

Eksamen STV1020 våren 2021 – sensorveiledning i blått

Poenggrenser:

A	fra 80 p
B	fra 65 p
C	fra 45 p
D	fra 35 p
E	fra 30 p
F	opptil 29 p

Del 1 (30 av 100 poeng)

Flervalgsoppgaver, rettes automatisk

Del 2 (70 av 100 poeng)

Du skal svare på alle oppgavene. Tallene i parentes viser antallet poeng du kan få for oppgaven. Hjelpemidler er tillatt, men det er ikke lov å samarbeide med andre om svarene, dette er en individuell eksamen som vil bli sjekket for plagiat både mot vitenskapelige tekster og hverandre.

Dette er vanligvis en skoleeksamen, men ble i år avviklet som en 4-timers hjemmeeksamen. Det betyr at studentene hadde tilgang på både pensum og egne notater: Samarbeid var ikke lov.

I mange land er stater opptatt av å få flere kvinner inn i lønnet arbeid. Men dette har vist seg å være utfordrende. Ifølge Verdensbanken var yrkesdeltakelsen til kvinner over 15 år omtrent 53% i 2014 (mot 73% blant menn). Du er interessert i ulike forklaringer på hvorfor færre kvinner enn menn er i lønnet arbeid, og hva staten kan gjøre med det.

Det første du ser på er hva slags ordninger for foreldrepermisjon man har i ulike land, hypotesen din er at bedre foreldrepermisjonsordninger er assosiert med høyere kvinnelig yrkesdeltakelse.

1. Hva er populasjonen, analyseenheten, den avhengige og den uavhengige variabelen for denne hypotesen? (4 poeng)

Populasjonen: Alle land i verden (1 poeng)

Analyseenheten: Land (1 poeng)

Avhengig variabel: Yrkesdeltakelse blant kvinner (1 poeng)

Uavhengig variabel: Foreldrepermisjonsordninger i land (1 poeng)

Du bestemmer deg for å bruke Verdensbankens estimat for prosent yrkesaktive kvinner som mål på yrkesdeltakelse blant kvinner (*Arbeid_kvinner*) og antallet dager med foreldrepermisjon som er helt eller delvis betalt av staten som mål på foreldrepermisjonsordninger (*Foreldrepermisjon*).

2. Hva er målenivået til disse variablene? Begrunn svaret ditt ved å vise til egenskapene til oppgitt målenivå (4 poeng)

Begge er kontinuerlige mål (2 poeng). Her kan verdier rangeres, man kan si noe om avstand/styrke og en enhets økning er det samme på alle nivåer av variabelen (2 poeng).

3. Hvilket mål på bivariat sammenheng og hvilken hypotesetest passer til å teste hypotesen din om en sammenheng mellom *Foreldrepermisjon* og *Arbeid_kvinner*? Begrunn svaret ditt kort. (2 poeng)

Pearsons r og t-test (1 poeng), fordi X og Y er kontinuerlige (1 poeng). Dersom studenten har svart med feil målenivå på en eller begge variabler i forrige oppgave og dermed har «følgefeil» her kan det gis opp til 1 poeng.

4. Synes du variabelen *Foreldrepermisjon* er en god måte å operasjonalisere foreldrepermisjonsordninger på? (4 poeng)

Her kan nevnes at målet antageligvis har høy grad av reliabilitet og begrepsklarhet ettersom det er snakk om en formell policy (2 poeng), men at det svikter på validitet ettersom man kun fanger opp antall dager som dekkes, men ikke f.eks. hvor mye av lønnen som dekkes, om det inkluderer kun mor, om det dekker alle som arbeider eller kun de i offentlig sektor, hvor fleksibel ordningen er, om ordningen faktisk etterleves, e.l. (2 poeng).

5. Foreslå en annen måte å operasjonalisere foreldrepermisjonsordninger på og forklar hvorfor du synes dette er et bedre mål. (4 poeng)

Her vil det være naturlig å spesifisere hvor mye av permisjonen som skal være dekket, eller på annen måte rette opp i problemene som ble diskutert i forrige svar (2 poeng for et ok mål og 2 poeng for en ok diskusjon)

Figur 1: Antall dager med foreldrepermisjon som er helt eller delvis betalt av staten



6. Figur 1 er en illustrasjon av variabelen *Foreldrepermisjon*. Her har variabelen blitt omkodet for å kunne illustrere den med et begrenset antall ulike farger på kartet. Hva gjør dette med målenivået og egenskapene til variabelen? (2 poeng)

Variabelen er nå ordinal (1 poeng), med 5 kategorier. Her kan verdier fremdeles rangeres, men man kan ikke si noe om avstand / styrke eller hva en økning til neste kategori betyr (1 poeng).

