

Sensurveiledning SOS1002 våren 2020

Merk: En del studenter hadde problemer med å laste ned eksamensoppgaven, og kom derfor litt seint gang (15-30 minutter)

Oppgave 1

Institutt for pedagogikk ønsker å vurdere hvor mye studentene lærer av ulike former for studieteknikk. Før de gjennomfører studien vil de gjerne ha innspill fra deg. De foreslår følgende design:

I tre uker skal bachelorstudentene kun bruke PC i alle emner, mens masterstudentene kun får lov til å skrive for hånd. Etter at de tre ukene har gått, blir begge gruppene bedt om å delta i en spørreundersøkelse der alle skal svare på hvor mye de synes de har lært i løpet av de tre ukene.

- a) Hva slags forskningsdesign er det som er foreslått? Vurder intern og ekstern validitet i det foreslåtte opplegget.

Ut fra figuren på side 133 i Ringdal (2018) er dette et kvasi-eksperiment med en eksperimentgruppe og en kontrollgruppe, der effekten bare måles etter at eksperimentet er gjennomført (Figur 7.1, gruppe 2 av 4). Etter som eksperimentet foregår i naturlige omgivelser må det videre karakteriseres som et felteksperiment. På side 138 presenterer Ringdal (2018) en figur 7.2 som viser at felteksperiment har lav indre validitet (kontroll over årsaksmekanismen) men høy ytre validitet (realisme).

Ekstern validitet: her venter vi en diskusjon av om funnene kan generaliseres til andre enn de som deltar. Hovedproblemet her er at forsøkspersonene ikke er tilfeldig fordelt på eksperiment- og kontrollgruppa (randomisering). De er heller ikke tilfeldig trukket, så spørsmålet er hva som er populasjonen her, og om pedagogikkstudenter er annerledes enn andre studenter (eller elever).

Intern validitet: Er målingen av årsak – virkning-forholdet riktig? Her kan studentene trekke fram seleksjon, som vil være et stort problem her, frafall, modning (som kan være relevant her, og et argument for å bruke en pre-test).

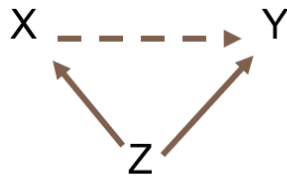
- b) Beskriv svakheter ved studien og kom med mulige forslag til forbedring.

På side 136 og 137 hevder Ringdal at de viktigste truslene mot indre validitet er seleksjonseffekter, frafallseffekter, instrumenteffekter, historisk relevante hendelser i samme periode og effekter av modning. Av disse effektene er det særlig seleksjonseffektene som er problemet i det skisserte designet, og da særlig de seleksjonseffektene som sannsynligvis virker inn når vi sammenligner en gruppe bachelorstudenter med en gruppe masterstudenter. Disse seleksjonseffektene hadde sannsynligvis blitt mindre hvis vi hadde valgt ut mer like studentgrupper. Her bør også studentene vurdere muligheten for å bruke to randomiserte grupper, en som eksperiment- og en som kontrollgruppe, og i tillegg vurdere mulighetene for å gjennomføre både en før- og etter-måling, og på den måte gjøre om designet til et ekte eksperiment.

Andre ting som kan trekkes fram er f.eks.:

- Reliabilitet: det er lite informasjon om hvordan de to gruppene skal følges opp underveis.

- Innholdsvaliditet: selvrapportert, opplevd læring som avhengig variabel har lav innholdsvaliditet. En eller flere prøver ville kunne gitt bedre informasjon om hvor mye studentene har lært.
- c) Figur 1 viser en kausalmodell med en spuriøs sammenheng. Forklar hva som menes med spuriøs sammenheng, og gi noen eksempler på når slike sammenhenger kan forekomme.



Figur 1. Spuriøs modell der Y er avhengig variabel, X er uavhengig variabel og Z er bakenforliggende variabel

