

Sensorveiledning SOS4020 Gammel ordning - høst 2004

a) Begrepsvaliditet handler om i hvilken grad vi har lykket med å måle/registrere det vi er interessert i å måle på en pålitelig måte. (Dessuten legger læreboken problemer utvalgsskjevhet og om eventuelle intervensjoner som skal evalueres svarer til det man har tilsiktet å gjøre.)

Siden de 7 spørsmålene skal slås sammen til et samlet mål på skoleholdninger er det viktig at alle spørsmålene skåres på sammenlignbar måte, eksempelvis slik at positive holdninger skårer høyt og negative holdninger lavt. De fire første utsagnene vil gi høy verdi (stemmer godt) hvis eleven har positive holdninger til skolen, og lave verdier hvis holdningene er mer negative (stemmer dårlig). Siden de tre siste spørsmålene er av negativ art, vil positive holdninger gi svaret ”stemmer dårlig”, mens negative holdninger gir svaret ”stemmer godt”. Førstnevnte må derfor kodes 3, mens sistnevnte må kodes 0, altså omvendt sammenlignet med de fire første spørsmålene. Anta at man ikke hadde valgt denne kodingen, men benyttet samme kode på alle 7 spørsmål. Da ville de med mest positive holdninger fått høyeste verdi på de fire første spørsmålene, samlet $3+3+3+3=12$ og laveste verdi på de tre siste, samlet $0+0+0=0$. Totalt ville de fått 12 poeng. De med mest negative holdninger ville fått laveste verdi på de tre første, samlet $0+0+0=0$, og høyeste på de tre siste, samlet $3+3+3=9$. Totalt ville de fått 9 poeng. De med litt over middels positive holdninger, eksempelvis 2 på de fire første og 1 de tre siste, ville fått $2+2+2+2+1+1+1=11$ poeng totalt. Mens de med litt under middel holdninger, eksempelvis 1 på de fire første og 2 på de tre siste, ville fått $1+1+1+1+2+2+2=10$ poeng totalt. Det ville med andre ord bli liten forskjell på de med positive og negative holdninger – indeksen ville ikke skille særlig godt mellom ulike holdninger.

Med hensyn til mulige begrensninger ved indeksens begrepsvaliditet kan kandidaten bringe inn mange forskjellige forhold. Eksempelvis er det et begrenset antall spørsmål som legges til grunn ved indekskonstruksjonen, og en kunne godt tenke seg en rekke andre aspekter som burde vært med. For det andre kan man problematisere noen av spørsmålene mht. om det bare er holdninger til skolen de måler. Man kan eksempelvis grue seg til skolen, ikke fordi en misliker skolen som sådan, men for man er utsatt for mobbing eller lignende.

En kan også reise spørsmålet om ikke noen av de 7 spørsmålene som her er stilt måler temmelig ulike dimensjoner. I så fall burde man kanskje vurdert å lage flere indekser, gjennom passende gruppering av spørsmålene.

b) En standardfeil måler graden av statistisk usikkerhet knyttet til et estimat – i dette tilfellet et (aritmetisk) gjennomsnitt. Standardfeilen er standardavviket i sannsynlighetsfordelingen til parameterestimatet.

Et konfidensintervall er en øvre og nedre grense for en parameter, konstruert slik at intervallet vil innholde den sanne parameterverdien med en valgt sannsynlighet. Når utvalget (som her) er stort, vil sannsynlighetsfordelingen ligge nær normalfordelingen og et 95% intervall konstrueres ved å multiplisere standardfeilen med 1.96.

$$SE(\hat{m}) = 3.4 / \sqrt{984} = 0.108$$

Øvre grense: $13.4 + 1.96 \cdot 0.108 = 13.61$

Nedre grense: $13.4 - 1.96 \cdot 0.108 = 13.19$

c)

$H_0: \hat{m}_A = \hat{m}_B$

$H_A: \hat{m}_A \neq \hat{m}_B$

$$SE(\hat{m}_A - \hat{m}_B) = \sqrt{[0.21]^2 + [0.18]^2} = 0.277$$

$$t = \frac{13.87 - 12.89}{0.277} = 3.54, \quad df = 298 + 254 - 2 = 550$$

Siden antallet frihetsgrader er stort kan en bruke normalfordelingen, og kritisk t-verdi på 5% nivået er da 1.96. Siden testobservatoren er vesentlig større enn dette forkaster vi nullhypotesen og trekker den konklusjon at det er forskjell mellom de to ungdomsskolene.

d) Standardavviket i en populasjonsfordeling er et mål på graden av spredning rundt gjennomsnittet av observasjonsenheter i populasjonen, dvs. på graden av individuelle ulikheter. Standardfeilen i en sannsynlighetsfordeling er et mål på statistisk usikkerhet knyttet til et estimat og ligger til grunn for beregning av statistiske usikkerhetsmarginer (konfidensintervall). En standardfeil er også et standardavvik, men ikke til en populasjonsfordeling, men til en sannsynlighetsfordeling over mulige utfall når man gjør en utvalgsundersøkelse.

