Løsning hjemmeregning 11

Oppgave 1

Gitt at X og Y har punktfordelinger som i tabellen:

x \ y	1	3	5
2	0,10	0,20	0,10
4	0,15	0,30	0,15

a) Skal finne punktfordelingene for hhv. X og Y. Gjør dette ved å summere tallene i kolonnene:

x/y	**1**	**3**	**5**	$P(X=x)$
2	0.10	0.20	0.10	0.40
4	0.15	0.30	0.15	0.60
$P(Y=y)$	0.25	0.50	0.25	

b) Finner forventningsverdiene for X og Y:

$$\begin{aligned}\mu_X &= E(X) = \sum_{\text{alle } x} xP(X=x) \\ &= 2\cdot P(X=2) + 4\cdot P(X=3) \\ &= 2\cdot 0.40 + 4\cdot 0.60 \\ &= \underline{\underline{3.2}}\end{aligned}$$

$$\begin{aligned}\mu_Y &= E(Y) = \sum_{\text{alle } y} yP(Y=y) \\ &= 1\cdot P(Y=1) + 3\cdot P(Y=3) + 5\cdot P(Y=5) \\ &= 1\cdot 0.25 + 3\cdot 0.50 + 5\cdot 0.25 \\ &= \underline{\underline{3.0}}\end{aligned}$$

Finner variansene for X og Y:

$$\begin{aligned}Var(X) &= \sum_{\text{alle } x} (x-\mu_X)^2 P(X=x) \\ &= \sum_{\text{alle } x} x^2 P(X=x) - \mu_X^2 \text{ (fra formelark)} \\ &= 2^2\cdot P(X=2) + 4^2\cdot P(X=3) - 3.2^2 \\ &= 2^2\cdot 0.40 + 4^2\cdot 0.60 - 3.2^2 \\ &= \underline{\underline{0.96}}\end{aligned}$$

$$
\begin{aligned}
Var\,(Y) &= \sum_{\text{alle } y} (y-\mu_Y)^2\, P\,(Y=y) \\
&= \sum_{\text{alle } y} y^2 P\,(Y=y) - \mu_Y^2 \text{ (fra formelark)} \\
&= 1^2 \cdot P\,(Y=1) + 3^2 \cdot P\,(Y=3) + 5^2 \cdot P\,(Y=5) - 3.0^2 \\
&= 1^2 \cdot 0.25 + 3^2 \cdot 0.50 + 5^2 \cdot 0.25 - 3.0^2 \\
&= \underline{\underline{2.0}}
\end{aligned}
$$

c) Finner kovariansen mellom X og Y:

$$
\begin{aligned}
Cov\,(X,Y) &= E\,(XY) - \mu_x\mu_y \text{ (fra formelark)} \\
&= \sum_{\text{alle } x}\sum_{\text{alle } y} xyP\,(X=x\cap Y=y) - \mu_x\mu_y \\
&= 2\cdot 1\cdot P\,(X=2\cap Y=3) + 2\cdot 3\cdot P\,(X=2\cap Y=3) + 2\cdot 5\cdot P\,(X=2\cap Y=5) \\
&\quad +4\cdot 1\cdot P\,(X=4\cap Y=1) + 4\cdot 3\cdot P\,(X=4\cap Y=3) + 4\cdot 5\cdot P\,(X=4\cap Y=5) - \mu_x\mu_y \\
&= 2\cdot 1\cdot 0.10 + 2\cdot 3\cdot 0.20 + 2\cdot 5\cdot 0.10 \\
&\quad +4\cdot 1\cdot 0.15 + 4\cdot 3\cdot 0.30 + 4\cdot 5\cdot 0.15 - 3.2\cdot 3.0 \\
&= \underline{\underline{0}}
\end{aligned}
$$

