TMA4240 Statistikk
Høst 2018

Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Anbefalt øving 9

Denne øvingen er basert på pensum forelest i 9. forelesningsuke. Oppgavene i denne øvingen handler blant annet om sannsynlighetsmaksimeringsestimator (SME), egenskaper til estimatorer og momentgenererende funksjoner.

Oppgave 1

Et problem med vindmøller til kraftproduksjon er at fugler kan kollidere med rotorbladene og dø. Som et prøveprosjekt monteres ei mølle (A) på et sted langs norskekysten, og det registreres kontinuerlig hvor mange fugler som blir funnet døde av kollisjonsskader rundt mølla. Erfaring fra Danmark med lignende møller tilsier at forventet antall døde fugler per uke er $\lambda = 1$. Anta at den tilfeldige variabelen Y, antall fugler som kolliderer og dør med mølle A per uke, er poissonfordelt.

a) Hvis vi antar like forhold i Norge som i Danmark og at etterfølgende uker er uavhengige, finn sannsynligheten for at det i løpet av de første 5 uker skal kollidere mer enn 10 fugler med vindmølle A.

Gitt at det kolliderer mindre enn 5 fugler, hva er sannsynligheten for at ingen fugler kolliderte?

b) I løpet av fire år blir det funnet 261 fugler rundt den norske mølla. Estimer parameteren λ med sannsynlighetsmaksimering (maximum likelihood) fra disse data.

Ei anna mølle (B) ble plassert litt lenger inn i landet, på en plass med mindre fugletetthet. Anta at X, antall fugler som kolliderer og dør med mølle B per uke, er poissonfordelt med parameter ν og uavhengig av Y. La så $Z = X + Y$ være summen av døde fugler ved de to møllene

c) Finn den momentgenererende funksjonen $M_Z(t)$ til den tilfeldige variabelen Z.

Bruk den momentgenererende funskjonen til å si hvilken fordeling Z har.

Oppgave 2

Det overraskende raset i Løsberga ved Steinkjer førte til stenging av vei og jernbane, med kostnad ca 1 million per dag. Etter dette (og andre ras) har nasjonal rassikringsgruppe levert krav til samferdselsminister Navarsete om en milliard kroner. I denne oppgaven skal vi studere en tenkt situasjon med kjent sannsynlighet for ras i forbindelse med sprengning.

Et firma har i oppdrag å utbedre en bilvei. Arbeidet medfører sprengningsarbeid, med risiko for ras. Fra basis kunnskap om geologien i området antar de sannsynlighet for ras $p = P(X = 1) = 0.15$. Her er stokastisk variabel $X = 1$ dersom det raser, mens $X = 0$ ellers.

a) Anta, kun i dette punktet, at det i sprengningsarbeidet langs veistrekningen finnes 4 slike mulige rassteder, og at eventuelle ras her vil skje uavhengige av hverandre.

S er antall ras av de fire mulige rasene. Argumenter for at antall ras er binomisk fordelt med parametre $n = 4$ og $p = 0.15$.

Hva er sannynligheten for at det går ingen ras?

Gitt at det blir minst ett ras, hva er sannsynligheten for at det blir flere enn ett ras?

Videre i oppgaven ser vi kun på et av de mulige rasstedene. Dersom det raser her, blir det uforutsett stenging av vei, omdirigering av trafikk, dekking av skader, etc. Dette har total kostnad 40 millioner kroner. Dersom det ikke raser, er ingen skade skjedd, og kostnad 0 kroner. Det er også mulig å legge om veien i anleggsperioden. Dette har en fast kostnad på 7 millioner kroner. Et ras vil da ikke gi ytterligere kostnad.
Strategi A er å sprenge uten omlegging av veien.
Strategi B er midlertidig omlegging av veien.

b) Beregn forventet kostnad Z under *Strategi A*. Finn også standardavviket til kostnad. Bør veien legges midlertidig om? Begrunn svaret.

For 5 millioner kroner kan det gjøres grundige geologiske undersøkelser som gir et sikkert svar om det vil rase eller ikke. Basert på resultatet av en slik undersøkelse vil man vite om *strategi A* eller *strategi B* bør velges.

Hva er forventet kostnad dersom man gjennomfører denne grundige undersøkelsen? Bør den grundige undersøkelsen gjennomføres? Begrunn svaret.

For 1 million kroner kan geologer undersøke området med enkle metoder og gi en kvalifisert uttalelse om ras ($Y = 1$) eller ikke ras ($Y = 0$). Denne enkle metoden er ikke sikker, og vi antar at de treffer med sannsynlighet $P(Y = 1|X = 1) = P(Y = 0|X = 0) = \gamma > 0.5$.

Vi skal først estimere sannsynligheten γ fra data der vi kjenner utfallet av X, om det gikk ras eller ikke. Firmaet har brukt geologene i lignende, uavhengige, situasjoner 15 ganger tidligere. I sju av tilfellene gikk det ras $X_i = 1$, $i = 1, \ldots, 7$. Uttalelsene til geologene var da som følger: $Y_1 = 0$, $Y_2 = 1$, $Y_3 = 0$, $Y_4 = 1$, $Y_5 = 1$, $Y_6 = 1$, $Y_7 = 1$. I åtte av tilfellene gikk det ikke ras $X_i = 0$, $i = 8, \ldots, 15$. Uttalelsene til geologene var da som følger: $Y_8 = 1$, $Y_9 = 0$, $Y_{10} = 0$, $Y_{11} = 0$, $Y_{12} = 1$, $Y_{13} = 0$, $Y_{14} = 1$, $Y_{15} = 0$.

En estimator for γ er:

$$\hat{\gamma} = \frac{\sum_{i=1}^{15} I_i}{15}$$

der $I_i = 1$ dersom $Y_i = X_i$, og $I_i = 0$ dersom $Y_i \neq X_i$.
Vi antar at I_i, $i = 1, \ldots, 15$ er uavhengige.

c) Er $\hat{\gamma}$ sannsynlighetsmaksimeringsestimatoren (SME) til γ? Også kalt maximum likelihood estimator. Svar ved å finne SME til γ.

Regn ut estimatet for γ basert på estimatoren $\hat{\gamma}$.

d) Beregn forventning og varians til $\hat{\gamma}$.

Firmaet vurderer nå å samle inn slike relativt billige data fra geologene.

e) Bruk i dette punktet den estimerte verdien av γ fra punkt c).

Bruk Bayes formel til å regne ut sannsynligheten for ras når geologene uttaler ikke ras. Regn videre sannsynligheten for ras når geologene uttaler ras.

Firmaet ønsker å regne forventet kostnad før geologene eventuelt kalles inn. Argumenter for at forventet kostnad er:

$$C = 1 + \sum_{y=0}^{1} \min[7, E(Z|Y=y)]P(Y=y),$$

der Z er kostnad ved *strategi A*. Videre er $\min[a,b] = a$ dersom $a < b$, og $\min[a,b] = b$, dersom $a > b$.

Hva blir forventet kostnad C? Bør undersøkelsen gjennomføres? Begrunn svaret.

Fasit

1. **a**) 0.014,0.015 **b**) 1.25

2. **a**) 0.77, 0.23 **c**) 0.67 **d**) $\mathrm{E}[\widehat{\gamma}] = \gamma$, $\mathrm{Var}[\widehat{\gamma}] = (1-\gamma)\gamma/15$ **e**) 0.08, 0.26, 5.65 millioner