

KANDIDAT

**2566**

PRØVE

**ME-100 1 Samfunnsvitenskapelig metode**

Emnekode	ME-100
Vurderingsform	Skriftlig eksamen
Starttid	24.11.2016 09:00
Sluttid	24.11.2016 14:00
Sensurfrist	15.12.2016 01:00
PDF opprettet	19.12.2017 14:12
Opprettet av	Digital Eksamen

## 1 ME-100, forside

- Ja
- Nei



Riktig. 0 av 0 poeng.

## 2 ME-100, oppgave 1

I denne oppgaven vil jeg først forklare hovedforskjellene mellom kvalitativ og kvantitativ metode for så å drøfte når et kvalitativt intervju er å foretrekke fremfor en kvantitativ spørreundersøkelse. Til slutt vil jeg i oppgave b) gjøre rede for bruk av kvalitativt intervju som metode.

### a)

I en kvalitativ metode har vi betydelig færre enheter enn i en kvantitativ, men vi har betydelig flere variabler. I kvalitativ metode opereres det gjerne med en maksimumsgrense på 20 enheter, mens kvantitativ kan ha flere tusen enheter.

En av hovedforskjellene mellom kvalitativ og kvantitativ metode er at vi som regel ønsker å finne sammenhenger i et utvalg som kan generaliseres til populasjonen i kvantitativ metode, mens vi i kvalitativ metode ofte har som formål å gå dypere inn i enkelte tema/case og ut i fra det vi finner der utvikle en teori. Kvalitativ metode har ofte en induktiv metode der vi går fra empiri til teori. Det vil si at vi sjekker først hvordan situasjonen er i virkeligheten f. eks. gjennom intervju med noen få, før vi ut i fra det utvikler en teori. I kvalitativ metode har vi ofte en teori på forhånd og så sjekker vi et stort utvalg av populasjonen for å sjekke om teorien stemmer med virkeligheten. Vi sier at vi har en deduktiv metode der vi går fra teori til empiri.

En annen viktig forskjell er hvordan vi velger ut enheter fra populasjonen vi er interessert i. I kvalitativ metode vil vi ofte velge personer som har opplevd det vi er interessert i eller har en spesiell kunnskap om temaet/casen vi vil undersøke, vi benytter oss av ikke-sannsynlighetsutvalg der vi kan velge personer vi tror er representative for det vi vil undersøke, eller bekvemmelighetsutvalg som består i å velge de som er lettest å få tak i eller gjøre det enklest mulig for oss selv og la utvalget bli bestemt ved at enkeltpersoner kan melde seg på frivillig. I kvantitativ metode vil vi som regel forsøke å trekke et så representativt utvalg fra populasjonen som mulig for å ha høyere sannsynlighet for at resultatet vi finner er generaliserbart for hele populasjonen, vi velger sannsynlighetsutvalg der vi kan trekke utvalget ved rent tilfeldig utvalg (trekker tilfeldig), strukturert tilfeldig utvalg (strukturere populasjonen og trekker enheter etter intervall f. eks. hver tiende enhet), klyngeutvalg (fordeler populasjonen i klynge, eks: kommune) eller stratifisert utvalg (fordeler populasjonen etter kategorier og trekker like mange prosent enheter fra hver kategori som det er i populasjonen).

Nå skal vi se på forskjellige fordeler med kvalitativt intervju og når kvalitativt intervju er å foretrekke fremfor en kvantitativ spørreundersøkelse.

Når vi ønsker å gå i dybden på et tema/case er et kvalitativt intervju å foretrekke fremfor en kvantitativ spørreundersøkelse. Dette er fordi i en kvantitativ spørreundersøkelse vil vi ha veldig mange enheter som svarer på forhåndsstrukturerte spørsmål med bestemte svaralternativer. I et kvalitativt intervju vil respondenten ha mulighet til å gå i dybden på enkelte temaer og utdype, mens i en kvantitativ spørreundersøkelse er svarene ofte bestemte på forhånd slik at man krysser av for det svaret som passer best. Dette er en god metode når vi vil vite hva mange personer mener om en sak, men ikke når vi ønsker å gå i dybden på dette temaet og forstå hvorfor personen mener det han/hun mener.