Oppgave 2

a) Vi skal estimere en regresjonslinje

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

for utslippsmengden Y som funksjon av år x. Regresjonsparametrene er gitt ved

$$
\begin{aligned}
\hat{\beta}_1 &= \frac{1}{M}\sum_{i=1}^{n}(x_i-\bar{x})\,Y_i, \text{ der } M = \sum_{i=1}^{n}(x_i-\bar{x})^2 \\
\hat{\beta}_0 &= \bar{Y} - \hat{\beta}_1\bar{x}
\end{aligned}
$$

Dersom vi bruker de ferdige utregningene, får vi:

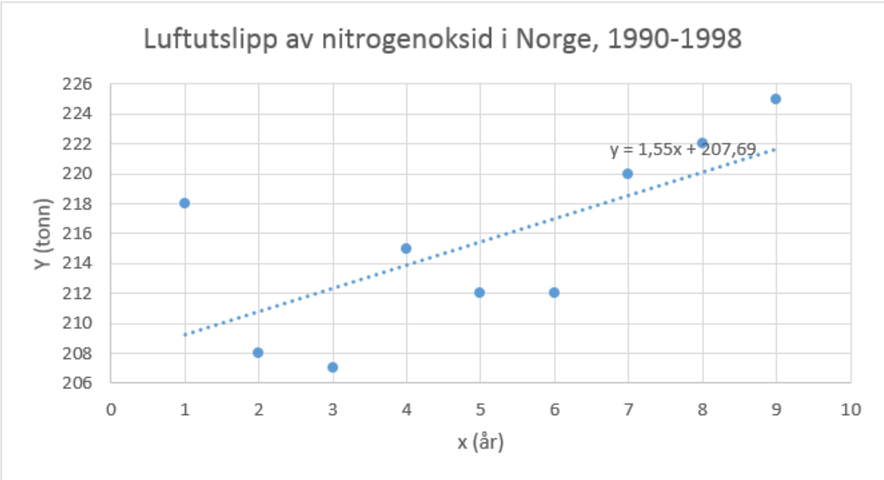
$$
\begin{aligned}
\hat{\beta}_1 &= \frac{1}{M}\sum_{i=1}^{9}(x_i-\bar{x})\,Y_i \\
&= \frac{1}{60}\cdot 93 \\
&= \underline{1.55}
\end{aligned}
$$

$$\begin{aligned}\hat{\beta}_0 &= \bar{Y} - \hat{\beta}_1 \bar{x} \\ &= 215.44 - 1.55 \cdot 5.0 \\ &= \underline{207.7}\end{aligned}$$

Likninga for den estimerte regresjonslinja blir altså:

$$\begin{aligned}\hat{Y} &= \hat{\beta}_0 + \hat{\beta}_1 x \\ &= \underline{\underline{207.7 + 1.55x}}\end{aligned}$$

Tegner inn observasjonene og den estimerte regresjonslinja i samme spredningsdiagram:



b) Vi skal formulere en hypotesetest for å undersøke om utslippene av nitrogenoksid har økt i perioden 1990-1998. Ettersom det er parameteren β_1 som angir økning per år, (dsv. antall tonn økning per år), vil hypotesetesten her bli:

$$\begin{aligned}H_0 &: \hat{\beta}_1 = 0 \\ H_0 &: \hat{\beta}_1 > 0\end{aligned}$$

Bruker testobservatoren

$$U_0 = \frac{\hat{\beta}_1 - \beta_1^0}{\frac{\sigma}{\sqrt{M}}},$$

der $\beta_1^0 = 0$ (den verdien av parameteren som påstås i nullhypotesen H_0). Vi får her:

$$\begin{aligned} U_0 &= \frac{\hat{\beta}_1 - \beta_1^0}{\frac{\sigma}{\sqrt{M}}} \\ &= \frac{1.55 - 0}{\frac{5.0}{\sqrt{60}}} \\ &= \underline{\underline{2.40}} \end{aligned}$$

Ettersom vi bruker et signifikansnivå på 1 %, sammenlikner vi testobservatoren med kvantilen $u_\alpha = u_{0.01} = 2.326$. Vi ser at $U_0 > u_\alpha$, altså konkluderer vi med at det **har vært en økning** i utslippene i perioden 1990-1998, ved et signifikansnivå på 1 %.