Figuren er hentet fra Figur 16.1 på side 341 i Ringdal (2018), og viser en spuriøs årsaksmodell der en bakenforliggende variabel (Z) påvirker både den uavhengige variabelen (X) og den avhengige variabelen (Y). Problemet med slike modeller er at de viser at den uavhengige variabelen (X) har effekt på den avhengige variabelen (Y) hvis vi ikke kontrollerer for den bakenforliggende variabelen (Z). Dette viser et meget sentralt problem når vi skal måle kausalitet, og er en av viktigste grunnene til bak påstanden om at korrelasjon ikke er det samme som kausalitet. Et klassisk eksempel på en slik spuriøs sammenheng er den tidsbestemte sammenhengen mellom nedgangen i antallet storkereir og nedgangen i fødselsraten i Danmark på 1960- og 70-tallet. Denne sammenhengen mellom antall storkereir og antall fødsler kan i stor grad forklares med økende urbanisering, og det faktum at familier på landsbygda fikk flere barn enn familier i byene. I tillegg har nedgangen i storkereir blitt forklart med økt drenering av våtmarker, mens nedgangen i fødselsratene har blitt forklart med introduksjonen av p-piller og økt yrkesaktivitet blant kvinner. I alle disse tilfellene kan vi slå sammen de bakenforliggende forklaringene og benevne den bakenforliggende variabelen (Z) for moderniserings-prosesser. Et eksempel som har vært brukt på forelesning er en mulig sammenheng mellom hårlengde og karakterer på eksamen, der kjønn (Z) kan tenkes å påvirke både hårlengde og karakterer. Et problem som ofte blandes sammen med spuriøsitetetsproblemet er modeller der det er Y som påvirker X (omvendt kausalitet) og modeller der X kan påvirke Y samtidig som Y også påvirker verdien av X. Et eksempel på en modell med omvendt kausalitet er jo flere politifolk det er i en by, jo mer kriminalitet blir det i byen. Et annet eksempel på omvendt kausalitet, som i pensum feilaktig fremstilles som en spuriøs effekt, er påstanden at jo flere brannbiler, jo større brann. I disse tilfellene er det ikke snakk om en bakenforliggende variabel (Z), men at kausaliteten går i motsatt retning. Et klassisk eksempel på vekslende kausalitet, der X påvirker Y samtidig med at Y påvirker X, er forholdet mellom et lands økonomi og grad av demokratisering. Hvis noen av studentene klarer å få fram slike nyansen, så bør de belønnes med beste karakter.

Merk: En del studenter har ikke sett at dette er et kvasi-eksperimentelt design. Dersom de likevel klarer å vise en del relevant pensumforståelse (f.eks. spørreundersøkelser, frafall, randomisering) bør de likevel få uttelling for det.

Oppgave 2

a) Tabell 1 viser et utdrag fra en datamatrix med to variabler, ppltrst sosial tillit (ingen sosial tillit = 0, høy sosial tillit = 10) og utdanning målt i antall år, eduysr. Gjennomsnittsverdiene for de to variablene er 4 og 11.

Bruk tabellen til å regne ut korrelasjonen mellom ppltrst og eduysr og forklar hva koeffisienten viser. Vis hele utregningen, ikke bare endelig svar.

Her blir studentene bedt om å regne ut og tolke Pearsons r , et korrelasjonsmål for kontinuerlige variabler som varierer fra $[-1, 1]$.

På forelesning er den fullstendige utregningen via kovarians og standardavvik gjennomgått, mens studentene har fått en forenklet framgangsmåte på øving der $(n-1)$ i kovariansen kan forkortes mot $(n-1)$ i standardavvikene. I tillegg har det blitt lagt ut en «kokebok» basert på Ringdal (2018, s. 316) på Blackboard, etter at det ble oppdaget en feil i pensum etter at eksamensoppgaven var bestemt. Hvis studenten følger en av disse oppskriftene, som de har tilgjengelig på Blackboard, er selve utregningen enkel. Oppgaven bør derfor ikke tillegges for stor vekt hvis studentene viser ellers liten forståelse for kvantitativ metode. Studenter som begrunner valget av Pearsons r (korrelasjonsmål for kontinuerlig variabler) og skriver mer om korrelasjonsmålet bør belønnes.

Riktig svar, uavhengig av framgangsmåte, skal være Pearsons $r = 0,54$, dvs. en relativt sterk positiv korrelasjon mellom sosial tillit og utdanningslengde.

Gjennomgangen her er basert på trinnvis framgangsmåte fra forelesning:

$$r = \frac{\text{kov}(X,Y)}{s_X s_Y} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{s_X s_Y (n-1)}, \text{ der } s_X = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n-1}} \text{ og } s_Y = \sqrt{\frac{\sum(Y_i - \bar{Y})^2}{n-1}}$$

1. Regner ut kovarians:

a. Regner ut avvik fra gjennomsnitt for begge variabler, beregner kryssproduktet til avvikene og summerer:

Ppltrst	eduysr	poltrst-4	eduysr-11	kryssprodukt
2	12	-2	1	-2
2	8	-2	-3	6
4	11	0	0	0
5	12	1	1	1
7	12	3	1	3
			Sum	8

b. Finner kovariansen ved å dele summen av kryssproduktet på $(n-1)$, dvs. $8/(5-1) = 2$

