Solution 1

(a) In the POT case with $\xi \neq 0$ and following the note to slide 102, we seek to solve the equation

$$1 - p = 1 - p_u(1 + \xi(x_p - u)/\sigma_u\}_+^{-1/\xi}, \quad x_p > u,$$

so if $1 + \xi(x_p - u)/\sigma_u > 0$ and $p_u > 0$ this gives

$$p/p_u = (1 + \xi(x_p - u)/\sigma_u\}_+^{-1/\xi} \implies x_p = u + \sigma_u\{(p_u/p)^\xi - 1\}/\xi = u + \sigma_u[\exp\{\xi \log(p_u/p)\} - 1]/\xi,$$

and $[\exp\{\xi \log(p_u/p)\} - 1]/\xi \to \log(p_u/p)$ as $\xi \to 0$, which gives the formula for $\xi = 0$.

(b) In the case of maxima and again following the argument in the notes, we aim to solve

$$1 - p = G^{1/m}(x_p) = \exp\{-\Lambda(x_p)/m\} \implies \Lambda(x_p) = -m \log(1 - p),$$

where for $\xi \neq 0$ we have $\Lambda(x_p) = \{1 + \xi(x_p - \eta)/\tau\}_+^{-1/\xi}$. Thus, provided $1 + \xi(x_p - \eta)/\tau > 0$,

$$x_p = \eta + \tau \left[\{-m \log(1 - p)\}^{-\xi} - 1\right] /\xi.$$

Since $\{-m \log(1 - p)\}^{-\xi} = \exp[-\xi \log\{-m \log(1 - p)\}]$, we see that

$$\lim_{\xi \to 0} \left[\{-m \log(1 - p)\}^{-\xi} - 1\right] /\xi = -\log\{-m \log(1 - p)\},$$

which gives the stated formula for $\xi = 0$.

Solution 2

(a) The return level x_p is the solution to the equation $1 - p = G(x_p)$, where G is the GEV distribution function, so we need to solve

$$1 - p = \exp\left\{-\left(1 + \xi\frac{x_p - \eta}{\tau}\right)_+^{-1/\xi}\right\},$$

which immediately gives the stated formula.

(b) The log likelihood function is defined as $\ell(\eta, \tau, \xi) = \log f(y_1, \ldots, y_n; \eta, \tau, \xi)$, so it is unchanged by 1–1 transformations of the parameters, such as setting $\eta = x_p - \tau a_p(\xi)$. Hence

$$\ell_\mathrm{p}(x_p) = \max_{\tau,\xi} \ell^*(x_p, \tau, \xi) = \max_{\tau,\xi} \ell\{x_p - \tau a_p(\xi), \tau, \xi\}.$$

The simplest approach to computation is to fix a grid of values of x_p and then optimise the function $\ell\{x_p - \tau a_p(\xi), \tau, \xi\}$ at each point of such a grid. Then join the dots ...

Solution 3

(a) Running the code below gives Figure 1.

```
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))
```

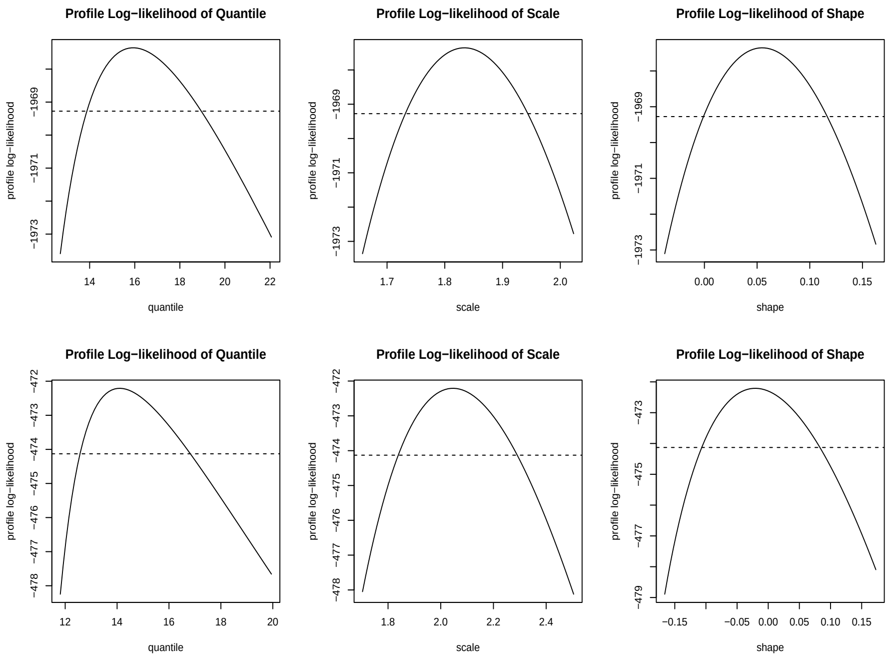

Figure 1: Profile log-likelihood plots for weekly maxima (top panel) and monthly maxima (bottom panel) of precipitation.

