

Eksamen juni 2001, Opgave 4

(a)

Udregninger:

Bemærk: Udregningerne i det følgende skal læses på nær afrundingsfejl. Resultaterne er skrevet af fra et WinT program, som tager udgangspunkt i de opgivne mellemregningsstørrelser, men regner meget mere nøjagtigt end antydnet nedenfor. Det er OK hvis man ved eksamen gør noget lignende. For eksempel er det en god idé at gemme $\hat{\beta}$ med fuld nøjagtighed i et register på lommeregneren og benytte dette resultat i senere udregninger hvor $\hat{\beta}$ indgår, men alligevel skrive $\hat{\beta}$ passende afrundet som facit (og i mellemregninger).

$$\bar{x} = 267.01/24 = 11.125$$

$$\bar{y} = 1077.65/24 = 44.902$$

$$\text{SSD}_x = 3089.09 - 267.01^2/24 = 118.4925$$

$$\text{SSD}_y = 51060.4 - 1077.65^2/24 = 2671.670$$

$$\text{SPD}_{xy} = 12511.6 - 267.01 \times 1077.65/24 = 522.2947$$

$$\hat{\beta} = 522.2947/118.4925 = \mathbf{4.4078}$$

$$\hat{\alpha} = 44.902 - 11.125 \times 4.4078 = \mathbf{-4.137}$$

$$\text{SSD}_{\text{res}} = 2671.670 - 4.4078^2 \times 118.4925 = 369.484$$

$$\hat{\sigma}^2 = 369.48/22 = \mathbf{16.79}$$

$$\hat{\sigma} = \sqrt{16.79} = \mathbf{4.10}$$

Standardafvigelsen på $\hat{\beta}$ estimeres ved

$$\sqrt{16.79/118.4925} = 0.376$$

og et 95% sikkerhedsinterval fås derfor, idet 2.074 er 97.5%-fraktilen i t -fordelingen med 22 frihedsgrader, som

$$\beta = 4.4078 \pm 2.074 \times 0.376 = 4.4078 \pm 0.7808$$

svarende til sikkerhedsintervallet [**3.627 , 5.189**] for hældningen.

For afskæringen har vi estimeret

$$\sqrt{16.79 \times \left(\frac{1}{24} + \frac{11.125^2}{118.4925} \right)} = 4.271$$

for standardafvigelsen, hvilket fører til sikkerhedsgrænser

$$\alpha = -4.137 \pm 2.074 \times 4.271 = -4.137 \pm 8.858$$

svarende til sikkerhedsintervallet [**-13.00 , 4.72**] for afskæringen.

(her er det tilstrækkeligt at angive enten estimeret med usikkerhed på formen “ $\pm \dots$ ” eller selve sikkerhedsintervallerne).

Her kommer en typisk 13-tals bemærkning, som viser at man har forstået hvad det hele drejer sig om:

Estimeret $\hat{\beta} = 4.4$ svarer til over 4 nye abonnenter pr. time. Det lyder imponerende!

(b)

Det væsentlige i dette spørgsmål er at man laver en tegning af y' er mod x' er med den estimerede regressionslinje korrekt indtegnet. Det er vigtigt at vælge en fornuftig skalering af akserne. x -aksen bør løbe fra 0 (så afskæringen kan indtegnes) til (ca.) 16. y -aksen kan man lade løbe fra -10 (eller 0, hvis man lader afskæringspunktet ligge lidt under $(0,0)$). Men tegningen skal selvfølgelig også kommenteres, f.eks. således:

Man ser at linien giver en rimelig beskrivelse af punktskyen, hvilket tyder på at vi har regnet rigtigt. Der er ikke antydning af krumning, som kunne tyde på et ikke-lineært forløb, og der er heller ikke noget som tyder på at variansen vokser eller aftager med x .

(c)

Hypotesen om at afskæringen er 0 svarer til at antallet af nytegnede abonnementer følger tid brugt på telefonisk markedsføring proportionalt. Det lyder umiddelbart rimeligt, forudsat at der ikke sideløbende er andre markedsføringsaktiviteter igang, som følger et andet mønster over tiden. Man kan sige at *hvis* der er sådanne aktiviteter igang, vil testet for $\alpha = 0$ kunne fortolkes som et test for om disse andre aktiviteter er uden betydning.

En lidt mere avanceret kommentar: Et andet fænomen, som kunne bevirke at afskæringen ikke kan sættes lig med 0, kunne være at man ved meget intense telefonkampagner nærmer sig et mætningspunkt, hvor man ringer til folk der er perifere i forhold til målgruppen, eller måske endda til de samme flere gange. Hypotesen $\alpha = 0$ kan altså også have andre fortolkninger, lidt afhængigt af oplysninger som ikke findes i opgaveteksten.

Konklusionen af selve testet er at hypotesen $\alpha = 0$ kan godkendes. Dette følger af at 95% sikkerhedsintervallet for α indeholder 0, hvilket jo er ækvivalent med at denne hypotese godkendes på niveau 95%.

Det er således ikke nødvendig direkte at regne teststørrelsen ud, men hvis man gør det alligevel får man

$$t = \frac{-4.137 - 0}{4.271} = -0.969$$

som skal vurderes tosidet i en t -fordeling med 22 frihedsgrader. Altså igen, hypotesen godkendes.