02424 Week 2

Exercise 1

The following are heart rate measurements (beats/minute) of one person measured throughout the day.

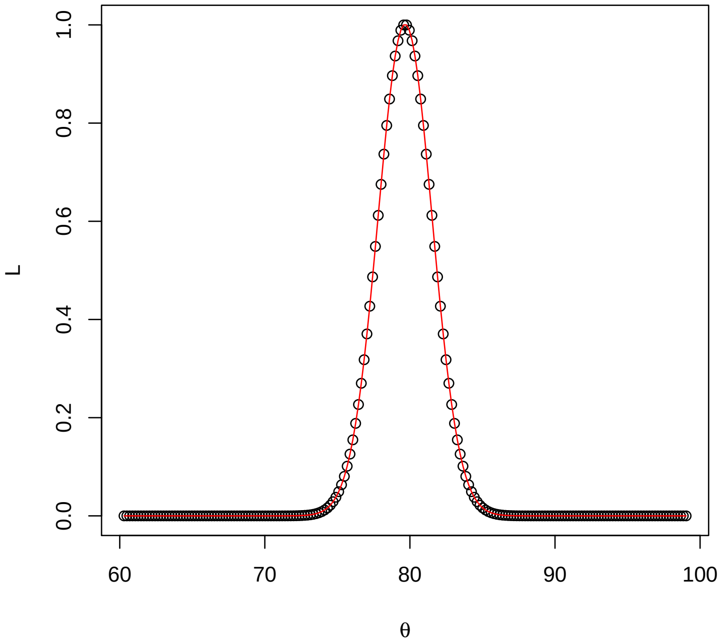
```
71  74  82  76  91  82  82  75  79  82  72  90
```

Assume that the data are an iid sample from $\mathrm{N}(\theta, \sigma^2)$, where σ^2 is assumed to be known at the observed sample variance. Sketch the likelihood function for θ if

a) the whole data are reported.

b) only the sample mean $\bar{y}$ is reported.

```
> x <- c(71, 74, 82, 76, 91, 82, 82, 75, 79, 82, 72, 90)
> s2 <- var(x)
> L.complete.data <- function(theta) {
+     prod(dnorm(x, mean = theta, sd = sqrt(s2)))
+ }
> x.ave <- mean(x)
> n <- length(x)
> L.ave <- function(theta) {
+     dnorm(x.ave, mean = theta, sd = sqrt(s2/n))
+ }
> th <- seq(x.ave - 3 * sqrt(s2), x.ave + 3 * sqrt(s2), length = 200)
> L <- sapply(th, L.complete.data)
> plot(th, L/max(L), ylab = "L", xlab = expression(theta))
> L <- sapply(th, L.ave)
> lines(th, L/max(L), col = "red")
```



Find the MLE, $\widehat{\theta}$, and the Hessian (for case a)) using the `optim` function in R (Note that the `optim` argument `method="Brent"` requires R 2.14.1 (or higher))

```
> nll.complete.data <- function(theta) {
+     -sum(dnorm(x, mean = theta, sd = sqrt(s2), log = TRUE))
+ }
> fit <- optim(x.ave, nll.complete.data, hessian = TRUE, method="Brent", lower=70, u
> fit[c("convergence", "par", "hessian")]

$convergence
[1] 0

$par
[1] 79.66667

$hessian
          [,1]
[1,] 0.2877907
```

Notice that the inverse hessian is equal to the observed sample variance divided by the number of observations.

```
> 1/fit$hessian

          [,1]
[1,] 3.474747

> s2/n

[1] 3.474747
```

Exercise 2

The measurements $y_1, y_2, ..., y_n$ are an iid sample from the Poisson distribution with density

$$f(y) = \frac{\lambda^y \exp(-\lambda)}{y!}.$$

a) Write down the combined likelihood function, the log-likelihood function, $l'_\lambda(\lambda; \mathbf{y})$ and $j(\lambda; \mathbf{y})$.
b) Derive the MLE, $\widehat{\lambda}$, and calculate the observed information.