The profile log-likelihood plots show certain differences between the return levels and their confidence intervals, with higher return levels estimated from the weekly maxima. In addition, the confidence intervals for the return levels are asymmetric, in particular for monthly maxima, so one should not compute symmetric confidence intervals based directly on the MLE (e.g., slide 24). For a closer inspection, we look at the estimates using the code

```
fit.w

Call: fgev(x = weekly.max, prob = 1/(m * 52))
Deviance: 3934.709

Estimates
quantile     scale     shape
15.92728   1.83354   0.05482

Standard Errors
quantile     scale     shape
 1.29123   0.05415   0.03082

Optimization Information
  Convergence: successful
  Function Evaluations: 49
  Gradient Evaluations: 13

fit.m

Call: fgev(x = monthly.max, prob = 1/(m * 12))
Deviance: 944.4168

Estimates
```

```
quantile     scale     shape
14.10006   2.04654  -0.02089

Standard Errors
quantile     scale     shape
 1.01697   0.11445   0.04842
```

The scale and shape parameter are very close to the MLEs from the previous sheet, shown in Table 1.

	$\widehat{\mu}_{\text{MLE}}$	$\widehat{\sigma}_{\text{MLE}}$	$\widehat{\xi}_{\text{MLE}}$
Weekly	2.25 (0.07)	1.83 (0.05)	0.05 (0.03)
Monthly	4.78 (0.16)	2.05 (0.11)	-0.02 (0.05)

Table 1: Parameter estimates of the fitted GEV models from the previous exercise sheet.

We can check whether the return levels estimated above correspond to those computed using the parameters in Table 1; recall from slide 102 that we can compute return levels from the parameters of the GEV via the formula

$$x_{p,c} = \mu_c + \tfrac{\sigma_c}{\xi_c}\left[\{-m_c \log(1-p_c)\}^{-\xi_c} - 1\right], \tag{1}$$

where we use the subscript c to denote the choice of time period of the maxima used for estimation, i.e., weekly or monthly.

Here one has to pay attention to the number of background observations. For the weekly estimates we apply (1) with $m_w = 24 \times 7 = 168$ hourly observations per week, and take $p_w = 1/(52 \times m \times m_w)$. Similarly, we take $m_m = 24 \times 31$ for the monthly background observations and set $p_m = 1/(12 \times m \times m_m)$. Plugging in the estimates from Table 1 we compute $\widehat{x}_{p,w} = 15.93$ and $\widehat{x}_{p,m} = 14.11$, using the code

```
(c(fit.weekly$estimate[1]+fit.weekly$estimate[2]/fit.weekly$estimate[3]
*(((-24*7*log(1-1/(24*7*52*m)))^-fit.weekly$estimate[3])-1),
fit.monthly$estimate[1]+fit.monthly$estimate[2]/fit.monthly$estimate[3]
*(((-24*31*log(1-1/(24*31*12*m)))^-fit.monthly$estimate[3])-1)))

15.93281    14.10830
```

Both estimates are very close to those from `fit.w` and `fit.m`.

We repeat the same procedure, but now for $m \in \{25, 50\}$ using the code below, which gives Figures 2 and 3.

```
m <- 25
fit.w25 <- fgev(weekly.max, prob=1/(m*52)) # fit to weekly maxima
fit.m25 <- 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.w25))
plot(profile(fit.m25))

m <- 50
fit.w50 <- fgev(weekly.max, prob=1/(m*52)) # fit to weekly maxima
fit.m50 <- 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.w50))
plot(profile(fit.m50))
```

