

NTNU  
Noregs teknisk-naturvitskaplege  
universitet

Fakultet for informasjonsteknologi,  
matematikk og elektroteknikk  
Institutt for matematiske fag



Bokmål

Faglig kontakt under eksamen  
Inge Olsen tlf. 73593538

**EKSAMEN I EMNE TMA4255 FORSØKSPLANLEGGING OG ANVENDTE  
STATISTISKE METODER**

Fredag 14. mai 2004  
Tid: 09.00 – 14.00

Hjelpemiddel: Alle trykte og håndskrevne hjelpemiddel er tillatte. Bestemt enkel kalkulator tillatt.

Sensuren faller 4. juni

**Oppgave 1.**

En bedrift var interessert i å finne ut hvorledes kvaliteten på en bestemt type tøy kunne forbedres. I første omgang ble det utført 8 enkeltforsøk der disse 4 faktorene var med, hver på 2 nivå.

	Nivå	
	-	+
A: Operatør	Operatør 1	Operatør 2
B: Maskin	Maskin 1	Maskin 2
C: Råmateriale	Type 1	Type 2
D: Fargestoff	Type 1	Type 2

Kvaliteten av tøystykket ble vurdert på en skala frå 1 til 10 der 10 er beste kvalitet. De 8 enkeltforsøkene sammen med resultatet av forsøket er gitt nedenfor:

A	B	C	D	Y
-	-	-	+	7.0
+	-	-	-	8.0
-	+	-	-	8.5
+	+	-	+	8.0
-	-	+	-	7.3
+	-	+	+	8.7
-	+	+	+	9.0
+	+	+	-	8.6

- a) Hvilken resolusjon har dette forsøket? Grunngi svaret. Regn ut estimatet for hovedeffekten av B og for samspillseffekten mellom A og B. De estimerte effektene av A, C, D, AC og BC ble 0.375, 0.525, 0.075, 0.125, 0.025 i gitt rekkefølge. Ut i fra dette forsøket ble det konkludert med at A, B og C var de faktorene som hadde mest betydning. Har du kommentarer til dette?

Etter litt vurdering bestemte de seg for å gjøre 8 tilleggsforsøk. Disse er synt nedenfor

A	B	C	D	Y
-	-	-	-	7.4
+	-	-	+	7.5
-	+	-	+	7.8
+	+	-	-	8.8
-	-	+	+	8.0
+	-	+	-	8.0
-	+	+	-	8.0
+	+	+	+	9.0

- b) Fra disse 8 forsøkene ble det estimert følgende 7 effekter for A, B, C, D, AB, AC og BC i gitt rekkefølge: 0.525, 0.675, 0.375, 0.025, 0.475, -0.025, -0.175. Bruk alle de 16 forsøkene til å finne et estimat for de 4 hovedeffektene og de 6 tofaktorsamspillene. Anta at de 4 største estimerte effektene er signifikante. Tolk resultatet av forsøket.

For å være litt sikrere i konklusjonen ønsker de å utføre en test for å se om de 4 største effektene virkelig er signifikante. Siden det ikke er gjort gjentak, vil de bruke samspill av orden høyere enn 2 i variansestimeringen, men de vil unngå at en eventuell blokkeffekt mellom de to fraksjonene blir tatt inn i feilen.

- c) Finn basert på dette et estimat for variansen til effektene og avgjør hvilke effekter som er signifikante. Bruk 5% signifikansnivå.

**Oppgave 2.**

For å møte konkurransen fra internasjonal industri, ønsket en bedrift å sette i gang et omfattende kvalitetsforbedringsarbeid. For et av produktene deres kan kvaliteten ut i fra gitte kriterier vurderes på en skala fra 0 til 100. For å få litt innsikt i kvaliteten til produktet bestemte de seg for å inspisere 5 tilfeldig valgte produkt hver dag i 20 dager. De 20 gjennomsnittverdiene er gitt nedenfor:

70.5 71.3 69.8 70.7 72.1 70.1 68.8 71.4 72.3 70.5 71.0 69.7 70.6 68.2 67.5 70.0 70.7 69.3 71.4 72.1

La  $\bar{X}_i$  være gjennomsnittsverdi for dag  $i$  og  $R_i$  den tilhørende variasjonsbredden. For de 20 dataene ovenfor får du opplyst at  $\sum_{i=1}^{20} \bar{x}_i = 1408$  og  $\sum_{i=1}^{20} r_i = 102$  (summen av de 20 variasjonsbreddene er 102).

