(Det tas forbehold om feil i løsningsforslaget.)

Oppgave 1

La O_1 = finne olje i reservoar 1, og O_2 = finne olje i reservoar 2. Vi har da: $P(O_1) = 0.5$, $P(O_2) = 0.6$, $P(\text{finne olje i begge reservoarene}) = P(O_1 \cap O_2) = 0.2$.

a) Finne olje i minst ett av reservoarene $= O_1 \cup O_2$;
$$P(O_1 \cup O_2) = P(O_1) + P(O_2) - P(O_1 \cap O_2) = 0.5 + 0.6 - 0.2 = 0.9.$$
Ikke finne olje i noen av de to $= (O_1 \cup O_2)^C$;
$$P((O_1 \cup O_2)^C) = 1 - P(O_1 \cup O_2) = 1 - 0.9 = 0.1.$$

b) Finne olje i kun reservoar 1 $= O_1 \cap O_2^C$;
$$P(O_1 \cap O_2^C) = P(O_1) - P(O_1 \cap O_2) = 0.5 - 0.2 = 0.3$$ — bruk Venndiagram !
Finne olje i kun reservoar 1 eller kun reservoar 2 $= (O_1 \cap O_2^C) \cup (O_1^C \cap O_2)$;
$$P\{(O_1 \cap O_2^C) \cup (O_1^C \cap O_2)\} = P(O_1) + P(O_2) - 2P(O_1 \cap O_2) = 0.5 + 0.6 - 2 \cdot 0.2 = 0.7$$
— bruk Venndiagram!

c) $P(O_2|O_1) = \dfrac{P(O_2 \cap O_1)}{P(O_1)} = \dfrac{0.2}{0.5} = 0.4.$
Siden $P(O_2|O_1) = 0.4 \neq P(O_2) = 0.6$, ser vi at begivenhetene O_1 O_2 ikke er uavhengige.

Oppgave 2

a*) La X være oljekonsentrasjonen i et tilfeldig valgt utslipp.
Da: $X \sim \mathcal{N}(21.6,\ 3.4)$.

$$\begin{aligned} P(X > 40) &= P(\frac{X - 21.6}{3.4} > \frac{40 - 21.6}{3.4}) \\ &= P(Z > 5.41) = 1 - P(Z < 5.41) \approx 1 - 1 = 0 \\ P(X < 25) &= P(\frac{X - 21.6}{3.4} < \frac{25 - 21.6}{3.4}) \\ &= P(Z < 1.0) = 0.8413 \end{aligned}$$

(Her er $Z \sim \mathcal{N}(1, 0)$.)

b) $P(X \in [\mu - 3\sigma, \mu + 3\sigma]) = P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = P(-3 \leq \frac{X-\mu}{\sigma} \leq 3) = P(-3 \leq Z \leq 3) = P(Z \leq 3) - P(Z \leq -3) = 0.9987 - 0.0013 = 0.9974$

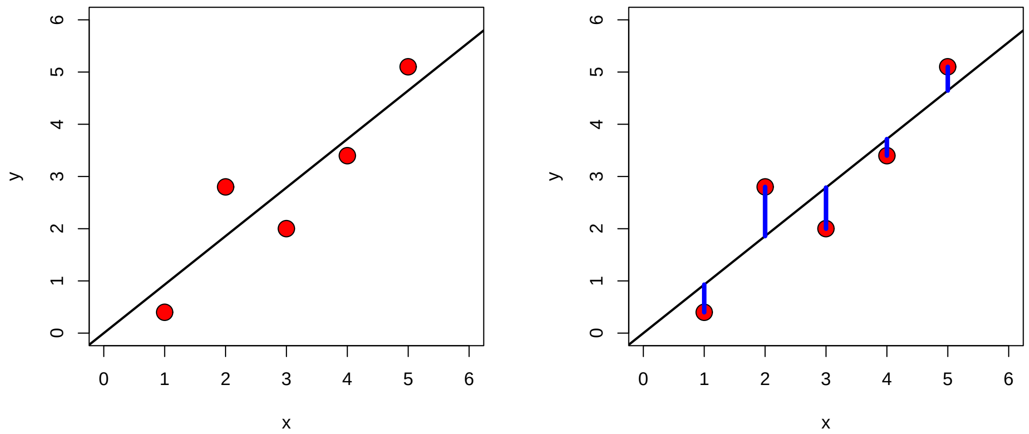
c*) Vi har 16 målinger: $x_1, \ldots, x_{16}$ med gjennomsnitt, $\overline{x} = 19.8$. Et 95 % konfidensintervall for μ er gitt ved: $\left(\overline{X} - z_{0.025}\sqrt{3.4^2/16}, \overline{X} + z_{0.025}\sqrt{3.4^2/16}\right)$. Innsatt data:
$\left(19.8 - 1.96\sqrt{3.4^2/16}, 19.8 + 1.96\sqrt{3.4^2/16}\right) = \left(18.13, 21.47\right)$.

