Problem 1 Verify the formulae for x_p given on slide 102 of the notes.

Problem 2 Suppose that the GEV distribution with parameters η, τ and ξ has been fitted to data $y_1, \dots, y_n$, and a profile log likelihood is required for the $1/p$-year return level x_p.

(a) Show that $x_p = \eta + \tau a_p(\xi)$, where for $p \in (0,1)$ we set $a_p(\xi) = [\{-\log(1-p)\}^{-\xi} - 1]/\xi$.

(b) The required profile log likelihood $\ell_{\rm p}(x_p)$ is defined as $\max_{\tau,\xi} \ell^*(x_p, \tau, \xi)$, where ℓ^* is the log likelihood function parametrised in terms of x_p, τ and ξ. Show that

$$\ell_{\rm p}(x_p) = \max_{\tau,\xi} \ell\{x_p - \tau a_p(\xi), \tau, \xi\},$$

where ℓ is the log likelihood parametrized in terms of η, τ and ξ, and hence explain how you would compute $\ell_{\rm p}(x_p)$.

Problem 3 This follows on from the previous `R` exercises using the hourly rainfall from Eskdalemuir, but now estimates return levels. Here's the code to read the data and make what you need:

```
library(evd)
library(lubridate)
load("eskrain.RData")
time.seq <- seq(from=min(date(eskrain)), to=max(date(eskrain)), length=149016)
precip_numeric <- as.numeric(eskrain)
esk.rain <- data.frame(date=as.Date(time.seq), precip=precip_numeric)

plot(esk.rain,type="h") # reality check - the maximum over 17 years is around 15 mm

daily.max <- apply(matrix(esk.rain$precip, ncol=24, byrow=T), 1, max)
weekly.max <- apply(matrix(daily.max, ncol=7, byrow=T), 1, max)
monthly.max <- apply(matrix(daily.max, ncol=30, byrow=T), 1, max)
```

(a) We first fit the GEV to weekly and monthly maxima with the return level for $m = 10$ years (set using `prob`) as one of the parameters:

```
m <- 10
fit.w <- fgev(weekly.max, prob=1/(m*52)) # fit to weekly maxima
fit.m <- fgev(monthly.max, prob=1/(m*12)) # fit to monthly maxima
#  compare profile log likelihoods for the two fits
par(mfrow=c(2,3))
plot(profile(fit.w))
plot(profile(fit.m))
```

You can see the estimates by typing (e.g.) `fit.w`. Do the return level estimates and confidence intervals for the two fits agree? Do you think one is more appropriate than the other, based on your previous analysis of the data?

Try refitting using return levels for higher values of m (25 years, 50 years, ...). What happens to the MLEs and confidence intervals? Why do you think this is? It might help to check the fits in the original parametrization:

```
(fgev(weekly.max))
(fgev(monthly.max))
```

and think about the fitted upper limit (if any) for rainfall.

(b) We now fit the generalized Pareto approximation with the return level for 10 years (set using `mper`) as one of the parameters. We use `npp` to set the number of background observations in each year:

```
u <- 5; m <- 10
(fit <- fpot(esk.rain$precip, threshold=u, mper=m, npp=365.25*24))
par(mfrow=c(1,2))
plot(profile(fit))  # profile log likelihoods
```

Is the profile plot for the return level consistent with the reality check and with the results in (a)? Try setting a larger value for `mper` (e.g., 25 or 100) and see how the profile changes. What is the effect of varying the threshold u?