

Oppgave 1

1 a) i. Vi skal finne $P(C) = P(A \cap B)$. Vi bruker addisjonsregelen, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, og kan regne ut

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

Da finner vi

$$P(C) = 0.80 + 0.45 - 0.89 = 0.36$$

1 a) ii. Hendelsene A og B er ikke disjunkte siden $P(A \cap B) = 0.36 > 0$. Alternativt kan vi vise dette ved å observere at $P(A) + P(B) = 1.25 \neq P(A \cup B)$. Hendelsene A og B er uavhengige siden $P(A) \cdot P(B) = 0.80 \cdot 0.45 = 0.36 = P(A \cap B)$.

1 b) Løsningsforslag 1: Vi vet at Pål har fått flere enn 50 henvendelser, altså at hendelsen A har inntruffet. Ekteparet går *ikke* på restaurant dersom Pia *ikke* får flere enn 50 henvendelser, det vil si hendelsen B' . Vi skal dermed regne ut $P(B'|A)$. Siden A og B er uavhengige så er $P(B'|A) = P(B')$. Fra komplementregelen følger det at

$$P(B') = 1 - P(B) = 1 - 0.45 = 0.55$$

1 b) Løsningsforslag 2: Vi vil finne $P(C'|A) = \frac{P(C' \cap A)}{P(A)}$. Fra f.eks. et venndiagram kan vi se at $P(C' \cap A) = P(A) - P(C)$. Dermed blir svaret

$$P(C'|A) = \frac{P(C' \cap A)}{P(A)} = \frac{P(A) - P(C)}{P(A)} = 1 - \frac{P(C)}{P(A)} = 1 - 0.45 = 0.55$$

1 c) i. Vi skal regne ut $P(X \leq 14)$, altså sannsynligheten for at tiden det går fra arbeidsdagen starter til Pål får sin første henvendelse er mindre eller lik 14 minutter. Vi vet X er eksponensialfordelt med forventning $E(X) = 8$, så med “kjent” notasjon kan vi skrive $X \sim \text{eksponential}(\frac{1}{8})$. Dermed er

$$P(X \leq 14) = 1 - e^{-\frac{14}{8}} \approx 0.826$$

1 c) ii. Vi skal regne ut $P(X \leq 14, Y \leq 14)$, som også kan skrives som $P(X \leq 14 \cap Y \leq 14)$. Siden X og Y er uavhengige, så er

$$P(X \leq 14, Y \leq 14) = P(X \leq 14) \cdot P(Y \leq 14)$$

Dermed er svaret på oppgaven

$$(1 - e^{-\frac{14}{8}}) \cdot (1 - e^{-\frac{14}{9}}) \approx 0.652$$

Oppgave 2

2 a) i. Summen av alle punktsannsynlighetene i en diskret sannsynlighetsfordeling må være lik 1. Derfor må vi ha at $k = 1 - 0.3 - 0.2 - 0.2 = 0.3$.

2 a) ii. $E(X) = 0 \cdot k + 1 \cdot 0.3 + 2 \cdot 0.2 + 3 \cdot 0.2 = 0.3 + 2 \cdot 0.2 + 3 \cdot 0.2 = 1.3$

2 b) i. $E(X + Y) = E(X) + E(Y) = 2.9 + 2.2 = 5.1$

2 b) ii. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) = 1.5 + 2.1 = 3.6$

2 c) i. Vi skal regne ut sannsynligheter i normalfordelingen, og kan for eksempel ta i bruk tabeller for standard normalfordeling. Da trenger vi $P(X \leq 6) = P\left(\frac{X-8}{4} \leq \frac{6-8}{4}\right) = P\left(Z \leq \frac{6-8}{4}\right) = \Phi(-0.5) = 0.3085 \approx 0.309$.

2 c) ii. Når det gjelder gjennomsnittet av 9 uavhengige normalfordelte stokastiske variabler så vil $\bar{X}$ være normalfordelt med forventning 8 og standardavvik $4/\sqrt{9} = 4/3$. For å ta i bruk tabeller trenger vi $P(\bar{X} \leq 6) = P\left(\frac{\bar{X}-8}{4/3} \leq \frac{6-8}{4/3}\right) = P\left(Z \leq \frac{6-8}{4/3}\right) = \Phi(-1.5) = 0.0668 \approx 0.067$

Oppgave 3

3 a) i. Vi antar at “vårt” kandidatnummer er 10307. Da er sannsynligheten for å vinne spillet i ett forsøk lik $p = \frac{1}{9} \approx 0.111$.

3 a) ii. Når vi gjentar spillet 5 ganger er utfallene uavhengige av hverandre, og sannsynligheten p for suksess (*vinn*) er den samme i hvert forsøk. Vi har derfor en binomisk forsøksrekke bestående av $n = 5$ forsøk og der sannsynligheten for suksess $p = 1/9$. Vi lar den stokastiske variabelen X representere antall suksesser i $n = 5$ forsøk, og dermed er $X \sim \text{binomisk}(5, 1/9)$ og

$$P(X = 2) = \binom{5}{2} \frac{1}{9}^2 \left(1 - \frac{1}{9}\right)^3 \approx 0.0867$$

3 b) i. Nullhypotesen er $H_0 : p = 0.125$ og alternativhypotesen er $H_1 : p > 0.125$. Vi vil altså teste om sannsynligheten for suksess, p , er større enn 0.125.

