



Oppgave 1 Empirisk gjennomsnittet er 1, empirisk median er 3, fjerde ordningsvariabel $X_{(4)} = 6$ og empirisk standardavvik er 6.89.

Oppgave 2 $P(0 < Z_1 < 1) = \Phi(1) - \Phi(0) = 0.84 - 0.5 = 0.34$.

$P(0 < Z_1 < 1 | Z_1 > 0) = P(0 < Z_1 < 1 \cup Z_1 > 0) / P(Z_1 > 0) = P(0 < Z_1 < 1) / P(Z_1 > 0) = 0.34 / 0.5 = 0.68$.

$P(Z_1 + Z_2 < 2) = \Phi(\frac{2-0}{\sqrt{1+1}}) = \Phi(\sqrt{2}) = 0.92$.

$E(V) = 10$ siden $V \sim \chi_{10}^2$.

$q = \chi_{0.05, 10}^2 = 18.31$.

$\text{Var}(V) = 2 \cdot 10 = 20$ siden $V \sim \chi_{10}^2$ (se f.eks. tabell).

Oppgave 3 Vi har at X er hypergeometrisk fordelt med parameter $n = 5$, $k = 8$ og $N = 20$. La $p = k/N = 0.4$

Dermed er $EX = np = n \frac{k}{N} = 5 \cdot \frac{8}{20} = 2$.

Videre er $\text{Var } X = np(1-p) \frac{N-n}{N-1} = 5 \cdot 0.4 \cdot 0.6 \frac{15}{19} = 0.95$.

Antall gutter blir $n - X$. Hendelsen at det er færre jenter enn gutter betyr at $X < n - X$ som betyr at $X < n/2 = 2.5$, d.v.s. hendelsen $X \leq 2$ når X er en diskret variabel som kun tar heltallige verdier. Fra tabell finner vi at $P(X \leq 2) = 0.70$.

Hvis antall gutter er lik antall jenter er $n - X = X$ og $n = 2X$ slik at $X = n/2 = 2.5$. Siden X er en diskret variabel som kun kan ta heltallige verdier er denne hendelsen lik den tomme mengde $\emptyset$. Det følger (fra sannsynlighetsaksiomene) at hendelsen $\emptyset$ har alltid sannsynlighet 0. Med andre ord har vi $P(X = n - X) = P(X = 2.5) = P(\{s \in S | X(s) = 2.5\}) = P(\emptyset) = 0$.

Antall ikke-ordnede utvalg på 5 elever blir $\binom{20}{5} = 15504$.

Oppgave 4 La C være hendelsen at den tilfeldig valgte personen har covid og T hendelsen at testen er positiv. Det gitt i oppgaven at $P(C) = 0.05$, $P(T|C) = .8$, og $P(T|C^c) = 0.01$.

Da er $P(C^c \cap T) = P(T|C^c)P(C^c) = 0.01 \cdot 0.95 = 0.0095$.

Lov om totalsannsynlighet gir at $P(T) = P(T|C^c)P(C^c) + P(T|C)P(C) = 0.0095 + 0.8 \cdot 0.05 = 0.0495$.

Sannsynligheten for at personen ikke er smittet gitt positiv test blir i følge Bayes setning,

$$P(C^c|T) = \frac{P(C^c \cap T)}{P(T)} = \frac{0.0095}{0.0495} = 0.19$$

Oppgave 5

a) Bruker vi substitusjonen $v = \frac{u^2}{2a^2}$ slik at $dv = \frac{u}{a^2} du$ får vi at kumulativ tetthet blir

$$\begin{aligned} F_X(x) &= \int_0^x \frac{u}{a^2} e^{-\frac{u^2}{2a^2}} du \\ &= \int_0^{\frac{x^2}{2a^2}} e^{-v} dv \\ &= 1 - e^{-\frac{x^2}{2a^2}}. \end{aligned}$$

