Calculations: see pages 3 and 4

B: Significant factors: A, B, E, AD (BC), BE

C: Residual plot: somewhat funnel shaped.

D: Resolution IV design has the drawback that two-factor interaction are aliased with each other.

It is better to use a 2^{5-1} design with resolution V, e.g. $D \rightarrow ABCE$

3/4

Examination 13-01-17 Problem 3

Calculation matrix and observations

exp	obs	M	A	B	C	D	E	AB	AC	AD	AE	BE	CE	DE
								CD	BD	BC				
(1)	7.81	1	-1	-1	-1	-1	-1	1	1	1	1	1	1	1
ad	7.88	1	1	-1	-1	1	-1	-1	-1	1	-1	1	1	-1
bd	7.50	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1
ab	7.75	1	1	1	-1	-1	-1	1	-1	-1	-1	-1	1	1
cd	7.88	1	-1	-1	1	1	-1	1	-1	-1	1	1	-1	-1
ac	8.06	1	1	-1	1	-1	-1	-1	1	-1	-1	1	-1	1
bc	7.44	1	-1	1	1	-1	-1	-1	-1	1	1	-1	-1	1
abcd	7.69	1	1	1	1	1	-1	1	1	1	-1	-1	-1	-1
e	7.12	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1
ade	7.44	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1
bde	7.50	1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1
abe	7.56	1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1
cde	7.44	1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1
ace	7.62	1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1
bce	7.25	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1
abcde	7.59	1	1	1	1	1	1	1	1	1	1	1	1	1

ANOVA

Source	Param	SS	df	MS	Fobs	sign
Model		0.8725	9	0.09695	26.33	*
M	7.596	923.1	1	923.1	2.508e+005	*
A	0.1031	0.1702	1	0.1702	46.22	*
B	-0.06063	0.05881	1	0.05881	15.97	*
C	0.02563	0.01051	1	0.01051	2.854	
D	0.01938	0.006006	1	0.006006	1.632	
E	-0.1556	0.3875	1	0.3875	105.3	*
AB(CD)	0.009375	0.001406	1	0.001406	0.382	
AC(BD)	0.01563	0.003906	1	0.003906	1.061	
AD(BC)	-0.06812	0.07426	1	0.07426	20.17	*
AE	0.009375	0.001406	1	0.001406	0.382	
BE	0.09563	0.1463	1	0.1463	39.74	*
CE	0.009375	0.001406	1	0.001406	0.382	
DE	0.03313	0.01756	1	0.01756	4.769	
Residual		0.02209	6	0.003681		
SStotcorr		0.8946	15			

SSReg = 0.87251 (sum of squares of A, B, C, D, E, AD, AE, BE, DE)

