Høgskolen i Østfold

SENSORVEILEDNING

Emnekode:	SFB12016
Emnenavn:	Metodekurs II: Samfunnsvitenskapelig metode og anvendt statistikk
Eksamensform:	Skriftlig 4 timers eksamen. Hjelpemilder: Godkjent kalkulator
Dato:	16.12.2019
Faglærer(e):	Bjørnar Karlsen Kivedal
Eventuelt:	

Oppgave 1 (10%)

a) Positivistisk/tilnærming i forskning
Virkeligheten: stabil og objekt virkeligheten. Lovmessigheter.
-Kunnskap: det generelle. Objektiv virkelighet som kan studeres gjennom objektive metoder og mål. Nøytrale forskere som beskriver virkeligheten. Kunnskap til kumulativ
-Metode: deduktive metoder. Individualistisk. Avstand. Tall (kvantitative metoder)

b) Fenomenologisk/fortolkningsbasert tilnærming i forskning
-Virkeligheten: dynamisk og menneskeskapt virkelighet
-Kunnskap: det unike og særegne (fenomena)
-Metode: Induktive metoder. Holistisk. Nærhet. Ord (kvalitative metoder)

c) **Ontologi**

Ontologi: læren om hvordan virkeligheten faktisk ser ut (Jacobsen, 2018: p. 22) eller læren om det å være, eksistens og virkelighet. Eksempler på sentrale ontologiske spørsmål basert på pensum-boken:

- «Er det handlende og kreative individer som driver økonomisk utvikling eller er det økonomiske utviklingen en prosess som i liten grad kan påvirkes» (Ibid)
- «Er mennesket rasjonelt, dvs. at det velger sine handlinger ut fra en vurdering av nytte og kostnad, eller er handlingene mer styrt av normer, følelser eller omgivelser?» (Ibid)

d) **Epistemologi** betyr «læren om kunnskap», omhandler i hvilken grad det er mulig å få kunnskap om denne verden (Ibid). Det er viktig at studentene forstår to forskjellige epistemologiske erkjennelser – skillet mellom virkeligheten som den er, og virkeligheten som oppfattes av forskeren. Eksempler på epistemologiske debatter som er beskrevet i pensum-boken:

- «I hvor stor grad styrer våre «før-dommer» vår oppfatning av virkeligheten, f.eks. hvordan vår oppvekst eller utdanning er med på å forme hvordan vi ser virkeligheten?» (Ibid)

Oppgave 2 (40%)

Studenten bør ha kjennskap til forskjellige FASER I UNDERSØKELSESPROSESSEN, og presentere de viktigste fasene som de skal bruke i undersøkelsen. Det er viktig at studentene gir gode begrunnelser av sine valg (det gjelder design, metoder, datainnsamling, analyse, etc.) og har en strukturert og logisk oppgave.

Fase 1: Problemstilling og utvalg av situasjoner til observasjon.

Fase2: Undersøkelsesdesign

Fase 3: Valg av informasjon

Fase 4: Hvordan skal informasjon samles inn

Fase 5: Utvalg av informanter og respondenter

Fase 6: Hvordan skal informasjon analyseres

Fase 7: Kritisk blikk over hvor gode funnene er og konklusjoner

Fase 8: Troverdighet av undersøkelse. Etiske dilemmaer

Oppgave 3 (50%)

a)

- $b_2 = 3{,}867$: Dersom lønnskostnader for helgen øker med \$1000, så øker omsetningen med \$3867, alt annet likt.
- ~~$b_3 = 0{,}408$: Dersom omsetningen samme uke i fjor økte med \$1000, så øker omsetningen i år med \$408, alt annet likt.~~
- $b_6 = 12{,}017$: Dersom det er hesteveddeløp, så er omsetningen i gjennomsnitt \$12017 høyere, alt annet likt.
- ~~$b_{10} = 37{,}07$: Dersom det er byggeaktivitet eller stengte gater i nærheten, så er omsetningen i gjennomsnitt \$37070 høyere, alt annet likt.~~
- $b_{14} = -1{,}557$: Dersom temperaturen øker med en grad Farenheit, så synker omsetningen med \$1557, alt annet likt.

b) $H_0: B_4 = B_5 = B_7 = B_8 = B_9 = B_{11} = B_{12} = B_{13} = B_{15} = B_{16} = 0$ mot H_A: én eller flere av påstandene i H_0 er gale. Her har vi en F-test. Testverdi der Modell 1 er modellen uten restriksjoner og Modell 2 er modellen med restriksjoner gir: $F = \frac{(R_{ur}^2 - R_r^2)/m}{(1-R_{ur}^2)/(n-k)} = \frac{(0{,}7766-0{,}7439)/10}{(1-0{,}7766)/(69-16)} = \frac{0{,}0327/10}{0{,}2234/53} = \frac{0{,}00327}{0{,}004215} = 0{,}7758$ Kritisk verdi blir $F_\alpha(Df_1, Df_2) = F_{0{,}10}(10,53) \approx F_{0{,}10}(10,50) = 1{,}729$. Siden testverdi er mindre enn kritisk verdi så beholder vi nullhypotesen. Vi har ikke funnet støtte for at noen av variablene som ikke er med i Modell 2 men som var med i Modell 1 har en signifikant effekt på omsetningen.