- a) Bruk informasjonen ovenfor til å lage et  $\bar{\bar{X}}$  – kontrollkort for kvaliteten til produktet. Plott inn de 20 observerte gjennomsnittene og vurder om framstillingsprosessen ser ut til å være i kontroll.

Flere kunder har klaget på kvaliteten til nettopp dette produktet. Bedriften undres på om det å øke temperaturen under framstillingsprosessen frå  $25^{\circ}\text{C}$  til  $30^{\circ}\text{C}$  kunne bedre kvaliteten. Det ble derfor samlet inn 5 observasjoner for hver av de to temperaturinnstillingene. Disse er synt nedenfor etterfølgt av en MINITAB utskrift.

$25^{\circ}\text{C}$	$x_i$	70.4	73.0	67.7	69.8	71.5
$30^{\circ}\text{C}$	$y_i$	71.7	70.1	73.4	69.1	72.0

**Two-Sample T-Test and CI: kval\_25; kval\_30**

Two-sample T for kval\_25 vs kval\_30

	N	Mean	StDev	SE Mean
kval_25	5	70,48	1,97	0,88
kval_30	5	71,26	1,68	0,75

Difference = mu kval\_25 - mu kval\_30

Estimate for difference: -0,78

95% upper bound for difference: 1,38

T-Test of difference = 0 (vs <): T-Value = -0,67 P-Value = 0,260 DF = 8

Both use Pooled StDev = 1,83

- b) Formuler problemstillingen til bedriften som en hypotesetest. Sett opp testobservator som er brukt i utrekningene fra MINITAB. Hvilke forutsetninger bygger denne testen på? Hvorledes bør en samle inn data for en slik test? Hva blir konklusjonen når signifikansnivået er satt til 5%? Grunngi svaret.

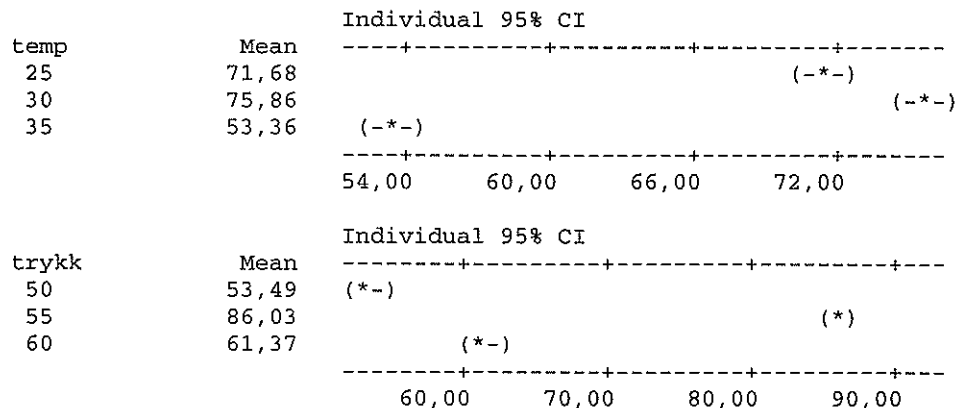
De bestemte seg for å gjøre et nytt forsøk der trykket under framstillingsprosessen ble variert sammen med temperaturen. Dataene fra dette forsøket sammen med en MINITAB utskrift er synt nedenfor.

Temperatur	Trykk								
	50		55		60				
25°C	50.8	50.7	49.4	93.7	90.9	90.9	74.5	73.0	71.2
30°C	63.4	61.6	63.4	93.8	92.1	97.4	70.9	68.8	71.3
35°C	46.6	49.1	46.4	69.8	72.5	73.2	38.7	42.5	41.4

### Two-way ANOVA: kvalitet versus temp; trykk

Analysis of Variance for kvalitet

Source	DF	SS	MS	F	P
temp	?	2578,22	1289,11	?	0,000
trykk	?	5188,93	2594,47	?	0,000
Interaction	?	643,35	160,84	?	0,000
Error	?	50,94	2,83		
Total	?	8461,44			



c) Hva slags forsøk er dette? Sett opp modellen for dette forsøket og gjør rede for hvilke forutsetninger en gjør for å kunne analysere det. Fyll inn verdier for de 8 spørsmålstegnene.

