TMA4245 Statistikk
Vår 2020

Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Innlevering 5

Dette er den andre av tre innleveringer i blokk 2. I denne innleveringen er det spesielt fokus på konfidens- og prediksjonsintervall. For å få godkjent innleveringen kreves det at minimum 40% av svarene er riktige, og at det kommer fram at det er gjort et ærlig forsøk på 75% av oppgavene. Alle deloppgaver teller like mye. Husk: Vi har statistikklab hvis du trenger hjelp med øvingen. Se fagets nettsider for mer info.

Oppgave 1

Man skal undersøke pH-verdien, μ, i et ferskvann. Til dette formål benyttes to ulike målemetoder, metode A og metode B. La X betegne målt verdi ved metode A og la Y tilsvarende betegne målt verdi med metode B. Anta at X og Y er uavhengige og at begge er normalfordelt med forventning μ. Målenøyaktigheten for de to målemetodene er forskjellige. Anta at variansen til X er $\sigma_A^2 = 0.2^2$, mens variansen til Y er $\sigma_B^2 = 0.1^2$. For å estimere μ er det foreslått tre estimatorer

$$\widehat{\mu} = Y \quad , \quad \widetilde{\mu} = \frac{1}{2}X + \frac{1}{2}Y \quad \text{og} \quad \mu^* = \frac{1}{5}X + \frac{4}{5}Y.$$

a) Hvilke to egenskaper kjennetegner en god estimator?

Hvilken av estimatorene $\widehat{\mu}$, $\widetilde{\mu}$ og μ^* vil du foretrekke? Begrunn svaret.

b) Hvilken sannsynlighetsfordeling har μ^*? Begrunn svaret.

Utled et 90%-konfidensintervall for μ ved å ta utgangspunkt i estimatoren μ^*.

Oppgave 2

Det vurderes å utbedre en lite trafikkert, men farlig veistrekning i Trondheimsområdet. I den sammenheng blir du bedt om å analysere hvor mye trafikk det er på veien. La X være antall biler som passerer et bestemt punkt på veistrekningen fra kl 16:00 til kl 18:00 på en tilfeldig valgt hverdag. Vi antar at X er poissonfordelt med parameter λ, dvs

$$f(x) = \frac{\lambda^x}{x!} \exp\{-\lambda\}, \quad x = 0, 1, 2, \ldots$$

a) Anta bare i dette punktet at $\lambda = 15$.

Regn ut $\mathrm{P}(X > 20)$ og $\mathrm{P}(10 \leq X < 20)$.

Finn E(X).

Anta at verdien til λ er ukjent i resten av oppgaven. Vi ønsker nå å finne realistiske verdier for λ. Vi observerer derfor antall passerende biler fra kl 16:00 til kl 18:00 på n tilfeldig valgte hverdager, $X_1, X_2, \ldots, X_n$. Vi antar at observasjonene er uavhengige og identisk fordelt med samme poissonfordeling som tidligere beskrevet. Resultatet av målingene for $n = 30$ tilfeldige dager, $x_1, x_2, \ldots, x_{30}$, gir $\sum_{i=1}^{30} x_i = 359$.

b) Definér estimatoren

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Bruk sentralgrenseteoremet til å utlede et tilnærmet $(1 - \alpha) \cdot 100\%$ konfidensintervall for λ basert på estimatoren $\hat{\lambda}$.

Regn ut konfidensintervallet for λ med de gitte dataene og $\alpha = 0.01$.

Kommunen ønsker å samle inn mer data om hvor trafikkert veien er. Derfor registreres det også hvor mange biler som passerer på veien fra kl 18:00 til kl 20:00 på m tilfeldig valgte hverdager, $Y_1, Y_2, \ldots, Y_m$. Vi antar at observasjonene er uavhengige og identisk poissonfordelte med parameter $\lambda/2$, altså halv intensitet i forhold til fra kl 16:00 til kl 18:00. Anta også at $Y_1, Y_2, \ldots, Y_m$ er uavhengige av $X_1, X_2, \ldots, X_n$.

c) Utled sannsynlighetsmaksimeringsestimatoren (maximum likelihood estimatoren) for λ basert på observasjonene $X_1, X_2, \ldots, X_n$ og $Y_1, Y_2, \ldots, Y_m$.

Oppgave 3

Levetiden, T, for en bestemt type elektroniske komponenter har vist seg å være eksponensialfordelt med parameter λ, dvs.

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & , t > 0 \\ 0 & , \text{ellers} \end{cases}$$

a) Vis at $Z = 2\lambda T$ er χ^2-fordelt med 2 frihetsgrader, dvs. at

$$f_Z(z) = \begin{cases} \frac{1}{2} e^{-\frac{z}{2}} & , z > 0 \\ 0 & , \text{ellers} \end{cases}$$

b) Bruk de 10 målte levetidene som er angitt nedenfor til å finne et 90% konfidensintervall for λ.

Komponent i	1	2	3	4	5	6	7	8	9	10
t_i	941.3	465.2	479.1	1071.9	504.7	933.4	1137.3	343.7	165.2	388.4

Fasit

1. **a**) Foretrekker μ^* **b**) $\mu^* \sim \mathrm{N}(\mu, 0.008)$, $[\mu^* - 0.15, \mu^* + 0.15]$

2. **a**) 0.0830, 0.8053 **b**) [10.34, 13.59]

3. **b**) $(8.438 \cdot 10^{-4}, 24.424 \cdot 10^{-4})$