Når vi skal undersøke et tema/case som vi vet lite om fra før er det ofte lurt å velge et kvalitativt intervju fremfor en kvantitativ spørreundersøkelse. Grunnen til dette er at dersom vi ikke har noen teori eller lite kunnskap og gå ut i fra er det vanskelig å stille spørsmål som kan kontrollere om teorien stemmer siden vi ikke har noen teori. Vi kan da gjennomføre et kvalitativt intervju der vi kan lære mye av respondentene og få en utdyping og gå i dybden på temaet og på bakgrunn av intervjuene danne en teori. Vi benytter en induktiv metode der vi går fra teori til empiri.

Når vi ikke har oversikt over alle mulige svar og det er vanskelig å velge noen enkle svaralternativer uten at vi utelukker noen eller glemmer viktige svaralternativer kan det være hensiktsmessig å velge et kvalitativt intervju. Dette henger sammen med hvor mye kunnskap/teori vi har om temaet fra før.

Ved å gjennomføre et kvalitativt intervju kan vi be om utdyping av enkelte temaer eller be respondenten forklare noe på en litt annen måte hvis det var noe vi ikke forstod helt. Vi kan helt enkelt fjerne usikkerhet ved svarene og sikre oss mer kunnskap om de temaene vi er interessert i.

Dersom vi har komplekse spørsmål kan det være aktuelt å velge kvalitativt intervju for å få muligheten til å forklare og presisere hva vi mener med spørsmålene, men dette kan også gjennomføres på kvantitative spørreundersøkelser ved å ta de over telefon der en person leser opp spørsmålene og han/hun kan da også presisere hva som menes dersom respondenten ikke skulle forstå spørsmålet.

Kvalitativt intervju gir respondenten mulighet til å ta opp egne temaer eller gå litt utenfor de temaene vi hadde planlagt. Hvorvidt dette skal godtas er opp til intervjueren. Kvalitativt intervju kan gjennomføres på en skala fra helt lukket der alle temaene er bestemt på forhånd eller helt åpen der det er åpning for å snakke om det meste. Det beste er antagelig å ha litt struktur, men la respondenten få muligheten til å ta opp egne temaer og gå litt utenfor temaet vi egentlig skulle snakke om. Dette kan hjelpe oss å forstå temaet/casen bedre og sette oss inn på ting vi ikke hadde sett som aktuelle.

### b)

Når vi skal gjennomføre et kvalitativt intervju må vi først få oversikt over det totale antallet vi ville intervjuet hvis vi hadde hatt uendelig med ressurser og tid, dette er populasjonen. Deretter må vi trekke et utvalg fra denne populasjonen. I kvalitativ metode regner vi ofte 20 enheter som en maksimumsgrense. Når vi skal trekke et utvalg til et kvalitativt intervju er det vanlig å trekke et ikke-sannsynlighetsutvalg. Dette er et utvalg som ikke er opplatt av at utvalget skal representere populasjonen. De vanlige måtene å trekke et utvalg på i ikke-sannsynlighetsutvalg er:

- Trekke personer som vi tror er representative for populasjonen - vi trekker enheter som vi tror er representative.
- Trekke personer som vi tror har mye kunnskap om temaet - vi trekker personer som vi tror har mye kunnskap om temaet vi skal undersøke og dermed kan tilføre mye under intervjuet, kan også være personer som har opplevd en ting vi skal undersøke.
- Bekvemmelighetsutvalg - vi trekker enheter som er lette å få tak i.
- Frivillig - vi lar utvelgelsesprosessen gå av seg selv ved at vi lar folk melde seg på selv frivillig.
- Trekker enheter etterhvert som vi får kunnskap - noen ganger kan vi starte med å intervju en person og benytte kunnskapen vi fikk av det først intervjuet til å velge neste respondent, slik kan vi fortsette.

