Løsningsforslag

Oppgavesettet og løsningsforslaget er utarbeidet av Jarle Møen. Det anbefales at sensorene gir inntil 10 poeng per delspørsmål.

Oppgave 1

a) Vi bruker F-test. Testobservatoren er $F = \frac{19,03^2}{16,31^2} = 1,36$. Fra tabell ser vi at kritisk grense på 5 %-nivå for tosidig F-test med 58 og 57 frihetsgrader er ca. 1,70. Følgelig kan vi ikke forkaste en nullhypotese om lik varians.

b) Vi bruker T-test for to uavhengige utvalg med lik varians. Testobservatoren er $= \frac{69,62-68,06}{3,28} = 0,48$ der nevneren er beregnet som $\sqrt{\frac{58\cdot 19,03^2+57\cdot 16,31^2}{59+58-2}\cdot\left(\frac{1}{59}+\frac{1}{58}\right)} = \sqrt{10,75} = 3,28$. Kritisk grense for tosidig T-test med 115 frihetsgrader er 1,98. Følgelig er vi ikke i nærheten av å kunne forkaste nullhypotesen om lik forventning for de to sensorene.

c) Vi bruker T-test for matchede par. Testobservatoren er $T = \frac{-7,04}{6,44/\sqrt{59}} = -8,40$. Kritisk grense på 5 %-nivå for tosidig T-test med 58 frihetsgrader er 2,00. Følgelig forkaster vi klart nullhypotesen om lik forventet poengsum når Bårdsen og Nansen retter.

d) (i) Feil – Størrelsen på utvalget påvirker ikke fordelingen til variabelen!
(ii) Rett – Gjennomsnittet inneholder en sum av alle observasjonene i utvalget, og en sum av uavhengig, identisk fordelte variabler vil nærme seg normalfordelingen når antall ledd i utvalget øker.

e) (i) Den første variansanalysen tester om de to førstesensorene har lik forventet poengsum. Den andre variansanalysen tester på samme måte om alle de tre andresensorene har lik forventet poengsum.
(ii) To av følgende tre forutsetninger bør være med: Normalfordelte variabler, lik varians og uavhengige trekninger.
(iii) Vi ser at det er ingen signifikant forskjell mellom de to førstesensorene, men at minst en av de tre andresensorene har en avvikende forventet poengsum når han eller hun retter. Nansen er den som skiller seg mest ut.

f) Testobservatoren er $T = r\cdot\sqrt{\frac{n-2}{1-r^2}} = 0,9306\sqrt{\frac{347-2}{1-0,9306^2}} = 47,22$. Kritisk grense for tosidig T-test på 5 %-nivå med 345 frihetsgrader er 1,96. Følgelig kan vi med svært god margin forkaste en nullhypotese om ingen korrelasjon. Dette er som forventet. Det ville vært uhyre oppsiktsvekkende om det ikke var noen samvariasjon mellom første- og andresensors bedømmelser.

g) (i) Rett – Korrelasjon måler graden av lineær samvariasjon slik at perfekt korrelerte variabler ligger på en rett linje. Sammenhengen mellom de to poengsummene vil derfor være på formen Poeng1=a+b*Poeng2, og den ene poengsummen er monotont stigende i den andre.

(ii) Feil – Konstantleddet, a, i formelen ovenfor trenger ikke være null og stigningstallet, b, trenger ikke være 1.

h) Vi vet at korrelasjonskoeffisienten er 0.9306. Forklaringsgraden til en enkel regresjon er kvadratet av korrelasjonskoeffisienten (derav navnet R^2). Følgelig har vi $R^2 = 0,87$.

