UNIVERSITETET I OSLO

Det matematisk-naturvitenskapelige fakultet

Eksamen i:	STK1100 — Sannsynlighetsregning og statistisk modellering.
Eksamensdag:	Fredag 31. mai 2019.
Tid for eksamen:	14.30 – 18.30.
Oppgavesettet er på 3 sider.	
Vedlegg:	Ingen.
Tillatte hjelpemidler:	Godkjent kalkulator. Formelsamling for STK1100.

Kontroller at oppgavesettet er komplett før
du begynner å besvare spørsmålene.

Oppgave 1

Ved en fødeavdeling kan tidspunktene for fødsler beskrives ved en Poissonprosess med rate $\alpha = 12$ per døgn. La X være antall fødsler i løpet av t timer. Da er X Poissonfordelt med parameter $\lambda = \alpha \cdot (t/24) = t/2$. (Du skal ikke vise dette.)

a) Hva er sannsynligheten for at det blir født minst ett barn i løpet av én time? Hva er sannsynligheten for at det blir født minst to barn i løpet av to timer?

b) La T være antall timer fra midnatt (kl. 00.00) til første fødsel skjer. Vis at for $t > 0$ er

$$P(T > t) = e^{-t/2}$$

og bestem sannsynlighetstettheten til T. Hva er forventet tidspunkt for første fødsel etter midnatt?

c) Finn sannsynligheten for at det ikke blir født noen barn i løpet av de to første timene etter midnatt. Anta så at det ikke har blitt født noen barn mellom midnatt og klokka 02.00. Finn sannsynligheten for at en da må vente minst to timer til før det skjer en fødsel.

Oppgave 2

En stokastisk variabel X er Rayleigh-fordelt hvis den har kumulativ fordeling

$$F(x) = \begin{cases} 1 - e^{-x^2/(2\theta^2)} & \text{for } x > 0 \\ 0 & \text{ellers} \end{cases} \tag{1}$$

der $\theta > 0$.

a) Vis at medianen i Rayleigh-fordelingen (1) er gitt ved $\widetilde{\mu} = \theta\sqrt{2\ln(2)}$.

(Fortsettes på side 2.)

b) Vis at sannsynlighetstettheten til X er

$$f(x) = \begin{cases} \frac{1}{\theta^2} x e^{-x^2/(2\theta^2)} & \text{for } x > 0 \\ 0 & \text{ellers} \end{cases}$$

c) Vis at $E(X) = \theta\sqrt{\frac{\pi}{2}}$ og bestem $V(X)$.
Vink: Husk at for gamma-funksjonen $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du$ har vi at $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$. Husk også at $\Gamma(1) = 1$ og $\Gamma(1/2) = \sqrt{\pi}$.

d) Vis at $Y = (X/\theta)^2$ er gamma(1,2)-fordelt, dvs. gamma-fordelt med formparameter $\alpha = 1$ og skalaparameter $\beta = 2$.

Oppgave 3

La $X_1, X_2, \ldots, X_n$ være høyden av n bølger ved en oljeplattform i Nordsjøen i en periode med stabile værforhold. Det er vanlig å anta at X_i-ene er uavhengige og Rayleigh-fordelte, og vi vil gjøre denne antagelsen her. Så X_i-ene har kumulativ fordeling gitt ved (1) i oppgave 2.

a) La $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$. Vis at $\hat{\theta} = \sqrt{\frac{2}{\pi}}\,\bar{X}$ er en forventningsrett estimator for θ og bestem variansen til estimatoren.

Vink: Husk resultatet i spørsmål c i oppgave 2.

b) Forklar at $\sum_{i=1}^n (X_i/\theta)^2$ er gamma(n,2)-fordelt, og vis at $E(\frac{1}{2n}\sum_{i=1}^n X_i^2) = \theta^2$.
Vink: Husk resultatet i spørsmål d i oppgave 2.

Det siste resultatet i spørsmål b viser at en annen rimelig estimator for θ er

$$\tilde{\theta} = \sqrt{\frac{1}{2n}\sum_{i=1}^n X_i^2} \tag{2}$$

c) Utled et 95% konfidensinterval for θ. Intervallet skal uttrykkes ved hjelp av n, $\tilde{\theta}$, a og b. Her er a og b henholdsvis 2.5% og 97.5% persentilene for gamma(n,2)-fordelingen.

Nedenfor er det gitt høyden (i meter) av 20 bølger ved oljeplattformen:

5.5	11.1	6.4	8.4	4.7	7.0	9.7	3.7	11.8	9.2
3.8	3.1	7.9	3.2	12.9	4.5	11.3	3.6	8.6	7.4

Hvis vi kaller disse observasjonene for $x_1, x_2, \ldots, x_{20}$ har vi at $\bar{x} = \frac{1}{20}\sum_{i=1}^{20} x_i = 7.19$ og $\frac{1}{40}\sum_{i=1}^{20} x_i^2 = 30.51$

d) Beregn $\hat{\theta}$ og $\tilde{\theta}$ ut fra de observerte bølgehøydene. Bestem også et 95% konfidensintervall for θ. Det opplyses at 2.5- og 97.5-persentilene i gamma-fordelingen med formparameter 20 og skalaparameter 2 er henholdsvis 24.4 og 49.3.

Ved å bruke parametrisk bootstrap med $B = 10000$ bootstrap simuleringer, fant vi at de estimerte standardfeilene til $\hat{\theta}$ og $\tilde{\theta}$ er henholdsvis $s_{\hat{\theta}} = 0.641$ og $s_{\tilde{\theta}} = 0.613$.

(Fortsettes på side 3.)

e) Beskriv hvordan bootstrap estimatene for standardfeilene er beregnet, og diskuter hva de estimerte standardfeilene sier deg om egenskapene til de to estimatorene.
(En kan vise at estimatoren $\tilde{\theta}$ ikke er forventningsrett. Men skjevheten er så liten at vi kan se bort fra denne.)

Oppgave 4

De stokastiske variabelene X og Y har simultan sannsynlighetstetthet

$$f(x,y) = \begin{cases} xye^{-(x+y)} & \text{for } x > 0 \text{ og } y > 0, \\ 0 & \text{ellers.} \end{cases}$$

a) Vis at den marginale sannsynlighetstettheten til X er gitt ved

$$f_X(x) = \begin{cases} xe^{-x} & \text{for } x > 0 \\ 0 & \text{ellers} \end{cases}$$

Hva er den marginale sannsynlighetstettheten til Y? Avgjør om X og Y er uavhengige.

Vi innfører nå nye stokastiske variabler $U = X + Y$ og $V = X/(X + Y)$.

b) Bestem den simultane sannsynlighetstettheten til U og V.

SLUTT