Norges teknisk–naturvitenskapelige universitet
Institutt for matematiske fag

ST0103 2014S – løsningsskisse

Oppgave 1

a) Antall partikler sendt ut på et halvt minutt er poissonfordelt med parameter $1{,}1 \cdot 0{,}5 = 0{,}55$. Sannsynligheten for at dette antallet er lik én, er $0{,}55^1 \cdot e^{-0{,}55}/1! = 0{,}32$. Sannsynligheten for at minst én partikkel sendes ut, er én minus sannsynligheten for at ingen sendes ut, $1 - 0{,}55^0 \cdot e^{-0{,}55}/0! = 0{,}42$.

b) Tida til neste utsendelse er eksponentielt fordelt. Kumulativ fordelingsfunksjon er gitt ved $1 - e^{-1{,}1t}$, slik at sannsynligheten for at tida er større enn ett minutt er $e^{-1{,}1 \cdot 1} = 0{,}33$. Alternativt kan vi finne svaret som sannsynligheten for at ingen utsendelser skjer i løpet av ett minutt, $1{,}1^0 \cdot e^{-1{,}1}/0! = 0{,}33$,

Oppgave 2

a) Det er rimelig å anta at antall dyr med parasitten er binomisk fordelt med parametre $n = 10$ og $p = 0{,}3$. Sannsynligheten for at to eller flere dyr har parasitten er én minus sannsynligheten for at null eller ett dyr har parasitten, $1 - \binom{10}{0}0{,}3^0 0{,}7^{10} - \binom{10}{1}0{,}3^1 0{,}7^9 = 0{,}85$.

b) Med normaltilnærmelse har et tilnærmet 95 %-konfidensintervall for p grenser $\hat{p} \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}$. Her er $\hat{p} = 29/70 = 0{,}414$, $\alpha = 0{,}05$, $z_{0{,}025} = 1{,}960$ (fra tabell) og $n = 70$, og vi får grenser $0{,}414 \pm 1{,}960\sqrt{0{,}414 \cdot 0{,}586/70} = 0{,}414 \pm 0{,}115$, og konfidensintervallet blir $[0{,}30, 0{,}53]$.

c) Lengden av intervallet er $2z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} \le z_{\alpha/2}/\sqrt{n}$ (ulikheten følger av at $\hat{p}(1-\hat{p})$ alltid er mindre enn eller lik $1/4$). Hvis vi velger n slik at $z_{\alpha/2}/\sqrt{n} < 0{,}2$, det vil si $n > (z_{\alpha/2}/0{,}2)^2 = 96{,}04$, altså $n \ge 97$, er vi sikret et intervall av lengde mindre enn 0,2.

d) Vi forkaster nullhypotesen for *store* verdier av testobservatoren

$$\frac{\hat{p} - 0{,}35}{\sqrt{0{,}35(1-0{,}35)/n}},$$

nærmere bestemt hvis den er større enn eller lik $z_{0{,}05} = 1{,}645$ (fra tabell) – da er nemlig sannsynligheten for feilaktig forkastning av nullhypotesen tilnærmet lik 0,05 hvis $p = 0{,}35$ (ved normaltilnærmelse av testobservatoren). Vi får verdien

$(0{,}414 - 0{,}35)/\sqrt{0{,}35 \cdot 0{,}65/70} = 1{,}13$, og vi forkaster dermed ikke nullhypotesen. Alternativt kunne vi ha funnet p-verdien, $P(Z \geq 1{,}13) = 0{,}13 > 0{,}05$, som også tilsier at vi ikke forkaster nullhypotesen.

e) Hvis $p = 0{,}35$ er sannsynligheten for forkastning tilnærmet lik 0,05, som beskrevet ovenfor. Hvis $p = 0{,}40$, er sannsynligheten for forkastning (teststyrken)

$$\begin{aligned} P\left(\frac{\hat{p} - 0{,}35}{\sqrt{0{,}35 \cdot 0{,}65/70}} \geq z_{0{,}05}\right) &= P\big(\hat{p} \geq 0{,}35 + 1{,}645\sqrt{0{,}35 \cdot 0{,}65/70}\big) \\ &= P\left(\frac{\hat{p} - 0{,}40}{\sqrt{0{,}40 \cdot 0{,}60/70}} \geq \frac{0{,}35 - 0{,}40 + 1{,}645\sqrt{0{,}35 \cdot 0{,}65/70}}{\sqrt{0{,}40 \cdot 0{,}60/70}}\right) \\ &= P(Z \geq 0{,}75) = P(Z \leq -0{,}75) = 0{,}227. \end{aligned}$$

Vi har antatt at $\hat{p}$ tilnærmet er normalfordelt med forventningsverdi p og varians $p(1-p)/70$.

