

Oppgave 1 (Møen)

- a) Analyse 1 viser at gjennomsnittskaracteren i masterstudiet for internt opptatte er 4,02 (ren B) mens gjennomsnittskaracteren for eksternt opptatte er 3,84 (en litt svak B). Forskjellen på 0,19 er klart signifikant.

Analyse 2 er en enveis faktoranalyse som tester om de tre gruppene, interne, eksterne fra Norden og andre eksterne har et felles forventet karakternivå. I lys av analyse 1 er det ikke overraskende at denne hypotesen forkastes. Interessant ny informasjon i denne analysen er at det er svært liten forskjell mellom nordiske eksterne studenter og ikke-nordiske eksterne studenter. Det er imidlertid større spredning rundt gjennomsnittet for de ikke-nordiske.

- b) Analyse 3 viser at interne og eksterne studenter er svært like mht. hvor mange studiepoeng de avlegger per semester, og den gjennomsnittlige studiepoengproduksjonen er rett i underkant av normen på 30 per semester. Det er ingen signifikant forskjell.

Analyse 4 viser at det heller ikke er signifikant forskjell med hensyn til andelen som er forsinket i studiet. Det er litt mindre enn en tredjedel i begge gruppene.

- c) Analysene viser at interne studenter gjør det litt bedre enn eksterne studenter, men det er ingen grunn åpenbar grunn til å endre opptakskvotene. Det er imidlertid større variasjon i prestasjonene til ikke-nordiske studenter enn til norske og nordiske. Det kan således være et forbedringspotensial i opptakskriteriene i det internasjonale opptaket.

- d) Fra analyse 2 ser vi at standardavviket for eksterne ikke-nordiske studenter er 0,5592 og for eksterne norske og nordiske studenter er det 0,4419. Vi gjør følgende F-test: $F = 0,5592^2 / 0,4419^2 = 1,60$. Kritisk grense på 5% nivå for $F_{175,569}$ er tilnærmet lik 1,18. Altså kan vi forkaste en nullhypotese om lik varians. (Dersom man har snudd brøken og får F-verdi lik 0,62 trenger man en nedre grense som er $1/1,18 = 0,85$.)

- e) Andelen som har sluttet er $68/893 = 0,0761$ for eksterne og $95/1335 = 0,0712$ for interne. Andelen som har sluttet totalt er $163/2228 = 0,0732$. Vi bruker en Z-test (U-testen) der
$$Z = \frac{0,0761 - 0,0712}{\sqrt{\left(\frac{1}{893} + \frac{1}{1335}\right)[0,0732 \cdot (1 - 0,0732)]}} = 0,443$$
. Kritisk grense for tosidig test på 5%-nivå er 1,96. Følgelig kan vi ikke forkaste en nullhypotese om like andeler.

- f) (i) Karakterdataene i vår analyse er gjennomsnittskaracterer. Disse vil være tilnærmet normalfordelte pga. sentralgrenseteoremet. Forutsetningen for de testene vi har brukt er derfor i orden. (For at det skal være rimelig å regne ut gjennomsnittskaracterer må vi imidlertid anta at tallkarakterene er målevariabler slik at f.eks. avstanden mellom A og B er den samme som avstanden mellom B og C på en underliggende «kunnskapsskala». Hvis ikke er for eksempel karakterene B, B, B, ikke likeverdig med A, B, C, selv om gjennomsnittet av tallkarakterene er det samme. Dette kan problematiseres fordi en karakterskala i realiteten er ordinal. Alt vi kan si med sikkerhet er at A er bedre enn B og B er bedre enn C. Vi vet ikke om skalaen praktiseres slik at avstanden mellom karakterene er lik langs en underliggende, målbar dimensjon.)

(ii) I analyse 1 og 3 kan vi bruke den fordelingsfrie Willcoxons toutvalgstest (Mann-Whitney-testen). I stedet for variansanalysen i analyse 2 kan vi bruke den fordelingsfrie Kruskal-Wallis-testen. Analyse 4 er det ikke aktuelt å erstatte med en fordelingsfri test. Den bygger ikke på en fordelingsantagelse, men på teori for binomiske forsøksrekker.

