

**Bokmål**

Institutt:	IKBM	
Eksamen i:	Stat100	Statistikk
Tid:	18. mai 2010	09.00-12.30
Emneansvarlig:	Trygve Almøy:	64 96 58 20

Tillatte hjelpemidler: C3: alle typer kalkulatorer, alle andre hjelpemiddel

Oppgaveteksten er på: 5 sider
Alle deloppgaver teller likt.

Alle steder der du utfører statistiske tester, skal du skrive opp både nullhypotesen og den alternative hypotesen og oppgi hvilket signifikansnivå du bruker.

Oppgave 1

Høyt kolesterolnivå er et stort problem, dette gjelder særlig for middelaldrende menn med en gjennomført usunn livsstil. Personene under er trukket ut fra denne (nitriste) populasjonen, og de fikk målt kolesterolinnholdet i blodet (måleenhet er mmol/l). Som et forsøk ble etter loddtrekning Odd, Kristian, Trygve, Lars, Tomas og Jan satt på en streng kostholdsdiett som ernæringsspesialister mente ville redusere kolesterolnivået, mens de resterende ble satt på et program med mye mosjon. Ett år etter at tiltakene ble satt i sving ble kolesterolnivået målt på nytt:

Navn	Kolesterol før behandling	Kolesterol etter behandling
Odd	6,7	6,0
Kåre	6,9	6,4
Solve	6,9	5,9
Kristian	5,8	5,4
Trygve	7,1	7,2
Lars	5,8	5,0
Tomas	6,4	6,2
Arne-Eri	5,6	4,8
Are	6,7	6,1
Ingar	6,1	6,0
Jan	7,0	6,0

Tabell 1: Kolesterolnivå før og etter behandling

- a) Kan du på bakgrunn av dette påstå at rett kosthold har positiv effekt?
Kan du på bakgrunn av dette påstå at mye mosjon har positiv effekt?
Sett opp modellene du bruker.

Svar ved hypotesetester med signifikansnivå 5 %. Med positiv effekt mener vi at det vil redusere gjennomsnittlig kolesterolnivå i hele populasjonen.

- b) Lag et 95 % konfidensintervall for forventet forskjell i effekt mellom kosthold og mosjon.
 Bruk dette intervallet til å avgjøre om disse to metodene er forskjellige.
 Hvilke antagelser gjør du?

Som en større oppfølger til dette forsøket, ville en se på tre forskjellige metoder:

- Bare rett Kosthold (K).
- Bare mye Mosjon (M).
- Både rett Kosthold og mye Mosjon samtidig (K-M)

Vi trakk ut 300 personer fra populasjonen beskrevet i innledningen og fikk følgende resultat:

90 av 100 med M fikk redusert kolesterolnivå
 75 av 90 med K fikk redusert kolesterolnivå
 100 av 110 med M-K fikk redusert kolesterolnivå.

- c) Konstruer et 95 % konfidensintervall for andelen som har positiv effekt av M.
 Gi en tolkning av dette intervallet.
- d) Kan du på bakgrunn av resultatene i dette forsøket påstå at metode og kolesterolreduksjon er avhengige variable?

Oppgave 2

I et slakteri har nyslaktede dyr typisk en høy pH-verdi. Vi ønsker å utføre en spesiell type kjøttbehandling på slaktet, men denne behandlingen er avhengig av at pH verdien er sunket til rundt 6,0. Siden det er veldig ressurskrevende å overvåke pH-verdien til hvert enkelt slakt, ønsker vi å kunne predikere pH basert på hvor lang tid det har gått fra slakting. For å finne sammenheng mellom pH og tid fra slakting tok en 10 forskjellige okse-slakt, lot disse henge et visst antall timer og målte pH verdier. Dette ga:

Tid etter slakt (i timer)	pH
1	7,02
1	6,93
2	6,42
2	6,51
4	6,07
4	5,99
6	5,59
6	5,80
8	5,51
8	5,36

Tabell 2: Tid etter slakting og pH-verdier for 10 forskjellige slakt.

Bruk Tabell 3 så mye du kan for å svare på punktene a) til d) nedenfor.

Coefficients:	Estimate	Std. Error
Intercept	7.0	0.097
Tid	-0.21	0.02
Residual standard error: 0.16 on 8 degrees of freedom		
Multiple R-squared: 0.93		

Tabell 3. Utskrift fra R. Forklaringsvariabel er tid etter slakt.

