

Løsningsforslag PrøveEksamen 1 i ISTx100y

Oppgave 1

Merk: Denne oppgaven er basert på Oppgave 1 på eksamen i TALM1005, kont 2018.

a) Hytta har vanntilførsel dersom en (eller) begge pumpene fungerer, altså $A \cup B$. Dermed blir sannsynligheten for hytta, en tilfeldig dag, har vanntilførsel gitt ved

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Sannsynligheten for snittet av A og B er ikke gitt i oppgaveteksten, men siden A og B er uavhengige hendelser er $P(A \cap B) = P(A) \cdot P(B)$. Dermed finner vi

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A) \cdot P(B) \\ &= 0.6 + 0.5 - 0.6 \cdot 0.5 \\ &= 0.8 \end{aligned}$$

b) Vi antar at X er binomisk fordelt med parametere $n = 7$ og $p = 0.8$. Vi skal regne ut sannsynligheten for at det *ikke* er vanntilførsel til hytte på minst én av de 7 dagene, altså sannsynligheten for minst én fiasko:

$$\begin{aligned} P(\text{minst en fiasko}) &= 1 - P(\text{ingen fiasko}) \\ &= 1 - P(X = 7) \\ &= 1 - \binom{7}{7} p^7 (1-p)^0 \\ &= 1 - p^7 \\ &= 0.7902848 \approx 0.79 \end{aligned}$$

En alternativ fremgangsmåte er å regne ut $P(X \leq 6)$ som vi fra tabell for kumulative sannsynligheter i binomisk fordeling ($n = 7$, $x = 6$, $p = 0.8$) kan lese av til å være 0.790.

c) Vi ønsker at sjansen for at minst én av pumpene fungerer skal være lik 0.99. Med andre ord ønsker vi $P(A \cup B \cup C) = 0.99$. Hva må da $P(C)$ være?. Det enkleste her er å se på komplementærhendelsen, nemlig at ingen av pumpene virker, som skal inntreffe med sannsynlighet 0.01. Altså, $P(A' \cap B' \cap C') = 0.01$. Siden hendelsene A' , B' og C' er uavhenige, så er

$$\begin{aligned} P(A' \cap B' \cap C') &= P(A') \cdot P(B') \cdot P(C') \\ &= (1 - 0.6) \cdot (1 - 0.5) \cdot (1 - P(C)) \\ &= 0.2 \cdot (1 - P(C)) \\ &= 0.01 \end{aligned}$$

Dermed kan vi løse for $P(C)$ og finner

$$\begin{aligned} 1 - P(C) &= 0.01/0.2 \\ \Downarrow \\ P(C) &= 1 - 0.01/0.2 = 0.95 \end{aligned}$$

c) Vi tenker oss at hyttene er nummerert som 1, 2, ..., 7. For den første hytta elektrikerens skal besøke har han 7 valg, for hver av disse har han 6 valg for andre hytta han skal besøke, for hver av disse har han 5 valg for den tredje hytta, osv. Det vil si at antall rekkefølger er gitt ved

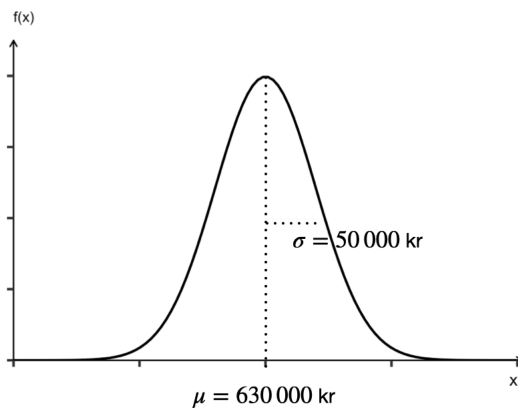
$$7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 7! = 5040$$

Oppgave 2

Merk: Denne oppgaven er basert på Oppgave 2 på eksamen i TALM1005, høst 2018.

a) X er en normalfordelt stokastisk variabel med forventning $\mu = 630\,000$ og standardavvik $\sigma = 50\,000$.

i. Sannsynlighetstettheten til X ser da omtrent slik ut:



ii. Vi skal regne ut $P(X > 640\,000)$ ved å gjøre om problemet til et problem i standard normalfordeling.

$$\begin{aligned} P(X > 640\,000) &= 1 - P(X \leq 640\,000) \\ &= 1 - P\left(\frac{X - 630\,000}{50\,000} \leq \frac{640\,000 - 630\,000}{50\,000}\right) \\ &= 1 - P(Z \leq 0.2) \end{aligned}$$