Mange land har ingen eller svært lite betalt permisjon. Du bestemmer deg for å omkode *Foreldrepermisjon* til en dummy-variabel som er 1 dersom landet har mer enn 90 dagers betalt permisjon og 0 ellers (*Foreldrepermisjon_dummy*).

7. Hvilket mål på bivariat sammenheng og hvilken hypotesetest bør du nå bruke for å teste hypotesen din om en sammenheng mellom *Foreldrepermisjon_dummy* og *Arbeid_kvinner*? Begrunn svaret kort. (2 poeng)

Sammenlikne gjennomsnitt ($\bar{Y}_1 - \bar{Y}_2$) og t-test (1 poeng), fordi X er kategorisk og Y kontinuert (1 poeng).

8. Det er også mulig å gjøre en regresjonsanalyse for å se på sammenhengen. Hva er noen av fordelene med å gjøre en regresjonsanalyse istedet for den bivariate analysen du akkurat har foreslått? (2 poeng)

Om man kjører en bivariat OLS modell så vil man få samme resultat, men OLS vil også gi deg informasjon om ting som forklart varians (1 poeng) og man kan inkludere kontrollvariabler (1 poeng). Andre valide argumenter kan også gi poeng.

Du kjører en OLS modell med *Arbeid_kvinner* som avhengig variabel og *Foreldrepermisjon_dummy* som uavhengig variabel, og får ut følgende resultat i R:

Call:

```
lm(formula = arbeid_kvinner ~ Foreldrepermisjon_dummy, data = dta)
```

Residuals:

Min	1Q	Median	3Q	Max
-40.668	-6.522	-0.172	8.854	33.532

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	52.9677	2.0573	25.746	<2e-16 ***
Foreldrepermisjon_dummy	-0.1959	2.7562	-0.071	0.943

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16.2 on 138 degrees of freedom

(1 observation deleted due to missingness)

Multiple R-squared: 3.662e-05, Adjusted R-squared: -0.007209

F-statistic: 0.005054 on 1 and 138 DF, p-value: 0.9434

9. Hva er konstantleddet, stigningsleddet, p-verdien til stigningsleddet, og R² i denne modellen? Gi en substansiell tolkning av hver av dem (8 poeng).

Konstantledd: 52.9677 (1 poeng)

Stigningsledd: -0.1959 (1 poeng)

P-verdi: 0.943 (1 poeng)

R²: 3.662e-05 / 0.00003662 (1 poeng, ½ poeng for å svare med Adjusted r²: -0.007209)

Tolkning (1 poeng for hver ok tolkning, den trenger ikke å være så omfattende som dette):

Konstantleddet på omtrent 53 gir uttrykk for den estimerte yrkesdeltakelsen (53%) blant kvinner i land der Foreldrepermisjon_dummy er 0, dvs. land med 90 dagers betalt permisjon eller mindre. I dette tilfellet gir dette substansiell mening ettersom det er en bivariat modell der referanse kategorien til forklaringsvariabelen finnes i dataene – dette er gjennomsnittet for yrkesdeltakelse for landene med 90 dagers betalt permisjon eller mindre.

Stigningsleddet gir den estimerte endringen i yrkesdeltakelse gitt at vi går fra 0 til 1 på forklaringsvariabelen. I dette tilfellet betyr dette en estimert reduksjon på kvinnelig yrkesdeltakelse på omtrent 0.2% om vi går fra 90 dager eller mindre til mer enn 90 dagers foreldrepermisjon. Dette er en svak substansiell effekt og den er heller ikke signifikant.

En p-verdi på 0.943 er langt fra å regnes som statistisk signifikant. Denne p-verdien betyr at selv om det ikke er noen sammenheng mellom uavhengig eller avhengig variabel så er det 94.3% sjans for å få en koeffisient som er så stor eller større – det er med andre ord svært sannsynlig å få dette estimatet på en sammenheng ved ren tilfeldighet.

R² går mellom 0 og 1 og en R² på 0.00003662 tyder på at den uavhengige variabelen forklarer svært lite av variasjonen i avhengig variabel.

10. Er hypotesen din styrket eller svekket av denne analysen? (2 poeng)

Hypotesen er svekket ettersom det ikke er noen signifikant eller substansiell stor sammenheng (2 poeng).