Formuler hypoteser og gjør rede for hvilke konklusjoner en kan trekke fra forsøket. Studer gjennomsnittene og finn ut hvilken funksjonssammenheng det ser ut til å være mellom kvalitet og trykk og temperatur?

Det ble og utført en regresjonsanalyse på dataene. Utskrift fra denne er synt nedenfor.

### Regression Analysis: kvalitet versus temp; temp\*temp; ...

The regression equation is

$$\text{kvalitet} = -4315 + 46,2 \text{ temp} - 0,534 \text{ temp*temp} + 135 \text{ trykk} - 1,14 \text{ trykk*trykk} - 0,291 \text{ temp*trykk}$$

Predictor	Coef	SE Coef	T	P
Constant	-4315,06	91,84	-46,98	0,000
temp	46,186	?	23,54	0,000
temp*temp	-0,53356	0,02741	?	0,000
trykk	135,382	3,072	44,07	0,000
trykk*trykk	-1,14422	0,02741	-41,74	0,000
temp*trykk	-0,29100	0,01938	-15,01	0,000

S = ?                      R-Sq = ?                      R-Sq(adj) = 99,1%  
 PRESS = 96,8373                      R-Sq(pred) = 98,86%

#### Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	?	?	?	0,000
Residual Error	21	59,2	2,8		
Total	26	8461,4			

Source	DF	Seq SS
temp	1	1510,7
temp*temp	1	1067,6
trykk	1	279,3
trykk*tr	1	4909,7
temp*try	1	635,1

#### Predicted Values for New Observations

New Obs	Fit	SE Fit	98,0% CI	98,0% PI
1	97,158	0,663	( 95,488; 98,828)	( 92,748; 101,568)

#### Values of Predictors for New Observations

New Obs	temp	temp*temp	trykk	trykk*tr	temp*try
1	28,4	807	55,6	3091	1579

- d) Fyll inn dei 7 verdiene der det står ? i utskriften ovenfor. Forklar forskjellen på T-testen og F-testen i regresjonsanalyseutskriften. Utfør en test for å teste om trykk kan bidra til å forklare variasjonen i kvalitet gitt at temperatur er med i modellen?
- e) Finn ut hvilke verdier av temperatur og trykk som ser ut til å maksimere kvaliteten til produktet ut i fra regresjonsanalysen. Hva blir estimert verdi for forventet kvalitet for disse verdiene av trykk og temperatur?  
 Bedriften ønsker å oppgi en verdi slik at de kan garantere at berre 1% av produktene har en kvalitet som er dårligere enn den. Hvilken verdi bør de oppgi dersom de bruker de optimale verdiene for trykk og temperatur? (Hint: Det er greit å finne en tilnærmet verdi fra eller ved hjelp av utskriften)

①

Løsningsforslag TMA 4255 Føroølsplanlegging og  
 anvendte statistiske metoder. 14/5-04

Oppgave 1

$$D = -ABC \Rightarrow I = -ABCD \Rightarrow \text{resolusjon 4}$$

$$\hat{l}_B = \frac{8.5 + 8 + 9 + 8.6}{4} - \frac{(7 + 8 + 7.3 + 8.7)}{4} = \frac{3.1}{4} = 0.775$$

$$\hat{l}_{AB} = \frac{7 + 8 + 7.3 + 8.6}{4} - \frac{(8 + 8.5 + 8.7 + 9)}{4} = -0.825$$

$AB = AB - CD \Rightarrow$  samspels effekten mellom A og B kan

skuldast CD samspelt.