For $a = 3 \text{ m/s}$ får vi dermed at en vindturbin ikke vil produsere strøm med sannsynlighet

$$P(X < 3) = F_X(3) = 1 - e^{-\frac{3^2}{2 \cdot 8^2}} = 0.068.$$

b) Bruker vi samme substitusjon som over får vi at forventet vindhastighet blir,

$$\begin{aligned} EX &= \int_{-\infty}^{\infty} x f_X(x) dx \\ &= \int_0^{\infty} x \frac{x}{a^2} e^{-\frac{x^2}{2a^2}} dx \\ &= \int_0^{\infty} \sqrt{2a^2 v} e^{-v} dv \\ &= \sqrt{2}a \int_0^{\infty} v^{\frac{3}{2}-1} e^{-v} dv \\ &= \sqrt{2}a \Gamma\left(\frac{3}{2}\right) \\ &= \sqrt{2}a \frac{\sqrt{\pi}}{2} \\ &= a \sqrt{\frac{\pi}{2}} \end{aligned}$$

- c) Minimum $X_{(1)}$ har kumulativ tetthet

$$\begin{aligned}
 F_{X_{(1)}}(x) &= P(X_{(1)} \leq x) \\
 &= 1 - P(X_{(1)} > x) \\
 &= 1 - P(X_1 > x \cap \dots \cap X_n > x) \\
 &= 1 - (1 - F_X(x))^n \\
 &= 1 - e^{-\frac{nx^2}{2a^2}}.
 \end{aligned}$$

Fordelingen er altså på samme form som fordelingen til hver X_i men med parameter a' gitt ved $1/a'^2 = n/a^2$, altså $a' = a/\sqrt{n}$. Dermed er, i følge resultatet i punkt b), $EX_{(1)} = a'\sqrt{\frac{\pi}{2}} = a\sqrt{\frac{\pi}{2n}}$

- d) Siden X kun tar positive verdier er transformasjonen $Y = u(X) = (X/a)^2$ med invers $X = w(Y) = aY^{1/2}$, en-til-en slik at tettheten til Y er gitt ved

$$f_Y(y) = f_X(w(y))|w'(y)| = \frac{ay^{1/2}}{a^2} e^{-y/2} \frac{a}{2} y^{-1/2} = \frac{1}{2} e^{-y/2}$$

som er tettheten til en kji-kvadratfordelt variabel med 2 frihetsgrader. Dermed er $\sum_{i=1}^n \left(\frac{X_i}{a}\right)^2$ (en sum av n uavhengige kji-kvadratfordelte variable, hver med 2 frihetsgrader) kji-kvadrat med $2n$ frihetsgrader. Dermed har vi

$$\begin{aligned}
 P\left(\chi_{1-\alpha/2, 2n}^2 < \frac{1}{a^2} \sum_{i=1}^n X_i^2 < \chi_{\alpha/2, 2n}^2\right) &= 1 - \alpha \\
 P\left(\frac{1}{\chi_{\alpha/2, 2n}^2} < \frac{a^2}{\sum_{i=1}^n X_i^2} < \frac{1}{\chi_{1-\alpha/2, 2n}^2}\right) &= 1 - \alpha \\
 P\left(\sqrt{\frac{\sum_{i=1}^n X_i^2}{\chi_{\alpha/2, 2n}^2}} < a < \sqrt{\frac{\sum_{i=1}^n X_i^2}{\chi_{1-\alpha/2, 2n}^2}}\right) &= 1 - \alpha \\
 P\left(\sqrt{\frac{\pi \sum_{i=1}^n X_i^2}{2\chi_{\alpha/2, 2n}^2}} < \underbrace{a\sqrt{\frac{\pi}{2}}}_{EX} < \sqrt{\frac{\pi \sum_{i=1}^n X_i^2}{2\chi_{1-\alpha/2, 2n}^2}}\right) &= 1 - \alpha
 \end{aligned}$$

som betyr at

$$\left(\sqrt{\frac{\pi \sum_{i=1}^n X_i^2}{2\chi_{\alpha/2, 2n}^2}}, \sqrt{\frac{\pi \sum_{i=1}^n X_i^2}{2\chi_{1-\alpha/2, 2n}^2}}\right)$$

er et $(1 - \alpha)$ -konfidensintervall for forventet vindhastighet EX .