11. Du bestemmer deg for å kode om foreldrepermisjonsvariabelen en gang til slik at land får verdien 1 dersom de har 90 dagers permisjon eller mindre og verdien 0 ellers. Du kaller variabelen *Lite_foreldrepermisjon*. Om du hadde kjørt modellen i Tabell 1 igjen, hva ville konstantleddet blitt? Hva ville koeffisienten til *Lite_foreldrepermisjon* blitt? Og hva ville R² blitt? Forklar endringene kort. (5 poeng)

Her snur vi bare variabelen, modellen er egentlig den samme, så konstantleddet blir 52.9677-0.1959 = 52.7718 ettersom vi nå endrer referansekategori til land med mer enn 90 dagers foreldrepermisjon; koeffisienten blir 0.1959 ettersom variabelen er snudd, og R² er uendret ettersom det er den samme modellen og samme data som er inkludert. 1 poeng for hvert riktige svar og 2 poeng for ok forklaringer på minst to av disse. Dette er en oppgave som ikke er så lett å slå opp i boken eller notatene, så den krever ganske stor forståelse, så ordlyden på forklaringene kommer antageligvis til å variere en del og vi bør være ganske snille på å belønne gode forsøk selv om de ikke kommer helt i mål.

En kollega påpeker at det er mulig sammenhengen du ser (eller mangelen på en sammenheng) kan være spuriøs, og at du ikke kan forkaste hypotesen din før du har kontrollert for andre forskjeller mellom land. Det er mange ting man kan kontrollere for, men du velger å kontrollere for yrkesdeltakelsen til menn (også målt i prosent, med variabelen *Arbeid_menn*).

Her er resultatet av Modell 2 i R:

```
Call:
lm(formula = arbeid_kvinner ~ Foreldrepermisjon_dummy + arbeid_menn,
    data = dta)
```

Residuals:

Min	1Q	Median	3Q	Max
-38.499	-7.624	2.372	9.175	31.377

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-10.8206	12.9417	-0.836	0.4046
Foreldrepermisjon_dummy	5.6655	2.8039	2.021	0.0453 *
arbeid_menn	0.8240	0.1654	4.983	1.86e-06 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 14.96 on 137 degrees of freedom
(1 observation deleted due to missingness)

Multiple R-squared: 0.1535, Adjusted R-squared: 0.1411

F-statistic: 12.42 on 2 and 137 DF, p-value: 1.106e-05

12. Hvordan kan vi tolke konstantleddet i Modell 2? (3 poeng)

Dette er estimatet for kvinners yrkesdeltakelse i land med mannlig yrkesdeltakelse på null og lite foreldrepermisjon (1 poeng). Tallet er negativt, noe som gir uttrykk for å at kvinners yrkesdeltakelse generelt er lavere enn menns i dataene (1 poeng). Gitt at kvinners yrkesdeltakelse er en prosent og menns yrkesdeltakelse aldri er 0, så er dette ikke et substansielt meningsfullt estimat (1 poeng). Man kan få poeng selv om ordlyden er litt annerledes såfremt man viser forståelse for hva som foregår i modellen.

13. Tolk endringen fra Modell 1 til Modell 2 i R2, i stigningsleddet til Foreldrepermisjon_dummy og p-verdien til dette stigningsleddet. (6 poeng)

1 poeng for en teknisk og 1 poeng for en substansiell tolkning av hver av de tre, tolkningene trenger ikke å være så omfattende som dette, og trenger ikke dekke akkurat denne ordlyden heller så lenge den er riktig og viser forståelse. Om man kun tolker Modell 2 og ikke sier noe om endringen fra Modell 1 til Modell 2 kan man få halv poenguttelling:

Endring i R2: Ved å inkludere arbeid_menn går vi fra en svært liten R2 til en R2 på 0.1535, det betyr at langt mer av variansen i arbeid_kvinner er forklart av denne modellen (1 poeng). Dette tyder på at arbeid_menn og arbeid_kvinner er sterk korrelert, noe som ikke er så rart ettersom både arbeidsforhold og måten man måler arbeid på antageligvis varierer i ulike land (1 poeng).

