

Oppgave 1: Normalfordelingen

a) En standard normalfordelt stokastisk variabel har forventning $E(Z) = 0$ og varians $\text{Var}(Z) = 1$.

b) Vi standardiserer og finner svaret i tabell

$$P(X \leq 1) = P\left(\frac{X-3}{2} \leq \frac{1-3}{2}\right) = P(Z \leq -1) = 0.1587$$

c) Y er normalfordelt med forventning

$$E(Y) = E(2 + 4Z) = 2 + 4E(Z) = 2 + 0 = 2$$

og varians

$$\text{Var}(Y) = \text{Var}(2 + 4Z) = 0 + 16 \cdot \text{Var}(Z) = 16$$

og dermed standardavvik

$$\text{SD}(Y) = \sqrt{\text{Var}(Y)} = 4$$

d) Fra sentralgrenseteoremet vet vi at en sum av uavhengige identisk fordelte variabler kan tilnærmes med en normalfordeling. Fra formelarket vet vi at for $X \sim \text{Ekspponential}(\lambda)$ så er $E(X) = \frac{1}{\lambda}$ og $\text{Var}(X) = \frac{1}{\lambda^2}$. Forventningsverdien til hver X_i , når $X_i \sim \text{Ekspponential}(2)$, er dermed $E(X_i) = 0.5$ og variansen er $\text{Var}(X_i) = 0.25$. Forventningsverdien til Y er dermed

$$E(Y) = E(X_1 + X_1 + \dots + X_{50}) = 50 \cdot 0.5 = 25$$

og variansen er

$$\text{Var}(Y) = \text{Var}(X_1 + X_1 + \dots + X_{50}) = 50 \cdot 0.25 = 12.5$$

Oppgave 2: Estimatorer for lengde

For å løse oppgaven kan det være lurt å skrive om estimatorene på denne måten:

$$\hat{\mu}_1 = Y, \quad \hat{\mu}_2 = \frac{\sum_i X_i}{5}, \quad \hat{\mu}_3 = \frac{1}{2}(\hat{\mu}_1 + \hat{\mu}_2), \quad \hat{\mu}_4 = \frac{1}{6}(\hat{\mu}_1 + 5\hat{\mu}_2), \quad \hat{\mu}_5 = 5\hat{\mu}_1 - 5\hat{\mu}_2$$

a) Vi sjekker først om estimatorene er forventningsrette. Siden $E(Y) = \mu$ så er $\hat{\mu}_1$ forventningsrett:

$$E(\hat{\mu}_1) = E(Y) = \mu.$$

Videre er

$$E(\hat{\mu}_2) = \frac{\sum_i E(X_i)}{5} = \frac{5\mu}{5} = \mu,$$

så $\hat{\mu}_2$ er også forventningsrett. Da følger det at

$$E(\hat{\mu}_3) = E\left(\frac{1}{2}(\hat{\mu}_1 + \hat{\mu}_2)\right) = \frac{1}{2}(\mu + \mu) = \mu$$

og

$$E(\hat{\mu}_4) = E\left(\frac{1}{6}(\hat{\mu}_1 + 5\hat{\mu}_2)\right) = \frac{1}{6}(\mu + 5\mu) = \mu,$$

så $\hat{\mu}_3$ og $\hat{\mu}_4$ er også forventningsrette. Da gjenstår bare $\hat{\mu}_5$, og denne har forventningsverdi

$$E(\hat{\mu}_5) = E(5\hat{\mu}_1 - 5\hat{\mu}_2) = 5\mu - 5\mu = 0.$$

Dermed er ikke $\hat{\mu}_5$ forventningsrett.

b) Vi skal regne ut varians til alle estimatorene. Vi finner først at

$$\text{Var}(\hat{\mu}_1) = \text{Var}(Y) = 0.2^2 = 0.04$$

og

$$\text{Var}(\hat{\mu}_2) = \frac{\sum_i \text{Var}(X_i)}{5^2} = \frac{5 \cdot 1}{5^2} = 0.2.$$

Videre har vi at

$$\text{Var}(\hat{\mu}_3) = \text{Var}\left(\frac{1}{2}(\hat{\mu}_1 + \hat{\mu}_2)\right) = \frac{1}{4}(0.04 + 0.2) = 0.06,$$

$$\text{Var}(\hat{\mu}_4) = \text{Var}\left(\frac{1}{6}(\hat{\mu}_1 + 5\hat{\mu}_2)\right) = \frac{1}{36}(0.04 + 5^2 \cdot 0.2) = 0.14$$

og

$$\text{Var}(\hat{\mu}_5) = \text{Var}(5\hat{\mu}_1 - 5\hat{\mu}_2) = 25 \cdot 0.04 + 25 \cdot 0.2 = 6.$$

Dermed er $\hat{\mu}_1$ estimatoren med lavest varians.

