

## **i Informasjon om eksamen**

### **SOS1120 - Kvantitativ metode**

Skriftlig skoleeksamen

29. mai 2018

4 timer

#### **Eksamensoppgaven**

Eksamenssettet består av 26 sider inkludert denne, med 22 oppgaver. Oppgavesettet er delt i 3 bolker. Les informasjonsdokumentet før hver bolk nøye.

Bolk 1: Flervalgsoppgaver

Bolk 2: Oppgaver i STATA

Bolk 3: Oppgaver med utregning og tekstbesvarelser

Du kan bla deg frem og tilbake ved hjelp av piltastene, eller ved hjelp av innholdsfortegnelsen nederst.

Lykke til!

14 **STATA 1**

Fra STATA åpner du det datasettet du skal bruke med følgende kommando:

**sysuse nlsw88**

Dette datasettet inneholder data fra en amerikansk surveyen National Longitudinal Survey of Young Women i 1988. Variabelen **wage** er timelønn i dollar.

Husk å fyll ut alle feltene.

Hva er gjennomsnittsverdien til variabelen **wage**? Skriv svaret her med 2 desimaler:

Hva er standardavviket til variabelen **wage**? Skriv svaret her med 3 desimaler:

Variabelen **union** sier om personen er organisert i en fagforening eller ikke.

Bruk STATA til å undersøke om fagforeningsmedlemmer og andre har lik eller forskjellig timelønn. Gjennomfør en hypotesetest der du bruker signifikansnivå 0.05.

a) Formuler nullhypotese og alternativ hypotese her:

b) Hva er t-verdien til testen? Oppgi resultatet fra STATA med 4 desimaler her:

c) Hva er korrekt tolkning av denne t-testen?

**Velg ett eller flere alternativer**

- p-verdien er over 0.05 og vi forkaster nullhypotesen
- p-verdien er mindre enn 0.0001 og det er svært lite sannsynlig at forskjellen i lønn skyldes tilfeldig variasjon. Fagforeningsmedlemmer har altså høyere lønn enn andre.
- p-verdien er 1 og vi kan være helt sikre på at forskjellen ikke skyldes tilfeldigheter.
- Vi forkaster nullhypotesen fordi t-verdien er større enn valgt p-verdi.
- p-verdien er mindre enn 0.05 og det er svært lite sannsynlig at forskjellen i lønn skyldes tilfeldig variasjon. Fagforeningsmedlemmer har derfor lavere lønn enn andre.

Kopier inn STATA-koden du brukte her:

(Linjeskift fungerer ikke i dette tekstfeltet, men ikke bry deg om det).

Maks poeng: 5

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**XXXXXXXX**

**15 STATA 2**

Bruk samme datasett som i forrige oppgave.

Variabelen *hours* måler normal arbeidstid og *married* om kvinnen er gift eller ikke.

a) Jobber gifte og ikke-gifte personer like mange timer? Her er det viktig å huske på eldre arbeidstakere oftere er gift enn de yngre arbeidstakerne, så du må kontrollere for alder. Undersøk dette med å bruke regresjonsanalyse. Hva er estimatet for differansen i arbeidstid for gifte og ugifte, kontrollert for alder? Skriv svaret her med 2

desimaler:

b) Hva er en rimelig tolkning av regresjonskoeffisienten for *married*?

**Velg ett alternativ**

- Det er ikke signifikant forskjell mellom gifte og ikke-gifte når man kontrollerer for alder (p-verdi 0.895)
- Gifte kvinner jobber mindre enn ugifte kvinner, og denne forskjellen er signifikant på 5% nivå.
- Det er ikke en signifikant forskjell mellom gifte og ikke-gifte fordi R-squared er lav (0.02).
- Gifte kvinner jobber mer enn ugifte kvinner, og denne forskjellen er signifikant på 5% nivå.

c) Hva er forventet arbeidstid for gifte kvinner på 40 år? Gjør utregning med 2 desimaler og oppgi svaret her:

d) Det er en mulighet for at forskjell i arbeidstid for gifte og ugifte kvinner er ulik for yngre og noe eldre kvinner. Estimer en modell der du lar betydningen av å være gift variere med alder. Gir resultatet grunn til å tro at det er en slik sammenheng?