Solution

a) The likelihood function is:

$$\prod_{i=1}^{n} \frac{\lambda^{y_i} \exp(-\lambda)}{y_i!}$$

and the log-likelihood:

$$\sum_{i=1}^{n} \log\left(\frac{\lambda^{y_i} \exp(-\lambda)}{y_i!}\right) = -n\lambda + \log(\lambda)\left(\sum_{i=1}^{n} y_i\right) - \sum_{i=1}^{n} \log(y_i!).$$

The score function is:

$$l'_\lambda(\lambda; \mathbf{y}) = -n + \frac{1}{\lambda}\left(\sum_{i=1}^{n} y_i\right)$$

and finally:

$$j(\lambda; \mathbf{y}) = \frac{1}{\lambda^2}\left(\sum_{i=1}^{n} y_i\right).$$

b) The MLE is:

$$\widehat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$

and the observed information:

$$j(\widehat{\lambda}; \mathbf{y}) = \frac{n}{\bar{y}}$$

Exercise 3

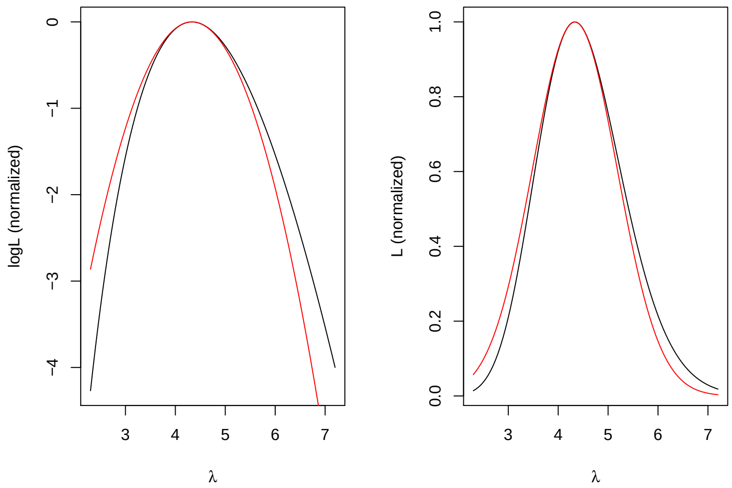
The following data are number of customers arriving at a cafe per 10 minutes:

```
4  6  3  7  2  4
```

Assume that the data are an iid sample from the Poisson distribution.
Plot the log-likelihood function and the quadratic approximation.

Set the maximum of the log-likelihood to zero and check a range of λ such that the log-likelihood is approximately between between -4 and 0. Do the same plot again but this time not on log-scale.

```
> par(mfrow = c(1, 2))
> x <- c(4, 6, 3, 7, 2, 4)
> l <- function(lambda) {
+     sum(dpois(x, lambda, log = TRUE))
+ }
> lam.hat <- mean(x)
> l.a <- function(lambda) {
+     l(lam.hat) - 0.5 * length(x)/lam.hat * (lambda - lam.hat)^2
+ }
> lam <- seq(2.3, 7.2, length = 100)
> l1 <- sapply(lam, l)
> plot(lam, l1 - max(l1), type = "l", xlab = expression(lambda),
+     ylab = "logL (normalized)")
> l2 <- sapply(lam, l.a)
> lines(lam, l2 - max(l2), col = "red")
> l1 <- sapply(lam, l)
> plot(lam, exp(l1 - max(l1)), type = "l", xlab = expression(lambda),
+     ylab = "L (normalized)")
> l2 <- sapply(lam, l.a)
> lines(lam, exp(l2 - max(l2)), col = "red")
```



Exercise 4

Question 1 and 2 from Exercise 2.1 in the textbook, but with the important change in Question 2 that T_w should be defined as:

$$T_w = wY_1 + (1-w)10Y_2$$

Solution

Q1:

$$E(T_1) = E(\frac{1}{2}Y_1 + \frac{10}{2}Y_2) = \frac{1}{2}E(Y_1) + \frac{10}{2}E(Y_2) = \frac{1}{2}\lambda + \frac{10}{2}\frac{\lambda}{10} = \lambda$$

$$V(T_1) = \frac{1}{4}V(Y_1 + 10Y_2) = \frac{1}{4}(\lambda + 10\lambda) = \frac{11}{4}\lambda$$

The factor 10 enters the expression for the variance squared, whereas the reduction due to dilution is only $\frac{1}{10}$.