2. Regner ut standardavvikene til *ppltrst* og *eduysr* ved å bruke avviksskårene i kolonne 3 og 4 over:

a. $s_{ppltrst} = \sqrt{\frac{\sum(-2^2 + -2^2 + 1^2 + 3^2)}{4}} = \sqrt{\frac{18}{4}} = 2,1213$

b. $s_{eduysr} = 1,7320$

3. Setter inn i formel:

$$r = \frac{2}{2,1213 * 1,7320} = 0,544$$

- b) Vedlegg 1 gjengir enn loggfil fra statistikkprogrammet Stata, og viser en analyse med data fra European Social Survey fra 2018. Beskriv oppbyggingen av modell 1-3 i vedlegg 1, og forklar hva modellene viser. Bruk modell 3 til å regne ut predikert tillit for 50 år gamle kvinner med 13 års utdanning i hvert av de tre landene.

I beskrivelsen av modellen bør studentene kommentere at avhengig variabel (AV) er en refleksiv indeks/skala, der reliabilitetsanalysen viser god reliabilitet (Cronbachs alfa > 0,70). Siden AV er kontinuerlig er det brukt lineær regresjon. Videre bør målenivå og behandling av de uavhengige variablene forklares, der det først legges til et dummysett for å sammenligne sosial tillit mellom tre land, og deretter et sett med forklaringsvariabler på individnivå. Spesielt flinke studenter vil kanskje se at de tre landene som er valgt er et vesteuropeisk demokrati (Norge), et tidligere østblokk-land (Estland) og et land fra det tidligere Jugoslavia (Serbia), og at ulik politisk og sosial historie, inkludert konflikt, kan forklare forskjellene mellom landene. Estland er valgt som referansekategori fordi dette er den største kategorien.

Flinke studenter vil kunne se at nesten all variansen her er mellom land, så selv om alle variablene unntatt kjønn er statistisk signifikante, går R^2 bare litt opp fra modell 1 til modell 2.

I modell 3 testes et samspill mellom land og utdanning og sosial tillit. Modellen viser at sammenhengen mellom sosial tillit og utdanning ikke er signifikant forskjellig i de tre landene, selv om forskjellen mellom Norge og Estland er svært nær statistisk signifikans.

For å predikere verdier bør studentene først sette opp den fulle regresjonsligningen, og deretter sette inn verdier for alder, kjønn og utdanning, og de ulike verdiene for land. Også ikke-signifikante variabler skal inkluderes i utregningen:

$$\hat{Y} = b_0 + b_1Norge + b_2Serbia + b_3Age + b_4Male + b_5Eduyrs + b_6Norge * Eduyrs + b_7Serbia * Eduyrs$$

$$\hat{Y} = 4,912 + 0,438*Norge - 2,12*Serbia + 0,003*agea - 0,069*male + 0,055*eduyrs - 0,021*Norge*Utdanning + 0,012*Serbia*eduyrs$$

$$\hat{Y} = 4,912 + 0,438*Norge - 2,12*Serbia + 0,003*50 - 0,069*0 + 0,055*13 - 0,021*Norge*13 + 0,012*Serbia*13$$

Riktig svar skal være:

$$\hat{Y}_{Estland} = 5,75$$

$$\hat{Y}_{Norge} = 5,91$$

$$\hat{Y}_{Serbia} = 3,79$$

- a) Se modell 4 i vedlegg 1. Bør politisk tillit være med i modellen? Begrunn svaret.

Denne går an å besvare på flere måter. Det er en tydelig sterk sammenheng mellom politisk og sosial tillit, og forklart varians gjør et hopp. Omvendt kausalitet kan være et stort problem, og de to variablene måler potensielt (deler av) samme fenomen. De som ser det, og argumenterer for at politisk tillit derfor ikke bør være med i modellen bør belønnes. Studenter som argumenterer med at variabelen er klart signifikant og bidrar til høy R^2 , og derfor bør være med, bør ikke trekkes for mye

for det. Politisk tillit bidrar til å forklare noe av forskjellen mellom landene (forskjellen mellom Serbia og Estland er klart mindre i modell 4 enn i modell 3). Studenter som ser dette, bør også belønnes.

Oppgave 3

Norge kjennetegnes av at folk har høy tillit til det offentlige. Noen grupper skiller seg ut med lavere tillit til at politikere arbeider for innbyggernes beste. Eksempel på lavere tillit finnes i grupper uten jobb eller som mottar trygd, med lav inntekt, har nedsatt funksjonsevne fysisk eller psykisk, eller er selv, eller barn av foreldre som er, født i et annet land enn Norge.