**Velg ett alternativ**

- Nei, interaksjonsleddet er statistisk signifikant på 5% nivå
- Ja, estimatet for begge variable er statistisk signifikante på 5% nivå
- Nei, estimatet for alder er ikke signifikant på 5% nivå
- Nei, interaksjonsleddet er ikke statistisk signifikant på 5% nivå
- Nei, estimatet for alder er negativ, mens estimatet for å være gift er positiv.

---

Maks poeng: 4

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**XXXXXXXXXX**

## 16 STATA + håndregning

Bruk samme datasett som i forrige oppgave.

Variabelen ***tfl\_exp*** er et mål på arbeidserfaring (antall år) og ***wage*** er timelønnen.

a) Estimer en enkel, lineær regresjonsmodell og beskriv med egne ord i hvilken grad timelønnen varierer med arbeidserfaring.

**Skriv svar på oppgave a) her:**

b) Undersøk om sammenhengen mellom timelønn og arbeidserfaring er rimelig å beskrive som en lineær eller ikke-lineær sammenheng. Begrunn svaret.

**Skriv svar på oppgave b) her:**

c) Undersøk om sammenhengen mellom timelønn og arbeidserfaring er rimelig å beskrive som en lineær eller ikke-lineær sammenheng hvis du også kontrollerer for variablene ***married***, ***race*** og ***union***. Begrunn svaret.

**Skriv svar på oppgave c) her:**

d) Bruk regresjonsmodellen i oppgave c) og regn ut for hånd forventet timelønn for en ugift, svart kvinne med 10 års arbeidserfaring som ikke er fagorganisert. Vis full utregning på eget ark.

**Skriv svar på oppgave d) her:**

---

Maks poeng: 5

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**XXXXXXXXXX**

## **i Håndregning**

### **Oppgaver med utregning og tekstbesvarelser**

I de etterfølgende oppgavene skal du vise full utregning der det bes om det. Utregninger gjøres for hånd på eget ark som leveres inn. Arket blir scannet og vedlagt den digitale besvarelsen.

For alle disse oppgavene gjelder følgende:

- Svaret på oppgaven skal skrives inn i Inspira.
- Utregninger er dokumentasjon som viser hvordan du har løst oppgaven. Du kan få noe uttelling for utregningen hvis det er en liten feil, men likevel viser god forståelse.
- Dersom oppgaven er **ubesvart** i skjemaet vil ikke håndregningen sjekkes. Du må derfor skrive inn svaret i svarboksen for at svaret skal bli vurdert.

Der det er tekstbesvarelser skal besvarelsene være korte og poengterte.

## 17 Standardavvik

Regn ut standardavviket (engelsk: "standard deviation") for følgende fem tall:

3, 5, 10, 2, 5

Skriv svaret med én desimal her:

Vis full utregning på eget ark.

Maks poeng: 1

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**XXXXXXXX**

**18 Er tiltaket effektivt?**

I en stor fabrikk ønsker de å effektivisere opplæring av nyansatte og tester ut to nye opplæringsopplegg for en spesifikk oppgave for å finne ut hvilken som er mest effektiv. Målet er hvor fort de gjennomfører oppgaven når opplæringen er ferdig. 20 nyansatte blir tilfeldig fordelt i to like store grupper som gjennomfører ulik opplæring. Gruppe 1 gjennomførte oppgaven på i gjennomsnitt 19.1 minutter med standardavvik 4.82, mens gruppe 2 gjennomførte oppgaven på i gjennomsnitt 23.3 minutter med standardavvik 5.56.