i) Formelen for stigningstallet i en enkel regresjon av *Poeng2* mot *Poeng1* kan skrives som $\hat{\beta} = \frac{\widehat{Cov}(Poeng1,oeng2)}{\widehat{Var}(Poeng2)}$ mens korrelasjonskoeffisenten mellom Poeng1 og Poeng2 er $\hat{\rho} = \frac{\widehat{Cov}(Poeng1,Poeng2)}{\widehat{SE}(Poeng1)\cdot\widehat{SE}(Poeng2)}$. Følgelig har vi at $\hat{\beta} = \hat{\rho}\frac{\widehat{SE}(Poeng1)}{\widehat{SE}(Poeng2)} = 0,9306 \cdot \frac{17.00}{17.53} = 0.90$.

j) $Cov(Poeng1,\ Poeng2) = E[(Poeng1 - EPoeng1) \cdot (Poeng2 - EPoeng2)]$

$$= E[(P + \varepsilon_1 - EP) \cdot (P + \varepsilon_2 - EP)] = E[(\{P - EP\} + \varepsilon_1) \cdot (\{P - EP\} + \varepsilon_2)]$$

$$= E[\{P - EP\}^2 + \{P - EP\} \cdot \varepsilon_2 + \{P - EP\} \cdot \varepsilon_1 + \varepsilon_1 \cdot \varepsilon_2] = E(P - EP)^2 = Var(P).$$

k) Det er ett sett med dummyvariabler for førstesensor, med Eriksen som utelatt kategori og ett sett med dummyvariabler for andresensor, med Nansen som utelatt kategori. Følgelig er det (5-2)=3 dummyvariabler tilsammen. Konstantleddet fanger opp de kandidatene som hadde Eriksen som førstesensor og Nansen som andresensor, altså de som fikk sin snittpoengsum satt av kommisjon 4.

l) Vi ser at regresjonen har svært lav forklaringskraft med R^2 nær null. F-testen er dessuten langt fra signifikant. Vi kan altså ikke forkaste en nullhypotese om at alle koeffisienter utenom konstantleddet er null samtidig. Da er det liten grunn til å legge vekt på at dummyvariabelen for at Trondsen er andresensor er nær ved å være signifikant på 5 %-nivå. Regresjonsanalysen gir grunnlag for å konkludere med at hvilken kommisjon en besvarelse blir rettet av ikke har vesentlig innflytelse på utfallet av sensuren.

m) Vi må bruke kjikvadrattesten for modelltilpasning (Goodness-of-fit).

Karakter	A	B	C	D	E
Observert	70	127	91	41	4
Forventet under H_0	333x0.10=33.30	333x0.25=83.25	333x30=99.90	333x0.25=83,25	333x10=33.30
Differanse	36.70	43.75	-8.90	-42.25	-29.30

Testobservatoren blir $\chi^2 = \frac{36,7^2}{33,30} + \frac{43,75^2}{83,25} + \frac{8,90^2}{99,90} + \frac{42,25^2}{83,25} + \frac{29,3^2}{33,30}$=111,45. Kritisk grense for kjikvadrattest med 4 frihetsgrader er 9,49. Vi kan altså med god margin forkaste en nullhypotese om at karakterfordelingen følger ECTS-normen. Fra de observerte frekvensene ser vi at karaktergivingen er for snill.

n) Vi kan betrakte andelen stryk som utfallet av en binomisk forsøksrekke. Forventet andel stryk er p=0,04. Standardavviket til denne andelen er $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0,04 \cdot 0,96}{400}} = 0,0098$. Pga. sentralgrenseteoremet vil andelen være tilnærmet normalfordelt. Et 95 % konfidensintervall er da på formen 0,04±1,96·0,0098≈0,04±0,02.

o) $P(Snitt < 30) = P\left(\frac{Snitt - \mu}{\sigma} < \frac{30 - \mu}{\sigma}\right) = P\left(Z < \frac{30 - 68}{17}\right) = P(Z < -2,24) = 0,0125.$
(Dette er lavt sammenlignet med den observerte strykprosenten og tyder på at den virkelige fordelingen har tykkere venstrehale enn normalfordelingen.)