

**EKSAMEN I
SOS4020
KVANTITATIV METODE
8. april 2010**

(4 timer)

Tillatte hjelpemidler: Ikke-programmerbar kalkulator

Liste med matematiske uttrykk/andeler i fordelinger (bakerst i oppgavesettet)

Sensur på eksamen faller torsdag 29. april 2010. Sensuren kan hentes på Studentweb fra ca kl. 14.30 samme dag.

Sensuren regnes som mottatt av studentene når den offentliggjøres. Vi minner om at kandidater som ønsker begrunnelse må søke om dette til instituttet senest 1 uke etter at sensur har falt. Klagefristen er tre uker fra sensuren er falt, eller tre uker fra begrunnelse er mottatt.

Oppgavesettet består av **5 sider** inkludert denne.

Kandidaten skal levere både originalen og kopien av besvarelsen.

Husk å notere deg kandidatnummeret ditt.

LYKKE TIL!

Oppgavesettet består av 12 delspørsmål. Alle spørsmålene skal besvares. Det siste arket inneholder formler som kan være til nytte ved besvarelsen av oppgaven.

NB: En sensor skal lese kopien av oppgaven din. Pass derfor på at gjennomslagskopien blir tydelig.

I forskningslitteraturen defineres gjerne *sosial tillit* som tillit til ukjente mennesker, for eksempel tiltro til at andre personer i det samfunnet vi lever i ikke vil oss noe vondt, men tvert er til å stole på og vil oss godt.

En komparativ spørreundersøkelsen ble gjennomført i en rekke europeiske land i 2008. I hvert land ble det trukket et representativt utvalg av personer i aldersgruppen 20-69 som fikk de samme spørsmålene. I undersøkelsen ble det spurt en rekke spørsmål om sosial tillit, og svarkategoriene for disse spørsmålene var en skala fra 0 til 10 hvor 0 svarer til lavest grad av tillit og 10 svarer til høyest grad av tillit. Ved å ta gjennomsnittet av enhetenes verdier på disse spørsmålene ble det konstruert et mål på sosial tillit.

Flere andre variable ble også registrert i alle undersøkelsene, og disse har følgende verdier:

- *Relativ inntekt* er kodet i prosentiler fra 0 til 100. Verdien 0 betyr at man er nederst i landets inntektsfordeling, verdien 50 betyr at man har inntekt som medianinntekten, og verdien 100 betyr følgelig at man er øverst i landets inntektsfordeling.
- *Alder* er en kontinuerlig variabel med 20 som nullpunkt, slik at 20-åringer har 0, 21-åringer 1, og så videre.
- *Utdanning* er en kontinuerlig variabel med verdier fra 0 til 9 som måler antall år utdanning utover grunnskolen.
- *Kjønn* er en dummyvariabel hvor kvinner er kodet 0 og menn er kodet 1.
- *Medlemskap* er en dummyvariabel hvor 0 betyr at respondenten ikke er aktivt medlem i en frivillig organisasjon, og 1 betyr at respondenten er medlem i en frivillig organisasjon.

Statistiske tabeller og enkelte opplysninger er gitt bakerst i oppgavesettet

OPPGAVER

- a) Gjør kort rede for fordeler og ulemper med eksperimentelle forskningsdesign.
- b) Autokorrelasjon omtales som et problem i forbindelse med regresjonsanalyse på tidsseriedata. Beskriv kort hva problemet består i, og hva som kan gjøres for å løse det.

Tabell 1. Mål på sentraltendens og spredning for sosial tillit-variabelen i utvalgte land.

	Norge	Finland	Italia	Tyskland	Frankrike	Spania	Russland
Gjennomsnitt	6,52	6,37	4,80	5,26	4,92	4,88	4,36
Standardavvik	1,47	1,52	1,57	1,75	1,62	1,64	2,15
N	223	212	276	245	279	223	289

- c) Tabell 1 viser utvalgte mål på sentraltendens og spredning for sosial tillit i utvalgte europeiske land. Beskriv kort variasjonen i sosial tillit mellom land basert på Tabell 1.
- d) Test om gjennomsnittet er statistisk signifikant høyere i Norge enn i Italia. Hva forteller testen? Anslå usikkerheten i forskjellen mellom gjennomsnittene med et konfidensintervall.
- e) Nevn kort hvordan man effektivt kan teste om *mange* gjennomsnitt er statistisk signifikant ulike hverandre.

