Løsningsforslag til eksamen ISTx100y høst 2023

Oppgave 1

a) $P(X \geq 2) = P(X = 2) + P(X = 3) = 0.1 + 0.4 = 0.5$

b) $\text{Var}(X - Y) = (1)^2 \cdot \text{Var}(X) + (-1)^2 \text{Var}(Y) + 2 \cdot (-1) \cdot (1) \cdot \text{Covar}(X, Y) = 1 + 2 - 2 \cdot 0.5 = 2$

c) $P(T \leq 1) = 1 - e^{-2^{0.4}} \approx 0.73$

d) $P(-1.645 \leq Z \leq 1.645) = P(Z \leq 1.645) - P(Z \leq -1.645) = 0.95 - 0.05 = 0.90$

e) Vi har $\hat{\lambda} = \frac{12}{2} = 6$, $t = 2$, $z_{\alpha/2} = 1.96$, og $\hat{\lambda}/t = 3$ og dermed er intervallet gitt ved $\left[6 - 1.96\sqrt{3}, 6 - 1.96\sqrt{3}\right] \approx [2.61, 9.39]$

f) Teststyrken er i dette tilfellet sannsynligheten for at $\bar{X} < 0.35$ der $\bar{X} = \frac{X_1 + X_2 + \cdots + X_{16}}{16}$ og $X_i \sim N(1.5, 4)$. I utregning av teststyrke skal vi altså anta at H_1 er sann slik at forventningsverdien er $\mu = 1.5$ og standardavviket er $\sigma = 4$. Vi vet f.eks. fra regneregler for normalfordelingen at $\bar{X} \sim N(\mu, \sigma/\sqrt{n})$. Merk at $\sigma/\sqrt{n} = 4/\sqrt{16} = 4/4 = 1$ som forenkler utregningene. Vi har altså at $\bar{X} \sim N(0.5, 1)$ og skal finne $P(\bar{X} < 0.35)$:

$$P(\bar{X} < 0.35) = P(\frac{\bar{X} - 1.5}{1} < \frac{0.35 - 1.5}{1}) = P(Z < -1.15) = 0.1251.$$

Her har vi brukt tabell for standard normalfordeling for å komme frem til svaret.

g) På formelarket står det hvordan vi kan tilpasse en lineær regresjonsmodell basert på koden smf.ols('y ~ x',data).fit() og da brukes standard notasjon med y som respons og x som forklaringsvariabel. I denne oppgaven blir regresjonsmodellen som estimeres fra dataene $w = \beta_0 + \beta_1 v + e$, altså er w responsen og v er forklaringsvariabelen (kovariaten).

Oppgave 2: Hendelser og sannsynlighet

a) Området viser hendelsen C, og siden C er en delmengde av B så er det riktige svaret $B \cap C$.

b) Vi regner først ut $P(A \cup B) = P(A) + P(B) - P(A \cap C) = 0.4 + 0.3 - 0.1 = 0.6$, og deretter finner vi at komplementærhendelsen er $P((A \cup B)^c)) = 1 - P(A \cup B) = 1 - 0.6 = 0.4$.

c) Vi skal trekke 7 personer fra 9, men vi bryr oss ikke om rekkefølgen. Vi har dermed en urnemodell der vi ser på et ikke-ordnet utvalg, og trekning uten tilbakelegging. Da er antall mulige lag-oppstillinger gitt ved $\binom{9}{7} = \frac{9!}{7!(9-7)!} = \frac{9!}{7! \cdot 2} = \frac{9 \cdot 8}{2} = 9 \cdot 4 = 36$.

d) Vi vet at det finnes 36 mulige lag-oppsett. Vi er ute etter å finne sannsynligheten for å trekke et lagoppsett som inneholder de to uvennene. Det finnes mange strategier man kan bruke for å telle antall gunstige utfall, og her er to av dem.

Forslag 1: Et lag-oppsett som *ikke* kan brukes må inneholde de to uvennene. Da er det 7 andre barn igjen på laget og 5 av disse skal være i lag-oppsettet. Det finnes $\binom{7}{5} = \frac{7!}{5!(7-5)!} = \frac{7!}{5! \cdot 2} = \frac{7 \cdot 6}{2} = 7 \cdot 3 = 21$ måter å trekke disse på. Dermed er det 21 "gunstige" utfall. Sannsynligheten for å trekke et slikt lagoppsett er $21/36 \approx 0.58$.

