

KRS4105 — Kvantitativ metode

Hjemmeeksamen høst 2020

Utlevering av oppgaven: 9. november kl. 09:00

Innleveringsfrist: 11. november kl. 15:00

Eksamensoppgaven

Du kan få maksimalt 42 poeng på eksamen. Hver oppgavetekst følges av en parentes som angir det maksimale antallet poeng du kan få på hver oppgave.

For å bestå må du ha minst 21 poeng.

Oppgave 1

Du har fått i oppdrag å undersøke utbredtheten/forekomsten av vold i den norske befolkningen. Du bestemmer deg for å gjennomføre en spørreundersøkelse der du vil finne ut hvor stor andel av befolkningen mellom 18 og 75 år som har vært utsatt for vold det siste året.

Svar på følgende spørsmål (a-e) på maksimalt 1 A4 til sammen (totalt 8 poeng).

- a. Angi en utvalgsmetode som gir gode forutsetninger for å undersøke hvor stor andel av befolkningen i aldersspennet 18-75 år som har vært utsatt for vold. For å få maksimal poenguttelling på denne oppgaven skal det framgå tydelig hva som skiller denne utvalgsmetoden fra andre utvalgsmetoder som gir dårligere forutsetninger (2 poeng).

Når du har samlet inn dataene dine får du i oppdrag å finne ut om utsatthet for vold varierer mellom ulike grupper i befolkningen. De deskriptive analysene viser blant annet at personer som har fullført videregående utdanning i betydelig mindre grad enn personer som ikke har fullført videregående rapporterer å ha blitt utsatt for vold: forskjellen er på hele 8 prosentpoeng. Du gjennomfører en



hypotesetest der du tester nullhypotesen om at det ikke finnes en sammenheng mellom utdanningsnivå og utsatthet for vold, og p-verdien fra signifikanstesten er på 0.16.

- b. Hvilken konklusjon om sammenhengen mellom utdanningsnivå og utsatthet for vold trekker du på bakgrunn av hypotesetesten dersom du tar utgangspunkt i et signifikansnivå på 5% ($\alpha=0.05$)? (1 poeng).
- c. Nevn ett avgjørende valg du tar i forbindelse med datainnsamlingen som kommer til å påvirke presisjonen i estimeringene dine, og forklar kort med egne ord hvorfor dette valget har noe å si (1 poeng).
- d. Oppgi én variabel du mener det vil være viktig å kontrollere for dersom du vil finne ut om det finnes en årsakssammenheng mellom utdanningsnivå og utsatthet for vold. Begrunn valget ditt kort (2 poeng).

De nordiske landene har gode forutsetninger for å gjøre registerbasert forskningsstudier. I en fordypende analyse planlegger du å gjennomføre en registerbasert kausalstudie av ulike årsaker til at folk utsettes for vold.

- e. Redegjør kort for én viktig metodisk styrke og én viktig metodisk svakhet ved å bruke registerdata til å undersøke årsakssammenhenger (2 poeng).

Oppgave 2

På starten av 1990-tallet formulerte Michael Gottfredsson og Travis Hirschi (G&H) den innflytelsesrike men kontroversielle teorien om at den viktigste forklaringsfaktoren bak individuelle forskjeller i tilbøyeligheten til å begå kriminalitet var individets underliggende evne til selvkontroll. De mente også at selvkontroll var nøkkelen til å forklare hvorfor gutter i større grad enn jenter begår kriminelle handlinger. Du bestemmer deg for å teste denne teorien, og benytter en spørreundersøkelse blant skoleelever, der elevene selv har fått svare på spørsmål om selvkontroll, kriminalitet og kjønn. Du lager én indeks for selvkontroll og én indeks for kriminalitet. Indeksen for selvkontroll (heretter kalt "Selvkontroll") varierer mellom 0 og 5, hvor 0 er den laveste graden av selvkontroll og 5 er den høyeste. Indeksen for kriminalitet (heretter kalt "Kriminalitet") angir totalsummen av antall selvrappørterte lovbrudd av ulike typer som respondenten har begått det sist året. Du lager også en dummy-variabel av for kjønn som tar verdien 1 for gutter og 0 for jenter (heretter kalt "Gutt"). Du estimerer flere lineære regresjonsmodeller; disse defineres i Modell 1-4 under.

