

EKSAMENSOPPGAVE I STA-2000

Eksamen i : **STA-2000 Brukerkurs i statistikk 2**

Eksamensdato : **Tirsdag 9. desember 2003**

Tid : **Kl 0900-1300**

Sted : **Samvirkegården 3. etg.**

Tillatte hjelpemidler : **Alle trykte og håndskrevne, samt
kalkulator**

Oppgavesettet er på 9 sider inkl. forside

Kontaktperson under eksamen:

Amanuensis Elinor Ytterstad, tlf 44 015

OPPGAVE 1

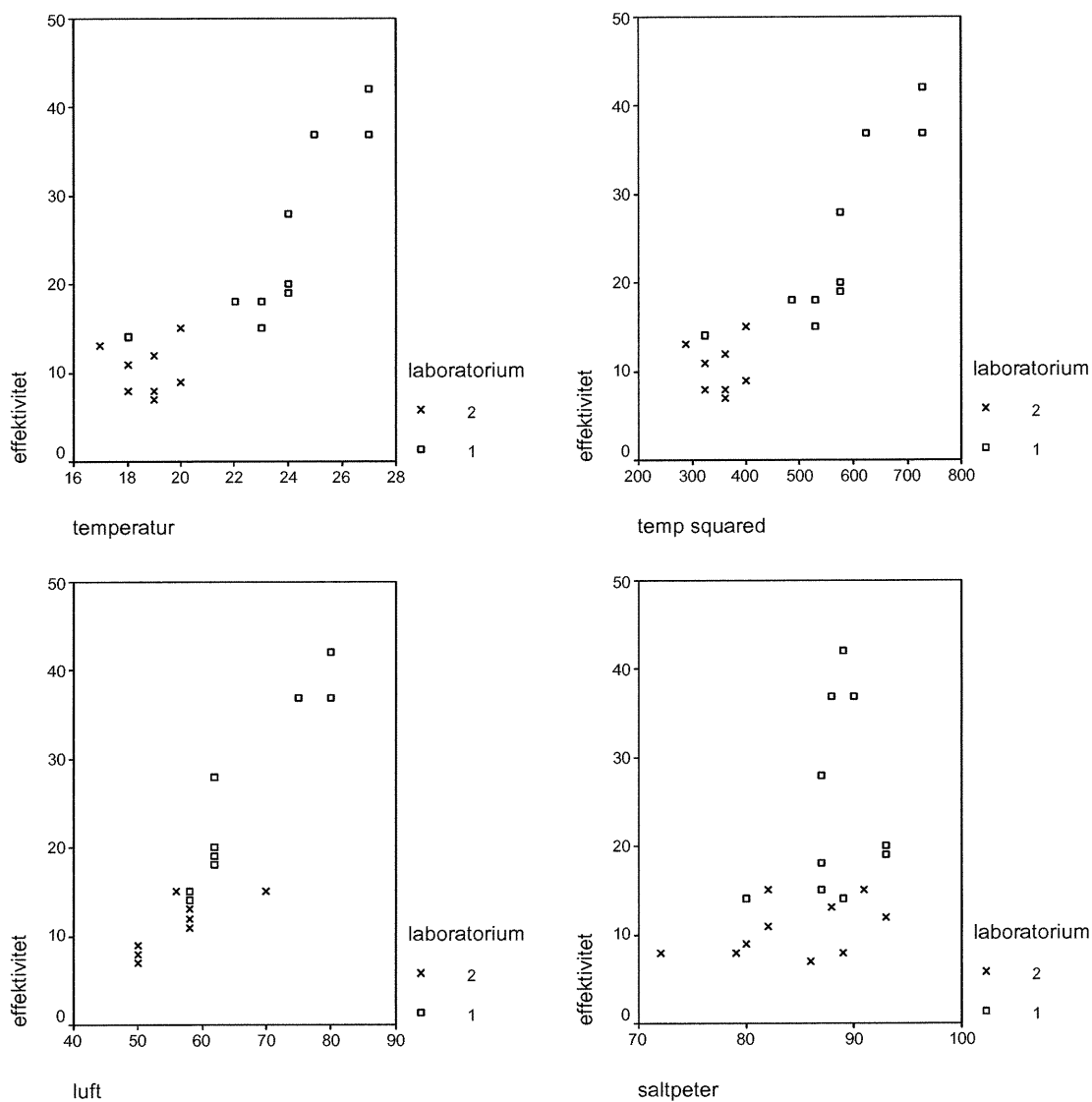
Består av punktene a) til h)

Dataene nedenfor stammer fra produksjon av saltpeter gjennom syreoksydasjon av ammoniakk. De uavhengige variablene er: **luft-tilførsel**, **temperatur** i kjølevannet, **saltpeterkonsentrasjonen** i absorpsjonsvæsken. Den avhengige variabelen er: produksjonsprosessens **effektivitet** (dvs. promille av ammoniakk som tapes som ikke-absorbert saltpetersyre). Eksperimentet er gjennomført på to ulike laboratorier. Det er angitt i variabelen **laboratorium** i tabellen nedenfor.