b)  $I = ABCD$

$$\Rightarrow \hat{A} = \frac{l_A + l_A'}{2} = \frac{0.375 + 0.525}{2} = 0.45$$

$$\hat{B} = \frac{l_B + l_B'}{2} = \frac{0.775 + 0.675}{2} = 0.725$$

$$\hat{C} = \frac{l_C + l_C'}{2} = \frac{0.525 + 0.375}{2} = 0.45$$

$$\hat{D} = \frac{l_D + l_D'}{2} = \frac{0.075 + 0.025}{2} = 0.05$$

$$\hat{AB} = \frac{-0.825 + 0.475}{2} = -0.175$$

$$\hat{BC} = \frac{0.025 - 0.175}{2} = -0.075$$

$$\hat{AC} = \frac{0.125 - 0.025}{2} = 0.05$$

$$\hat{BD} = \frac{l_{BC}^i - l_{BC}^j}{2} = \frac{-0.025 - 0.125}{2} = -0.075$$

$$\hat{AD} = \frac{l_{BC}^i - l_{BC}^j}{2} = \frac{-0.175 + 0.025}{2} = -0.075$$

$$\hat{CD} = \frac{l_{AD}^i - l_{AD}^j}{2} = \frac{0.475 + 0.025}{2} = 0.25$$

②

Estimeret Effekt av å bytte operator 1 med operator 2 = 0.45

Estimeret - u - - a - maskin 1 med maskin 2 = 0.725

For C og D

C	D	
	-	+
-	8.175	7.575
+	7.975	8.675

Som synes at med råmateriale type 1 er det en negativ effekt av å bruke fargestoff 2 istedenfor 1.

Med råmateriale type 2 er det motsatt. Best resultat med råmateriale type 2 og fargestoff type 2.

c)  $\hat{\mu}_{BC} = \frac{l_D' - l_D}{2} = -0.025$        $\hat{\mu}_{CD} = \frac{l_B' - l_B}{2} = -0.05$

$\hat{\mu}_{BD} = \frac{l_C' - l_C}{2} = -0.075$        $\hat{\mu}_{AD} = \frac{l_A' - l_A}{2} = 0.075$

$\hat{\sigma}_{\text{effekt}}^2 = \frac{(-0.025)^2 + (-0.075)^2 + (-0.05)^2 + (0.075)^2}{4} = 0.0036$

$\hat{\sigma}_{\text{effekt}} = 0.06$

kritisk verdi =  $t_{0.025, 4} \cdot 0.06 = 2.7764 \cdot 0.06 = 0.167$

=> at alle de 4 største effektene er signifikante på 5%

I tillegg kan AA samspillet vurderes til å være signifikant.

(3)

## Oppgave 2

a)

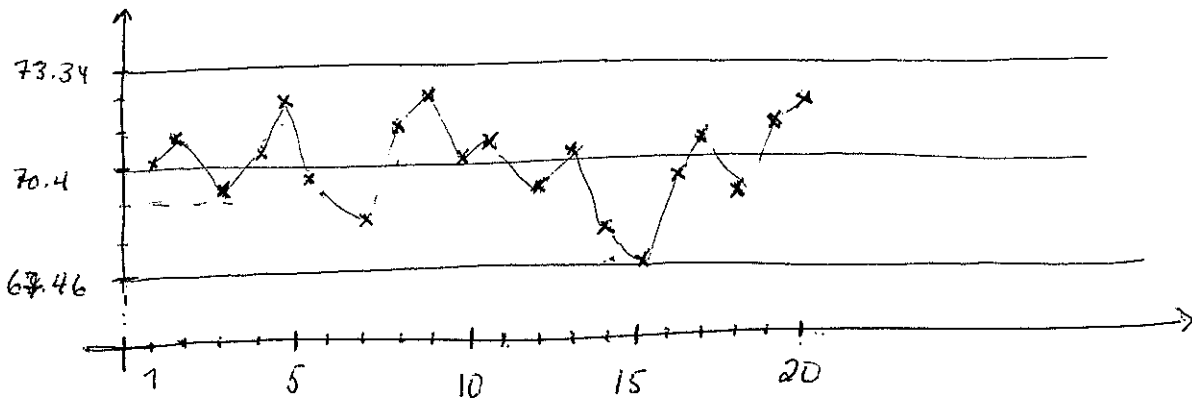
$$SL = \frac{1408}{20} = 70.4$$

$$\bar{\bar{X}} = \frac{102}{20} = 5.1$$

$$\hat{s} = \frac{\bar{s}}{d_2(5)} = \frac{5.1}{2.326} = 2.19$$

$$UKG = 70.4 + 3 \cdot \frac{2.19}{\sqrt{5}} = 73.34$$

$$LKG = 70.4 - 3 \cdot \frac{2.19}{\sqrt{5}} = 67.46$$



Ingen av observasjonene fell utenfor kontrollgrensene.