Når vi har valgt ut hvem vi vil intervju må vi bestemme hvordan intervjuet skal gjennomføres. Det beste for å få god flyt i samtalen/intervjuet er ofte ansikt-til-ansikt intervju. Dette er imidlertid veldig kostnadskravende fordi det krever at respondent og intervjuer er på samme sted til samme tid. Vi kan da risikere å ikke få intervjuet enkelte enheter vi ønsket å intervju fordi det ikke er praktisk mulig. Da er alternativer som telefon,

chat eller e-post alternative muligheter. Ved telefon vil vi fortsatt beholde mange av fordelene med ansikt-til-ansikt intervju samtidig som vi kan være på forskjellige steder og dermed gjør det lettere å få gjennomført intervjuet. Chat og e-post vil gjøre at svarene på spørsmålene ikke kommer med en gang, respondenten får bedre tid til å tenke seg om, dette kan både være positivt og negativt.

Når det gjelder selve gjennomføringen av intervjuet kan vi velge hvor strukturert vi ønsker at intervjuet skal være. Vi kan ha et helt åpent intervju der det ikke er noen begrensning på hva vi kan snakke om eller vi kan ha et helt lukket intervju der intervjuer har bestemt alle spørsmålene på forhånd og det ikke er åpning for å gå utenfor dette. Det beste er nok en mellomting der vi har en viss struktur slik at vi ikke beveger oss langt unna det intervjuet var ment og handle om, men der vi samtidig lar respondenten gå litt utenfor spørsmålet og også kan ta opp egne temaer litt på siden av hovedtemaet.

De fleste spørsmålene er skrevet ned på forhånd, men under selve intervjuet kan det være lurt å be om utdypning på temaer du synes virker interessant for problemstillingen og be om forklaring på ting du ikke forstår helt. Her har det kvalitative intervju en fordel som ikke er mulig i kvantitativ spørreundersøkelse. Det er også mulig å stille spørsmål dersom du kommer på noe underveis som kan være aktuelt. Det kan være sannsynlig at man kommer på flere aktuelle spørsmål om temaer man ikke har så mye kunnskap om etterhvert som du lærer mer om temaet.

Når man gjennomfører intervjuer med flere personer kan det være at samme ting blir sagt flere ganger og vi når et punkt der det ikke er mer informasjon å hente om temaet vi undersøker. Vi sier da at vi har nådd et metningspunkt. Når vi mener at vi har nådd et metningspunkt kan det være hensiktsmessig å ikke gjennomføre flere intervjuer med personer med samme bakgrunn og opplevelse som de vi allerede har snakket med, men heller fokusere på de områdene der vi ikke har nådd metningspunktet enda.

Under gjennomføringen av intervjuet er det beste om vi kan ta opp intervjuet, men noen kvier seg for dette og vi skal derfor alltid spørre om tillatelse først. Selv om vi tar opp alt som blir sagt bør vi ta notater underveis for å gjøre jobben lettere for oss senere når vi skal finne fram i opptaket, i tillegg gjør det at vi blir mer fokusert på det som blir sagt. Rett etter intervjuet bør man skrive et lite sammendrag av det man husker, dette er antagelig det viktigste som ble sagt. Hele intervjuet bør transkriberes slik at man har alt som ble sagt skrevet ned, dersom man gjennomfører chat eller e-post intervju vil dette allerede være gjort.

Når dataene er samlet inn gjennom intervju, transkribert og skal analyseres er ikke dette like enkelt som en kvantitativ analyse der vi kan la et dataprogram gjøre jobben for oss. Vi må selv forsøke å få oversikt over den informasjonen vi har samlet inn. Dette kan vi gjøre på forskjellige måter:

- Meningsstetting - vi forkorter den informasjonen vi har ned til det viktigste som ble sagt.
- Meningskategorisering - vi kategoriserer meningene inn under flere underkategorier. Da kan vi lettere få oversikt over hva som er det viktigste og vi kan om flere av respondentene snakket om samme temaer.
- Meningsstrukturering med narrativer - vi setter handlingen inn i en sosialistisk og tidsrekkefølge. Fokuset er på historier.
- Vi fokuserer på historier som er fortalt under intervjuet. Dersom det ikke ble fortalt noen historier forsøker vi å fortelle handlingen i intervjuet som en historie.
- Ad-hoc metode for meningsstrukturering

Meningen med kvalitativt intervju er at vi skal kunne utarbeide ny eller dypere kunnskap om et tema/case. Etter at vi har analysert resultatene kan vi forsøke å utarbeide en teori basert på det vi har funnet ut.