Vi skal i denne oppgaven undersøke sammenhengen mellom effektivitet og de andre variablene ved hjelp av regresjons- og variansanalyse.

effektivitet	temperatur		laboratorium	
	luft	saltpeter		
Y	X_1	X_2	X_3	Z
42	80	27	89	1
37	80	27	88	1
37	75	25	90	1
28	62	24	87	1
18	62	22	87	1
18	62	23	87	1
19	62	24	93	1
20	62	24	93	1
15	58	23	87	1
14	58	18	80	1
14	58	18	89	1
13	58	17	88	2
11	58	18	82	2
12	58	19	93	2
8	50	18	89	2
7	50	19	86	2
8	50	19	72	2
8	50	19	79	2
9	50	20	80	2
15	56	20	82	2
15	70	20	91	2

Nedenfor finner du spredningsplott av **effektivitet** mot de uavhengige variablene, der også **(temperatur)²** er tatt med.



- a) Skriv kort hva vi ser etter i slike plott.
Hva finner du her som kan være nyttig i den videre analysen?

Resultater fra fem modeller er presentert nedenfor.

Modell A: $Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_3 \cdot X_2^2 + \beta_4 \cdot X_3 + \beta_5 \cdot Z + E$

Modell B: $Y = \beta_0 + \beta_1 \cdot X_1 + \beta_3 \cdot X_2^2 + \beta_5 \cdot Z + \beta_6 \cdot Z \cdot X_1 + \beta_7 \cdot Z \cdot X_2^2 + E$

Modell C: $Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_3 \cdot X_2^2 + \beta_5 \cdot Z + E$

Modell D: $Y = \beta_0 + \beta_1 \cdot X_1 + \beta_3 \cdot X_2^2 + \beta_5 \cdot Z + E$

Modell E (full modell):

$$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_3 \cdot X_2^2 + \beta_4 \cdot X_3 + \beta_5 \cdot Z + \beta_6 \cdot Z \cdot X_1 + \beta_7 \cdot Z \cdot X_2^2 + E$$

b) Fyll inn tallene som mangler i tabellen nedenfor.

Modell	R^2	Mallow's C_p	Antall parametre i modell, inklusive β_0	Største Jackknife-residual i absoluttverdi
A	.937	9.88		2.480
B				4.405
C				2.574
D	.916	12.26		3.103
E	.957	8		3.800

- c) Skriv opp regresjonsmodell A med estimerte koeffisienter. Skriv også opp de to separate regresjonsligningene for hvert laboratorium i denne modellen.
- d) Utfør en added-last-test av laboratorium i modell A. Skriv opp hypotesen. Hva er resultatet av testen?
- e) Se igjennom analyseresultatene fra alle modeller. Er det signifikant forskjell mellom de to laboratoriene, eller signifikant interaksjonen mellom laboratorium og luft eller laboratorium og (temperatur)² i noen av modellene?
- f) Hvilken modell, av de fem som er presentert her, vil du si forklarer best variasjonen i *effektivitet*. Begrunn svaret. Hvilke kriterium har du benyttet for å komme frem til denne modellen.
- g) Regn ut justerte $\bar{Y}$ i modell C for laboratorium 1 og laboratorium 2. For å lette beregningen oppgis:

$\bar{Y}$	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_2^2$
17.52	60.43	21.14	456.29

- h) Dersom en overser alle uavhengige variable unntatt laboratorium og analyserer materialet med en en-veis variansanalyse, blir resultatet:

ANOVA

effektivitet

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	915,202	1	915,202	15,068	,001
Within Groups	1154,036	19	60,739		
Total	2069,238	20			

Den avhengige variabelen er som tidligere *effektivitet*. Det er 11 observasjoner fra laboratorium 1 og 10 observasjoner fra laboratorium 2. Bruk opplysninger fra tabellen over og test om det er forskjell mellom laboratorier.