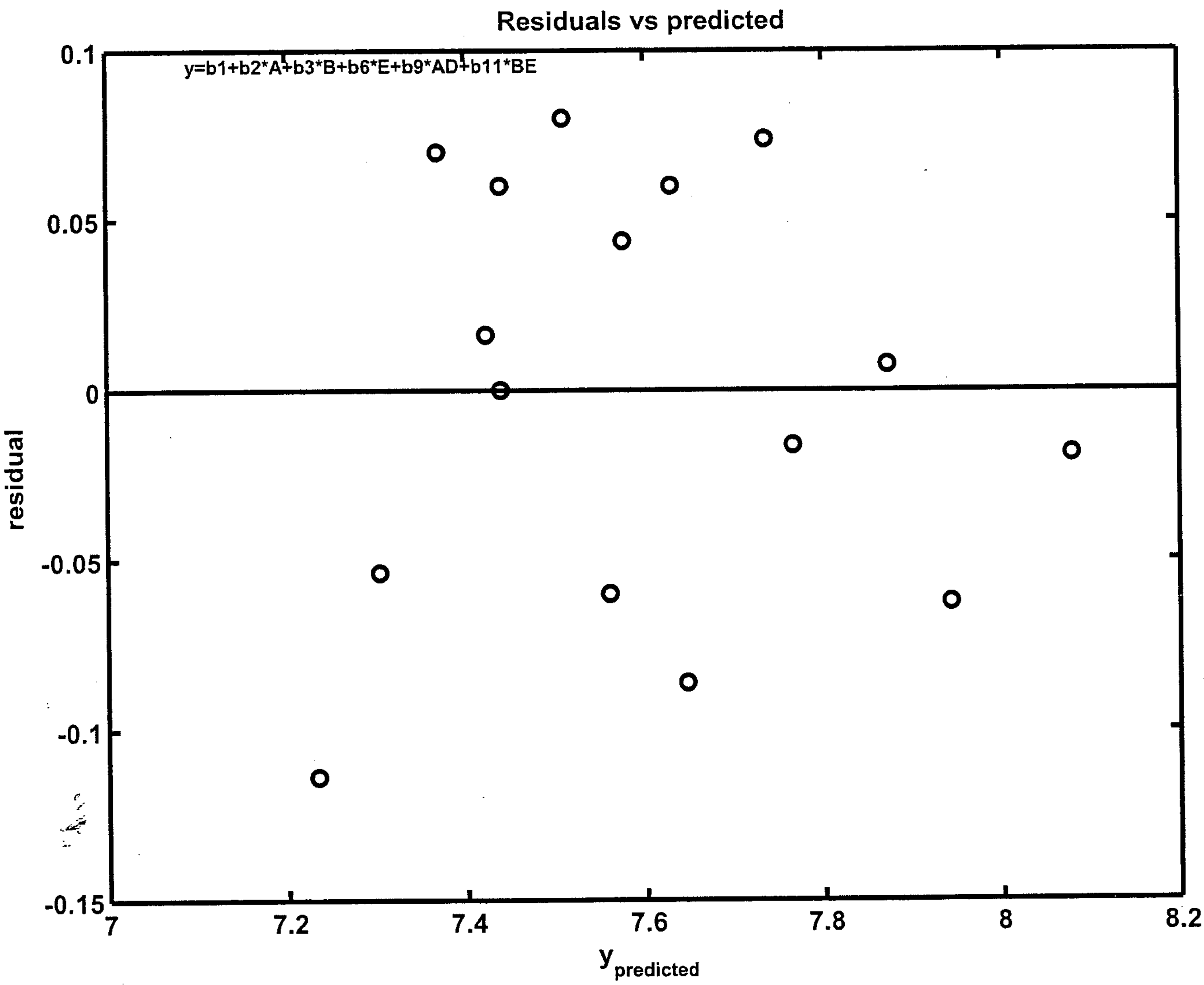
SSE = SStotcorr-SSReg = 0.022088

Fstatparam=F(0.95,1,6) = 5.99 FstatReg=F(0.95,9,6) = 4.10

Model: y=b1+b2*A+b3*B+b6*E+b9*AD+b11*BE

Residuals : 0.07375 -0.0625 -0.06 -0.01625 0.0075 -0.01875 0.01625 0.06 ....

-0.11375 2.2204e-016 0.06 -0.08625 0.07 0.04375 -0.05375 0.08



Residuals vs predicted
y=b1+b2*A+b3*B+b6*E+b9*AD+b11*BE
0.1
0.05
0
-0.05
-0.1
-0.15
residual
7
7.2
7.4
7.6
7.8
8
8.2
$y_{predicted}$

Problem 4 13-01-17

Combined array design

Response model:

$$y(\underline{x}, z) = \beta_0 + \gamma_1 z_1 + \beta_2 x_2 + \beta_3 x_3 + \delta_{21} x_2 z_1 + \varepsilon$$

y = fill height deviation

process variables: x_2 = pressure

x_3 = line speed

noise variable: z_1 = percent carbonation

Parameters:

$\beta_0 = 1.00$ $\gamma_1 = 1.50$

$\beta_2 = 1.13$ $\delta_{21} = 0.38$

$\beta_3 = 0.88$

$s^2 = 0.66$ $s_z^2 = 1$

Task: Determine the conditions to get $y \simeq 0$ with minimum standard deviation (POE)

Mean model: Expected value of y

$$E_z(y) = \beta_0 + \beta_2 x_2 + \beta_3 x_3$$

$$E_z(y) = 1 + 1.13 x_2 + 0.88 x_3$$

Variance model:

$$V_z(y) = s_z^2 \cdot \sum \left[\frac{\partial y(x, z)}{\partial z_i} \right]^2 + s^2$$

$$V_z(y) = s^2_z \, (\gamma_1 + \delta_{21} x_2)^2 + s^2$$

Calculate $[V_z(y)]_{min}$

Set process variable X_2 at its lower bound.