Når vi skal presentere det vi har funnet må vi være obs på at dersom vi har mye personlige detaljer med kan det være mulig å finne ut hvem vi har pratet med dermed vil det være lurt å kun ta med det nødvendige for å beskytte enkeltpersonene du har intervjuet og holde deg innenfor på de etiske punktene. Vi må også huske at alle som er med på intervjuet må delta etter egen vilje. De skal vite hva undersøkelsen skal brukes og være beviste på hva de deltar på.

Besvart

### 3 ME-100, oppgave 2

a)

I tabellen ser vi at vi har tre variabler. Den avhengige variabelen er "Presidentkandidat", mens de to uavhengige variablene er "Utdanning" og "Alder". Enhet er de vi er interessert i å undersøke, i dette tilfellet er det befolkningen i delstaten Florida. Det er sannsynlig å anta at det kun er befolkningen med stemmerett i Florida som har blitt spurt, men dette sier hverken krysstabellen eller teksten noen ting om. I denne analysen er vi altså ute etter å se hvorvidt utdanning og alder har innvirkning på valget av presidentkandidat på befolkningen i delstaten Florida.

Først skal jeg se på hele krysstabellanalysen for å se om det finnes noen sammenhenger og ulikheter. Så vil jeg sjekke effekten av kun en uavhengig variabel av gangen kontrollert for den andre uavhengige variablene for å se om begge påvirker den avhengige variabelen og eventuelt hvor mye de påvirker. I tolkingen av tallene vil jeg kun nevne prosenttallene for Trump, siden tallene er prosentuert horisontalt vil resultatet bli det samme om vi ser på tallene til Clinton eller Trump, effekten er den samme bare med motsatt retning.

Når vi ser på hele modellen ser vi at Trump har flertall på begge aldersgruppene på lav utdanning, mens Clinton har flertall på begge aldersgruppene på høy utdanning. Vi ser at lav utdanning og under 45 år er der Trump har størst oppslutning og dermed Clinton lavest oppslutning. Clinton har høyest oppslutning på gruppen med høy utdanning og under 45 år, her har da Trump lavest oppslutning.

For å sjekke hvilken av de uavhengige variablene som påvirker valget av presidentkandidat og se hvor sterk påvirkningen er skal jeg nå sjekke en uavhengig variabel av gangen kontrollert for den andre variabelen.

Først tar vi bort effekten av alder for å kunne se effekten av utdanning alene. Dette gjør vi ved å sammenligne lav utdanning med alder under 45 år med høy utdanning med alder under 45 år, deretter sammenligner vi lav utdanning med alder 45 år eller eldre med høy utdanning med alder 45 år eller eldre. Da ser vi at de med lav utdanning og alder under 45 år sa 58% at de skulle stemme på Trump. På høy utdanning med alder under 45 år sa 44% at de skulle stemme på Trump. Forskjellen mellom lav og høy utdanning med alder under 45 år på hvor mange prosent som sa de skulle stemme på Trump er altså  $58\% - 44\% = 14\%$

Nå skal vi sammenligne lav utdanning 45 år eller eldre med høy utdanning 45 år eller eldre for så å ta gjennomsnittseffekten av utdanning på presidentvalget. På lav utdanning med 45 år eller eldre ser vi at 53% sa de skulle stemme på Trump. På høy utdanning med 45 år eller eldre sa 48% at de skulle stemme på Trump. Forskjellen mellom lav og høy utdanning med alder 45 år eller eldre på hvor mange prosent som sa de skulle stemme på Trump er  $53\% - 48\% = 5\%$

Gjennomsnittseffekten av utdanning på hvor mange prosent som sa de skulle stemme på Trump i Florida når vi har kontrollert for alder er altså  $(14\% + 5\%) / 2 = 19\% / 2 = 9,5\%$ . Det vi kan se ut i fra dette er at befolkningen i Florida med lav utdanning vil 9,5% flere stemme på Trump enn befolkningen i Florida med høy utdanning. Effekten er lik men motsatt for Clinton.