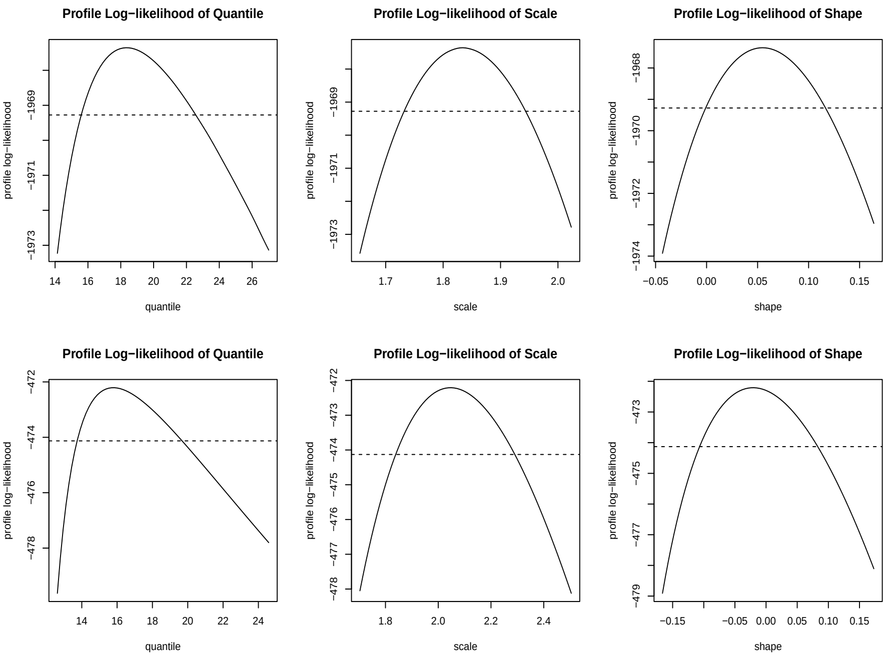

Figure 2: Profile log-likelihood plots for weekly maxima (top panel) and monthly maxima (bottom panel) of precipitation for the 25 year return period.

Both figures show higher return levels, as expected also from (1), since p_c is smaller, but there is a larger increase in the return level from the weekly maxima; we recall that the estimated shape parameter $\widehat{\xi}_w > 0$, whereas $\widehat{\xi}_m < 0$, giving infinite and finite upper bounds, respectively. Therefore, for larger m we would expect $x_{p,w}$ to approach ∞ as $m \to \infty$, whereas $x_{p,m}$ should approach the upper bound $\mu_m - \sigma_m/\xi_m$ (recall Problem 1 from Sheet 6.)

The increase in estimated return levels is accompanied by the increased uncertainty from extrapolating into more extreme regions of the tail. For instance, one can argue by taking into the account the uncertainty of the MLEs in formula (1); if $\xi_w > 0$, we have the term

$$\frac{\widehat{\sigma}_c}{\widehat{\xi}_c}\left[\{-m_w \log(1-p_w)\}^{-\xi_w} - 1\right],$$

where $\left[\{-m_w \log(1-p_w)\}^{-\xi_w} - 1\right] \to \infty$ as $m \to \infty$, and therefore the variance of the estimated $\widehat{x}_{p,w}$ will also increase, thus giving larger confidence intervals.

(b) We now follow the Poisson process approach and use the following code, which gives Figure 4.

```
 u <- 5; m <- 10
 (fit <- fpot(esk.rain$precip, threshold=u, mper=m, npp=365.25*24))
 Call: fpot(x = esk.rain$precip, threshold = u, npp = 365.25 * 24, mper = m)
Deviance: 1058.954

Threshold: 5
Number Above: 356
Proportion Above: 0.0024

Estimates
  rlevel     shape
14.78335   0.06699
```