Sammenlign konklusjon her med det du kom frem til under punkt e).

Dersom du kom frem til ulike konklusjoner her og i punkt e), hva skyldes det?

Model A, ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1939,457	5	387,891	44,832	,000 ^a
	Residual	129,781	15	8,652		
	Total	2069,238	20			

a. Predictors: (Constant), laboratorium, saltpeter, luft, temperatur, temp squared

b. Dependent Variable: effektivitet

Model A, Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	70,893	49,067		1,445	,169
	luft	,530	,149	,478	3,563	,003
	temperatur	-8,234	4,285	-2,525	-1,921	,074
	temp squared	,222	,101	2,990	2,194	,044
	saltpeter	-,115	,146	-,061	-,791	,441
	laboratorium	-1,819	1,832	-,092	-,993	,337

a. Dependent Variable: effektivitet

Model B, ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1965,443	5	393,089	56,807	,000 ^a
	Residual	103,795	15	6,920		
	Total	2069,238	20			

a. Predictors: (Constant), lab x temp_squared, lab x luft, luft, temp squared, laboratorium

b. Dependent Variable: effektivitet

Model B, Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-176,838	92,853		-1,904	,076
	luft	3,255	1,600	2,934	2,035	,060
	temp squared	9,755E-02	,082	1,314	1,187	,254
	laboratorium	64,829	32,284	3,262	2,008	,063
	lab x luft	-1,150	,549	-2,847	-2,096	,053
	lab x temp_squared	-1,51E-03	,001	-1,921	-1,052	,310

a. Dependent Variable: effektivitet

Model B, ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1934,049	4	483,512	57,225	,000 ^a
	Residual	135,189	16	8,449		
	Total	2069,238	20			

a. Predictors: (Constant), laboratorium, luft, temperatur, temp squared

b. Dependent Variable: effektivitet

Model C, Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	69,758	48,468		1,439	,169
	luft	,486	,136	,438	3,565	,003
	temperatur	-8,874	4,159	-2,721	-2,134	,049
	temp squared	,238	,098	3,200	2,423	,028
	laboratorium	-1,637	1,796	-,082	-,911	,376

a. Dependent Variable: effektivitet

Model D, ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1895,575	3	631,858	61,853	,000 ^a
	Residual	173,663	17	10,215		
	Total	2069,238	20			

a. Predictors: (Constant), laboratorium, luft, temp squared

b. Dependent Variable: effektivitet

Model D, Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-32,614	7,590		-4,297	,000
	luft	,637	,128	,574	4,965	,000
	temp squared	2,919E-02	,010	,393	2,975	,008
	laboratorium	-1,143	1,959	-,057	-,583	,567

a. Dependent Variable: effektivitet

Modell E, ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1979,879	7	282,840	41,148	,000 ^a
	Residual	89,359	13	6,874		
	Total	2069,238	20			