Din forskningsgruppe har fått i oppdrag å utvikle og gjennomføre en undersøkelse for å bli bedre kjent med grupper med lav tillit til det offentlige:

- a) Metoden for undersøkelsen er deltagende observasjon. Skisser et opplegg for å gjennomføre undersøkelsen.

Oppgaven **avgrens**er metoden til **deltagende observasjon**. På pensum er det en egen bok dedikert til denne metoden (Fangen, 2010 (2. utgave (1. utgave 2004)/ trykt opp 3. opplag 2017)). Tjora (2017) har også eget kapittel dedikert til observasjon som metode.

Gode oppgaver holder seg innenfor metodens avgrensing (deltagende observasjon). Det skal ikke trekkes for å vise kunnskap om andre metoder som kan bidra til å gjennomføre undersøkelsen, men slike forslag må ikke gå på bekostning av å svare på det oppgaven spør om. Gode oppgaver henviser også til pensum. Oppgaven ber detaljer innenfor de følgende tre kulepunkt. Det er stor frihet til hvordan selve undersøkelsen kan gjennomføres:

- Beskriv utvalg og avgrensing av undersøkelsen.

Oppgavetekstens innledning viser til ulike grupper som har lav tillit. Det vil være fornuftig å avgrense undersøkelsen til en eller få av disse gruppene heller enn å søke representasjon fra alle. Gode besvarelser gir en god og konkret beskrivelse av hvilken/e gruppe som skal studeres (utvalg) og omfang i tid (når/varighet) og rom (hvor/hvordan).

- Gjør rede for forskerens deltagerrolle i undersøkelsen, hvordan data samles inn, registreres og planer for klargjøring av data for analyse.

Både Tjora og Fangen diskuterer ulike deltagerroller. Fanger i mer detalj. Beskrivelse av deltagerrollen er grad av deltagelse (passiv/aktiv), grad av åpenhet om deltagelsen og evt. veksling mellom ulike roller. Ekstra gode besvarelser problematiserer deltagerrollen/e og dens utfordringer/begrensninger.

Dette spørsmålet ber også om redegjørelse for innsamlingsmetode og registrering. Stikkord er Skriftlig: observasjons/feltnotat, visuelt opptak (lyd&bilde) osv, evt at observasjon er gjort på nett (f.eks diskusjonsfora).

Planer for klargjøring for analyse handler om i hvilken form data skal analyseres. Her har alle i semesteroppgave gjort om debattprogram på tv til tekst for analyse. Det kan være en måte å besvare denne delen på.

Det forventes at studenten reflekterer over sine valg.

- Hvilke etiske utfordringer står dere over for, og hvordan sikrer dere at etiske retningslinjer følges?

Her er det en rekke etiske utfordringer. Personvern er åpenbart (informert samtykke, trygg håndtering av data, avklaring om anonymisering (obs unntak kan argumenteres for hvis informant har offentlige rolle / er kjent med / samtykker til offentliggjøring). Dette er nøye gjennomgått på

forelesning. For noen av gruppene som studeres kan det være ekstra etiske utfordringer (psykisk syke mm). Observante studenter har fått med seg at innsamling av personsensitiv informasjon skal godkjennes av NSD mm.

- b) Hvilke kvalitetskriterier tar dere i bruk for å sikre at undersøkelsen blir best mulig? Gi eksempler på hvordan dere kan styrke kvalitet i undersøkelsen.

Her viser pensum (og forelesninger) til kriterier som validitet og reliabilitet, transparens, refleksjon og refleksivitet, former for generalisering (konseptuell, ikke statistisk). Refleksjon premieres fremfor evnen til å gjengi definisjoner fra pensumlitteraturen.

- c) Diskuter hvordan forskernes for- (Tjora) eller før- (Fangen) forståelse håndteres i undersøkelsen. Hva skjer i veksling mellom forståelse og for(før)forståelse analyse av data?

Forskerens for-/førforståelse kan påvirke alle deler av en undersøkelse. Redegjørelse for egen forforståelse kan være nødvendig i et prosjekt, men er også nyttig for å styrke kvaliteten i eget arbeid (transparens/refleksivitet).

I analysesituasjonen foregår en veksling mellom å bygge forståelse og konferere med forforståelse. Heidegger (i Fangen s.47) kaller denne vekslingen den hermeneutiske sirkel (Det forventes ikke at studentene gjør akkurat denne koblingen). I analysesituasjonen vil førforståelsen gradvis modifiseres i møte med ny kunnskap. Bevissthet om egne forestillinger, antagelser og kunnskap, muliggjør aktiv bruk og kontrastering mot funn i forskningen.