$(x_2)_{min} = -1$

$$[V_z(y)]_{min} = 1(1.50 - 0.38)^2 + 0.66$$

$$[V_z(y)]_{min} = 1.9144$$

$$[\sigma_z(y)]_{min} = 1.38 = POE$$

Calculate X_3:

$$E_z(y) = 1 + 1.13X_2 + 0.88X_3$$

$E_z(y) = 0 \qquad X_3 = 0.15$

Process conditions:

$X_2 = -1 \quad X_3 = 0.15$

$$y = 0 \pm 1.38$$

Problem 5

Design matrix and calculation results are given on page 4

Here is given a summary of the calculations.

A: Fit a model with three regressor variables and three interaction terms.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3 + \varepsilon$$

x_1 = spray volume x_2 = belt speed

x_3 = brand.

y = coating uniformity

Number of:

observations $N = 16$

factorial points $n_f = 8$, two replicates

parameters $p = 7$

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

$$b_j = \frac{\sum x_{ij} y_i}{\sum x_{ij}^2} = \frac{\sum x_{ij} y_i}{N}$$

$$\underline{b} = [\,37 \quad 1.375 \quad 3 \quad 0.875 \quad 8.375 \quad -0.875 \quad 1\,]$$
$$\quad \underline{b_0} \quad \underline{b_1} \quad \underline{b_2} \quad b_3 \quad \underline{b_{12}} \quad b_{23} \quad b_{123}$$

Significance test:

Test variable: $F_{obs} = \dfrac{SS(b_j)/1}{SS_E/(N-p)}$

$SS(b_j) = N \cdot b_j^2$

$SS_E = \underline{y}^T \underline{y} - \underline{b}^T \underline{X}^T \underline{y} = \sum (y_i - \hat{y}_i)^2$

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

Significant parameters: b_0, b_1, b_2, b_{12}

B: Significant model

Test variable: $F_{obs} = \dfrac{SS_{mod}/(p-1)}{SS_E/(N-p)}$

$SS_{mod} = \underline{b}^T \underline{X}^T \underline{y} - N\bar{y}^2 = \sum (\hat{y}_i - \bar{y})^2$

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

$F_{obs} > F_{stat} \implies$ significant model c: Prediction interval in the point

$\underline{X} = [1 \quad 1 \quad 1 \quad 1]$

$$y_0 = \hat{y}(\underline{X}_0) \pm t(0.975, N-P) \cdot \sqrt{s^2 (1 + \underline{X}_0^T (\underline{\underline{X}}^T \cdot \underline{\underline{X}})^{-1} \cdot \underline{X}_0}$$

$$\underline{X}_0 = [1 \quad X_{10} \quad X_{20} \quad X_{30} \quad X_{10} X_{20} \quad X_{20} X_{30} \quad X_{10} X_{20} X_{30}]^T$$

$$y_0 = 50.75 \pm 5.50$$

$$45.25 < y_0 < 56.25$$

4/4

Examination 13-01-17 Problem 5

Calculation matrix and observations

exp	obs	M	A	B	C	AB	AC	BC	ABC
(1)	40	1	-1	-1	-1	1	1	1	-1
a	25	1	1	-1	-1	-1	-1	1	1
b	30	1	-1	1	-1	-1	1	-1	1
ab	50	1	1	1	-1	1	-1	-1	-1
c	45	1	-1	-1	1	1	-1	-1	1
ac	25	1	1	-1	1	-1	1	-1	-1
bc	30	1	-1	1	1	-1	-1	1	-1
abc	52	1	1	1	1	1	1	1	1
(1)	36	1	-1	-1	-1	1	1	1	-1
a	28	1	1	-1	-1	-1	-1	1	1
b	32	1	-1	1	-1	-1	1	-1	1
ab	48	1	1	1	-1	1	-1	-1	-1
c	43	1	-1	-1	1	1	-1	-1	1
ac	30	1	1	-1	1	-1	1	-1	-1
bc	29	1	-1	1	1	-1	-1	1	-1
abc	49	1	1	1	1	1	1	1	1

ANOVA

Source	Param	SS	df	MS	Fobs	sign
Model		1337	6	222.8	54.2	*
b0	37	2.19e+004	1	2.19e+004	5328	*
b1	1.375	30.25	1	30.25	7.358	*
b2	3	144	1	144	35.03	*
b3	0.875	12.25	1	12.25	2.98	
b12	8.375	1122	1	1122	273	*
b13(incl Res)	-0.25	1	1	1	0.2432	
b23	-0.875	12.25	1	12.25	2.98	
b123	1	16	1	16	3.892	
Residual		37	9	4.111		
Lack of fit		1	1	1	0.2222	
Pure error		36	8	4.5		
Stotcorr		1374	15			

Fstatmod = F(0.95,p-1,N-p) = 3.37 Fstatparam = F(0.95,1,N-p) = 5.12

FstatLOF = F(0.95,m-p,N-m) = 5.32

R2 = 0.9731 R2adj = 0.9551

Prediction inteval in the point: [1 1 1]

b = [37 1.375 3 0.875 8.375 -0.875 1]

ypredict = 50.75

preint = 5.50

45.25 < y < 56.25

Problems 6.

