

**EKSAMEN I
SOS4020 – KVANTITATIV METODE**

**SKOLEEKSAMEN
2. mars 2006
(4 timer)**

Bruk av ikke-programmerbar kalkulator er tillatt under eksamen. Ingen andre hjelpemidler er tillatt.

Oppgavesettet består av 4 sider.

Sensur på eksamen faller torsdag 23. mars kl. 14.00. Sensuren blir slått opp på tavlen utenfor Aud. 7, 1.et. i Eilert Sundts hus. Sensuren kan også hentes på Studentweb ca 1 time senere.

Sensuren regnes som mottatt av studentene når den slås opp, med mindre gyldig fravær kan dokumenteres. (Husk å notere deg kandidatnummeret ditt.)

Kandidaten skal levere både originalen og kopien av besvarelsen.

LYKKE TIL!

Alle spørsmål skal besvares. Hver deloppgave teller like mye.

I en undersøkelse ble det innhentet registerbaserte opplysninger fra 6 000 avgangselever fra grunnskolen i 2004. Disse var trukket ut tilfeldig og utgjorde ca. 10 prosent av elevmassen det året. En var interessert i å studere variasjoner i elevenes grunnskolepoeng. Grunnskolepoeng er basert på elevenes vitnemål og beregnes som summen av elevenes karakterer i 11 fag. Antallet grunnskolepoeng varierer således mellom 11 (en i alle fag) og 66 (seks i alle fag).

I krysstabellen under er elevene rangert etter antall grunnskolepoeng og deretter delt inn i fem like store grupper. Tabellen viser hvordan gutter og jenter fordeler seg når det gjelder antall grunnskolepoeng.

Tabell 1 Jenter og gutters fordeling av grunnskolepoeng. Prosent

Antall grunnskolepoeng	Jenter	Gutter	Totalt
Mindre enn 35,5	13	27	20
35,5 – 42,0	17	23	20
42,1 – 47,3	20	20	20
47,4 – 52,5	22	18	20
Mer enn 52,5	28	12	20
Totalt	100	100	100
N=	2 990	3 010	6 000

a)

Hva forteller denne tabellen? Ved hjelp av kjiqvadrat-testen kan en undersøke om forskjellen mellom gutter og jenter skyldes tilfeldigheter. Gjør rede for hva kjiqvadrat-testen går ut på. Hvor mange frihetsgrader har denne tabellen? Kjiqvadratet for tabellen er 363,3. Gitt antallet frihetsgrad gir dette en signifikanssannsynlighet som er lavere enn 0,0001 ($p < 0,0001$). Hva blir konklusjonen av testen?

b)

Hva innebærer Type-I og Type-II-feil? Redegjør for sammenhengen mellom disse.

c)

For elevene som inngår i undersøkelsen er gjennomsnittet 44 grunnskolepoeng. Standardavviket er 9,2 og standardfeilen er 0,12. Hva er standardavvik og standardfeil uttrykk for og hva er forskjellen mellom dem? Konstruer et 5-prosent konfidensintervall rundt gjennomsnittet i utvalget. Forklar hva konfidensintervallet kan brukes til i denne sammenheng.

Tabell 2 viser elevenes gjennomsnittlige grunnskolepoeng, brutt ned etter kjønn og innvandringsstatus. Tabellen gir også informasjon om standardfeil. Minoritets elever er alle som har to utenlandsfødte foreldre, majoritets elever har en eller to norskfødte foreldre.

Tabell 2 Gjennomsnittlig antall grunnskolepoeng brutt ned etter henholdsvis kjønn og innvandringsstatus. Standardfeil og antall (N)

	Gjennomsnitt	Standardfeil	N
Kjønn			
Jenter	46,3	0,16	2 990
Gutter	41,8	0,17	3 010
Innvandringsstatus			
Majoritet	44,3	0,12	5 582
Minoritet	40,4	0,48	418

d)

Hvorfor er standardfeilen til minoritets elevene så mye høyere enn de andre gruppens standardfeil? Redegjør kort for t-testen. Hva sier tabell 2 om karakterforskjeller i norsk skole? Still opp nullhypoteser og alternative hypoteser og bruk t-testen til å undersøke om forskjellen i grunnskolepoeng mellom henholdsvis gutter og jenter og mellom majoritets- og minoritets elever er statistisk signifikant på 5-prosentnivå.

(Hint: Standardfeilen (SE) til differansen mellom to gjennomsnitt

kan beregnes ut fra formelen $SE(\bar{x}_1 - \bar{x}_2) = \sqrt{SE_{x_1}^2 + SE_{x_2}^2}$).

e)

I lineær regresjonsanalyse forutsettes lineære sammenhenger mellom avhengig og uavhengige variabler. Hvordan kan en løse dette dersom det viser seg at forutsetningen ikke er tilstede i datamaterialet? Redegjør kort for andre forutsetninger som lineær regresjonsanalyse bygger på.

Tabell 3 viser resultater fra lineære regresjonsanalyser av elevenes grunnskolepoeng. De uavhengige variablene er kodet på følgende måte:

--- **Kjønn** er kodet slik at jenter har verdien 0 og gutter har verdien 1

--- **Innvandringsstatus** er kodet slik at majoritets elever har verdien 0 og minoritets elever har verdien 1

--- **Kjønn * Innvandringsstatus** er produktet av de to variablene

--- **Foreldres utdanningsnivå** angir antall års utdanning utover grunnskole, for den av foreldrene med høyest utdanning. Varierer fra 0 år (kun grunnskole) til 12 år (doktorgradsutdanning)

Tabell 3 Lineære regresjonsanalyser av elevenes grunnskolepoeng (N=6 000)

	Modell 1		Modell 2		Modell 3	
	b	se (b)	b	se (b)	b	se (b)
Kjønn	-4,5	0,23	-4,5	0,21	-4,6	0,22
Innvandringsstatus	-2,9	0,50	-1,3	0,49	-1,6	0,66
Foreldres utdanningsnivå	-	-	1,3	0,04	1,3	0,04
Kjønn * Innvandringsstatus	-	-	-	-	0,6	0,91
Konstant	46,6	0,17	40,8	0,23	40,8	0,23

f)

I modell 1 er kjønn og innvandringsstatus trukket inn i analysen. Forklar hva resultatene viser. Sammenlikn med resultatene fra tabell 2. Hvorfor avviker/avviker ikke disse to tabellene fra hverandre?

g)

I modell 2 kontrolleres for foreldrenes utdanningsnivå. Fortolk koeffisienten for foreldrenes utdanning. Er denne statistisk signifikant? Vurder ut fra modell 2 hvorvidt foreldrenes utdanningsnivå kan være en årsak til forskjeller i grunnskolepoeng mellom elever med ulik kjønn og innvandringsstatus. Hvorfor endrer konstantleddet seg i forhold til modell 1?

h)

Hva er samspill? Er det ut fra modell 3 rimelig å konkludere med at det foreligger samspill mellom kjønn og innvandringsstatus på elevenes grunnskolepoeng?

Lykke til!