Oppgave 3: Produksjon av børsteløse motorer

a) La hendelsen A betegne at en motor er feilfri, la B betegne en feil som kan repareres og la C betegne at en motor må kasseres (har en feil som ikke kan repareres). Vi skal finne $P(B|A')$, der A' er komplementærhendelsen til A. Vi merker oss at snittet av hendelsen B og A' er hendelsen B. Det er fordi A' er unionen av de to disjunkte hendelsene B og C, så det som er felles i A' og B er hele hendelsen B. Vi setter dette inn i formel for betinget sannsynlighet:

$$P(B|A') = \frac{P(B \cap A')}{P(A')} = \frac{P(B)}{1 - P(A)} = \frac{0.11}{1 - 0.85} = 0.733$$

b) Vi ser på problemet som en binomisk forsøksrekke der $n = 25$ og sannsynligheten for suksess (feilfri motor) $p = 0.85$ i alle forsøk. Vi kan la X betegne den binomisk fordelte variabelen som teller antall suksesser i $n = 25$ forsøk, og vi er da ute etter å finne $P(X \geq 23)$. Denne sannsynligheten finner vi ved å summere tre punktsannsynligheter

$$P(X = 23) + P(X = 24) + P(X = 25)$$

. Vi setter inn i formel for binomiske punktsannsynligheter og finner

$$\binom{25}{23} p^{23} (1-p)^2 + \binom{25}{24} p^{24} (1-p)^1 + p^{25} = 0.254$$

c) La hendelsen A betegne at en motor er feilfri, la hendelsen B betegne en feil som kan repareres og la hendelsen C betegne at en motor må kasseres. Videre lar vi D betegne hendelsen at en motor blir sendt til reparasjon. Vi vet fra oppgaveteksten at $P(D|A) = 0.02$, $P(D|B) = 0.90$, og $P(D|C) = 0.40$. Fra loven om total sannsynlighet finner vi da

$$P(D) = P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C) = 0.02 \cdot 0.85 + 0.90 \cdot 0.11 + 0.40 \cdot 0.04 = 0.132$$

d) Vi skal regne på pålitelighet i Weibullfordelingen. Vi har $R(300) = e^{-(\lambda \cdot 300)^\alpha} = e^{-(3/4)^{1.5}} = 0.522$

e) I denne situasjonen ser vi på levetid med en seriekobling. Sannsynligheten for at bilen fungerer i mer enn 300 timer er $R(300) \cdot R(300) = 0.273$

Oppgave 4: Lakselus ved oppdrettsanlegg

a) Vi ser på en Poissonprosess der tellinger gjøres over et visst antall fisk. I denne del-oppgaven teller den stokastiske variabelen X antall lus over $t = 10$ fisk, slik at $X \sim \text{Poisson}(10\lambda)$. Dermed følger estimatet $\hat{\lambda} = \frac{x}{t} = \frac{8}{10} = 0.8$.

b) Under nullhypotesen er X Poissonfordelt med forventning $10\lambda_0 = 10 \cdot 0.5 = 5$, og vi skal regne ut $P(X \geq 8) = 1 - P(X \leq 7)$ (dette uttrykket følger av at X bare kan ta heltallsverdier). Vi kan slå opp i tabell og finner $1 - P(X \leq 7) = 1 - 0.8666 = 0.1334$. Siden p -verdien er 0.1334 forkaster vi ikke nullhypotesen.

c) Teststyrken er sannsynligheten for å forkaste nullhypotesen dersom alternativhypotesen er sann. Testen forkastes for observasjoner større enn eller lik 9. For $\lambda = 0.85$ er $X \sim \text{Poisson}(8.5)$ og $P(X \geq 9) = 1 - P(X \leq 8) = 1 - 0.5231 = 0.4769$, der vi igjen har brukt tabell for kumulative sannsynligheter i Poissonfordelingen.

d) Vi lar nå Y betegne antall lus på 20 fisk, der $Y \sim \text{Poisson}(20\lambda)$. Når nullhypotesen er sann er $Y \sim \text{Poisson}(20 \cdot 0.5 = 10)$. Vi finner $P(Y \geq 16) = 1 - P(Y \leq 15) = 1 - 0.9513 = 0.0487$. Siden p -verdien er lavere enn α kan nullhypotesen forkastes til fordel for alternativhypotesen ved signifikansnivå $\alpha = 0.10$.

Oppgave 5: Økonomisk ulikhet og pressefrihet

a) Medianen er omtrent 35

b) Den empiriske korrelasjonen er omtrent -0.5. Dette ser vi fordi det er en lineær negativ trend i dataene (høy WPFI opptrer samtidig med lav GINI). Dermed må den empiriske korrelasjonen være et negativt tall. Av de to negative svaralternativene har vi -0.5 og -1. Siden det er mye "støy" i dataene (observasjonene ligger ikke på en rett linje) så må -0.5 være det riktige svaret.

c) Regresjonslinja er $\hat{y} = 98.8508 - 0.7399x$. Dette svaret kan du dobbelsjekke ved å se på regresjonslinja som plottet (ingen andre svaralternativer passer med plottet).

d) Predikert WPFI ved Gini 42.6 blir $\hat{y} = 98.8508 - 0.7399 \cdot 42.6 = 67.33106 \approx 67.33$. Dette svaralternativet kan man også komme frem til ved å studere plottet: for GINI litt større enn 40 gir linja en WPFI-verdi som er litt mindre enn 70.

e) Et 90% konfidensintervall finner vi ved

$$\left[\hat{\beta}_1 - t_{0.1/2, n-2} \text{SE}(\hat{\beta}_1), \hat{\beta}_1 + t_{0.1/2, n-2} \text{SE}(\hat{\beta}_1) \right]$$

Den kritiske verdien fra t-fordelingen $t_{0.1/2, n-2} = t_{0.05, 50} = 1.676$ finner vi fra tabell, mens $\hat{\beta}_1 = -0.7399$ og $\text{SE}(\hat{\beta}_1) = 0.168$ finner vi fra utskriften. Dermed får vi intervallet $[-1.021, -0.458]$