For 95%-konfidensnivå blir det observerte intervallet

$$\left(\sqrt{\frac{3.14 \cdot 11250}{2 \cdot 129.56}}, \sqrt{\frac{3.14 \cdot 11250}{2 \cdot 74.22}}\right) = (11.68 \text{ m/s}, 15.43 \text{ m/s})$$

Oppgave 6

a) Rimelighetsfunksjonen blir

$$L(\beta) = \prod_{i=1}^n \frac{1}{2\pi\sigma^2} e^{-\frac{(y_i - \beta x_i)^2}{2\sigma^2}} = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta x_i)^2}$$

og log-rimelighetsfunksjonen

$$l(\beta) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta x_i)^2.$$

I maksimum er

$$\frac{dl}{d\beta} = -\frac{1}{2\sigma^2} \sum_{i=1}^n 2(y_i - \beta x_i)(-x_i) = 0.$$

Altså er

$$\sum_{i=1}^n y_i x_i - \sum_{i=1}^n \beta x_i^2 = 0$$

som betyr at SME av β er

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}.$$

Estimatoren $\hat{\beta}$ er en lineærkombinasjon av uavhengige normalfordelte variabler og dermed normalfordelt med

$$E\hat{\beta} = \frac{\sum_{i=1}^n x_i EY_i}{\sum_{i=1}^n x_i^2} = \frac{\sum_{i=1}^n x_i \beta x_i}{\sum_{i=1}^n x_i^2} = \beta,$$

og

$$\text{Var } \hat{\beta} = \frac{\sum_{i=1}^n x_i^2 \text{Var } Y_i}{(\sum_{i=1}^n x_i^2)^2} = \frac{\sigma^2}{\sum_{i=1}^n x_i^2}.$$

Innsatt observerte verdier får vi $\hat{\beta} = 60\,743 \text{ kr/km}^2$ og $\widehat{\text{SE}} \hat{\beta} = \sqrt{\frac{300000^2}{41668}} = 1470 \text{ kr/km}^2$.

b) Forventning og varians til $\hat{\beta}$ blir

$$E\hat{\beta} = \frac{1}{n} \sum_{i=1}^n \frac{EY_i}{x_i} = \frac{1}{n} \sum_{i=1}^n \frac{\beta x_i}{x_i} = \beta$$

og

$$\text{Var } \hat{\beta} = \frac{1}{n^2} \sum_{i=1}^n \frac{\text{Var } Y_i}{x_i^2} = \frac{\sigma^2}{n^2} \sum_{i=1}^n \frac{1}{x_i^2}.$$

Begge estimatorene er dermed forventningsrette men $\hat{\beta}$ har minst varians siden

$$\frac{\text{Var } \hat{\beta}}{\text{Var } \hat{\beta}} = \frac{n^2}{\sum_{i=1}^n x_i^2 \sum_{i=1}^n 1/x_i^2} = \frac{10^2}{0.00301 \cdot 41668} = 0.79 < 1,$$

og er derfor å foretrekke.

- c) Vi ønsker å teste $H_0 : \beta = \beta_0 = 60000$ kr vs. $H_1 : \beta > \beta_0$. Det følger at testobservatoren

$$Z = \frac{\hat{\beta} - \beta_0}{\sigma / \sqrt{\sum x_i^2}}$$

er standard normalfordelt under H_0 . Stor verdier av Z vil indikere at H_1 er sann så vi forkaster hvis $Z > z_\alpha$. For $\alpha = 0.05$ blir kritisk verdi $z_\alpha = 1.64$. Observert verdi av Z blir

$$z = \frac{60743 - 60000}{1470} = 0.505$$

Vi beholder derfor H_0 . Testens p-verdi blir

$$P(Z > z | H_0) = 1 - \Phi(z) = 1 - \Phi(0.505) = 0.31.$$

- d) For $\beta = 65000$ blir teststyrken

$$\begin{aligned} P(Z > z_\alpha | H_1) &= P\left(\frac{\hat{\beta} - \beta_0}{\sigma / \sqrt{\sum x_i^2}} > z_\alpha | H_1\right) \\ &= P\left(\frac{\hat{\beta} - \beta}{\sigma / \sqrt{\sum x_i^2}} > z_\alpha + \frac{\beta_0 - \beta}{\sigma / \sqrt{\sum x_i^2}} | H_1\right) \\ &= 1 - \Phi\left(z_\alpha + \frac{\beta_0 - \beta}{\sigma / \sqrt{\sum x_i^2}}\right) \\ &= \Phi\left(\frac{\beta - \beta_0}{\sigma / \sqrt{\sum x_i^2}} - z_\alpha\right) \\ &= \Phi\left(\frac{5000}{1470} - 1.64\right) = \Phi(1.76) = 0.9609. \end{aligned}$$