

**UTSATT EKSAMEN I
SOS4020 – KVANTITATIV METODE**

**SKOLEEKSAMEN
10. januar 2007
(4 timer)**

Ikke-programmerbar kalkulator er tillatt under eksamen. Ingen andre hjelpemidler er tillatt.

Sensuren faller onsdag 31. januar kl. 14.00. Sensuren blir slått opp på tavlen utenfor Aud. 7, 1. etasje i Eilert Sundts hus. Sensuren kan også hentes på Studentweb fra ca kl. 14.30 samme dag.

Sensuren blir regnet som mottatt av studentene når den blir slått opp, med mindre det kan dokumenteres gyldig fravær. Vi minner om at kandidater som ønsker begrunnelse må søke til instituttet om dette **senest 1 uke** etter at sensur er mottatt. Frist for å klage på karakteren er tre uker etter at sensuren har falt.

Oppgavesettet er på **4 sider** inkludert denne.

Kandidaten skal levere både originalen og kopien av besvarelsen.

NB! Skriv så hardt at kopien blir leselig. Tusj penn kan ikke brukes.

Husk å notere deg kandidatnummeret ditt.

LYKKE TIL!

Alle spørsmål (1-12) skal besvares. Hver deloppgave teller like mye.

I en undersøkelse av 1000 tilfeldig utvalgte husholdninger ønsket man blant annet å kartlegge hvor stor byrde det er å betjene den samlede gjelden i husholdningen. Til dette formålet utviklet man en gjeldsbelastningsvariabel ved å dele den samlede gjelden i husholdningen på den samlede inntekten og gange med 100. Hvis gjelden er like stor som inntekten, vil denne variabelen få verdien 100. Er gjelden dobbelt så stor, blir gjeldsbelastningen lik 200, osv.

Krysstabellen nedenfor (tabell 1) viser sammenhengen mellom gjeldsbelastning, om en bor i egen eller leid bolig og husholdningsinntekt. Gjeldsbelastningsvariabelen er omkodet til en dummyvariabel hvor laveste verdi er til og med 250 prosents gjeldsbelastning og høyeste verdi er gjeldsbelastning over 250 prosent.

Tabell 1: Gjeldsbelastning avhengig av inntekt og om en eier egen bolig.

			Inntektsgruppe			
			Inntil 200 000 kroner		200 000 kroner eller	
			Eier egen bolig?		Eier egen bolig?	
			Nei	Ja	Nei	Ja
Gjeldsbelastning over 250 prosent?	Nei	Antall	52	70	278	501
		Prosent	66,7%	92,1%	91,1%	92,8%
	Ja	Antall	26	6	27	39
		Prosent	33,3%	7,9%	8,9%	7,2%
Total		Antall	78	76	305	540
		Prosent	100,0%	100,0%	100,0%	100,0%

Kjikkvadrattest

Inntektsgruppe		Value	df	Asymp. Sig. (2-sided)
Inntil 200 000 kroner	Pearson Chi-Square	15,132	1	,000
200 000 kroner eller mer	Pearson Chi-Square	,719	1	,396

Spm. 1: Gi en substansiell forklaring på sammenhengen mellom gjeldsbelastning, inntekt og det å eie egen bolig. Baser fortolkningene både på fordelingene i krysstabellen og på kjikkvadrattestene. Illustrer resonnetet med en kausalmodell.

Spm. 2: Forklar begrepene odds og oddsrate.

Beregn odds for å ha gjeldsbelastning over 250 prosent for personer med lav inntekt (inntil 200 000 kroner) som henholdsvis bor i egen og leid bolig.

Beregn oddsraten mellom de to foregående oddsene og fortolk denne med ord.

Tabell 2 nedenfor viser resultatet av regresjonsanalyser med gjeldsbelastningsvariabelen som avhengig variabel (ikke kodet til dummyvariabel). De uavhengige variablene er:

- *Egen bolig* med verdien 1 for dem som eier egen bolig (selveierbolig eller andelsbolig) og 0 for dem som leier boligen.
- *Alder* i antall år over 20 år, dvs. 0 = 20 år, 1 = 21 år etc.
- *Alder kvadrert*
- *Utdanning* som er kodet om til to dummyvariabler – en for videregående og en høyskole. Referanse kategorien er grunnskoleutdanning.
- *Inntekt*, samlet husholdningsinntekt i hele tusen kroner, dvs. 100 = 100 000 kroner, 101 = 101 000 kroner etc.
- Produktledd mellom variablene *egen bolig* og *inntekt*.

Tabell 2: Regresjonsanalyser med gjeldsbelastning som avhengig variabel.

	Modell 1		Modell 2		Modell 3		Modell 4		Modell 5	
	b	SE(b)	b	SE(b)	b	SE(b)	b	SE(b)	b	SE(b)
Konstant	126,751	7,793	183,208	10,084	183,409	15,381	210,737	12,507	233,327	14,601
Egen bolig	-34,690	9,924	-17,058	9,819	-17,031	9,952	-11,120	9,890	-54,364	17,595
Alder			-2,697	0,320	-2,717	1,242	-2,875	0,322	-2,764	0,323
Alder kvadrert					0,001	0,021				
Inntekt							-0,068	0,019	-0,140	0,030
Inntekt* egen bolig									0,113	0,038
Justert R ²	0,011		0,076		0,075		0,087		0,095	
N	999		999		999		999		999	

Spm. 3: Gi en substansiell fortolkning av resultatet i modell 1.

Spm. 4: Fortolk konstant og koeffisienter i modell 2.

Spm. 5: Gjør rede for endringene for variabelen *egen bolig* fra modell 1 til modell 2. Forklar hva endringene skyldes.

Spm. 6: Gir modell 3 støtte for en antakelse om at sammenhengen mellom alder og gjeldsandelen er kurvelineær? Begrunn svaret.

Spm. 7: Et problem ved polynomregresjon er kolinearitet.

Gjør rede for hva kolinearitet er.

Hvordan kan sterk kolinearitet påvirke resultater av regresjonsanalyse?

Hvordan kan kolinearitetsproblemet avhjelpes?

Spm. 8: Gir modell 4 støtte til en hypotese om at forskjellen i gjeldsbyrde mellom ulike aldersgrupper primært skyldes inntektsforskjeller? Begrunn svaret.

Spm 9: Hva er definisjonen på samspill (interaksjon)?

Framstill grafisk sammenhengen mellom inntekt og gjeldsbelastning blant 20-åringer for henholdsvis de som eier og ikke eier egen bolig.

Gi en substansiell fortolkning av samspillet mellom boligvariabelen og inntektsvariabelen.

Spm. 10: I boken til Skog skilles det mellom ulike målsetninger ved multivariat analyse. Den ene går ut på å forklare mest mulig variasjon i avhengig variabel mens den andre går ut på å bestemme mest mulig nøyaktig effekten av en årsaksfaktor.

Gjør rede for forskjellen mellom disse to målsetningene.

Hva har etter din mening vært målsetningen for analysene ovenfor? Begrunn svaret.

Spm 11: Gjør rede for skillet mellom usystematiske (tilfeldige) og systematiske målefeil.

Hvilke konsekvenser har usystematiske målefeil i henholdsvis avhengig og uavhengig variabel for den interne validiteten i en undersøkelse?

Spm. 12: En av forutsetningene for regresjonsanalyse er at restleddene skal være homoskedastiske.

Hva innebærer denne forutsetningen?

Hvordan kan en avdekke eventuell heteroskedastisitet?

Hva er konsekvensene for analyseresultatene av heteroskedastiske restledd?