b)

$$H_0: \mu_{25} = \mu_{30}$$

$$H_1: \mu_{30} > \mu_{25}$$

$$T = \frac{\bar{X} - \bar{Y}}{s_p \sqrt{\frac{1}{5} + \frac{1}{5}}}$$

$$s_p^2 = \frac{\sum_{i=1}^5 (x_i - \bar{x})^2 + \sum_{i=1}^5 (y_i - \bar{y})^2}{8}$$

$x_1, \dots, x_5$  }  $N(0, \sigma^2)$  og uavh.  
 $y_1, \dots, y_5$  }

En test utført mellom temperaturmålingene i randomisert rekkefølge.



(4)

P-verdi er 0.26  $\Rightarrow$  ingen grunn til å forkaste  $H_0$ .

c) Jo-veis variansanalyse med 3 gjentak.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk} \left. \begin{array}{l} N(0, \sigma^2) \\ \text{uavh.} \end{array} \right\}$$

$$\sum \alpha_i = \sum \beta_j = \sum_i \alpha\beta_{ij} = \sum_j \alpha\beta_{ij} = 0$$

$$DF' = [2, 2, 4, 18, 26]$$

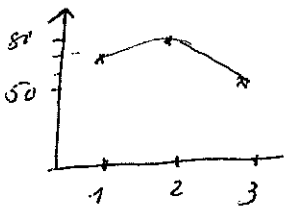
$$H_0: \alpha\beta_{11} = \alpha\beta_{12} = \dots = \alpha\beta_{33} = 0$$

$H_1$ : minst ein  $\neq 0$ .

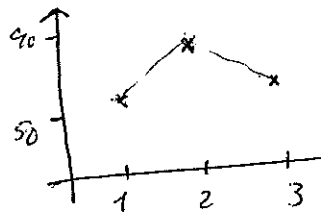
$P(F_{4,18} > 56.83) = 0.000 \Rightarrow$  forkast  $H_0$  som betyr at

begge faktorene har effekt.

Gjennomsnitt temp.



Gjennomsnitt trykk.



Det ser ut til å være en kvadratisk sammenheng.

d)

$$SD \hat{\beta}_{temp} = \frac{46.186}{23.54} = 1.962$$

$$s = \sqrt{2.8} = 1.679$$

$$T \hat{\beta}_{temp \times temp} = \frac{-0.53356}{0.02741} = -19.46$$

$$SSR = 8461.4 - 59.2 = 8402.3$$

$$MSR = \frac{8402.3}{5} = 1680.5$$

$$R^2 = \frac{8461.4 - 59.2}{8461.4} = \frac{8402.3}{8461.4} = 99.3$$

$$F = \frac{1680.5}{2.8} = 596.32$$

(5)

T-testen testar om ein regressjonsvariabel er signifikant gitt at dei andre er med i modellen.

F-testen testar om regressjonen er signifikant eller om ein eller fleire av regressjonsvariabelane kan fjerdere variasjonen i responsvariabelen.

$$H_0: \beta_{trykk} = \beta_{trykk \times trykk} = \beta_{temp \times trykk} = 0 \quad H_1: \min 1 \text{ er } \neq 0$$

$$F = \frac{\frac{SSR_F - SSR_{R}}{3}}{\frac{SSE}{21}} = \frac{8402 - (1510.7 + 1067.6)}{3} = \frac{5824.1}{2.8}$$

$$= 693.45 > f_{0.05, 3, 21} = 3.07 \Rightarrow \text{forkast } H_0$$

$\Rightarrow$  trykk kan bidra aktivt.

t = temp.

T = trykk

$$\hat{y} = -4315 + 46.2t - 0.53t^2 + 135T - 1.14T^2 - 0.29tT$$

$$\frac{\partial \hat{y}}{\partial t} = 46.2 - 1.06t - 0.29T = 0 \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \Rightarrow \begin{array}{l} t \approx 28.4 \\ T \approx 55.6 \end{array}$$

$$\frac{\partial \hat{y}}{\partial T} = 135 - 2.28T - 0.29t = 0$$

$$\hat{y}(28.1, 55.2) \approx 96.8$$

Et 98.0 prediksjonsintervall er gitt ved (92.75, 101.57)

0; dei to oppgje verdier 92.75 og 92.7 eller 92.5