Oppgave 3

a) La X være tida reaksjonen tar. Kumulativ fordelingsfunksjon til X er gitt ved $1 - e^{-x/\mu}$, slik at $P(X > 10) = 1 - (1 - e^{-10/10}) = 0{,}37$. Den betingede sannsynligheten for at reaksjonen tar lenger tid enn 20 ms gitt at den tar lenger tid enn 10 ms er

$$\begin{aligned} P(X > 20 \mid X > 10) &= \frac{P(X > 20 \cap X > 10)}{P(X > 10)} \\ &= \frac{P(X > 20)}{P(X > 10)} = \frac{e^{-20/\mu}}{e^{-10/\mu}} = e^{-10/\mu} = e^{-10/10} = 0{,}37. \end{aligned}$$

Vi kunne også ha innsett det siste ved hjelp av hukommelsesløsheten av eksponentiell fordeling, $P(X > s + t \mid X > s) = P(X > t)$, med $s = t = 10$.

b) Simultantettheten er gitt ved $L = (\frac{1}{\mu}e^{-x_1/\mu})(\frac{1}{\mu}e^{-x_2/\mu}) \cdots (\frac{1}{\mu}e^{-x_n/\mu}) = e^{-\sum x_i/\mu}/\mu^n$, som også er likelihoodfunksjonen sett som en funksjon av μ. Da er $\ln L = -\sum x_i/\mu - n \ln \mu$, som har maksimum når $0 = d \ln L/d\mu = \sum x_i/\mu^2 - n/\mu$, altså $\sum x_i = n\mu$, eller $\mu = \sum x_i/n = \bar{x}$. Sannsynlighetsmaksimeringsestimatoren er altså $\bar{X}$.

c)

$$\begin{aligned} E\left(\frac{1}{11}\left(\sum_{i=1}^{5} X_i + \frac{1}{0{,}9}\sum_{j=1}^{6} Y_j\right)\right) &= \frac{1}{11}\left(\sum_{i=1}^{5} EX_i + \frac{1}{0{,}9}\sum_{j=1}^{6} EY_j\right) \\ &= \frac{1}{11}\left(\sum_{i=1}^{5} \mu + \frac{1}{0{,}9}\sum_{j=1}^{6} 0{,}9\mu\right) = \frac{1}{11}(5\mu + 6\mu) = \mu \end{aligned}$$

og

$$E\left(\frac{1}{2}\bar{X}+\frac{1}{2\cdot 0{,}9}\bar{Y}\right)=\frac{1}{2}E\bar{X}+\frac{1}{2\cdot 0{,}9}E\bar{Y}=\frac{1}{2}\mu+\frac{1}{2\cdot 0{,}9}0{,}9\mu=\mu,$$

så begge estimatorene er forventningsrette.

$$\begin{aligned}
\operatorname{Var}\left(\frac{1}{11}\left(\sum_{i=1}^{5}X_i+\frac{1}{0{,}9}\sum_{j=1}^{6}Y_j\right)\right)&=\frac{1}{11^2}\left(\operatorname{Var}\sum_{i=1}^{5}X_i+\operatorname{Var}\left(\frac{1}{0{,}9}\sum_{j=1}^{6}Y_j\right)\right)\\
&=\frac{1}{11^2}\left(\operatorname{Var}\sum_{i=1}^{5}X_i+\frac{1}{0{,}9^2}\operatorname{Var}\sum_{j=1}^{6}Y_j\right)=\frac{1}{11^2}\left(\sum_{i=1}^{5}\operatorname{Var}X_i+\frac{1}{0{,}9^2}\sum_{j=1}^{6}\operatorname{Var}Y_j\right)\\
&=\frac{1}{11^2}\left(\sum_{i=1}^{5}\mu^2+\frac{1}{0{,}9^2}\sum_{j=1}^{6}(0{,}9\mu)^2\right)=\frac{1}{11^2}\left(5\mu^2+6\mu^2\right)=\frac{1}{11}\mu^2=0{,}0909\mu^2
\end{aligned}$$

mens

$$\begin{aligned}
\operatorname{Var}\left(\frac{1}{2}\bar{X}+\frac{1}{2\cdot 0{,}9}\bar{Y}\right)&=\frac{1}{2^2}\operatorname{Var}\bar{X}+\frac{1}{(2\cdot 0{,}9)^2}\operatorname{Var}\bar{Y}\\
&=\frac{1}{4}\frac{\mu^2}{5}+\frac{1}{3{,}24}\frac{(0{,}9\mu)^2}{6}=\left(\frac{1}{4}\cdot\frac{1}{5}+\frac{1}{3{,}24}\frac{0{,}9^2}{6}\right)\mu^2=0{,}0917\mu^2,
\end{aligned}$$

og siden den første estimatoren har noe mindre varians, foretrekker vi den.