Nå skal vi se på effekten av alder kontrollert for utdanning, dvs. at vi tar vekk effekten av utdanning og kun ser på effekten av alder når det gjelder hvem man sier man vil stemme på av Clinton og Trump i Florida.

Vi tar først å sammenligner under 45 år med 45 år eller eldre på lav utdanning. Der ser vi at de under 45 år sa 58% at de ville stemme på Trump, mens av de som var 45 år eller eldre sa 53% at de ville stemme på Trump. Forskjellen mellom under 45 år og 45 år eller eldre blant befolkningen i Florida med lav utdanning er altså  $58\% - 53\% = 5\%$ .

Så skal vi sammenligne de med høy utdanning under 45 år med de med høy utdanning som er 45 år eller eldre. Vi ser at de med høy utdanning under 45 år sier 44% at de vil stemme på Trump, mens de med høy utdanning som er 45 år eller eldre sier 48% at de vil stemme på Trump. Forskjellen mellom de med høy utdanning under 45 år og de med høy utdanning som er 45 år eller eldre er  $44\% - 48\% = -4\%$ . Det vil si at av de med høy utdanning vil 4% flere av de som er 45 år eller eldre stemme på Trump enn de som er under 45 år, retningen er altså motsatt av de med lav utdanning der det var flest av de under 45 år som ville stemme på Trump.

Gjennomsnittseffekten av alder blir da

$$(5\% + (-4\%)) / 2 =$$

$$(5\% - 4\%) / 2 =$$

$$1\% / 2 =$$

$$0,5\%$$

Vi ser at gjennomsnittseffekten av alder på befolkningen i Florida når det gjelder om de skal stemme på Trump eller Clinton er 0,5%. Det vil si at det er 0,5% flere av de som er under 45 år som vil stemme på Trump enn de som er 45 år eller eldre. Denne forskjellen er så liten vi kan se bort i fra denne effekten.

Ut i fra denne analysen ser vi at det kun er utdanningsnivået som har betydning for om man vil stemme på Trump eller Clinton i Florida. Den gjennomsnittlige effekten av utdanningsnivå var 9,5%. For Trump betyr dette at 9,5% flere på lav utdanning som ville stemme på han enn det var på høy utdanning i Florida. For Clinton er det da motsatt 9,5% flere på høy utdanning vil stemme på hun enn det var på lav utdanning i Florida. Gjennomsnittseffekten av alder så vi var på kun 0,5% og derfor velger vi å se bort i fra denne siden det er en såpass svak sammenheng mellom alder og hvem befolkningen i Florida ville stemme på.

**b)**

Når vi skal sikre et representativt utvalg av en populasjon må vi først finne hele populasjonen. I dette eksempelet hvor vi skal gjennomføre en meningsmåling på tvers av hele USA må vi først finne ut hvor mange stemmeberettigede det er i USA for å finne populasjonen vi er interessert i.