d) Vi vil teste $H_0: \mu = 21.6$ mot $H_1: \mu < 21.6$ på signifikansnivå 0.01. Teststørrelse: $Z = \frac{\overline{X}-21.6}{\sqrt{\frac{3.4^2}{16}}}$.
Test: Forkast H_0 dersom $Z \leq -z_{0.01} = -2.326$
Data: Observert verdi av Z: $\frac{19.8-21.6}{\sqrt{\frac{3.4^2}{16}}} = -2.12$; dvs. observert verdi av Z er ikke i forkastningsområdet; konklusjon: behold H_0 (den nye metoden er ikke bedre enn den gamle).

e) Vi må bruke t-test. Teststørrelse: $T = \frac{\overline{X}-21.6}{\sqrt{\frac{S^2}{16}}}$, der $S^2 = \frac{1}{15}\sum_{i=1}^{16}(X_i - \overline{X})^2$ (og med dataene beregnet til $s^2 = 1.9^2$).
Test: Forkast H_0 dersom $T \leq -t_{0.01,15} = -2.602$
Data: Observert verdi av T: $\frac{19.8-21.6}{\sqrt{\frac{1.9^2}{16}}} = -3.79$; dvs. observert verdi av T er i forkastningsområdet; konklusjon: forkast H_0 (den nye metoden er bedre enn den gamle).

Oppgave 3

a) Spredningsdiagram med inntegnet regresjonslinje:



Når regresjonslinjen skal tegnes inn "på øyemål", må vi prøve å velge en linje som gjør avvikene mellom linje og datapunktene (indikert i figur til høyre) samlet minst mulig.

b) Data: $(x_1, y_1), \ldots, (x_n, y_n)$; modell: $Y_i = \beta x_i + \epsilon_i, i = 1, \ldots, n$ ($\epsilon_1, \ldots, \epsilon_n$: uif. $N(0, \sigma^2)$).

Parameteren β kan estimeres ved å finne den verdien av b som minimerer de kvadrerte avvikene mellom måling, y_i, og tilsvarende verdi på regresjonslinje, bx_i. Dvs. vi definerer SSE(b) ved:

$$\mathrm{SSE}(b) = \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 = \sum_{i=1}^{n}(y_i - bx_i)^2, \; (\widehat{y}_i = bx_i)$$

Velg stigningstallet b slik at SSE blir minimert.
Betrakt SSE som en funksjon av b, deriverer funksjonen SSE(b) (mht. b), og sett den deriverte lik null.

$$\frac{d}{db}\mathrm{SSE}(b) = \sum_{i=1}^{n} \frac{d}{db}\{y_i - bx_i\}^2 = -2\sum_{i=1}^{n}\{y_i - bx_i\}x_i == -2\Big\{\sum_{i=1}^{n} y_i x_i - b\sum_{i=1}^{n} x_i^2\Big\} = 0$$

$$\Rightarrow b = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2}.$$

Derfor er minstekvadratersestimatoren for β: $\widehat{\beta} = \dfrac{\sum_{i=1}^{n} Y_i x_i}{\sum_{i=1}^{n} x_i^2}$

c) Estimat: $\dfrac{0.4 + 2\cdot 2.8 + 3\cdot 2.0 + 4\cdot 3.4 + 5\cdot 5.1}{1 + 2^2 + 3^2 + 4^2 + 5^2} = \dfrac{51.1}{55} = 0.929.$

Forventing til estimator:

$$\mathrm{E}(\widehat{\beta}) = \mathrm{E}\Big(\frac{\sum_{i=1}^{n} Y_i x_i}{\sum_{i=1}^{n} x_i^2}\Big) = \frac{\sum_{i=1}^{n} \mathrm{E}(Y_i) x_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} \overbrace{\beta x_i}^{E(Y_i)}\, x_i}{\sum_{i=1}^{n} x_i^2} = \beta\frac{\sum_{i=1}^{n} x_i x_i}{\sum_{i=1}^{n} x_i^2} = \beta,$$

dvs. estimatoren $\widehat{\beta}$ er forventingsrett for β.