

Veiledningsoppgaver

Oppgave 1.

Tabellen nedenfor viser antall abonnementer solgt og antall kaffekopper drukket i løpet av en dag for syv tilfeldig valgte selgere:

Kaffekopper drukket:	2	1	4	3	2	0	4
Abonnement solgt:	4	12	5	2	20	21	4

- Lag et spredningsdiagram for datasettet, og finn korrelasjonskoeffisienten. Estimér regresjonslinjen med antall kaffekopper drukket som forklaringsvariabel og antall abonnement solgt som responsvariabel, og tegn den inn i spredningsdiagrammet.
- Finn forventet antall abonnementer solgt for en tilfeldig valgt selger som har drukket 2 kaffekopper.
- Hva er den kvadratiske feilfunksjonen SS_E ? Gi definisjonen, regn ut SS_E , og gi en tolkning. Bruk SS_E til å estimere standardfeilen til $\hat{\beta}$, estimatoren for helningen til regresjonslinjen.
- Finn et 95% konfidensintervall for helningen β til regresjonslinjen.
- Undersøk om det er sammenheng mellom antall kaffekopper drukket og antall abonnement solgt. Du kan bruke signifikansnivå 5%.
- Hvilke forutsetninger må vi gjøre for at resonnementene ovenfor skal holde?

Oppgave 2.

Tabellen nedenfor viser andelen skilsmisser (skilsmisser per 1000 bestående ekteskap) for utvalgte år:

År:	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010
Skilsmisser:	2.9	3.4	4.9	6.5	7.9	9.5	11.7	10.9	12.6	11.6

- Estimér regresjonslinjen, og undersøk om skilsmisseraten har økt ved å utføre en hypotesetest. Du kan bruke at $SS_E = 7.870$.
- Lag et spredningsdiagram med regresjonslinjen tegnet inn. Hvordan vil du beskrive utviklingen i skilsmisseraten? Hvilken skilsmisserate vil du forvente i 2025?

Oppgave 3.

Bruk dataene for skilsmisseraten i perioden 1965-1995 i Oppgave 2, og finn et 92% konfidensintervall for helningen β til regresjonslinjen.

Oppgave 4.

Vis at minimum for den kvadratiske feilfunksjonen

$$f(\alpha, \beta) = \varepsilon_1^2 + \dots + \varepsilon_n^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$$

er gitt ved

$$\hat{\beta} = \frac{S_{XY}}{S_X^2} = R \cdot \frac{S_Y}{S_X}, \quad \hat{\alpha} = \bar{Y} - \hat{\beta} \cdot \bar{X}$$

Oppgave 5. Eksamen MET1190 11/2018

Tabellen nedenfor viser antall timer forberedelser i løpet av semesteret og antall poeng på statistikkeksamen for fem tilfeldig valgte studenter:

Antall timer forberedelse:	150	170	200	170	100
Poeng på eksamen:	45	67	99	35	46

- Beregn gjennomsnitt, median og modus for antall timer forberedelser.
- Beregn og tolk varianser og kovarianser fra utvalget. Regn om til andre sprednings- og samvariasjonsmål hvis det er nyttig for å tolke.
- Estimer regresjonlikningen med antall poeng som responsvariabel og antall timer forberedelser som forklaringsvariabel. Illustrer dataene i et spredningsdiagram og tegn inn den estimerte regresjonslinjen.
- Tolk de estimerte koeffisientene. Beregn og tolk regresjonens R^2 .
- Forklar hva et 95% konfidensintervall er, og finn et 95% konfidensintervall for helningen β til regresjonslinjen.

Oppgave 6.

Oppgaver fra læreboken [L]: 7.1, 7.5, 7.8, 7.10, 7.11

Svar på veiledningsoppgaver

Oppgave 1.

- $r = -0.743$, $\hat{\alpha} = 18.8$, $\hat{\beta} = -3.98$
- $\hat{y} = 10.85$ når $x = 2$
- $SS_E = 172.85$, $SE(\hat{\beta}) = 1.604$
- $-3.98 \pm 2.57 \cdot 1.604 = [-8.10, 0.15]$
- Nullhypotese $\beta = 0$, forkastningsområde $|T| > 2.57$, datasettet gir $T = -2.48$, ingen sammenheng
- $\varepsilon_1, \dots, \varepsilon_n$ er uavhengige og normalfordelte med $\varepsilon_i \sim N(0, \sigma^2)$ for en konstant σ

Oppgave 2.

- $\hat{\alpha} = -449$, $\hat{\beta} = 0.230$, raten har økt på signifikansnivå 5% siden $T = 10.54$ er i forkastningsområdet $T > 1.86$.
- Skilsmisseraten ser ut til å ha økt i 1965-1995, og flatet ut i 1995-2010, og basert på estimert regresjonslinje fra dataene i den siste perioden anslår vi skilsmisseraten til 12.3 i 2025

Oppgave 3.

$$0.297 \pm 2.191 \cdot 0.0176 = 0.297 \pm 0.039 = [0.259, 0.336]$$