

Ekstraoppgave 6

Eksempel 2 på sidene 2-3 i Storviks notat om “Numerical optimization of likelihoods” omhandler data fra Illinois om mengden av nedbør (i inches) for 227 “summer storms” i årene 1960-1964. Vi skal i denne oppgaven se litt nærmere på disse dataene.

Du kan lese dataene inn i vektoren \mathbf{x} ved kommandoene:

```
x=scan("http://www.uio.no/studier/emner/matnat/math/STK2120/v13/illrain.dat",
       na.strings="*")
x=x[!is.na(x)]
```

- a) Lag et histogram over dataene og kommenter histogrammets utseende.

Som i Storviks notat vil vi anta at dataene er gamma-fordelte. Vi antar altså at de observerte nedbørmengdene er uavhengige observasjoner av gamma-fordelte stokastiske variabler med sannsynlighetstetthet

$$f(x | \alpha, \lambda) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} \quad (1)$$

Merk at gamma-fordelingen (1) har en annen parameterisering enn den som er benyttet i avsnitt 4.4 i D&B.

- b) Finn uttrykk for likelihooden og log-likelihooden. (Disse er gitt i notatet til Storvik. Men prøv selv før du ser på uttrykkene der.)
- c) Deriver log-likelihooden med hensyn på α og λ . Angi to ligninger som maksimum likelihood estimatene må tilfredsstille. (Prøv selv før du ser på uttrykkene i notatet til Storvik.) Har disse ligningene en eksplisitt løsning?

For å bestemme maksimum likelihood estimatene må vi finne de verdiene av α og λ som maksimerer likelihooden (eller ekvivalent maksimerer log-likelihooden). Vi skal senere se hvordan det kan gjøres. Nå vil vi nøye oss med å lage plott av likelihooden og bruke det til å gi et grovt anslag av de maksimerende verdiene.

Kommandoene nedenfor gir et plott av likelihooden. (Kommandoene er en lett modifisert versjon av de som er gitt på side 44 i notatet til Storvik.)

```

alpha = seq(0.35,0.55,0.005)
lambda = seq(1,3,0.05)
loglik = matrix(nrow=length(alpha),ncol=length(lambda))
n = length(x)
sumx=sum(x)
sumlogx = sum(log(x));
for(i in 1:length(alpha))
for(j in 1:length(lambda))
loglik[i,j] = n*alpha[i]*log(lambda[j])+(alpha[i]-1)*sumlogx-
              lambda[j]*sumx-n*log(gamma(alpha[i]))
lik=exp(loglik)
par(mfrow=c(1,2))
persp(alpha,lambda,lik,theta=330,phi=45,shade=1,zlab="lik")
image(alpha,lambda,lik,col=gray((0:32)/32))

```

- d) Utfør kommandoene. Pass på at du forstår hva de ulike kommandoene gjør.
- c) Bruk plottet du får til å gi et grovt anslag for verdiene til maksimum likelihood estimatene.