Dersom vi har ressurser nok ville det beste vært å spørre alle i populasjonen, men dette er tidkrevende og kostbart så det er meget lite sannsynlig at det ville være mulig. Derfor må vi gjennomføre et utvalg av populasjonen og utvalget må være et sannsynlighetsutvalg som vil si at vi trekker utvalget slik at det er stor sannsynlighet for at utvalget representerer populasjonen. Blant sannsynlighetsutvalgene kan vi velge mellom:

- Rent tilfeldig utvalg - vi trekker så mange vi ønsker tilfeldig fra populasjonen.
- Strukturert tilfeldig utvalg - vi strukturerer utvalget (f. eks. alfabetisk) og trekker så enheter ved et intervall. Hvor stort intervall blir mellom hver vi trekker fra populasjonen kommer an på størrelsen på populasjonen i forhold til størrelsen vi ønsker på utvalget. Eks: populasjonen er 100 personer og vi ønsker et utvalg med 10 personer, vi trekker da nummer 5, 15, 25, 35 osv.
- Klyngeutvalg - populasjonen fordeles inn i klynger (eks: klasser, kommuner).
- Stratifisert utvalg - vi deler inn populasjonen i kategorier og så trekker vi like mange prosent fra kategoriene som det er i populasjonen. Eks: hvis det er 40% menn og 60% kvinner i populasjonen deles disse i to kategorier og så trekker vi 40% av utvalget fra gruppen med menn og 60% av utvalget fra gruppen med kvinner. På denne måten sikrer vi oss at vi får like mange prosent menn og kvinner i utvalget som i populasjonen.

Når man skal trekke et representativt utvalg for en meningsmåling i USA er det viktig å ta hensyn til hvordan presidenten velges i USA. Dette gjøres ved at presidentkandidaten som får flertall i en stat får alle valgmenne som tilhører denne staten (antall valgmann varierer fra stat til stat). Dermed er det ikke hensiktsmessig å ta et rent tilfeldig utvalg eller et strukturert tilfeldig utvalg av hele befolkningen og spørre dem hvem de vil stemme på fordi det er ikke viktig hvem som har flertall på landsbasis det er antall valgmann som teller. Et alternativ når man skal trekke et representativt utvalg av befolkningen til et presidentvalg i USA er derfor å først stratifisere populasjonen etter hvilken stat de har stemmerett og deretter trekke like mange prosent enheter fra hver stat slik at det er likt som antall prosent valgmann i den staten. Da kan man få et representativt utvalg av befolkningen i USA som egner seg til et presidentvalg. Et annet alternativ er å fordele populasjonen i klynger etter i hvilken stat de har stemmerett for så å trekke like mange fra hver stat (f. eks. 1000 personer fra hver stat). Deretter kan vi se på hvem som får flertall i de forskjellige statene for så å tildele dem antallet valgmann som det er i de statene de får flertall på meningsmålingen. Når vi har gjort dette med alle statene ser vi hvilken presidentkandidat som har fått flest valgmann.

Det er viktig å være klar over at det er ikke sikkert vi kan stole på det svaret vi får fra alle i utvalget, vi må vurdere reliabiliteten til meningsmålingen. Reliabilitet er hvorvidt vi kan stole på de svarene vi får. Det er ikke lett å si noe konkret om hvorvidt vi kan stole på det folk sier eller ikke, men vi kan anta at når et av alternativene er kontroversielle, slik som Trump, vil det være sannsynlig at det er flere som ikke tør/vil si hvem de egentlig vil stemme på og reliabiliteten blir dårligere.

Besvart

#### 4 ME-100, oppgave 3

**a)**

Enhet er det vi undersøker, i dette tilfelle individer i USA. Variabler er det som enhetene svarer på. I dette tilfellet er avhengig variabel presidentvalg, og uavhengig variabel etnisitet og arbeidsledighet. Verdi er tallene vi får på variablene i dette tilfellet prosentandel som stemte på Donald Trump, prosentandel hvite og prosentandel arbeidsledige.

**b)**

Stigningstallene er helningskoeffisienten og viser til økningen per X på den avhengige variabelen (konstanten). I denne regresjonsanalysen er tallene oppgitt i prosent og det vil si at en økning i X = 1% økning. Stigningstallet på etnisitet er på 1,4, det vil si at for hver prosent økning av antall hvite i en stat så øker antall stemmer til Trump med 1,4%. Stigningstallet på arbeidsledighet er 2,6, det vil si at for hver prosent økning i arbeidsledigheten så øker antall stemmer til Trump med 2,6%.