Forslag 2: I dette forslaget teller vi først lagoppsett som faktisk fungerer (inneholder ikke uvennene). Siden det totalt er 9 barn, er det bare ett lagoppsett som *ikke* inneholder noen av de to uvennenene (de to sitter sammen på benken, men det må vi anta at går bra). Videre finnes det 7 mulige lagoppsett som bare inneholder uvenn nr 1, og så finnes det 7 mulige lagoppsett som bare inneholder uvenn nr 2. Totalt er det dermed $1 + 7 + 7 = 15$ utfall som går bra, og $36 - 15 = 21$ lagoppsett som ikke vil fungere. Sannsynlighen for at vi har trukket et lagoppsett med de to uvennenne er derfor $21/36 \approx 0.58$.

Oppgave 3: Aluminiumsplater

a) Forventet vekt av pappesken med ti plater er

$$\mathrm{E}(X_1 + X_2 + \cdots + X_{10} + 50) = \mathrm{E}(X_1) + \mathrm{E}(X_2) + \cdots + \mathrm{E}(X_{10}) + 50 = 10 \cdot 100 + 50 = 1050.$$

b) Vi regner først på varians og finner at variansen til pappesken med ti (uavhengige) plater er

$$\mathrm{Var}(X_1 + X_2 + \cdots + X_{10} + 50) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_{10}) = 10 \cdot (0.8)^2.$$

Dermed er standardavviket $\sqrt{10} \cdot 0.8 = 2.5298 \approx 2.53$

c) Antallet defekte plater i en eske med ti plater vil være binomisk fordelt med parametere $n = 10$ og $p = 0.21$. Sannsynligheten $p = 0.21$ kommer fra sannsynligheten for at en plate er defekt:

La X betegne vekten til en tilfeldig valgt aluminiumsplate. Vi vet at $X \sim N(100, 0.8)$ og vi vet at vi aksepterer vekt i intervallet [99, 101]. Vi finner først sannsynligheten for at X tar en verdi innenfor intervallet:

$$P(99 \leq X \leq 101) = P\left(\frac{99 - 100}{0.8} \leq Z \leq \frac{101 - 100}{0.8}\right) = P(-1.25 \leq Z \leq 1.25) =$$

$$P(Z \leq 1.25) - P(Z \leq -1.25) = 0.8944 - 0.1056 = 0.7888$$

Deretter finner vi sannsynligheten for at vekten er *utenfor* det aksepterte intervallet ved

$$1 - 0.7888 = 0.2112 \approx 0.21.$$

Videre vet vi at $P(X \geq 2) = 1 - P(X < 2) = 1 - (P(X = 0) + P(X = 1))$. Vi regner ut $P(X = 0) = (1 - 0.21)^{10} = 0.0947$ og $P(X = 1) = 10 \cdot 0.21 \cdot (1 - 0.21)^9 = 0.2517$. Dermed er svaret $1 - 0.0947 - 0.2517 \approx 0.65$

d) Vi regner ut gjennomsnittlig vekt $\bar{x} = 2512/25 = 100.48$ og testobservatoren $z = \frac{100.48-100}{0.8/\sqrt{25}} = 3$. Den kritiske verdien for testen er $z_{0.001} = 3.090$. Siden $z < z_{0.001}$ kan vi ikke forkaste H_0 til fordel for H_1 ved signifikansnivå 0.001.

Oppgave 4: Kobbertråd

a) $E(X) = \lambda s = 0.2 \cdot 3 = 0.6$ svakheter/millimeter

b) $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{0.6^0}{0!}e^{-0.6} + \frac{0.6^1}{1!}e^{-0.6} + \frac{0.6^2}{2!}e^{-0.6} = (1 + 0.6 + \frac{0.6^2}{2})e^{-0.6} = 0.9768847 \approx 0.98.$

(Her kan må sjekke om svaret virker rimelig ved å slå opp i tabell for kumulative sannsynligheter i Poissonfordelingen. Forventning 0.6 står ikke der, men ved $\lambda t = 0.5$ ser vi at $P(X_t \leq 2) = 0.9856$ som er veldig nære vårt svar.)

c) S vil ha en eksponensialfordeling, og dermed er $E(S) = 1/\lambda = 1/0.2 = 5$ millimeter.

d) Siden eksponensialfordelingen er "hukommelsesløs" kan vi regne ut $P(S \leq 1)$ uten å ta hensyn til at vi allerede har 2 mm feilfri kobbertråd. Vi finner $P(S \leq 1) = 1 - e^{-\lambda} = 1 - e^{-0.2} = 0.1812692 \approx 0.18$

Oppgave 5: Lineær regresjon

a) Hypotesetesten som utføres er om stigningstallet β_1 er signifikant ulikt fra 0, altså $H_0 : \beta_1 = 0$ mot $H_1 : \beta_1 \neq 0$.

b) Fra modellen finner vi prediksjonen $\hat{y} = -8.4156 + 4.4416 \cdot 43 = 182.5732 \approx 182.6$

c) Prediksjonsintervallet kan vi lese av fra figuren for skostørrelse 40 og vi ser at [162 cm, 177 cm] må være det riktige svaret.