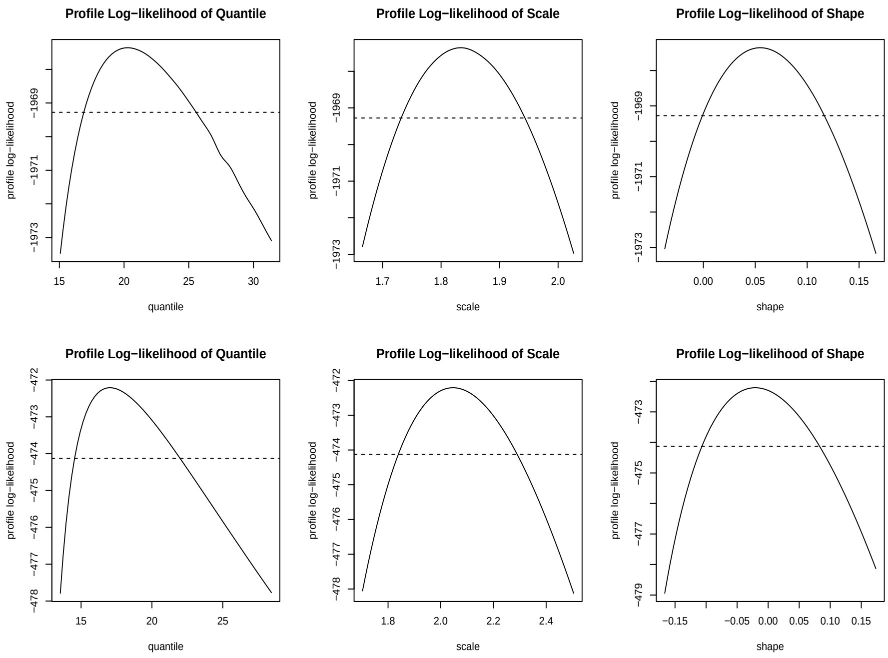

Figure 3: Profile log-likelihood plots for weekly maxima (top panel) and monthly maxima (bottom panel) of precipitation for the 50 year return period.

```
Standard Errors
 rlevel    shape
1.14293  0.05382

 par(mfrow=c(1,2))
 plot(profile(fit))
```

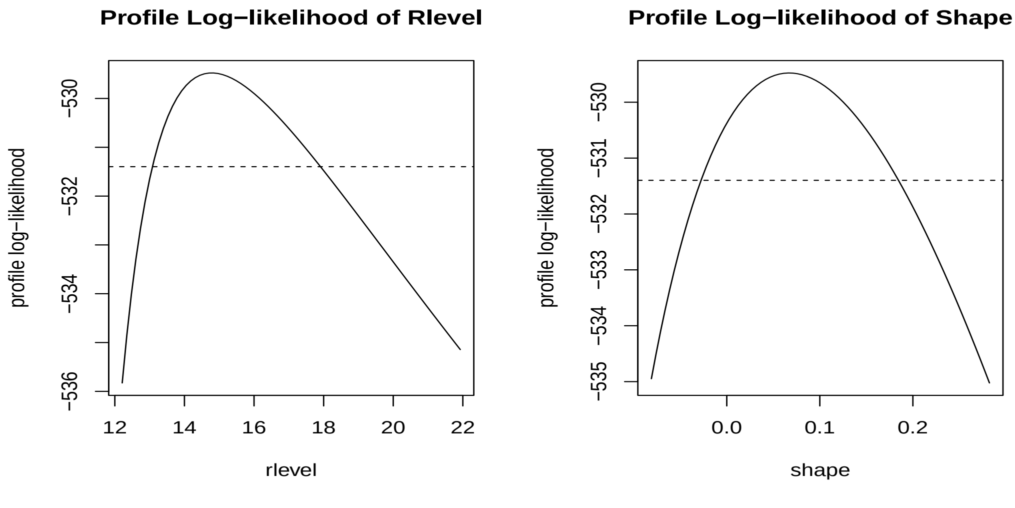

Figure 4: Profile log-likelihood of the return level (left) and shape parameter (right) for the Poisson process.

We also run the following code and, to compare the results, illustrate in Figure 5 only the profile log-likelihood plots of the return levels for different periods m.

```
m <- 25
```

```
(fit25 <- fpot(esk.rain$precip, threshold=u, mper=m, npp=365.25*24))

m <- 50
(fit50 <- fpot(esk.rain$precip, threshold=u, mper=m, npp=365.25*24))
```

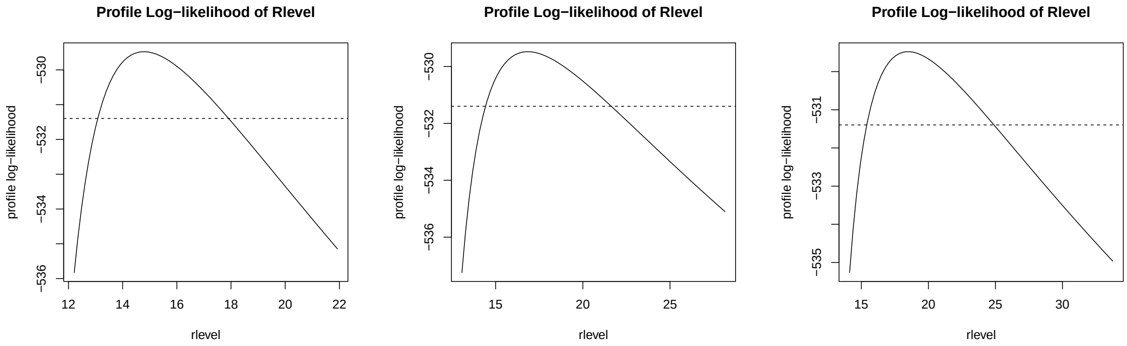

Figure 5: Profile log-likelihood plots for 10 (left), 25 (center) and 50 (right) year return level for the Poisson process.