- a) Sett opp den lineære regresjonsmodellen der pH er respons, og tid er forklaringsvariabel. Gi en tolkning av parametrene i modellen.
Hvordan vil du estimere forventet pH i nyslakta dyr?
- b) Lag et 95 % konfidensintervall for regresjonskoeffisienten (β).
Gi en forklaring på intervallet til slakteriarbeiderne.
Bruk intervallet til å teste om pH endrer seg med tiden etter slakting.
Hvilket nivå tester du på?
- c) Prediker pH for et slakt som har hengt i 4 timer.
For et slikt slakt vil et 99 % prediksjonsintervall være (5,6; 6,7). Gi en forklaring på hva dette betyr.
Hvis antallet observasjoner i Tabell 2 ble svært stort, hva ville du estimert bredden til prediksjonsintervallet til?
Hvor lenge vil du foreslå at et slakt bør henge, dersom slakteriet vil at det skal ha en pH på 6?
- d) Hva er et residual? Svar konkret for denne undersøkelsen?
Finn residualene for de to slaktene som har hengt i en time (Tabell 2).
Hva vil du si om modellen basert på Figur 1 (side 5)?

En prøvde seg deretter med logaritmen til tid etter slakting (med naturlig grunntall) som forklaringsvariabel (kall dette Modell 2). Dette ga et residualplot som er vist i Figur 2 (side 5) og resultater vist i Tabell 4. Basert på tilgjengelig informasjon (Tabell 3 og Tabell 4, samt begge figurer):

- e) Vis at R^2 ved Modell 2 er 0,98.
Gi to begrunnelser for å fortrekke Modell 2.

Coefficients	Estimate	Std. Error
Intercept	6.98	0.048
Log (tid)	-0.73	0.03
Residual standard error: 0.08 on 8 degrees of freedom		
Tabell 4. Utskrift fra R. Forklaringsvariabel er logaritme til tid etter slakt.		

- f) Vanskelig oppgave (denne deloppgaven ble tatt ut til eksamen, men prøv likevel)
Anta Modell 2.

Hvorfor kan du ikke estimere pH for nyslakta dyr?
Hva blir predikert pH for dyr som har hengt 1 time?
Estimer pH reduksjon pr. time.
Hvor lenge bør i følge denne modellen et slakt henge for å oppnå en pH på 6?
For alle delspørsmål: Sammenlign med tilsvarende resultat dersom du bruker tid som forklaringsvariabel.

En bestemte seg også for å undersøke om det var forskjell i pH mellom dyreslag. En tok 10 slakt fra ku (storfe), 12 slakt fra okse (storfe) og 8 slakt fra lam (småfe), og målte pH. Alle slakt hadde hengt i 5 timer. En valgte å bruke enveis-variansanalyse for å analysere data.

$$\text{Modell 3: } Y_{ij} \sim N(\mu_i, \sigma)$$

Basert på innsamlet data fikk vi:

	Df	Sum Sq	Mean Sq	F
Dyreslag		0.1349		
Residuals		0.4036		

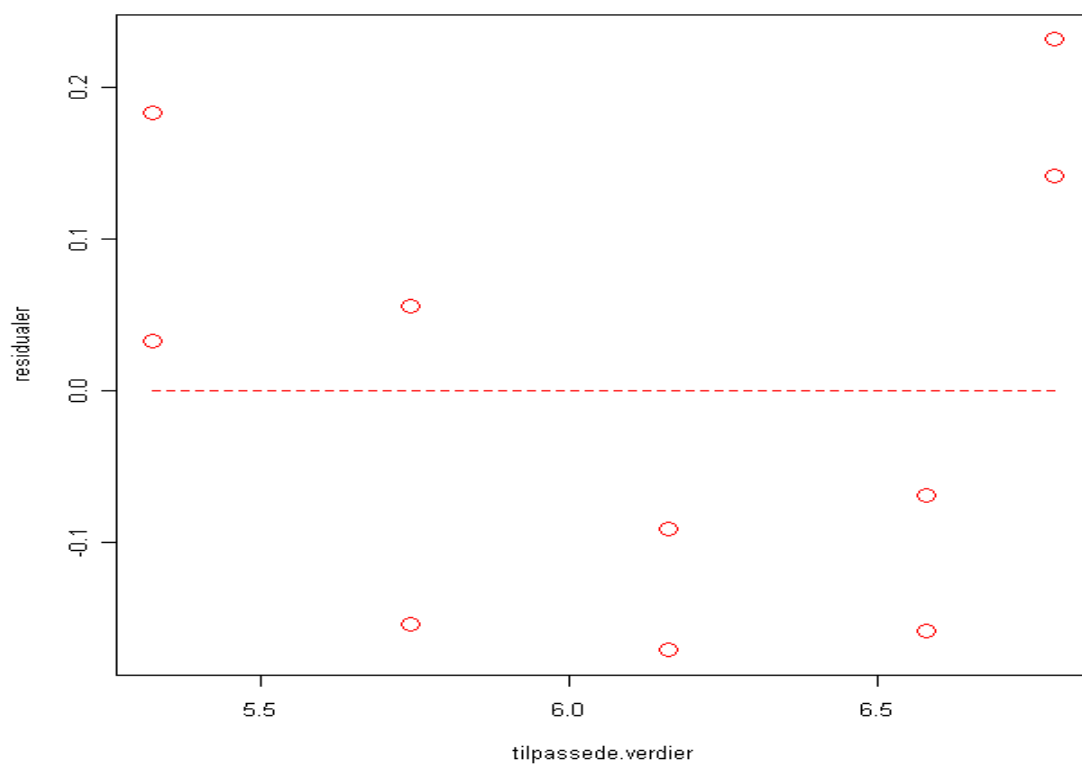
	Mean
ku	5.91
lam	5.74
okse	5.82

Tabell 5: Noen resultater fra enveis variansanalyse.

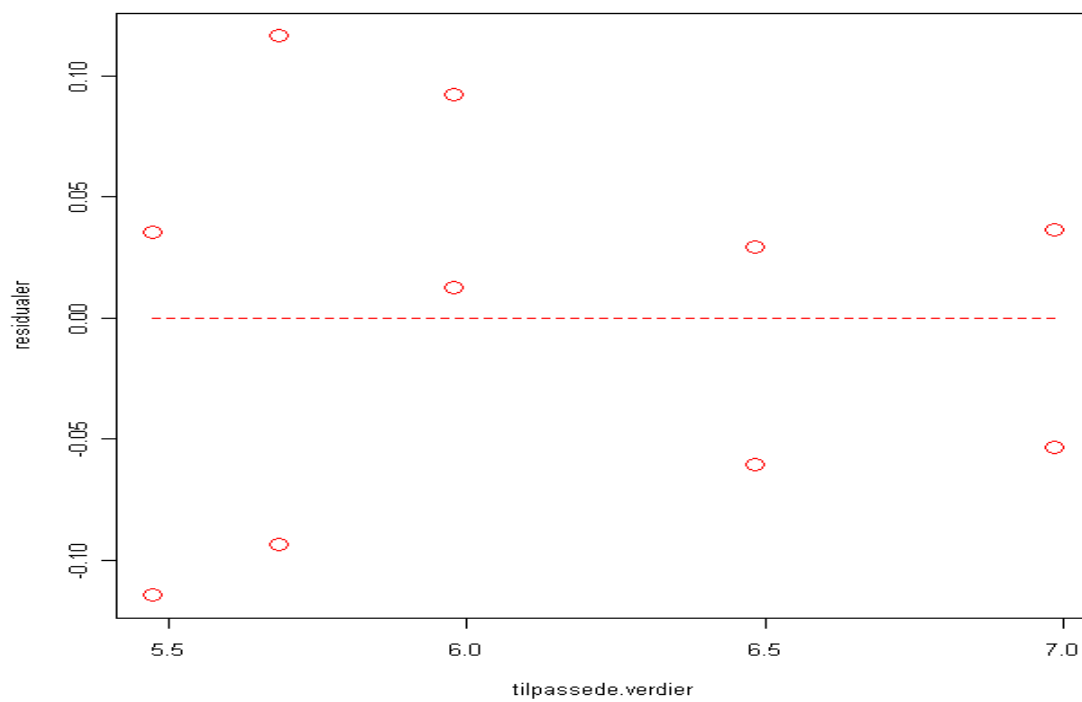
- g)** Forklar **alle** uttrykk som forekommer i Modell 3. (Vær konkret og svar i forhold til denne undersøkelsen)
 Estimer alle parametre i Modell 3.
 Test om det er forskjell mellom dyreslagene.
- h)** For slakt som henger i 5 timer, kan du påstå:
 At det er signifikant forskjell i pH mellom storfe og småfe?
 At det er signifikant forskjell i pH av mellom okse og ku?
 Utfør nødvendige tester på 10 % nivå.

Data for oppgave 2g og 2h (disse trenger du strengt tatt ikke)

pH_ku	pH_lam	pH_okse
5,83	5,75	5,98
5,86	5,73	5,93
5,66	5,94	5,70
5,76	5,67	5,76
5,94	5,78	5,74
5,98	5,64	5,76
6,16	5,75	5,78
5,81	5,66	5,77
6,00		5,90
6,13		5,98
		5,72
		5,82



Figur 1: Residualplot med tid etter slakt som forklaringsvariabel



Figur 2: Residualplot med logaritme av tid etter slakt som forklaringsvariabel