Svar på følgende spørsmål (a-g) på maksimalt 1 A4 til sammen (totalt 8 poeng).

Innledningsvis tester du hypotesen om en bivariat sammenheng mellom selvkontroll og kriminalitet, og estimerer følgende modell:

$$\text{Modell 1:} \quad \text{Kriminalitet} = 10 - 1.3\text{Selvkontroll}$$

p-verdi (F-test): 0.000

- a. Er resultatet fra Modell 1 i tråd med G&Hs teori? Forklar kort hvorfor/hvorfor ikke (1 poeng).

For å teste hypotesen om betydningen av selvkontroll for å forklare kjønnsforskjeller i kriminalitet, estimerer du de følgende to modellene (Modell 2 og Modell 3). Alle regresjonskoeffisientene i de to modellene er signifikante på 5%-nivå, med unntak av estimatet for *Gutt* i Modell 3 (som ikke er statistisk signifikant på dette signifikansnivået).

$$\text{Modell 2: } \text{Kriminalitet} = 2 + 6_{\text{Gutt}}$$

$$\text{Modell 3: } \text{Kriminalitet} = 12 + 0.5_{\text{Gutt}} - 2_{\text{Selvkontroll}}$$

- b. Tolk regressionskoeffisienten for *Gutt* i Modell 2 (1 poeng).
c. Tolk regressionskoeffisienten for *Selvkontroll* i Modell 3 (1 poeng).
d. Gjør en sammenligning av regressionskoeffisienten for *Gutt* i Modell 2 og Modell 3. Er disse resultatet i linje med G&Hs teori? Forklar kort hvorfor/hvorfor ikke (2 poeng).

Du vil videre nyansere analysen din av sammenhengen mellom selvkontroll og kriminalitet, og estimerer følgende modell (Modell 4). Alle regresjonskoeffisientene i modellen er signifikante på 5%-nivå.

$$\text{Modell 4: } \text{Kriminalitet} = 15 + 0.7_{\text{Gutt}} - 4.8_{\text{Selvkontroll}} + 0.4_{\text{Selvkontroll} * \text{Selvkontroll}}$$

- e. Hvilken egenskap ved sammenhengen mellom selvkontroll og kriminalitet er det du først og fremst får kunnskap om i Modell 4, og hvilken teknikk har du benyttet deg av for å få fram denne egenskapen? (1 poeng).
f. Beskriv sammenhengen mellom selvkontroll og kriminalitet med utgangspunkt i informasjonen fra Modell 4 (1 poeng).
g. Hvor mange lovbrudd kan vi forvente at en jente som scorer 3 på selvkontroll-variabelen har begått ifølge Modell 4? Oppgi svaret med én desimal (1 poeng).

Oppgave 3

I debatten om kriminalitet og utsatthet for kriminalitet framstår det ofte som at det finnes én gruppe individer som *begår* lovbrudd og en annen gruppe individer som *utsettes* for lovbrudd. Samtidig finnes det en del forskning som viser at det er en betydelig overlapp mellom disse gruppene, og enkelte kriminologiske teorier hevder at både utsatthet for kriminalitet og egen kriminalitet har felles

forklaringsfaktorer som finnes i en persons bosted, familie, vennekrets og skolemiljø, i tillegg til i individuelle egenskaper.