Figures 1 and 4 show a smaller difference between the estimated return levels from the Poisson process and the monthly maxima for $m = 10$; however, note that the estimated shape parameter under the Poisson process approach is $\widehat{\xi}_{PP} = 0.07$, whereas it is negative for `fit.m`. The confidence intervals are also similar. However, as we consider a longer return period, we notice that both the return level and the confidence intervals are larger for the Poisson process approach relative to the fit from the monthly maxima in (a).

We have from the reality-check that the hourly maximum over 17 years is 15 mm, so the return level $\widehat{x}_{PP} = 14.8$ mm for $m = 10$ is a fairly reasonable estimate.

We now let the threshold u vary using the code below. We observe in Figure 6 differences in the estimated return levels and the shape parameters, as well as the resulting confidence intervals due to the additional or fewer exceedances, an example of bias-variance tradeoff.

```
u <- 2; m <- 10
(fitu2 <- fpot(esk.rain$precip, threshold=u, mper=m, npp=365.25*24))

Call: fpot(x = esk.rain$precip, threshold = u, npp = 365.25 * 24, mper = m)
Deviance: 9080.691

Threshold: 2
Number Above: 3426
Proportion Above: 0.023

Estimates
  rlevel     shape
13.47979   0.02993

Standard Errors
 rlevel    shape
0.58314  0.01597

par(mfrow=c(1,2))
plot(profile(fitu2))
```

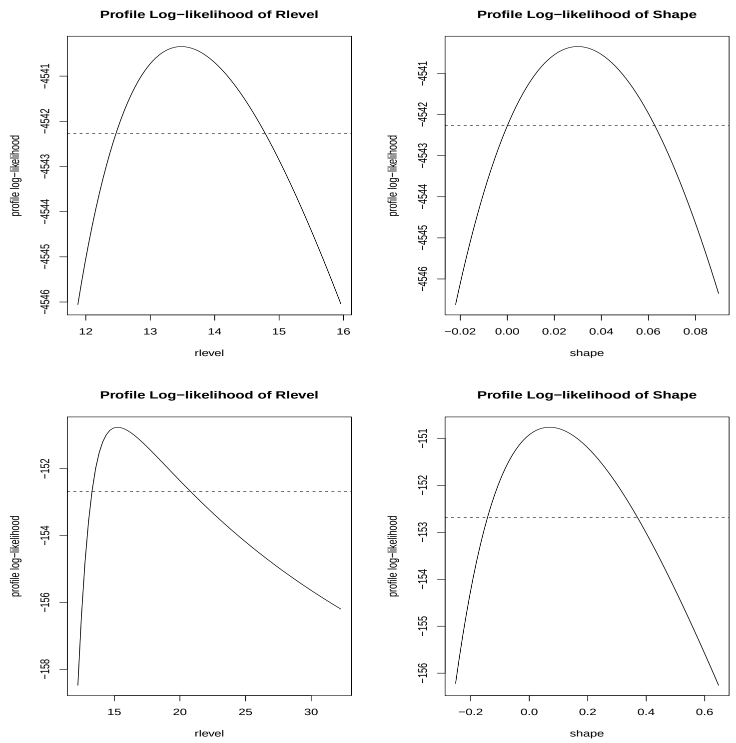

Figure 6: Profile log-likelihood plots for $u = 2$ (top) and $u = 7$ (bottom) for the Poisson process.

```
u <- 7; m <- 10
(fitu7 <- fpot(esk.rain$precip, threshold=u, mper=m, npp=365.25*24))

Call: fpot(x = esk.rain$precip, threshold = u, npp = 365.25 * 24, mper = m)
Deviance: 301.5227

Threshold: 7
Number Above: 91
Proportion Above: 6e-04

Estimates
  rlevel     shape
15.24709   0.06932

Standard Errors
rlevel   shape
1.5124  0.1285

par(mfrow=c(1,2))
plot(profile(fitu7))
```