Til en multivariat regresjonsanalyse har vi den generelle multivariate regresjonslikningen:

$$Y = a + b_1X_1 + b_2X_2 \dots + e$$

der a er konstanten,  $b_1$ ,  $b_2$  osv. er de uavhengige variablene og e er restleddet som modellen ikke dekker.

I denne regresjonsanalysen er

a = ukjent (prosentandel stemmer på Trump)

$b_1 = 1,4$  (etnisitet)

$b_2 = 2,6$  (arbeidsledige)

$e =$  ukjent

den generelle multivariate regresjonslikningen til denne regresjonsanalysen blir da:

$$Y = a + 1,4X + 2,6X + e$$

Ettersom  $a$  ikke er kjent kan vi ikke regne ut hvor mange prosent stemmer modellen forventer per delstat men vi kan se på hvordan forskjellen med forskjellig andel arbeidsledighet og etnisitet påvirker antall prosent stemmer på Trump.

I en delstat med 50% hvite og en arbeidsledighet på 5%, vil effekten av regresjonsmodellen bli:

$$Y = a + 1,4 * 50\% + 2,6 * 5\% + e$$

$$Y = a + 70\% + 13\% + e = 83\%$$

I en delstat med 70% hvite og en arbeidsledighet på 3% vil effekten av regresjonsmodellen bli:

$$Y = a + 1,4 * 70\% + 2,6 * 3\% + e$$

$$Y = a + 98\% + 7,8\% + e = 105,8\%$$

Ettersom dette er prosent vil en prosentandel stemmer til Trump på 105,8% som i siste eksempel være umulig, men det er viktig å huske at dette viser kun effekten av andelen hvite og andelen arbeidsledige. Siden vi ikke vet  $a$  (konstanten) kan vi ikke få et nøyaktig tall på hvor mange prosent stemmer han ville fått i en delstat med 70% hvite og 3% arbeidsledighet, men det virker naturlig at konstanten vil være negativ. Det kan også være at modellen vår er svak og at etnisitet og arbeidsledighet forklarer lite, mens  $e$  forklarer mye. Hvis dette er tilfellet vil vi si at vi har en svak modell.

c)

Signifikans er et mål på hvorvidt vi kan stole på resultatene. Dersom vi har en signifikans på 5% har vi en sikkerhet på 95%, det vil si at vi vil ta feil 1 av 20 ganger. Signifikansen tar hensyn til hvor sterk sammenhengen mellom variablene er og hvor godt utvalget er. Vi kan selv velge hvilket signifikansnivå vi ønsker, det er vanlig å velge 10%, 5% eller 1%. I samfunnsvitenskapen er det vanligst og velge 5% eller 1%. Noen mener at vi bør strebe etter 1% altså at vi tar feil kun 1 av 100 ganger, men det virker som det er relativt bred konsensus om at 5% signifikans er tilstrekkelig.

Når det gjelder signifikans så er vi ute etter å unngå å gjøre to typer feil. Feil 1 er å forkaste en nullhypotese som stemmer, mens feil 2 er å ikke forkaste en nullhypotese som ikke stemmer. Jo strengere signifikansnivå du setter desto mindre blir sannsynligheten for å gjøre en feil type 1, men samtidig øker sannsynligheten for at vi gjør en feil type 2. Dette gjelder også andre veien bare motsatt retning. Ved en høyere signifikans minsker sannsynligheten for å gjøre en type 2 feil, mens sannsynligheten for en type 1 feil øker.

Jo strengere signifikans vi har desto større blir feilmarginen og jo mildere signifikans vi har desto mindre blir feilmarginen. Dette gjelder også for utvalget; jo færre vi har i utvalget desto større blir feilmarginen og motsatt større utvalg = lavere feilmargin. En tommelfingerregel er at for å få en halvvering i feilmarginen så må vi fireoble utvalgsstørrelsen.

Når stigningstallene er signifikante, slik vi har i denne analysen, vil det si at vi har en lavere signifikans enn det signifikansnivået vi satt før vi begynte analysen.

Besvart