3 b) ii. *Løsningsforslag 1:* Vi lar den stokastiske variabelen X representere antall suksesser (*vinn*) i 50 forsøk (*spill*), slik at $X \sim \text{binomisk}(50, p)$. Som testobservator kan vi bruke

$$Z = \frac{\hat{p} - 0.125}{\sqrt{0.125(1 - 0.125)/50}} = \frac{\hat{p} - 0.125}{0.0468}$$

der $\hat{p} = \frac{X}{50}$. Testobservatoren vil være tilnærmet standard normalfordelt, $N(0, 1)$, under nullhypotesen.

3 b) ii. *Løsningsforslag 2:* Vi lar den stokastiske variabelen X representere antall suksesser (*vinn*) i 50 forsøk (*spill*), slik at $X \sim \text{binomisk}(50, p)$. Som testobservator kan vi benytte $\hat{p} = \frac{X}{n}$. Når H_0 er sann er $X \sim \text{binomisk}(50, 0.125)$ og $\hat{p}$ er da tilnærmet normalfordelt med forventning 0.125 og standardavvik $\sqrt{0.125(1 - 0.125)/50} = 0.0468$.

3 b) ii. *Løsningsforslag 3:* Vi lar den stokastiske variabelen X representere antall suksesser (*vinn*) i 50 forsøk (*spill*), slik at $X \sim \text{binomisk}(50, p)$. Som testobservator kan vi benytte X s. Når H_0 er sann er

$X \sim \text{binomisk}(50, 0.125)$ og da tilnærmet normalfordelt med forventning $50 \cdot 0.125 = 6.25$ og standardavvik $\sqrt{50 \cdot 0.125(1 - 0.125)} = 2.3385$.

3 b) iii. Løsningsforslag 1: Ved signifikansnivå $\alpha = 0.05$ skal vi forkaste nullhypotesen dersom vi observerer $z > 1.645$.

3 b) iii. Løsningsforslag 2: Ved signifikansnivå $\alpha = 0.05$ skal vi forkaste nullhypotesen dersom vi observerer $\hat{p} > k$, der $k = 0.125 + 1.645 \cdot 0.04677072 = 0.2019378 \approx 0.2020$.

3 b) iii. Løsningsforslag 2: Ved signifikansnivå $\alpha = 0.05$ skal vi forkaste nullhypotesen dersom vi observerer $x > k_x$, der $k_x = 6.25 + 1.645 \cdot 2.3385 = 10.09683 \approx 10.10$. I praksis vil det si å forkaste H_0 dersom vi observerer 11 eller flere suksesser.

3 c) Løsningsforslag 1: Vi regner ut $\hat{p} = 9/50 = 0.18$ og $z = 1.175949 \approx 1.1759$. Siden z ikke ligger i forkastningsområdet ($z \leq 1.645$) så skal vi ikke forkaste nullhypotesen. Basert på 50 observasjoner (spill) har ikke Mons tilstrekkelig bevis for å påstå at det finnes en programmeringsfeil i koden til Petter.

3 c) Løsningsforslag 2: Vi regner ut $\hat{p} = 9/50 = 0.18$. Siden $\hat{p}$ ikke ligger i forkastningsområdet ($\hat{p} \leq 0.2020$) så skal vi ikke forkaste nullhypotesen. Basert på 50 observasjoner (spill) har ikke Mons tilstrekkelig bevis for å påstå at det finnes en programmeringsfeil i koden til Petter.

3 c) Løsningsforslag 2: Siden $x = 9$ ikke ligger i forkastningsområdet ($x < 11$) så skal vi ikke forkaste nullhypotesen. Basert på 50 observasjoner (spill) har ikke Mons tilstrekkelig bevis for å påstå at det finnes en programmeringsfeil i koden til Petter.

Oppgave 4

4 a) Den empiriske korrelasjonen er $r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = -0.9229926 \approx -0.923$

4 b) Estimert stigningstall er $\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = -0.9604048 \approx -0.96$. Estimert krysningspunkt på y -aksen er $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 2.6325 - (-0.9604048) \cdot 3.362 = 5.8613808 \approx 5.86$. Dermed blir $\hat{y} = 5.86 - 0.96x$.

4 c) Vi skal regne ut et 90% konfidensintervall for β_1 . Vi antar at forventningsverdien til Y , gitt x , er $\beta_0 + \beta_1 x$, og at standardavviket til Y , gitt x , er σ . Vi trenger i første omgang et estimat på dette standardavviket σ , som vi finner fra formelen

$$s = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Vi vet (se bok eller formelark) at godhetsmålet for regresjonsmodellen, r^2 , er gitt ved

$$r^2 = \frac{SS_T - SS_E}{SS_T} = 1 - \frac{SS_E}{SS_T} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Vi løser denne ligningen for $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ og finner at

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = (1 - r^2) \cdot \sum_{i=1}^n (y_i - \bar{y})^2$$

Ved å bruke svaret fra oppgave a ($r = -0.9229926$) og tall i oppgaveteksten, finner vi dermed at

$\sum_{i=1}^n (y_i - \hat{y}_i)^2 = (1 - (-0.9229926)^2) \cdot 24.61 = 3.6443643$, og dermed vil estimatet på σ bli

$$s = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2} = 0.4499608.$$

For å komme frem til et 90% konfidensintervall trenger vi den kritiske verdien $t_{0.05,18} = 1.734$ og standardfeilen

$$SE(\hat{\beta}_1) = \frac{s}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} = 0.0943789. \text{ Dermed kan vi regne ut } [\hat{\beta}_1 - t_{0.05,18} \cdot SE(\hat{\beta}_1), \hat{\beta}_1 + t_{0.05,18} \cdot SE(\hat{\beta}_1)]$$

$$= [-0.9604048 - 1.734 \cdot 0.0943789, -0.9604048 + 1.734 \cdot 0.0943789] \approx [-1.12, -0.8]$$