Det er gjort lineære regresjonsanalyser av spørreskjemadataene på utvalget trukket i Norge.

Resultatene fra analysene er gjengitt i Tabell 2.

Tabell 2. Resultater fra lineær regresjonsanalyse av sosial tillit for Norge. $N = 223$.

	Modell 1		Modell 2		Modell 3	
	B	SE(b)	b	SE(b)	B	SE(b)
Konstant	5,32	0,33	4,97	0,27	4,93	0,23
Alder	-0,01	0,02	-0,01	0,01	-0,02	0,01
Relativ inntekt	0,013	0,006	-0,007	0,04	-0,003	0,04
Kjønn	0,11	0,13	0,07	0,11	0,23	0,09
Utdanning			0,22	0,01	0,083	0,03
Utdanning*Kjønn					0,174	0,01
R^2	0,17		0,22		0,25	

- f) Gi substansielle tolkninger av konstantledd og regresjonskoeffisienter i modell 1. Hvilken gruppe har lavest tillit og høyest tillit ifølge modell 1?
- g) I modell to er det tatt inn ytterligere en variabel. Diskuter hva endringen i koeffisienten for relativ inntekt fra modell 1 til modell 2 kan skyldes.
- h) Hva skiller modell 3 fra modell 2 og hvorfor er endringen gjort? Test om den nye koeffisienten i modell 3 er statistisk signifikant forskjellig fra null. Hva forteller testen? Hvilken annen metode kunne vi brukt for å besvare samme spørsmål?

- i) Prediker forventet sosial tillit hos kvinner og menn henholdsvis 0, 3 og 6 år utdanning utover grunnskole og som tilhører referansegruppen på øvrige variable. Skisser de predikerte verdiene i et diagram. En konklusjon som ble trukket i den komparative analysen av sosial tillit i Europa var at "*utdanning skaper tillit mellom innbyggerne i et samfunn*". Hvor godt mener du dette passer for Norge?

I tabell 3 rapporteres resultatene fra en logistisk regresjonsanalyse med variabelen *Medlemskap* som avhengig variabel. Sosial tillit og alder er uavhengige variable. Her er aldersvariabelen kategorisert og tatt med i modellen som et sett av dummyvariable.

Tabell 3. Logistisk regresjon av medlemskap i frivillig organisasjon. Norge. $N = 223$.

	b	SE(b)	exp(b)
Konstant	-2,5	0,22	0,082
Sosial tillit	0,05	0,02	1,051
Alder: 20-29	(ref)		
Alder: 30-39	-0,12	0,08	0,89
Alder: 40-49	-0,17	0,11	0,84
Alder: 50-69	0,12	0,07	1,13

- j) Gi substansielle tolkninger av de logistiske regresjonskoeffisientene og exp(b) for de uavhengige variablene. Er de logistiske regresjonskoeffisientene statistisk signifikante? Vis testen(e) du bruker.
- k) Prediker andelen som skårer 1 på den avhengige variabelen for de som har hhv. 3 og 6 på sosial tillit-variabelen. Hvordan vurderer du muligheten for å gi koeffisienten for sosial tillit en kausal fortolkning?
- l) Redegjør for Hosmer-Lemeshows goodness-of-fit-test. For modellen vist i tabell 2 et Hosmer-Lemeshow-testen er beregnet til 9,89 med 8 frihetsgrader. Hva forteller dette oss?

NOEN OPPLYSNINGERStandardfeilen til forskjellen på to uavhengige størrelser

$$SE(\hat{m}_1 - \hat{m}_2) = \sqrt{[SE(\hat{m}_1)]^2 + [SE(\hat{m}_2)]^2}$$

Kritiske verdier i statistiske fordelinger:Normalfordelingen / Student's t
med svært mange frihetsgrader:

<u>Andel</u>	<u>z/t</u>
95%	±1,96
99%	±2,58

Kji-kvadratfordelingen:

<u>Andel ved df = 1</u>	<u>χ^2</u>
95%	3,84
99%	6,63
<u>Andel ved df = 8</u>	<u>χ^2</u>
95%	15,5
99%	17,5