Oppgave 2 (Møen)

- a) Vi ser at karaktersnittet på masteroppgavene ligger på rundt 4,27 mens karaktersnittet for kurs ligger rundt 3,75. I gjennomsnitt får studentene altså en halv karakter bedre på oppgaven enn på kursdelen. Doble utredninger får omkring 1/3 bedre karakter enn enkle utredninger, men det er liten forskjell i gjennomsnittskarakter i kursdelen mellom studenter som skriver enkle og doble utredninger. Det er imidlertid større spredning på karakterene til de som skriver enkle utredninger. (En mulig forklaring på det er at de virkelig svake har problemer med å finne noen å skrive sammen med eller er ikke tilstrekkelig til stede i studiet til å kunne samarbeide, og en del virkelig gode velger å skrive alene.)
- b) Like gode studenter kan forvente å få 0,28 bedre karakter om de skriver dobbel utredning sammenlignet med om de skriver enkel utredning. Dette kan skyldes at studenter som skriver sammen presterer bedre fordi de kan diskutere og hjelpe hverandre. Det kan også skyldes at sensorene ikke stiller strenge nok krav til doble utredninger med hensyn til arbeidsomfang. En tredje mulighet kan være at studenter som velger å skrive dobbel utredning er mer motivert for å legge ned ekstra arbeid i masteroppgaven. En fjerde mulighet er at den beste av to forfattere på en dobbel utredning vanligvis hever oppgaven opp til sitt kvalitetsnivå.
- c) (i) Kolonne 3: $1.957 + 0.883 + 0.427*4 = \underline{4,548}$
Kolonne 4: $1.957 + 0.883 + 0.264*4 + 0.163*4 = \underline{4,548}$
(ii) Kolonne 3: $1.957 + 0.883 + 0.427*4 + 0.0506*(4,5-3,5) = \underline{4,5986}$
Kolonne 4: $1.957 + 0.883 + 0.264*4,5 + 0.163*3,5 = \underline{4,5985}$
- d) Vi ser fra punkttestimatene i kolonne (4) at den beste studenten påvirker karakteren litt mer enn den dårligste studenten. Standardavvikene tilsier imidlertid at de to koeffisientene ikke er signifikant forskjellige fra hverandre. Det ser vi også direkte ved at karakterdifferansen i kolonne (3) ikke er signifikant i tillegg til snittkarakteren. Vi kan således ikke påvise noe omfattende gratispassasjerproblem.
- e) Spesifikasjonene i kolonne (3) og (4) representerer nøyaktig samme regresjonsligning. Man kan relativt enkelt regne seg fra den ene til den andre og vise den nøyaktige sammenhengen mellom koeffisientene. (Inferensen i de to regresjonsanalysene har imidlertid forskjellige nullhypoteser.)
- f) Dersom vi flytter studenter fra veileder 10 til veileder 30 og holder alt annet likt, øker forventet karakter på masteroppgaven deres med ca. 1,5! Dette kan skyldes ulikt skjønn ved karaktersettingen eller at veileder 30 er dyktigere i å veilede studentene fram til å skrive gode oppgaver. [Nærmere analyser viser at forskjellen mellom veilederne blir vesentlig mindre om man utelater oppgaver med dårligere karakterer enn C. Spriket drives altså i stor grad av bruken av disse relativt sjeldne karakterene.]

- g) Med 65 dummier og 5% signifikansnivå forventer vi å finne $0,05 \cdot 65 \approx 3$ signifikante koeffisienter selv om det ikke er sanne forskjeller mellom veilederne. Når vi finner 11 er det rimelig å anta at det finnes reelle forskjeller.

Oppgave 3 (Lillestøl)

Merk at $E(I) = 0 \cdot (1-p) + 1 \cdot p = p$ og at $I^2 = I$ for 0-1 variable, slik at også $E(I^2) = p$.

- a) $E(T) = E(Y \cdot I) = E(Y) \cdot E(I) = E(Y) \cdot p$
 $\text{var}(T) = E(T^2) - (ET)^2 = E(Y^2 \cdot I^2) - (EY \cdot p)^2 = E(Y^2) \cdot E(I^2) - (E(Y))^2 \cdot p^2$
 $= (\text{var}Y + (EY)^2) \cdot p - (E(Y))^2 \cdot p^2 = \text{var}Y \cdot p + (EY)^2 p(1-p)$
- b) $D(z)$ er mellom 0 og 1 og økende med z . (Dette ses ved en enkle grensebetraktning der en lar z gå mot $-\infty$ og $+\infty$..
- c) Det er bare å sette inn og regne ut. Merk at $P(Y=0)=1-P(Y=1)$.