Øving 8, ST1301

Oppgave 1 Hva gjør følgende funksjon? Hvilken fordeling har variabelen `n` som returneres som funksjonsverdi? Forklar hvorfor. Forutsett at to enkle positive tall blir oppgitt som argumenter.

```
xxx <- function(t,lambda) {
  tidspunkt <- 0
  n <- 0
  repeat {
     ventetid <- rexp(n=1,rate=lambda)
     tidspunkt <- tidspunkt + ventetid
     if (tidspunkt > t)
        break()
     else
        n <- n + 1
  }
  n
}
```

Oppgave 2 Anta at X er en triangulært fordelt stokastisk variabel med sannsynlighetstetthet

$$f_X(x) = 2x, \quad 0 \leq x \leq 1. \tag{1}$$

Utled kumulativ sannsynlighetstetthet til X, $F_X(x) = P(X \leq x)$. Programmer så en funksjon som simulerer n realisasjoner fra $f_X(x)$ ved bruk av inversjonsmetoden. Lag et histogram av 1000 realisasjoner fra fordelingen.

Estimer også forventningen til X basert på 1000 simulerte realisasjoner av X og kontroller resultatet ved å sammenligne med den teoretisk forventingen $E(X)$ (utled denne).

Oppgave 3 I øving 7 tilpasset vi gammefordeling til et sett observerte levetider og beregnet sannsynlighetsmaksimeringsestimatene av parameterne i modellen til $\hat{\lambda} = 1.07$ og $\hat{r} = 1.75$. Vi trenger å kontrollere at estimatene er riktige (d.v.s. lik parameterverdiene som maksimaliserer likelihoodfunksjonen). Lag et tredimensjonalt overflateplot av likelihoodfunksjonen (L) som funksjon av parameterne λ og r ved hjelp av funksjonen `persp`. Denne funksjonen trenger tre hovedargumenter; to vektorer `x` og `y` som spesifiserer punkter på et rutenett (i dette tilfelle sekvenser av λ og r verdier, og en matrise `z` som inneholder funksjonsverdier (i dette tilfelle $L(\lambda, r)$) i de ulike punktene på rutenettet.

Programmer en funksjon som lager et slikt overflateplot ved å bygge videre på funksjonen `lnL` fra øving 7. La max og min verdier for hver av de to parameterne være innargument. Bruk en nøstet for-løkke for å beregne matrisen av funksjonsverdier før kallet til `persp`.

En teknikalitet: Merk at likelihoodfunksjonen ($L = e^{\ln L}$) blir mindre enn minste tall vi kan jobbe med i R slik at `exp(lnL( ))`vil få verdien 0 i beregningene (numerisk underflow). Derfor er det bedre å ta vare på log-verdiene av L i nevnte matrise og så trekke fra maximum av alle verdiene før vi beregner L ved ekponentiering (dette vil i praksis bare innebære en reskalering av L som gjør at vi ungår underflow.)

Undersøk de forskjellige tilleggsargumentene til `persp` som `expand`, `theta`, og `phi`, `shade`, `border`, og `ticktype=''detailed''` (se hjelpesiden).

Programmer en funksjon som simulerer 1000 tilfeldig utvalg fra den tilpassede modellen (bruk `rgamma` og anta at parameterverdiene er lik estimatene beregnet ovenfor) og returner tilhørende estimat i form av en 1000×2 matrise. Estimer standardfeilen til estimatene på grunnlag av de simulerte verdiene. Lag også et spredningsplot av de simulerte estimatene. Er det noen likhet mellom spredningsplottet og overflateplottet av likelihoodfunksjonen?