- Din overordnede oppgave er gjennomføre et ”mini-forskningsprosjekt” hvor du skal undersøke sammenheng mellom utsatthet for kriminalitet (X) og egen kriminalitet (Y) empirisk, samt å undersøke om denne sammenhengen helt eller delvis kan forklares av faktorer knyttet til ett av følgende tema/områder: bosted, familie, vennekrets, skolemiljø *eller* individuelle egenskaper. Merk at du kun skal velge å fokusere på ett av disse temaene/områdene, ikke flere (totalt 26 poeng).

Din studiepopulasjon er ungdomsskoleelever i 9. klasse og videregående skoleelever i 2. klasse (dvs. elever i Vg2), og du anvender Stockholmsenkäten 2014 som ditt datamateriale. Denne spørreundersøkelsen inneholder spørsmål om utsatthet for kriminalitet og egen deltakelse i kriminalitet, samt en del andre spørsmål om sosial bakgrunn, holdninger, venner, skole, familie og bosted. Dette datamaterialet, som du skal bruke til å besvare Oppgave 3, heter ”SE2014 KRS4105 ht20 eksamen” og skal lastes ned fra Canvas (se under Filer/Eksamen). I samme mappe ligger det en variabelliste du kan bruke hvis du vil.

Som avhengig variabel (Y) skal du bruke en eksisterende sum-indeks (kalt ”brott_ind”), som er basert på flere spørsmål om hvorvidt en respondent har begått ulike typer lovbrudd. Det er obligatorisk å undersøke den overordnede sammenhengen mellom utsatthet for kriminalitet og egen kriminalitet i en lineær regresjonsanalyse, men du står relativt fritt til å bygge opp selve analysen slik du selv mener er hensiktsmessig. Du skal blant annet selv velge hvordan du ønsker å måle utsatthet, samt hvilke andre uavhengige variabler du ønsker å kontrollere for. Det finnes med andre ord mange ulike innfallsvinkler for å besvare denne oppgaven, og det er mulig å gjøre både mer og mindre komplekse/nyanserte analyser. Det fremgår av hver deloppgave (a-h under) hva som kreves for å få maksimal poenguttelling. Uansett hva du velger å gjøre skal det framgå klart av besvarelsen hvorfor du har valgt å besvare oppgaven slik du har gjort. Valgene dine skal begrunnes metodisk og/eller teoretisk.

Merk at du skal inkludere minst ett mål på utsatthet for kriminalitet, og at andre kontrollvariabler skal ha forankring i ditt valg av tema (bosted, familie, vennekrets, skolemiljø eller individuelle egenskaper). Hvilket tema du velger må fremgå av oppgavebesvarelsen.

I alle spørsmål som handler om statistisk signifikans kan du ta utgangspunkt i et signifikansnivå på 5% ($\alpha=0.05$).

Besvar oppgave 3 ved å gi en kort introduksjon til ditt valg av tema (bosted, familie, vennekrets, skolemiljø eller individuelle egenskaper), for deretter å besvare alle deloppgavene (a-h). Hele oppgave 3 skal besvares på maksimalt 8 A4-sider, inkludert tabeller og figurer, men ekskludert evt. referanseliste og STATA-kode. Kopier STATA-koden du har brukt direkte fra do-filen og lim den inn som et vedlegg til slutt i oppgaven.

Struktur oppgavebesvarelsen i følgende deler:

- a. Operasjonalisering av den sentrale uavhengige variabelen ”utsatthet”. Her skal du benytte variablene som begynner med ”f78”. Du står fritt til å velge om du vil benytte én eller flere av disse variablene for å måle utsatthet for kriminalitet (2 poeng).
 - For å få full poenguttelling for denne oppgaven må du konstruere en variabel som baserer seg på minst to av variablene som måler utsatthet.
- b. Operasjonalisering av én eller flere kontrollvariabler som kan brukes til å undersøke noen av de kriminologiske forklaringene på sammenhengen mellom utsatthet for kriminalitet og egen kriminalitet, dvs. bosted, familie, vennekrets, skolemiljø eller individuelle egenskaper. Merk at uansett valg av operasjonalisering av kontrollvariabler skal du holde deg innenfor ett av disse temaene (2 poeng).
 - For å få full poenguttelling for denne oppgaven må du enten lage en indeks basert på minst to variabler, eller velge to eller flere variabler som inkluderes som separate kontrollvariabler i den multivariate modellen. Du må også angi hvilket målnivå de(n) nye variabelen(e) er på.
- c. En univariat analyse av den avhengige variabelen (2 poeng).
 - For å få full poenguttelling for denne oppgaven må du oppgi minst ett mål på sentraltendens og ett mål på spredning, og beskrive med egne ord hvordan fordelingen ser ut.
 - Du kan presentere resultatet ved å kopiere tabellen og/eller figuren direkte fra STATA og lime den inn i besvarelsen.
- d. En bivariat lineær regresjonsmodell (Modell 1) med resultatene presentert i en tabell og med påfølgende tolkning av innholdet i tabellen (4 poeng).
 - For å få full poenguttelling for denne oppgaven skal modellen være korrekt spesifisert, du må oppgi en tydelig tolkning av et mål som angir hvor godt modellen passer til data, du må gi en korrekt tolkning av regresjonskoeffisienten for den uavhengige variabelen, og du må oppgi hvorvidt det finnes en statistisk signifikant sammenheng mellom den uavhengige og den avhengige variabelen i modellen.
 - Vurder om du i STATA-koden skal spesifisere hvorvidt den uavhengige variabelen skal inkluderes som en faktorvariabel (dvs. et dummy-sett) eller som en kontinuerlig variabel.
 - Du kan presentere resultatet fra Modell 1 ved å kopiere tabellen direkte fra STATA-outputet og lime den inn i besvarelsen (se tips i ” Generell instruks”).
- e. En multivariat lineær regresjonsmodell (Modell 2) med resultatene presentert i en tabell og med påfølgende tolkning av innholdet i tabellen (5 poeng).
 - For å få full poenguttelling for denne oppgaven skal modellen være korrekt spesifisert, du må oppgi en tydelig tolkning av et mål som angir hvor godt modellen passer til data, du må gi en korrekt tolkning av samtlige regresjonskoeffisienter for de uavhengige variablene, og du må gi en vurdering av hvorvidt samtlige regresjonskoeffisienten for de uavhengige variablene er statistisk signifikante.

- Vurder for hver av de uavhengige variablene om du i STATA-koden skal spesifisere hvorvidt denne skal inkluderes som en faktorvariabel (dvs. et dummy-sett) eller som en kontinuerlig variabel.
- Du kan presentere resultatet fra Modell 2 ved å kopiere tabellen direkte fra STATA-outputet og lime den inn i besvarelsen (se tips i ” Generell instruks”).
- f. En analyse av hva som har hendt med sammenhengen mellom utsatthet for kriminalitet og egen kriminalitet etter at du har lagt til din/dine valgte kontrollvariabel(er) (3 poeng).
 - For å få full poenguttelling for denne oppgaven må din sammenligning av den bivariate (Modell 1) og multivariate modellen (Modell 2) være korrekt og tydelig. Videre må tolkningen gjøres i lys av de kriminologiske teoriene som sier at sammenhengen mellom utsatthet for kriminalitet og egen kriminalitet har felles forklaringsfaktorer. Gir resultatene empirisk støtte til slike teorier eller ikke, og hvordan kan du basert på Modell 1 og Modell 2 se dette?
- g. En tredje modell (Modell 3) der du tester forutsetningen om linearitet eller additivitet. Merk at du kun skal undersøke én av disse forutsetningene. Om du velger linearitet skal du kun teste forutsetningen for én uavhengig variabel i modellen, og om du velger additivitet skal du kun teste forutsetningen for én mulig kombinasjon av de uavhengige variablene. Begrunn kort hvorfor du velger den/de variablene(e) du gjør (4 poeng)
 - For å få full poenguttelling for denne oppgaven må du spesifisere en modell som undersøker enten linearitet eller additivitet, du må gjøre en tolking av om forutsetningen du velger holder eller ikke (dvs. at du må avklare om det finnes en form for ikke-linearitet eller en interaksjon), og du må gjøre en korrekt avlesning og tolkning av modellen.
 - Vurder for hver av de uavhengige variablene om du i STATA-koden skal spesifisere hvorvidt denne skal inkluderes som en faktorvariabel (dvs. et dummy-sett) eller som en kontinuerlig variabel.
 - Du kan presentere resultatet fra Modell 3 ved å kopiere tabellen direkte fra STATA-outputet og lime den inn i besvarelsen (se tips i ” Generell instruks”).
 - I tolkningen av linearitet/additivitet kan du velge å illustrere sammenhengen med en figur, som også kan kopieres direkte fra STATA og inn i besvarelsen.
- h. En problematisering av din multivariate modell og muligheten for å trekke kausale konklusjoner på bakgrunn av analysen du har gjennomført, samt ett forslag til hvordan studien kunne blitt videreutviklet for å bedre kunne belyse årsakssammenhenger (4 poeng).
 - For å få full poenguttelling for denne oppgaven skal både problematiseringen og forslaget ditt være tydelig relatert til skillet mellom korrelasjon og kausalitet.

