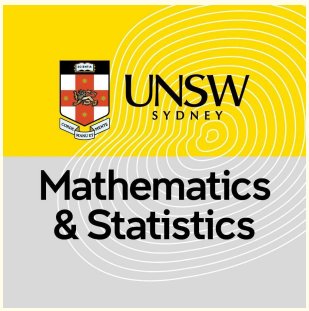


MATH5805
A Modern Introduction to Extreme Value Theory
Term 1, 2025

Dr. Boris Béranger

Week 1 - Lab Solutions

Exercise 1

Approximately 40% of the Netherlands is below sea level. Much of it has to be protected against the sea by dykes. These dykes have to withstand storm surges that drive the seawater level up along the coast. The Dutch government, balancing considerations of cost and safety, has determined that the dykes should be sufficiently high that the probability of a flood (which occurs when the seawater level exceeds the top of the dyke) in a given year is 10^{-4}. The question is then: how high should the dykes be built to meet this requirement?

a) Using `R`, simulate weekly observations of the sea level (in metres) at the town of Delfzijl between the January 1st 2002 and January 1st 2022, with each observation being drawn from the Exponential distribution with rate 2. For reproducibility use `set.seed(123)`.

Solution:

```
set.seed(123)
n.weeks <- 52
n.years <- 20
data <- matrix(rexp(1040, rate=2), ncol=n.years, nrow=n.weeks)
```

b) Assuming that the Exponential distribution is the true underlying distribution ($F_X(x)$), and using the simulated data, calculate the height that the dyke should be built at.

Solution:
Let $p = 1 - 10^{-4}$, we are looking for the p-th quantile from the distribution of the yearly maxima. Let λ denote the rate parameter. The distribution function of the maxima is $(1 - \exp\{-\lambda x\})^n$, where $n = 52$ since we are taking the maximum of 52 observations. An estimate of the rate parameter is $\hat{\lambda} = nm \left(\sum_{i=1}^{nm} x_i\right)^{-1}$ where x_i represents an observed value of the sea level and m is the number of years (such that $nm = 52 \times 20 = 1040$).
The height of the dykes such that the probability of the sea level to exceed this height with probability 10^{-4} is thus $-\hat{\lambda} \log\left(1 - p^{1/n}\right) \approx 6.76$ meters.

```
mean.dat <- mean(data)
p <- 1-1e-4
dyke.height <- -log(1- p^(1/n.weeks)) * mean.dat
dyke.height
```

```
## [1] 6.760443
```

c) Repeat the previous question assuming that the Gamma distribution is the true underlying distribution.

Solution:
The height of the dykes such that the probability of the sea level to exceed this height with probability 10^{-4} is ≈ 6.70 meters.

```
llik <- function(par){
  sum(dgamma(data, shape=par[1], rate=par[2], log=TRUE))
}
est <- optim(c(1,1), llik, control=c(fnscale=-1))
dyke.height2 <- qgamma(p^(1/n.weeks), shape=est$par[1], rate=est$par[2])
dyke.height2
```

```
## [1] 6.707218
```

d) Display a histogram of the observed yearly maxima of the sea level and add the dyke heights calculate in the previous questions.

Solution:

```
yearly.max <- apply(data,2,max)

hist(yearly.max, xlim=c(0, max(yearly.max, dyke.height, dyke.height2)),
     xlab="Sea level (metres)", main="Histogram of observed sea levels")
abline(v=c(dyke.height, dyke.height2), col=c(2,3), lwd=2)
```

