

Sensorveiledning PSYC2104, vår 2020

OPPGAVE 1.

Psykologiske tester kan brukes til å treffe beslutninger. Forklar hvordan baseraten (baserate=BR) og seleksjonsraten (selection rate=SR) sammen med testens validitet, influerer på hvorvidt beslutninger blir riktige eller gale.

BR, SR og kriterievaliditet må defineres. I læreboka i teststatistikk er disse begrepene knyttet til rekruttering av medarbeidere, men det kunne like gjerne vært knyttet til andre kontekster, slik som det å bruke tester til å avdekke diagnoser (baseraten ville i så fall vært hvor mange som i virkeligheten har en gitt diagnose).

Det er en fordel om man også relaterer dette til falske positive og negative, sammen med sanne positive og negative. Dette kan igjen relateres til hit rate og miss rate.

Baseraten og seleksjonsraten bør sees i sammenheng. Jo mer like disse er, jo bedre er utgangspunktet for å treffe riktige beslutninger (dvs. best mulig «hit-rate»). Videre er det ikke hensiktsmessig å bruke en test i denne sammenheng hvis baseraten er høy, og/eller hvis seleksjonsraten er høy. Med moderat til lav baserate kan imidlertid bruk av tester være fordelaktig, og nytteverdien stiger i takt med validitetskoeffisienten (r_{xy}). Samtidig må man ta hensyn til seleksjonsraten. Andelen sanne positive (SP) man kan forvente å få kan estimeres ved å ta utgangspunkt i BR, SR og testen kriterievaliditet, presentert i denne formelen:

$$p(\text{SP}) = \text{BR} * \text{SR} + r_{xy} \sqrt{\text{BR}(1 - \text{BR})\text{SR}(1 - \text{SR})}$$

Bruke gjerne formelen til å illustrere noen eksempler med ulike validitetskoeffisienter og ulike BR og SR.

OPPGAVE 2

a) Gjør kort rede for ulike måter å estimere reliabilitet på.

Først må reliabilitet defineres. Deretter må de fire hovedformene for reliabilitet forklares: test-retest, alternativ form, split-half og intern konsistensbasert reliabilitet (Cronbach's alfa). Det bør forklares litt om styrker og svakheter ved de enkelte metodene. Inter-rater reliabilitet kan nevnes i tillegg

OPPGAVE 3.

a) Forklar hva en målings standardfeil (standard error of measurement =SEM) er.

SEM er et mål på statistisk usikkerhet som indikerer hvor store feilmarginer vi opererer med mht. testskårer. SEM er basert på klassisk testteori som hviler på

forutsetningen om at en persons observerte skåre (x) er et produkt av «sann skåre»(T) og målefeil (e): $X=T+e$.

SEM estimeres på grunnlag av reliabiliteten til skårene på en test og standardavviket for skårene på testen. Jo større standardavvik, jo større blir SEM. Men samtidig har reliabiliteten til testskåtene også mye å si, jo lavere reliabiliteten er, jo større usikkerhet og dermed større SEM. Motsatt gir høy reliabilitet lavere SEM – en perfekt reliabilitet på 1 ville gitt en SEM på 0. Det er en fordel om konfidensintervall for SEM forklares (her eller i de neste to deloppgavene) og at dette relateres til normalfordelingskuven.

Test-retest reliabiliteten til skårene på WAIS-V fullskala IQ er estimert til å være på 0.945 ($r_{xx} = 0.945$). Skårene på testen har videre et gjennomsnitt på 100 og et standardavvik på 15.

b) *Estimer standardfeilen for måling (SEM) for fullskala IQ. Vis hvordan du kommer frem til svaret.*

SEM beregnes ut ifra denne formelen: $\sigma\sqrt{1 - r_{xx}} = 15\sqrt{1 - 0.945} = 15 \cdot 0.2345 = 3.52$

c) *Du har gjennomført testing av to personer på WAIS-V. Person 1 fikk en fullskala IQ-skåre på 105, mens person 2 fikk en skåre på 90.*

- Estimer 95 % konfidensintervall for sann skåre på fullskala IQ for hver av de to personene. Vis hvordan du kommer frem til svaret

Et 95 % konfidensintervall estimeres gjennom å gangen SEM med 1.96 (tallet 2 godtas også):

Person 1: $105 \pm 1.96 \cdot 3.52 = 105 \pm 6.9$. Dermed blir 95% KI: [98.1, 111.9]

Person 2: $90 \pm 1.96 \cdot 3.52 = 90 \pm 6.9$. Dermed: [83.1, 96.9]

- Estimer standardfeilen for forskjell i skåre mellom personer for fullskala IQ. Kan du med 95% sikkerhet si at de to personene har en skåre som er forskjellig fra hverandre? Vis hvordan du kommer frem til svaret.

$$\begin{aligned} SEM \text{ diff} &= \sigma\sqrt{2 - r_{xx} - r_{xx}} \\ &= 15\sqrt{2 - 0.945 - 0.945} = 15 * 0.332 \\ &= 4.98 \end{aligned}$$

95% SEM for differanse = $1.96 * 4.98 = 9.96$

Differansen er på 15 poeng, noe som er større 9.96. Det er dermed mindre enn 5% sannsynlighet for tilfeldig forskjell, og vi antar derfor at det er en forskjell i sann skårer mellom de to personene.