c) $H_0: B_6 = 0$ mot $H_A: B_6 \neq 0$. Testverdi blir da $\frac{12{,}0166}{5{,}78583} = 2{,}077$. Kritiske verdier blir $\pm t_{\alpha/2}(n-k) = \pm t_{0{,}05}(69-6) = \pm t_{0{,}05}(63) \approx \pm t_{0{,}05}(60) = \pm 1{,}671$. Testverdien er utenfor forkastningsområdet og vi forkaster nullhypotesen. Vi har funnet støtte for påstanden om at hesteveddeløp har en effekt på ukentlig omsetning.

d) Vi kan belyse dette ved å beregne to varianter av Modell 3: En der $X6_i = 0$ og en der $X6_i = 1$, hhv. ikke hesteveddeløp og hesteveddeløp i uke i. I uker der det ikke er hesteveddeløp blir den beregnede modellen da $\hat{Y}_i = 136{,}5 + 5{,}9X2_i$ og i uker der det er hesteveddeløp blir den beregnede modellen $\hat{Y}_i = 136{,}5 + 6{,}5X2_i$. Altså er stigningstallet høyere hvis det er hesteveddeløp. Det vil si at lønnskostnadene i helgen har større effekt på omsetningen hvis det er hesteveddeløp. En økning i lønnskostnadene i helgen (som f.eks. følge av at man øker bemanningen) vil øke omsetningen mer dersom det er hesteveddeløp. Dette kan f.eks. være fordi det blir flere kunder på kjøpesenteret som da kan betjenes lettere når det er flere ansatte på jobb og at omsetningen dermed øker.

e) Her kan vi beregne forskjellen mellom $\hat{Y}$ for Modell 4 dersom

- $X2 = 22$ og $X2 = 23$ og alle andre variabler er like. Vi får da $\hat{Y}_{|X2=23} - \hat{Y}_{|X2=22} = (b_1 + b_2 \cdot 23 + b_3 \cdot 23^2) - (b_1 + b_2 \cdot 22 + b_3 \cdot 22^2) = b_2(23-22) + b_3(23^2 - 22^2) = b_2 + b_3 \cdot (529 - 484) = 60{,}09 - 0{,}99 \cdot 45 = 60{,}09 - 44{,}55 = 15{,}54$. Ukesomsetningen øker med \$15 540.
- $X2 = 29$ og $X2 = 30$ og alle andre variabler er like. Vi får da da $\hat{Y}_{|X2=23} - \hat{Y}_{|X2=22} = (b_1 + b_2 \cdot 30 + b_3 \cdot 30^2) - (b_1 + b_2 \cdot 29 + b_3 \cdot 29^2) = b_2(30-29) + b_3(30^2 - 29^2) = b_2 + b_3 \cdot (900 - 841) = 60{,}09 - 0{,}99 \cdot 59 = 60{,}09 - 58{,}41 = 1{,}68$. Ukesomsetningen øker med \$1 680

Det er også mulig å bruke den deriverte her for å beregne forskjellen: $\frac{\partial \hat{Y}}{\partial X2} = b_2 + 2 \cdot b_3 \cdot X2^2$

f) I Modell2 hadde vi en økning i ukentlig omsetning med \$3 867 dersom lønnskostnaden økte med \$1000. Dette vil da i Modell 2 gjelde både ved en økning fra 22 til 23 tusen dollar og fra 29 til 30 tusen dollar. I Modell 4 er effekten av økte lønnskostnader på omsetningen høyere dersom lønnsutgiftene allerede er lave (minimumsverdien på lønnskostnadene er \$22064 og maksimum er \$32 499 – se deskriptiv statistikk). Ved å ha med kvadrert forklaringsvariabel muliggjør vi at forklaringsvariabelen har en ikke-lineær effekt på den avhengige variabelen slik som her. Effekten av lønnskostnader i helgen er mindre jo høyere lønnskostnadene i utgangspunktet er.

g) For å sammenligne modeller som har samme avhengige variabel (Y) men ulikt antall forklaringsvariabler må vi sammenligne justert R-kvadrat som sier noe om modellens forklaringskraft. Denne er ikke oppgitt for Modell 1 og må derfor beregnes. Først beregner vi $R^2 = \frac{ESS}{TSS} = \frac{23579{,}6}{30362{,}1} = 0{,}7766$ og så kan vi beregne justert R-kvadrat:

$$\bar{R}^2 = 1 - \left[(1 - R^2) \cdot \left(\frac{n-1}{n-k}\right)\right] = 1 - \left[(1 - 0{,}7766) \cdot \frac{69-1}{62-16}\right]$$
$$= 1 - \left[(0{,}2234) \cdot \frac{68}{53}\right] = 1 - 0{,}2866 = 0{,}7134$$

Vi ser dermed at justert R-kvadrat er høyest for Modell 2 (0,7235) som da har høyest forklaringskraft og bør brukes for å anslå ukesomsetning.

Siden vi fant en signifikant samspillseffekt for X6 på X2 i oppgave d) og en signifikant ikke-lineær effekt av X2 i oppgave e) og f), så kan en utvidelse av modellen til å også inkludere leddene X6X2 og X2sq.