Er det grunnlag for å si at den ene opplæringen er mer effektiv enn den andre?

a) Formuler nullhypotese og alternativ hypotese her:

b) Hvilken statistisk test skal du bruke? Skriv svaret her:  og skriv inn antall

frihetsgrader her:

c) Regn ut testen og finn p-verdien. Bruk tre desimaler i utregning. Vis full utregning på eget ark og rapporter t-verdien med én desimal her:

d) Denne t-verdien tilsvarer en p-verdi i intervallet:

**Velg ett alternativ**

- Mellom 0.1 og 0.05
- Mellom 0.05 og 0.025
- Mindre enn 0.025
- Større enn 0.1

e) Hva er konklusjonen på testen? Skriv svaret her:

---

Maks poeng: 4

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**XXXXXXXX**

## 19 Sentralgrenseteoremet

Forklar med **egne ord** hovedpoenget i sentralgrenseteoremet (engelsk: "central limit theorem"). Hva sier teoremet og hva kan vi bruke det til? (Unngå så langt det er mulig å bruke de samme formuleringene i pensum). Skriv kort, men i fulle setninger. Svaret gir max 2 poeng.

**Skriv ditt svar her...**

---

Maks poeng: 2

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**XXXXXXXX**



## 20 **Konfidensintervall**

Et konfidensintervall (engelsk: "confidence interval") omtales gjerne som om man er f.eks. "95% sikker". Forklar med egne ord hva som egentlig menes med dette. Med andre ord: hva er den presise betydningen av et slik konfidensintervall?

**Skriv ditt svar her...**

---

Maks poeng: 2

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**XXXXXXXXXX**

## 21 Diskriminering?

I 1973 ble universitetet i Berkley saksøkt for kjønnsdiskriminering. Bakgrunnen var at på de mest populære studieprogrammene fikk 52% av de mannlige søkerne studieplass, mens tilsvarende andel blant kvinner bare var 42%. Tabellen nedenfor viser fordelingen fordelt på hvilket studieprogram de søkte på.

**Tabell: Andelen av søkere som fikk studieplass (to øverste rader) og antall søknader til hvert studieprogram (to nederste rader).**

	A	B	C	D	E	F	Antall
Menn	62%	63%	37%	33%	28%	6%	52%
Kvinner	82%	68%	34%	35%	24%	7%	42%
Antall menn	825	560	325	417	171	373	2691
Antall kvinner	108	25	395	375	393	341	1835

Ut fra disse tallene, er det grunnlag for å si at det var kjønnsdiskriminering ved opptaket til universitetet? Gi en begrunnelse for svaret ditt som også viser til et kjent statistisk fenomen og hvordan dette slår ut i dette tilfellet.

Skriv ditt svar her...

Maks poeng: 2

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**XXXXXXXXXX**

## 22 Multippel regresjon

Tabellen nedenfor er hentet fra Allison (1999) og viser resultater fra en multippel regresjonsanalyse der yrkesinntekt er utfallsvariabelen (i dollar), og forklaringsvariablene er alder, utdanning og sivilstand. Variabelen **Married** har verdien 1 hvis gift og 0 hvis ikke gift. **Age** og **Schooling** er kontinuerlige variable i hele år. Den andre kolumnen er de estimerte regresjonsparametrene, og den siste kolumnen er p-verdier for parametrene. (SE er ikke oppgitt).

Gjør utregninger på eget ark.

**TABLE 8.5** Regression of Income (Dollars) on Schooling, Marital Status, and Their Interaction

<i>Variable</i>	<i>Coefficient</i>	<i>p Value</i>
Age	652	.003
Schooling	-980	.57
Married	-48,592	.04
Schooling × Married	3,912	.04
Intercept	8,404	

a) Har gifte og ugifte med samme utdanningnivå lik lønn? Begrunn svaret.

**Skriv svar på oppgave a) her:**

b) Hvor mye er forventet inntekt for en gift person som er 35 år og har 11 års utdanning? Hva er forventet inntekt for en tilsvarende person med 5 års utdanning?

Vis utregninger på eget ark.

**Skriv svar på oppgave b) her:**

c) Gir dette grunn til å tro at man kan øke egen yrkesinntekten sin ved å gifte seg? Begrunn svaret.

**Skriv svar på oppgave c) her:**

Maks poeng: 10

**Knytte håndtegninger til denne oppgaven?**

Bruk følgende kode:

**XXXXXXXXXX**

