

STV1020 Politisk analyse 2: Forskningsdesign og kvantitative metoder

– eksamen våren 2021

Oppgavetekst fra UiOs eksamen våren 2021, STV1020 (overskriften er eksamenssett.no sin; utsnittene under er klippet uendret ut av UiOs publiserte dokument)

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.

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)

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

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)

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)

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)

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

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)

13. Tolk endringen fra Modell 1 til Modell 2 i R^2 , i stigningsleddet til Foreldrepermisjon_dummy og p-verdien til dette stigningsleddet. (6 poeng)
14. Vil du si at intern eller ekstern validitet er den største utfordringen i denne analysen? Begrunn svaret (3 poeng)

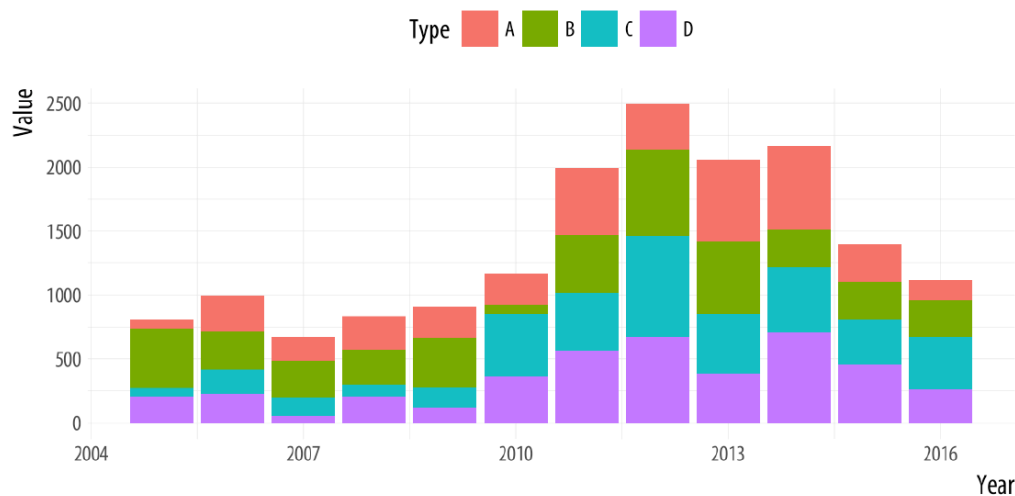
Du blir nå særlig opptatt av landene som har gode foreldrepermisjonsordninger, men som allikevel har lav kvinnelig yrkesdeltagelse. Du bestemmer deg for å gjøre en case-studie av Japan for å forstå mer om hva som skjer i disse landene.

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

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)
17. Er dette et eksempel på en teoritestende eller en teoriutviklende case-studie? Forklar hvorfor. (2 poeng)
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 plottet 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.



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)
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)

Kilde: Oppgaveteksten er gjengitt ordrett fra Universitetet i Oslo, Institutt for statsvitenskap: «Eksamen STV1020 våren 2021 – sensorveiledning i blått» (eksamenstv1020_v21_sensorveiledning.pdf), publisert på uio.no under STV1020 – Tidligere eksamensoppgaver. UiO har ikke publisert oppgavesettet separat; dette arket inneholder bare oppgaveteksten, klippet ut av dokumentet uten endringer. Sensorveiledningens tekst er utelatt. Flervalgsdelen (Del 1) er ikke publisert av UiO; her gjengis Del 2. Satt opp av eksamenssett.no, som ikke er tilknyttet UiO.