

Løsningsforslag eksamen ISTx100y H24

December 19, 2024

Oppgave 1: Mål på spredning

a) $\bar{x} = 3.4$ timer og $m_x = 3$ timer

b) Det empiriske standardavviket blir

$$s = \sqrt{\frac{1}{9} \sum_{i=1}^{10} (x_i - \bar{x})^2} = 1.9550504398153572 \approx 1.96 \text{ timer.}$$

Oppgave 2: Hendelser og sannsynlighet

a) La A betegne PostIBox brukes, og A' at NordPost brukes. La videre B betegne ‘fremme i tide’. Det er 60% sjanse for at PostIBox er ledige, og dermed sannsynlighet $1 - 0.6 = 0.4$ for at NordPost må brukes; $P(A') = 0.4$. Vi vet fra oppgaven at $P(B|A') = 0.80$, slik at $P(B'|A') = 0.20$. Vi skal regne ut

$$P(A' \cap B') = P(B'|A')P(A') = 0.2 \cdot 0.4 = 0.08$$

b) La A betegne PostIBox brukes, og A' at NordPost brukes. La videre B betegne ‘fremme i tide’. Vi vet fra oppgaven at $P(B|A) = 0.99$ og $P(B|A') = 0.80$. Fra setningen om total sannsynlighet finner vi

$$P(B) = P(B|A)P(A) + P(B|A')P(A') = 0.99 \cdot 0.6 + 0.80 \cdot 0.4 = 0.914.$$

c) Vi lar S_N betegne ‘sendt med NordPost’, S_P betegne ‘sendt med PostIBox’ og S' betegne ‘ikke sendt’. Som over er B betegnelsen på ‘fremme i tide’, og B' komplementærhendelsen. Vi bruker Bayes regel og skriver

$$P(S_N|B') = \frac{P(B'|S_N)P(S_N)}{P(B')} = \frac{P(B'|S_N)P(S_N)}{P(B'|S_N)P(S_N) + P(B'|S_P)P(S_P) + P(B'|S_I)P(S_I)}$$

I uttrykket setter vi inn $P(B'|S_N) = 1 - 0.80 = 0.20$, $P(S_N) = 0.4$, $P(B'|S_P) = 1 - 0.99 = 0.01$, $P(S_P) = 0.55$, $P(B'|S_I) = 1$, $P(S_N) = 0.05$. Dermed blir

$$P(S_N|B') = \frac{0.20 \cdot 0.4}{0.20 \cdot 0.4 + 0.01 \cdot 0.55 + 1 \cdot 0.05} = 0.59040590405 \approx 0.590$$

Oppgave 3: Diskrete sannsynlighetsfordelinger

- a) $P(X = 3) = \binom{6}{3} 0.7^3 0.3^3 \approx 0.185$
- b) Bruker tabell og finner $P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.070 = 0.930$
- c) I binomisk fordeling er $E(X) = np$. Bruker regneregler for forventningsverdi og finner $E(2X + 3) = 2np + 3 = 2 \cdot 6 \cdot 0.7 + 3 = 11.4$
- d) Utrykker først $P(Y > 5 | Y > 4) = \frac{P(Y > 5)}{P(Y > 4)} = \frac{1 - P(Y \leq 5)}{1 - P(Y \leq 4)}$. Henter kumulative sannsynligheter fra tabell for Poissonfordelingen og finner $P(X > 5 | X > 4) = \frac{1 - 0.7029}{1 - 0.5321} = 0.6349647 \approx 0.635$
- e) Fra Poisson-fordeling følger det at $Var(Y) = E(Y) = 4.5$. Bruker regneregler for varians og finner $Var(2X + 3) = 2^2 \cdot 4.5 + 0 = 4 \cdot 4.5 = 18$

Oppgave 4: Normalfordeling

- a) $P(V \leq 2) = P(Z \leq 0) = 0.5$
- b) $P(0 \leq V \leq 4) = P(-1 \leq Z \leq 1) \approx 0.68$
- c) I standard normalfordeling er $z_{0.01} = 1.282$. Vi regner om til 'vår' normalfordeling; $P(V > k) = P(Z > \frac{k-2}{2}) = 0.1$ der $\frac{k-2}{2} = 1.282$ som gir $k = 2 + 2 \cdot 1.282 = 4.564$
- d) Under null-hypotesen er $\bar{X} \sim N(0, 1/\sqrt{5})$ og da er $P(\bar{V} \leq -0.48) = P(Z \leq -1.07)$ der Z er standard normalfordelt. Fra tabell følger det at $P(Z \leq -1.07) = 0.1423$.