Generell instruks

Du leverer oppgaven ved å laste opp ett PDF-dokument. Vær oppmerksom på at kun én fil skal leveres inn.

Alle hjelpemidler er tillatt under eksamen, inkl. lærebok, notater forelesningsslides og andre ressurser. Regler for kildehenvisning er avgjørende for om anvendelsen av hjelpemidlene er lovlig, mer om dette finner du her:

<http://www.uio.no/studier/program/kriminologi/kildehenvisning.html>

Både oppgaveløsningen og selve besvarelsen skal være et individuelt arbeid, og alle former for samarbeid er derfor forbudt. Plagiatkontroll vil bli gjennomført for å sikre at alle leverer en selvstendig besvarelse.

Hele oppgavesettet skal til sammen besvares med maksimalt 4000 ord fordelt på inntil 10 A4-sider. Sideantallet inkluderer tabeller og figurer, men ekskluderer forside (valgfritt), referanseliste (obligatorisk dersom du henviser til pensum/annen litteratur) og vedlegg med STATA-kode (obligatorisk). Følgende formatteringer skal benyttes i selve besvarelsen (dvs. de 10 A4-sidene):

- Skrifttype Times New Roman, skriftstørrelse 12
- 1 ½ linjeavstand
- Valgfritt referansesystem (ett system må benyttes i hele besvarelsen)

Forsiden, referanselisten og vedlegget med STATA-kode kan formatteres slik du ønsker, men sørg for at du leverer en tydelig og lesbar besvarelse.

Tips til STATA-output: Dersom du klipper tabeller som ren tekst (Copy eller Ctrl+C) direkte fra Output-vinduet i STATA til Word, vil disse se best ut dersom du benytter enkel linjeavstand, skrifttype Courier New og en liten nok skriftstørrelse til at tabellen kan limes inn i Word-dokumentet uten at det legges inn linjeskift. Dersom du benytter denne formatteringen kun for tabellene dine, vil ikke dette bryte med formatteringskravene over. Dersom du kopierer tabellene som tabeller (Copy table) fra Output-vinduet i STATA, hender det (litt avhengig av Windows-oppsettet på maskinen din) at man må lime inn i Excel *før* man kan kopiere videre fra Excel og lime inn i en Word-tabell. Pass på at dere lager tabeller som er oversiktlige og lesbare.

Eksamen besvares anonymt med *kandidatnummer*. Kandidatnummeret ditt finner du i Studentweb.

Kontaktinformasjon under eksamen:

Kontaktpunkt: Hvis det oppstår problemer under eksamen, ta kontakt med, Info@jus.uio.no, merk mailen: **HASTER-EKSAMEN-KRS4105**