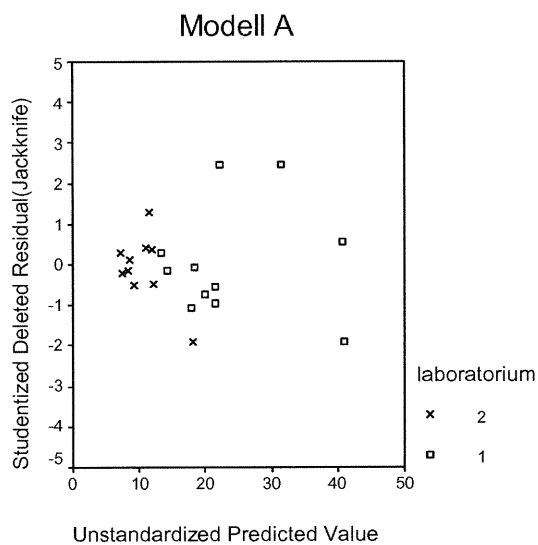
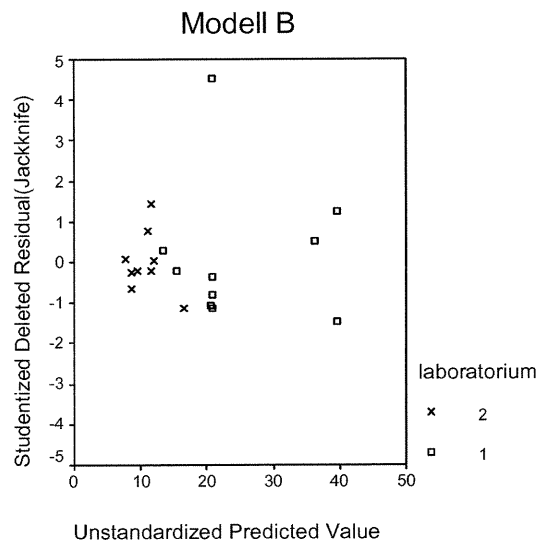
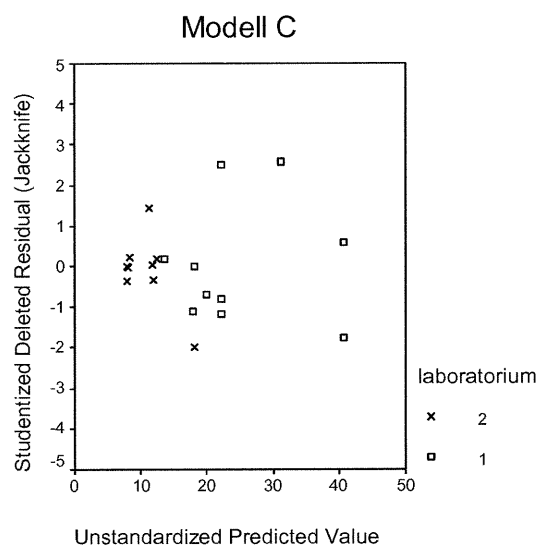
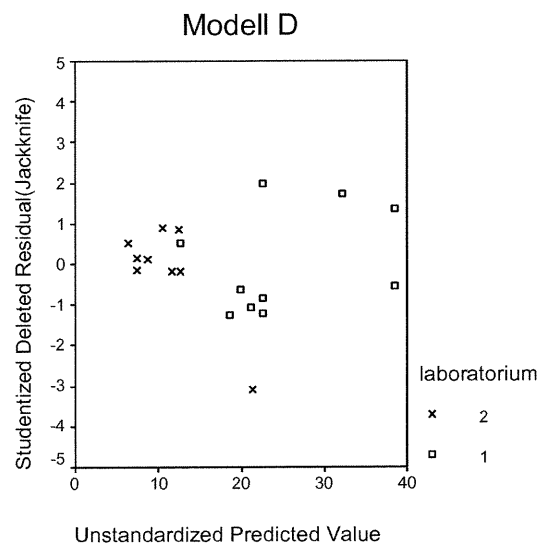
a. Predictors: (Constant), lab x temp_squared, lab x luft, saltpeter, luft, temperatur, laboratorium, temp squared

b. Dependent Variable: effektivitet

Modell E, Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-129,890	98,112		-1,324	,208
	luft	4,570	1,841	4,119	2,483	,027
	temperatur	-11,602	8,165	-3,558	-1,421	,179
	temp squared	,506	,297	6,812	1,701	,113
	saltpeter	-9,66E-02	,144	-,051	-,671	,514
	laboratorium	79,324	34,014	3,991	2,332	,036
	lab x luft	-1,377	,574	-3,411	-2,399	,032
	lab x temp_squared	-3,63E-03	,002	-4,634	-1,768	,101

a. Dependent Variable: effektivitet



OPPGAVE 2

Består av punktene a) og b)

I prosessen med å fornikle støtfangere, skal en undersøke om tykkelsen av nikkel er avhengig av mengde nikkel og pH-verdi. I eksperimentet brukes tre pH-verdier (5, 5.5 og 6) og to ulike mengder nikkel (10 og 18 gram) i blandingen støtfangerne dyppes i. Totalt 18 støtfangere randomiseres til disse seks kombinasjonene av nikkel og pH, tre i hver blanding. Tykkelse av nikkelbelegg er målt i $\text{cm} \times 10^{-3}$, og resultatet ble:

	pH 5	pH 5.5	pH 6
18 gram nikkel	250	211	221
	195	172	150
	188	165	170
10 gram nikkel	115	88	69
	165	112	101
	142	108	72

- a) Er noen av faktorene nikkel og pH random-faktorer. Begrunn svaret. Fyll inn de verdiene som mangler i ANOVA-tabellen nedenfor.
- b) Test alle tre hypoteser angående nikkel, pH, eller interaksjonseffekt på tykkelsen av belegget på støtfangerne. Utfør testen i samsvar med svaret ditt under punkt a).

	df	SS	MS	F
nikkel		31250.000		
pH				
interaksjon		670.333		
residual	12	8423.333	701.944	
total	17	46950		