Endring i stigningsledd: Stigningsleddet for Foreldrepermisjon_dummy går fra å være substansielt svært lite til å være ganske stort. I Modell 2 ser vi at det å gå fra å være et land med lite foreldrepermisjon til et land med mer enn 90 dagers permisjon er assosiert med en økning i kvinners yrkesdeltakelse på omtrent 6 prosentpoeng (1 poeng). Dette tyder på en positiv sammenheng mellom foreldrepermisjon og kvinners yrkesdeltakelse, men denne sammenhengen kan fremdeles være spuriøs dersom det er andre egenskaper til land som

påvirker både om et land har gode permisjonsordninger og forskjellen på menns og kvinners yrkesdeltakelse (1 poeng).

Endring i p-verdi: Vi går også fra en p-verdi på 0.943 til en på 0.0453, det betyr at vi går fra en ikke-signifikant til en signifikant p-verdi (1 poeng). Det er vanlig å kalle et estimat statistisk signifikant dersom det har en p-verdi som er mindre enn 5%, noe som betyr at det er mindre enn 5% sjanse for å få et estimat (en koeffisient) som er så stort eller større enn det man fikk gitt at det ikke er noen sammenheng mellom variablene. En p-verdi 0.0453 betyr altså at det er 4.53% sjanse for å få dette estimatet ved ren tilfeldighet selv om det ikke er noen sammenheng mellom variablene. Dette styrker hypotesen om at det er en sammenheng (1 poeng).

14. Vil du si at intern eller ekstern validitet er den største utfordringen i denne analysen?
Begrunn svaret (3 poeng)

1 poeng for å si intern validitet. Her er intern validitet et åpenbart problem ettersom det fremdeles er mange ting vi burde kontrollert for i modellen. Det er også vanskelig å kontrollere for alle relevante faktorer i analyser på land-nivå (1 poeng). Ekstern validitet er ikke et like stort problem ettersom svært mange land i verden er inkludert i studien (man kan se i tabellene at det er 140 land i dataene, men det er ikke nødvendig å spesifisere dette), men man kan argumentere for at landene som er inkludert antageligvis er systematisk forskjellige fra de som ikke er inkludert (vi mangler som regel data på land som er fattigere, med mer konflikt osv.) (1 poeng). Også ok å påpeke at intern validitet er en forutsetning for ekstern validitet og at man dermed har problemer med begge.

Du blir nå særlig opptatt av landene som har gode foreldrepermisjonsordninger, men som allikevel har lav kvinnelig yrkesdeltakelse. Du bestemmer deg for å gjøre en case-studie av Japan for å forstå mer om hva som skjer i disse landene.

15. Hva er universet av mulige caser i dette tilfellet? (2 poeng)

Alle land med «gode» foreldrepermisjonsordninger og også lav kvinnelig yrkesdeltakelse. Gitt operasjonaliseringen over det naturlig å tenke at det er land med mer enn 90 dagers permisjon som også har lav kvinnelig yrkesdeltakelse. Et poeng for å kun si land med «gode» foreldrepermisjonsordninger eller å kun si land med lav kvinnelig yrkesdeltakelse. Ingen uttelling for å si bare land.

Du er særlig interessert i å teste teorien om at tradisjonelle holdninger til kjønnsroller er det som begrenser kvinners arbeidsdeltakelse.

16. Hva er to testbare empiriske konsekvenser (hypoteser) av denne teorien? (4 poeng)

Her er det mange å ta av, men noen eksempler er (1 poeng for hver EK som er empirisk testbar og 1 poeng for at den tester teorien):

- Sjefer med mer tradisjonelle holdninger til kjønnsroller ansetter færre kvinner*
- Kvinner med mer tradisjonelle holdninger til kjønnsroller jobber mindre*
- Kvinner med menn som har mer tradisjonelle holdninger til kjønnsroller jobber mindre*