Model I

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

$$E(\underline{y}) = \underline{\underline{X}}_1 \cdot \underline{\beta}_1$$

$$\underline{\beta}_1 = [\beta_0 \quad \beta_1 \quad \beta_2 \quad \beta_3]^T$$

$$\underline{\underline{X}}_1 = \begin{array}{cccc} x_0 & x_1 & x_2 & x_3 \\ \end{array}$$
$$\underline{\underline{X}}_1 = \begin{bmatrix} 1 & 0 & \sqrt{2} & -1 \\ 1 & -\sqrt{2} & 0 & 1 \\ 1 & 0 & -\sqrt{2} & -1 \\ 1 & \sqrt{2} & 0 & 1 \end{bmatrix}$$

$$\underline{\underline{X}}_1^T \cdot \underline{\underline{X}}_1 = \begin{bmatrix} 4 & & & \underline{\underline{0}} \\ & 4 & & \\ \underline{\underline{0}} & & 4 & \\ & & & 4 \end{bmatrix}$$

$$\underline{b}_1 = (\underline{\underline{X}}_1^T \underline{\underline{X}}_1)^{-1} \cdot \underline{\underline{X}}_1^T \underline{y}$$

Expected values:

$$E(\underline{b}_1) = (\underline{\underline{X}}_1^T \underline{\underline{X}}_1)^{-1} \cdot \underline{\underline{X}}_1^T E(\underline{y}) = (\underline{\underline{X}}_1^T \underline{\underline{X}})^{-1} \cdot \underline{\underline{X}}_1^T \cdot \underline{\underline{X}}_1 \cdot \underline{\beta}_1$$

$$E(\underline{b}_1) = \underline{\beta}_1$$

Model II (True model)

$$E(\underline{y}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2$$

$$E(\underline{y}) = \underline{\underline{X}}_1 \underline{\beta}_1 + \underline{\underline{X}}_2 \underline{\beta}_2$$

$$\underline{\beta_2} = [\beta_{11} \quad \beta_{22} \quad \beta_{33}]^T$$

$$\underline{\underline{X}}_2 = \begin{array}{ccc} x_1^2 & x_2^2 & x_3^2 \\ \end{array}$$
$$\left[\begin{array}{ccc} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 0 & 2 & 1 \\ 2 & 0 & 1 \end{array}\right]$$

Expected values:

$$E(\underline{b}_1) = (\underline{\underline{X}}_1^T \underline{\underline{X}}_1)^{-1} \underline{\underline{X}}_1^T (\underline{\underline{X}}_1 \underline{\beta}_1 + \underline{\underline{X}}_2 \underline{\beta}_2)$$

$$E(\underline{b}_1) = \underline{\beta}_1 + (\underline{\underline{X}}_1^T \underline{\underline{X}}_1)^{-1} \underline{\underline{X}}_1^T \underline{\underline{X}}_2 \underline{\beta}_2$$

$$E(\underline{b}_1) = \underline{\beta}_1 + \underline{\underline{A}} \; \underline{\beta}_2$$

$\underline{\underline{A}}$ = alias matrix

gives the alias structure for the parameters in the vector $\underline{\beta}_1$

$$\underline{\underline{X}}_1^T \underline{\underline{X}}_2 = \left[\begin{array}{ccc} 4 & 4 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 4 & -4 & 0 \end{array}\right]$$

$$\underline{\underline{A}} = (\underline{\underline{X}}_1^T \underline{\underline{X}}_1)^{-1} \cdot \underline{\underline{X}}_1^T \underline{\underline{X}}_2 = \left[\begin{array}{ccc} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{array}\right]$$

$$E(\underline{b_1}) = \underline{\beta_1} + \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \beta_{11} \\ \beta_{22} \\ \beta_{33} \end{bmatrix}$$

$$E(b_0) = \beta_0 + \beta_{11} + \beta_{22} + \beta_{33}$$

$$E(b_1) = \beta_1$$

$$E(b_2) = \beta_2$$

$$E(b_3) = \beta_3 + \beta_{11} - \beta_{22}$$