Signifikanssannsynligheten til en testobservator er sannsynligheten for at vi skal få den testobservatoren vi har fått, eller en som er enda større i absoluttverdi, bare som resultat av tilfeldigheter. En liten signifikanssannsynlighet er med andre ord en indikasjon på at nullhypotesen har liten troverdighet og vil føre til at vi forkaster nullhypotesen.

e) Alpha-koeffisienten er et mål knyttet til indekser som er dannet ved å summere flere variable til et samlet mål. De enkelte spørsmålene er antatt å måle den samme underliggende dimensjonen, i dette tilfellet holdninger til skolen. Alpha-koeffisienten kalles ofte et generaliserbarhetsmål og er et mål på hvilket samsvar man kan forvente mellom den indeksen man faktisk har konstruert, og andre indekser på det samme som kunne vært konstruert ved å stille et like stort antall beslektede spørsmål om samme tema. Koeffisienten tolkes som en korrelasjonskoeffisient. Koeffisienten størrelse vil være avhengig dels av antall spørsmål som inngår i indeksen, og dels av hvor sterk sammenheng det er mellom informantenes svar på de enkelte spørsmålene. Jo flere spørsmål indeksen er basert på, desto større blir alpha. Og jo bedre samsvar det er mellom svarene på de enkelte spørsmålene, desto større blir alpha.

f) Koeffisientene testes med t-test, og er signifikante på 5% nivået dersom testobservatoren er større enn 1.96 i tallverdi (to-halet test), siden antallet observasjoner er stort.

$$\text{Kjønn:} \quad t = \frac{0.04}{0.22} = 0.18, \text{ ikke signifikant}$$

$$\text{Ungdomsskole A:} \quad t = \frac{0.99}{0.29} = 3.41, \text{ signifikant}$$

$$\text{Privat skole:} \quad t = \frac{0.59}{0.40} = 1.48, \text{ ikke signifikant}$$

$$\text{Videregående skole:} \quad t = \frac{0.52}{0.28} = 1.86, \text{ ikke signifikant}$$

Det er ingen påviselig kjønnseffekt, kontrollert for skoletype, med hensyn til skoleholdninger. Elevene ved ungdomsskole A har mer positive skoleholdninger enn elevene ved ungdomsskole B (referanseskolen). Derimot er det ikke signifikante forskjeller mellom referanseskolen på den ene siden og henholdsvis privat skole og videregående skole på den annen side.

Ungdomsskole B har lavest skåre i gjennomsnitt, mens ungdomsskole A har høyest skåre.

g)

$$\text{Alder:} \quad t = \frac{-0.94}{0.29} = -3.24$$

Alder kvadrert: $t = 0.15 / 0.05 = 3.00$

Begge koeffisienter er signifikante.

Hensikten med andregradsleddet er å ta vare på eventuell kurvelinearitet. Siden koeffisienten til andregradsleddet er statistisk signifikant forskjellig fra null, må nullhypotesen om at sammenhengen er lineær forkastes.

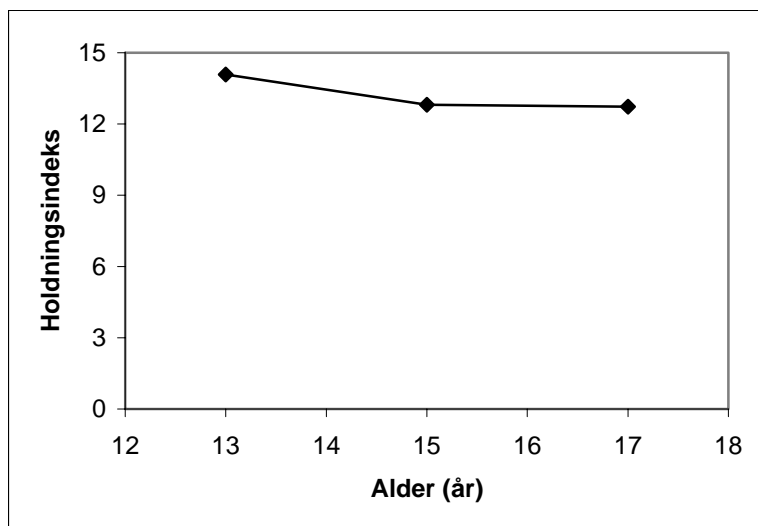
Man kan også innse at siden andregradsleddet er positivt, står vi overfor en u-formet sammenheng og et minimumspunkt når aldersvariabelen er $X^* = -(-0.94) / (2 * 0.15) = 3.1$, dvs. i aldergruppen $13 + 3.1 = 16.1$ år.