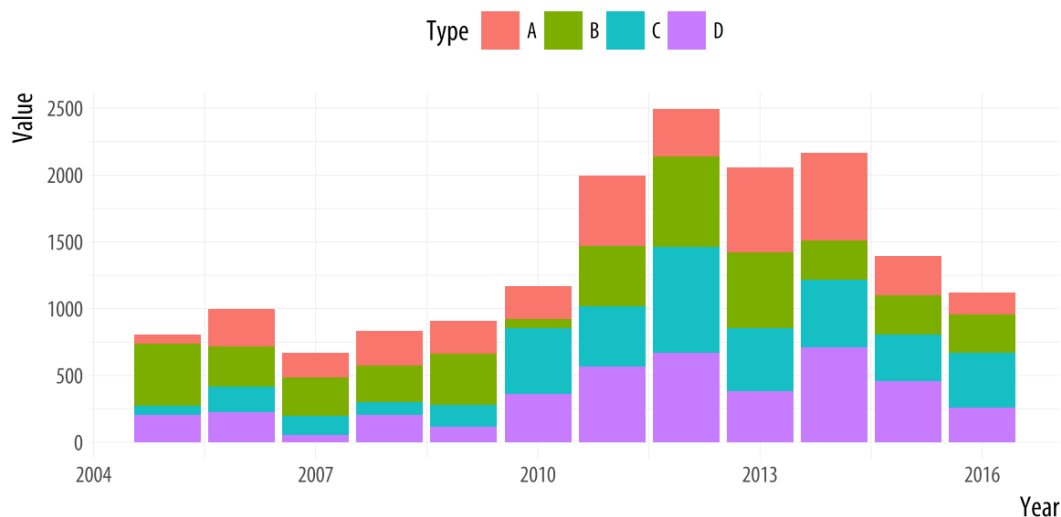
17. Er dette et eksempel på en teoritestende eller en teoriutviklende case-studie? Forklar

hvorfor. (2 poeng)

Det er mest nærliggende å svare at dette er en teoritestende case-studie (1 poeng) ettersom vi har sagt at målet er å teste en teori om konservative holdninger (1 poeng). Men man kan også få full uttelling for å argumentere for at et så vagt teoretisk utgangspunkt lett kan føre til teoriutvikling eller at dette er en teoriutviklende studie fordi de uttalte motivet for studien var å forstå hva som foregår på bakken, men her får man kun poeng om argumentet henger sammen med svaret.

18. Du får tak i survey-data på hvor mange kvinner som jobber i ulike sektorer (A, B, C, D), som du tenker er nyttig for å forstå mer av konteksten du studerer. Du plotter dataene i Figur 2. Hva er problematisk med denne grafen, hvis målet ditt er å vise endringer over tid i andelen som jobber i de ulike sektorene? (3 poeng)

Figur 2: Antallet kvinner som jobber i ulike sektorer.



Her får man 1 poeng for hvert ok argument man har, maks 3 poeng. Her er noen ting man kan påpeke, men andre valide poenger kan også gi uttelling:

- *Det er ingen stabil grunnlinje, noe som gjør at det er vanskelig å sammenligne endringer fra år til år innad i kategoriene.*
- *Det er lett å se på endringer over tid for kategori D, siden den har en stabil grunnlinje, men ikke for de andre kategoriene.*
- *Fargene er problematiske ettersom de ikke vil synes dersom man printer grafen i svart-hvitt.*
- *X-aksen er ok beskrevet, men Y-aksen og beskrivelsen av variabelen («Type») er lite informative.*
- *Det er også problematisk å plote antall når målet er å se på andeler.*

19. Du bestemmer deg for å gjøre formelle intervjuer for å skjønne mer om hvordan individuelle holdninger kan påvirke valgene til kvinner. Hva kjennetegner denne typen intervjuer og hva er noen fordeler og ulemper ved dem? (4 poeng)

Kjennetegn: Det kjennetegnes av at følger en intervjuguide der ordlyd og rekkefølge på spørsmålene ligger fast. Det er lite fleksibilitet i strukturen på samtalen og alt gjøres så likt som mulig (2 poeng for å nevne minst to av disse tingene).

Fordeler er høy grad av reliabilitet, man vil få ganske like svar selv om det er ulike personer som gjennomfører intervjuene (1 poeng for å nevne minst en av disse tingene). Ulemper er at det er vanskelig å bygge tillit eller følge opp interessante nye ideer som dukker opp (1 poeng for å nevne minst en av disse tingene).

20. Du velger å gjennomføre intervjuene i Tokyo, der du har mange kontakter og kjenner konteksten godt. Hva er noen fordeler og ulemper ved denne tilnærmingen? (4 poeng)

Fordeler: lett tilgang, bruke mindre tid/ressurser på å finne folk, lettere å sammenligne svarene til respondentene ettersom de er fra samme kontekst, lettere å skape tillit og forstå svarene om man kjenner konteksten godt. Alt dette kan bidra til bedre intern validitet i studien (opp til 2 poeng for å nevne to av disse tingene eller noe annet fornuftig).

Ulemper: man får et skjevt utvalg av befolkningen, noe som truer ekstern validitet. Ettersom alle er fra samme kontekst er det mulig man får færre perspektiver representert (opp til 2 poeng for å nevne to av disse tingene eller noe annet fornuftig).