Q2:

$$\begin{aligned} V(T_w) &= V(wY_1 + (1-w)10Y_2) \\ &= w^2V(Y_1) + (1-w)^2 100V(Y_2) \\ &= w^2\lambda + (1-w)^2 100\frac{\lambda}{10} \\ &= \lambda(11w^2 - 20w + 10) \end{aligned}$$

The second degree polynomial takes its minimum at $w = \frac{10}{11}$ and the variance of the estimater becomes $V(T_w) = \frac{10}{11}\lambda$.

Exercise 5

Q1

As stated on page 268 the $G(\nu/2, 2)$-distribution is the χ^2-distribution with ν degrees of freedom.

Q2

The gamma distribution has the density (see page 268)

$$\begin{aligned} g(y;\alpha,\beta) &= \frac{1}{\beta\Gamma(\alpha)}\left(\frac{y}{\beta}\right)^{\alpha-1}\exp(-y/\beta) \\ &= \frac{1}{\beta^{\alpha}\Gamma(\alpha)}y^{\alpha-1}\exp(-y/\beta) \end{aligned}$$

with $\alpha \in \mathbb{R}_+$ and $\beta \in \mathbb{R}_+$.
Since the sample consists of i.i.d. observations the likelihood function is

$$\begin{aligned} L(\alpha,\beta) &= \prod_{i=1}^{n} g(y_i;\alpha,\beta) \\ &= \prod_{i=1}^{n} \frac{1}{\beta^{\alpha}\Gamma(\alpha)}y_i^{\alpha-1}\exp(-y_i/\beta) \end{aligned}$$

Hence the log-likelihood function is

$$\begin{aligned} \ell(\alpha,\beta) &= \log L(\alpha,\beta) \\ &= -n\alpha\log\beta - nlog\Gamma(\alpha) + (\alpha-1)\sum\log y_i - \sum(\frac{y_i}{\beta}) \end{aligned}$$

Q3

For α known:

$$\ell(\beta) = -n\alpha\log\beta - \sum(\frac{y_i}{\beta})$$

and the score function becomes

$$\ell'(\beta) = -\frac{n\alpha}{\beta} + \frac{1}{\beta^2}\sum y_i$$

and it's readily seen that the MLE is

$$\begin{aligned} \widehat{\beta} &= \frac{1}{n\alpha}\sum_{i=1}^{n} y_i \\ &= \frac{\bar{y}}{\alpha} \end{aligned}$$

Exercise 6

$$L(\lambda) = \prod_{i=1}^{n} \lambda \exp(-\lambda y_i) = \lambda^n \prod \exp(-\lambda y_i)$$

Hence the log-likelihood function is

$$\ell(\lambda) = n \log \lambda - \lambda \sum y_i$$

The score function becomes

$$\ell'(\lambda) = n\frac{1}{\lambda} - \sum y_i$$

By putting the score function to zero, we see that the MLE is

$$\widehat{\lambda} = \frac{n}{\sum y_i} = \frac{1}{\bar{y}}$$

Exercise 7

The likelihood function is

$$L(\sigma) = \frac{1}{(2\pi)^{n/2}\sigma^n} \prod_{i=1}^{n} \exp(-\frac{1}{2}\frac{\tilde{y}_i^2}{\sigma^2})$$

where $\tilde{y}_i = y_i - \mu$.

Q1

The log-likelihood function is then

$$\ell(\sigma) = -n \log \sigma + \sum(-\frac{1}{2}\frac{\tilde{y}_i^2}{\sigma^2}) + const$$

Q2

The score function is

$$\ell'(\sigma) = S(\sigma) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum \tilde{y}_i^2$$

Q3

The observed information is

$$j(\sigma; y) = -l''(\sigma; y) = -\frac{n}{\sigma^2} + \frac{3}{\sigma^4}\sum \tilde{y}_i^2$$

Q4

The expected information is

$$\begin{aligned} i(\sigma) &= E(j(\sigma;y)) \\ &= E(-\frac{n}{\sigma^2} + \frac{3}{\sigma^4}\sum \tilde{y}_i^2) \\ &= -\frac{n}{\sigma^2} + \frac{3}{\sigma^4} n\sigma^2 \\ &= \frac{2n}{\sigma^2} \end{aligned}$$

Q5

The Cramer-Rao Lower Bound (CRLB) is

$$CRLB = \frac{1}{i(\sigma)} = \frac{\sigma^2}{2n}$$