h) Predikert (forventet) verdi for gutter (kjønn=0) i ungdomsskole B (0 på alle dummyvariable) er:

13 år (alder=0): $14.09 + (-0.94)*0 + 0.15*0*0 = 14.09$

15 år (alder=2): $14.09 + (-0.94)*2 + 0.15*2*2 = 12.81$

17 år (alder=4): $14.09 + (-0.94)*4 + 0.15*4*4 = 12.73$



Kurvene for andre skoleslag blir parallelle med kurven for ungdomsskole B og vil gjennomgående ligge på et høyere nivå. Det forutsettes ikke at kandidaten har gjort utregningen for alle skoleslag.

Kandidaten vil kanskje påpeke et problem med den substansielle tolkbarhet av kurvene, idet hver enkelt av disse kurvene (med mulig unntak av kurven for privatskole) for noen aldersgrupper vil referere til grupper med svært få (eller ingen) observasjoner: Det er få 17-åringer i grunnskolen og ingen 13- og 15-åringer i videregående skole.

De positive skoleholdningene **avtar** med andre ord med stigende alder blant de yngste. Mellom 15- og 17-åringene er det imidlertid liten forskjell. Enkelte kandidater vil sannsynligvis innse at laveste holdningsverdi gjenfinnes i aldersgruppen 16 år, enten ved å beregne predikert verdi også for denne gruppen, eller som angitt ovenfor. I så fall vil de også innse at kurven er økende fra 16 til 17 år.

i) Økning i standardfeilen når nye variable tas inn i modellen tolkes i læreboken som tegn på kolinearitet. Det er et tegn på at det er forholdsvis høy korrelasjon mellom variabelen ”Videregående skole” og de to aldersvariablene som er tatt inn i modellen. Grunnen til dette er at det bare er de eldste som går på videregående.

j) Variabelen ”Sosial status” er ikke signifikant ($t = -0.20$), mens ”God foreldrekontakt” er klart signifikant ($t = 12.6$). Det er med andre ord ingen klare tegn til at barn fra lavstatushjem har andre holdninger til skolen enn barn fra høystatushjem (når en som her kontrollerer for kjønn, alder, foreldrekontakt og skoleslag). Derimot er det en klar sammenheng mellom graden av foreldrekontakt og skoleholdninger, kontrollert for de øvre variablene i modellen: Barn som sier de har god kontakt med sine foreldre har gjennomgående atskillig mer positive holdninger enn barn med dårlig foreldrekontakt. En legger merke til at R^2 øker fra under 2% til over 14% når denne variabelen tas inn i modellen. Det ser med andre ord ut til at gode relasjoner til foreldre har en gunstig innvirkning på de unges holdninger til skolen.

k) Det er ikke grunnlag for å hevde at forskjellen mellom de to ungdomsskolene er resultat av ulikheter mht. sosial status og elevenes foreldrekontakt, siden parameterverdien til ungdomsskole A ikke har forandret seg nevneverdig og fortsatt er signifikant forskjellig fra null i modell nr. 3.

l) Dersom sosial status bare påvirker skoleholdninger blant de i videregående skole, og ikke blant andre, vil det foreligge samspill mellom variablene ”Sosial status” og ”Videregående skole”. Én måte å undersøke dette på vil bestå i å gjøre partielle analyser, der en først ser på de som går i videregående skole separat, og deretter de øvrige. Hvis hypotesen er riktig vil en finne en effekt av sosial klasse blant elevene i videregående, men ikke blant de andre. Den andre tilnærmingen vil bestå i å bruke et samspillsledd, definert som produktet av de nevnte variablene. Et signifikant samspillsledd vil peke i retning av forskjellig effekt av sosial status i ulike skoleslag.