I tabell for standard normalfordeling (eller ved å bruke programvare) kan vi slå opp $P(Z \leq 0.2) = \Phi(0.2) = 0.5793$. Dermed blir

$$P(X > 640\,000) = 1 - 0.5793 = 0.4207$$

b) Vi skal finne en kritisk verdi k slik at $P(X > k) = 0.05$. Ved å gjøre om problemet til en standard normalfordeling finner vi

$$P\left(\frac{X - 630\,000}{50\,000} > \frac{k - 630\,000}{50\,000}\right) = 0.05,$$

og fra kritiske verdier i standard normalfordelingen vet vi at

$$P(Z > 1.645) = 0.05$$

slik at vi må ha $\frac{k-630\,000}{50\,000} = 1.645$. Dermed kan vi løse for k , og finner

$$\begin{aligned}\frac{k - 630\,000}{50\,000} &= 1.645 \\ \Downarrow \\ k - 630\,000 &= 1.645 \cdot 50\,000 \\ \Downarrow \\ k &= 630\,000 + 1.645 \cdot 50\,000 \\ &= 712\,250\end{aligned}$$

For å være blant de 5% med høyest inntekt må man altså tjene minst 712 250 kr i året.

c) Vi skal se på den samlede inntekten til 4 ingeniører. Siden disse møtes tilfeldig på bussen så antar vi at de ikke har noe med hverandre å gjøre og at deres respektive inntekter kan ansees som uavhengige. Vi definerer total inntekt som $Y = X_1 + X_2 + X_3 + X_4$, der X_1, X_2, X_3 og X_4 er uavhengige normalfordelte stokastiske variabler som alle har forventningsverdi 630 000 og standardavvik 50 000. Da er $E(Y) = 4 \cdot 630\,000 = 2\,520\,000$ og $\text{Var}(Y) = 4 \cdot 50\,000^2$, slik at $\text{SD}(Y) = \sqrt{4} \cdot 50\,000 = 100\,000$. Her skal faktisk finne sannsynligheten for at Y faller innenfor $\pm$ ett standardavvik fra forventningsverdien, så vi kan umiddelbart skrive ned at

$$P(2\,420\,000 < Y \leq 2\,620\,000) \approx 0.68$$

Vi kan også regne ut dette ved

$$\begin{aligned}P(2\,420\,000 < Y \leq 2\,620\,000) &= P\left(\frac{2\,420\,000 - 2\,520\,000}{100\,000} < Y \leq \frac{2\,620\,000 - 2\,520\,000}{100\,000}\right) \\ &= P(-1 < Z \leq 1) \\ &= P(Z \leq 1) - P(Z \leq -1) \\ &= 0.8413 - 0.1587 = 0.6826\end{aligned}$$

d) Vi ser på et tilfeldig utvalg $X_1, X_2, \dots, X_{50}$ av uavhengige normalfordelte variabler med forventning $E(X_i) = 630\,000$ og standardavvik $\sigma = 50\,000$. Da er

$$\bar{X} \sim N\left(630\,000, \frac{50\,000}{\sqrt{50}}\right)$$

og vi kan regne ut

$$\begin{aligned}P(\bar{X} > 640\,000) &= 1 - P(\bar{X} \leq 640\,000) \\ &= 1 - P\left(\frac{\bar{X} - 630\,000}{50\,000/\sqrt{50}} \leq \frac{640\,000 - 630\,000}{50\,000/\sqrt{50}}\right) \\ &= 1 - P(Z \leq 1.414214) \approx 1 - P(Z \leq 1.41) = 1 - 0.9207 = 0.0793\end{aligned}$$

Oppgave 3

Merk: Denne oppgaven er basert på Oppgave 5 på eksamen i TALM1005, høst 2018.

a) Vi skal regne ut et 95% konfidensintervall for forventningsverdien μ i en normalfordeling med kjent standardavvik $\sigma = 10$ g/km. Vi har $n = 6$ målinger og et gjennomsnitt av disse på $\bar{x} = 118.5$ g/km. For intervallet trenger vi kritisk verdi $z_{\alpha/2} = z_{0.05/2} = z_{0.025} = 1.960$. Da kan vi sette inn i formel og regne ut

$$\left[\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right] = [110.4983, 126.5017] \approx [110.50, 126.50]$$

b) Dersom vi *veldig mange ganger* velger ut 6 biler, måler gjennomsnittlig utslipp og regner ut et konfidensintervall for μ , så vil 95% av disse (veldig mange) intervallene inneholde/dekke den faktiske parameterverdien. Når vi gjør 6 uavhengige målinger av utslipp er det altså 95% sjanse for at vi ender opp med et numerisk intervall der nedre grense i intervallet faller til venstre for μ på tallinja, samtidig som øvre grense faller til høyre for μ .

Oppgave 4

a) Vi ser en positiv lineær trend slik at høye sanne verdier for gjennomstrømning (x) korresponderer med høye sensormålinger av gjennomstrømning (y). Dette tyder på en positiv korrelasjon mellom x og y .

b) I figuren ser vi at punktene ligger rundt hjelpelinja $y = x$. Dette tyder på et krysningspunkt β_0 ("beta0") omtrent lik 0, og et stigningstall β_1 ("beta1") omtrent lik 1.

c) Vi ser en lineær sammenheng mellom x og y , men vi ser at variansen i sensormålingene (y) blir større jo større den sanne gjennomstrømningen (x) er. Det bryter med antagelsen om at $\text{Var}(Y|x) = \sigma^2$ uansett verdi av x .