OPPGAVE 4

- a) *Undersøk om effektene av holdning, subjektiv norm og oppfattet atferdskontroll har en effekt på intensjon om å trene. Presenter resultatet av analysen og gi også din tolkning av resultatene – og hvordan du vil beskrive sammenhengene du eventuelt finner. Hvilken av de tre prediktorene ser ut til å ha sterkest effekt på den avhengige variabelen? Forklar.*

Resultatene av analysen må presenteres i form av en tabell som viser både standardiserte og ustandardiserte regresjonskoeffisienter, med tilhørende p-verdi som benyttes til å avgjøre om koeffisientene er statistisk signifikante på 5%-nivå. R^2 må også kommenteres.

Det er viktig at man tolker:

- retning på koeffisientene. Her er det viktig å forklare at når holdning til trening øker, så øker intensjon om å trene. Likeså øker intensjon om trene når skåre på subjektiv norm øker. Videre betyr regresjonskoeffisienten på 0.15 til oppfattet atferdskontroll at de som har høy kontroll (kodet 1) har i gjennomsnitt 0.15 poengs høyere intensjon om å trene enn de som har lav kontroll (kodet 0).
- om effektene er statistisk signifikant eller ikke på 5% nivå
- hvor sterk effekten er. Her er det kun meningsfylt å sammenligne effekten til holdning og subjektiv norm, gjennom å sammenligne størrelsen på de standardiserte koeffisientene. Holdning har en langt sterkere effekt enn subjektiv norm.

- b) *Undersøk om forutsetningene om linearitet, fravær av multikollinearitet og normalfordelte residualer er oppfylt i den multiple regresjonsmodellen du estimerte i oppgave a.*

Linearitet. Denne forutsetningen er kun relevant for prediktorene holdning og subjektiv norm (oppfattet atferdskontroll er dikotom og det er ikke mulig å undersøke linearitet da). Linearitet kan undersøkes på flere måter. Man kan se visuelt på de såkalte partial plots, og ut i fra denne ser det ut som en rett linje beskrivende for data. Men det finnes mer eksakte tester som viser om en ikke-lineær kurve gir bedre tilpasning til data, dvs. høyere andel forklart varians (R^2). Eksempelvis kan man undersøke om en kvadratisk eller en logaritmisk modell vil passe bedre til data (gi signifikant høyere forklart varians). En rett linje beskriver sammenhengen mellom holdning og intensjon godt, og ingen andre ikke-lineære modeller vil passe bedre til data. Når det gjelder subjektiv norm, så vil en kvadratisk modell passe litt (men ikke så mye) bedre. I dette tilfellet betyr det at effekten av sosial norm er liten til å begynne med, for så å øke mer og mer.

Multikollinearitet: Alle korrelasjonene mellom de tre prediktorene er relativt lave, dvs. under .80. En bedre test er å se på toleranseverdiene og/eller VIF. Disse tilsier at multikollinearitet ikke er et problem i regresjonsmodellen.

Normalfordelte residualer: Ut i fra en visuell inspeksjon av histogrammet over residualene estimert ut i fra regresjonsmodellen kan man slutte at residualene er noenlunde normalfordelte.

- c) *Estimer den direkte, indirekte og totale effekten av holdning til trening på atferd (forekomst av trening).*

Merk: I oppgave c kan dere bruke makroen indirect.sbs, men det er ikke noe krav om det siden flere har problemer med å få indirect.sbs til å virke. Det er helt i orden å bruke fremgangsmåten basert på Baron & Kenny's fire trinn, som ikke krever at man bruker indirect.sbs.

Gjennom å benytte indirect sbs, vil man kunne lese av effekten av holdning på atferd i stor grad er mediert av intensjon om å trene. Likeledes vil man kunne komme fram til samme konklusjon gjennom å benytte Baron & Kenny's fire trinn for mediatoranalyse.

Hvis man benytter subjektiv norm og oppfattet atferdskontroll som kovariater, vi resultatet bli slik:

Den direkte effekten (sti c') er lik: 0.039 (p = 0.441)

Den indirekte effekten (sti a*sti b) er lik: $(0.45 * 0.67) = 0.30$

Den totale effekten er (sti c) er lik: 0.339 (p <.001)

Det godtas også at man beregner de ovennevnte effektene uten å ta med subjektiv norm og oppfattet atferdskontroll som kovariater.

Plusspoeng hvis modellen tegens opp grafisk. Det er viktig at kandidaten ikke benytter andel forklart varians som en indikator på om en mediatoreffekt er tilstede eller ikke, det er ikke riktig.

Benyttes makroen indirect.sbs, kan man undersøke om den indirekte effekten er signifikant forskjellig fra null. Det mest pålitelige er å se på konfidensintervallet for den indirekte effekten, og om dette krysser null. Det gjør ikke konfidensintervallet, og den indirekte effekten er følgelig signifikant forskjellig fra null og vi har en mediatoreffekt til stede.