Oppgave 5: Lyslenker og levetid

La T_i representere levetiden til en pære lyslenken til Anna. Da er T_i eksponensialfordelt med parameter $\lambda = 1/125$.

- a) $P(T_i > 100) = e^{-100/125} = e^{-0.8} = 0.449329 \approx 0.449$
- b) Vi lar T representere levetiden til hele lyslenken som er koblet i serie. Da er $P(T > 100) = \prod_{i=1}^n P(T_i > 100) = (e^{-0.8})^n$
- c) Sannsynligheten for at en pære går i stykker er $P(T \leq 200) = 1 - e^{-200\lambda}$.

d) Vi estimerer $\hat{p} = 8/30$ og henter ut fra tabell $z_{\alpha/2} = 1.96$. Dermed er et tilnærmet 95% gitt ved

$$\left[\hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right] = [0.1084215, 0.4249119] \approx [0.108, 0.425]$$

e) Vi finner forventning og varians til begge estimatorene:

$$E(\hat{p}_1) = \frac{np + mp}{n + m} = p, \text{ } Var(\hat{p}) = \frac{np(1-p) + mp(1-p)}{(n+m)^2} = \frac{1}{n+m} \cdot p(1-p)$$

$$E(\hat{p}_2) = 0.5\left(\frac{np}{n} + \frac{mp}{m}\right) = p, \text{ } Var(\hat{p}_2) = 0.5^2\left(\frac{np(1-p)}{n^2} + \frac{mp(1-p)}{m^2}\right) = 0.5^2 \frac{n+m}{nm} \cdot p(1-p)$$

Begge estimatorene er forventningsrette. Innsatt $n = 30$ og $m = 20$ blir variansen til $\hat{p}_1$ er $0.02p(1-p)$ mens variansen til $\hat{p}_2$ er omtrent $0.0208p(1-p)$. Vi kan også løse oppgaven ved å regne

$$\frac{Var(\hat{p}_1)}{Var(\hat{p}_2)} = \frac{1/(n+m)}{(0.5^2(n+m)/(nm))} = \frac{1}{0.5^2} \frac{nm}{(n+m)^2} = \frac{1}{0.5^2} \frac{30 \cdot 20}{(30+20)^2} = 0.96$$

Siden $\hat{p}_1$ har noe lavere varians enn $\hat{p}_2$, og begge er forventningsrette, så foretrekker vi $\hat{p}_1$.

Oppgave 6: Enkel lineær regresjon

a) Riktig påstand: Et 95% konfidensintervall for stigningstallet β_1 er $[0.149, 1.100]$.

b) $R^2 = \frac{SST - SSE}{SST}$ som gir $SSE = SST \cdot (1 - R^2)$. Videre er $SST = \sum_{i=1}^n (y_i - \bar{y})^2 = (n-1) \cdot 0.748$. Til slutt kan vi finne estimatet for varians i støyledd ved

$$\frac{SSE}{n-2} = \frac{1}{n-2} SST \cdot (1 - R^2) = \frac{n-1}{n-2} 0.748 \cdot (1 - 0.579) = 0.3598949 \approx 0.36$$

c) For å gjennomføre testen må vi sammenligne observert verdi av testobservatoren vår med kritisk verdi fra t -fordelingen med $n-2 = 7$ frihetsgrader. Vi bruker $\alpha/2 = 0.005$ fordi vi skal ha en grense i hver hale av fordelingen. Fra tabell finner vi $t_{7,0.001} = 3.499$. Dermed skal vi forkaste dersom vi observerer $T < -3.499$ eller $T > 3.499$.