Problem 1 If $X \sim \text{GPD}(\xi, \sigma)$, then $\text{P}(X > x) = (1 + \xi x/\sigma)_+^{-1/\xi}$ for $x > 0$, $\sigma > 0$ and $\xi \in \mathbb{R}$.

(a) Verify that $\lim_{\xi \to 0} \text{P}(X > x) = \exp(-x/\sigma)$ for $x > 0$.

(b) Find the support of X in the three cases $\xi > 0$, $\xi < 0$ and $\xi = 0$.

(c) Find $\text{E}(X)$.

(d) Show that if $x > 0$ and $x + u$ lies in the support of X, then $\text{P}(X > u + x \mid X > u) = (1 + \xi x/\sigma_u)_+^{-1/\xi}$, where $\sigma_u = \sigma + \xi u$, and deduce that $\text{E}(X - u \mid X > u) = (\sigma + \xi u)/(1 - \xi)$.

Hint for (c): if X is a positive-valued random variable with a finite mean, then $\text{E}(X) = \int_0^\infty \text{P}(X > x)\,\text{d}x$.

Problem 2 To run this and the next exercise, first load the packages:

```
load(evd,mev,ismev,scales,lubridate,gridExtra,ggplot2,dplyr,tidyr,ggdist,ggpubr,xts)
```

The problem below should familiarise you with some of the R tools used to fit the GEV and GPD models to data. The dataset used contains hourly precipitation (mm) from 1 January 1970 to 31 December 1986 at Eskdalemuir, in southern Scotland. To load it use

```
load("eskrain.RData")
```

This is a time series object, and it may be easier to work with a dataframe, for instance, using the code

```
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)
```

Try to plot the precipitation time series, and highlight the points corresponding to the exceedances for a threshold u using (e.g.)

```
u<-5
plot_esk <- plot(esk.rain, type="h", ylab="Hourly rainfall (mm)", xlab="Time")
points(esk.rain[esk.rain$precip > u,], col="red", cex=.25, pch=20)
abline(u, 0, col="red")
```

(a) We first fit the GEV to the daily, weekly and monthly maxima, made using the code

```
### start by taking daily maxima
daily.max<-apply(matrix(esk.rain$precip, ncol=24, byrow=T), 1, function(x)max(x))

### then use daily maxima to compute weekly maxima
weekly.max<-apply(matrix(daily.max, ncol=7, byrow=T), 1, function(x)max(x))

### finally, take monthly maxima, with months of 30 days (!)
monthly.max<-apply(matrix(daily.max, ncol=30, byrow=T), 1, function(x)max(x))
```

Now try to model each of these series and investigate the resulting fit. Do you expect problems with the daily maxima? You could use the code

```
fit.weekly<- fgev(weekly.max)
par(mfrow=c(2, 2))
plot(fit.weekly)
```

(b) For each fitted model, give the MLEs and their standard errors, and use them to compute 95% confidence intervals for the parameters. Profile log likelihood confidence intervals could instead be computed using

```
par(mfrow=c(1,3))
plot(profile(fit.monthly))
```

What do you conclude? Are these intervals similar to those from the MLEs?

(c) To test whether the shape parameter for the monthly maxima might be zero, use a formal test based on the likelihood ratio statistic between a GEV with three parameters and the Gumbel model, which sets $\xi = 0$:

```
fit.monthly0<- fgev(monthly.max, shape=0)
ratio<-fit.monthly0$dev-fit.monthly$dev
qchisq(0.95,1)
p<-1-pchisq(ratio1,1)
```

What do you conclude if you work with a significance level $\alpha = 0.05$?

Problem 3 Consider the Eskdalemuir data again, now considering the exceedances of a threshold u.

(a) The choice of u can be difficult., but a good choice should lead to a stable model for higher thresholds. The mean residual life plot (mean excess plot) is often used:

```
mrlplot(esk.rain$precip)
```

Comment on the stability (or not) for various thresholds u.

One can also look at the estimated parameters over a certain range of u under a point process approach:

```
tcplot(esk.rain$precip, tlim=c(0.1,7), model="pp", nt = 20)
```

Comment on the resulting stability plots. Do the results agree with those for the mean excess plot?

(b) For a chosen threshold u, fit the Poisson process likelihood to the exceedances and comment on the resulting estimates. To investigate the resulting fit you may use the code

```
### here we take u=5, but you can choose u based on part (a)
thresh.fit<-fpot(esk.rain$precip, threshold=5, model="pp", start=list(loc=10,
 scale=1.2, shape=0.1), npp=365.25*24)
### next, we look at the resulting diagnostic plots
par(mfrow=c(1,3))
plot(thresh.fit)
```

Comment on the fit.

(c) For the same threshold u as in (b), fit a GPD model to the exceedances and comment on the resulting estimates. Do you notice any differences from the estimates in (b)? You can use the code

```
### here we take a fixed threshold u=5, but you can choose u based on part (a)
thresh.fit_gpd<-fpot(esk.rain$precip, threshold=5, start=list(scale=1.2,shape=0.1))
### next, we look at the resulting diagnostic plots
par(mfrow=c(1,2))
plot(thresh.fit_gpd)
```