Solution 1

(a) As $Y_2 < Y_1$, knowledge that $Y_1 = y_1$ means that $Y_2 < y_1$. Hence

$$\mathrm{P}(Y_2 < y_2 \mid Y_1 = y_1) = \mathrm{P}(Y_2 < y_2 \mid Y_2 < y_1) = \mathrm{P}(Y_2 < y_2)/\mathrm{P}(Y_2 < y_1), \quad y_2 < y_1.$$

Now Y_2 is the limiting variable corresponding to the maximum of an infinite number of rescaled variables $(X_j - b_m)/a_m$, so it must have the same distribution as Y_1, except that $Y_2 < Y_1$. (Note for later use that the same argument applies to all the Y_j, with the ordering imposing constraints.) Hence

$$\mathrm{P}(Y_2 < y_2 \mid Y_1 = y_1) = \mathrm{P}(Y_2 < y_2)/\mathrm{P}(Y_2 < y_1) = e^{-\Lambda(y_2)}/e^{-\Lambda(y_1)} = \exp\{\Lambda(y_1) - \Lambda(y_2)\}, \quad y_2 < y_1.$$

Differentiating the conditional distribution with respect to y_2 yields

$$f(y_2 \mid y_1) = \{-\dot{\Lambda}(y_2)\} \exp\{\Lambda(y_1) - \Lambda(y_2)\}, \quad y_2 < y_1,$$

and clearly $f(y_1) = \{-\dot{\Lambda}(y_1)\} \exp\{-\Lambda(y_1)\}$. Hence

$$f(y_1, y_2) = f(y_2 \mid y_1) f(y_1) = \{-\dot{\Lambda}(y_2)\} \exp\{\Lambda(y_1) - \Lambda(y_2)\} \times \{-\dot{\Lambda}(y_1)\} \exp\{-\Lambda(y_1)\},$$

which reduces to the given density. For the induction, suppose the expression in the question holds for the first $r-1$ order statistics. Then for $y_1 > \cdots > y_r$ we have

$$\begin{aligned}\mathrm{P}(Y_r < y_r \mid Y_1 = y_1, Y_2 = y_2, \ldots, Y_{r-1} = y_{r-1}) = \mathrm{P}(Y_r < y_r \mid Y_r < y_{r-1}) &= \mathrm{P}(Y_r < y_r)/\mathrm{P}(Y_r < y_{r-1}) \\ &= \exp\{\Lambda(y_r) - \Lambda(y_{r-1})\}\end{aligned}$$

and

$$f(y_r \mid Y_1 = y_1, \ldots, Y_{r-1} = y_{r-1}) = \{-\dot{\Lambda}(y_r)\} \exp\{\Lambda(y_{r-1}) - \Lambda(y_r)\}.$$

The inductive hypothesis gives

$$f(y_1, \ldots, y_{r-1}) = \exp\{-\Lambda(y_{r-1})\} \prod_{i=1}^{r-1} \{-\dot{\Lambda}(y_i)\},$$

which yields the required expression, i.e.,

$$f(y_1, \ldots, y_r) = f(y_r \mid Y_1 = y_1, \ldots, Y_{r-1} = y_{r-1}) f(y_1, \ldots, y_{r-1}) = \exp\{-\Lambda(y_r)\} \prod_{i=1}^{r} \{-\dot{\Lambda}(y_i)\}. \quad (1)$$

(b) To obtain the marginal density of Y_r we must integrate the joint density over the set $\mathcal{S} = \{(y_1, \ldots, y_{r-1}) : y_r < y_{r-1} < \cdots < y_1\}$. So we first integrate over $y_1 \in (y_2, \infty)$, then over $y_2 \in (y_3, \infty)$, and so on up to integration over $y_{r-1} \in (y_r, \infty)$. Note that $\Lambda(\infty) = 0$, since $\mathrm{P}(Y_1 \leq \infty) = \exp\{-\Lambda(\infty)\} = 1$. Using the density in (**??**) we obtain

$$\begin{aligned} f(y_r) &= \int_{y_r}^{\infty} \cdots \int_{y_3}^{\infty} \int_{y_2}^{\infty} \exp\{-\Lambda(y_r)\} \prod_{j=1}^{r} \{-\dot{\Lambda}(y_j)\} \, \mathrm{d}y_1 \cdots \mathrm{d}y_{r-1} \\ &= \exp\{-\Lambda(y_r)\}\{-\dot{\Lambda}(y_r)\} \int_{y_r}^{\infty} \cdots \int_{y_3}^{\infty} \int_{y_2}^{\infty} \prod_{j=1}^{r-1} \{-\dot{\Lambda}(y_j)\} \, \mathrm{d}y_1 \cdots \mathrm{d}y_{r-1}. \end{aligned} \quad (2)$$

The innermost integral (over y_1) gives

$$\prod_{j=2}^{r-1}\{-\dot\Lambda(y_j)\}\int_{y_2}^{\infty}\{-\dot\Lambda(y_1)\}\,\mathrm{d}y_1 = \prod_{j=2}^{r-1}\{-\dot\Lambda(y_j)\}\times[-\Lambda(u)]_{y_2}^{\infty} = \prod_{j=2}^{r-1}\{-\dot\Lambda(y_j)\}\times\Lambda(y_2).$$

The next integral (over y_2) gives

$$\prod_{j=3}^{r-1}\{-\dot\Lambda(y_j)\}\int_{y_3}^{\infty}\{-\dot\Lambda(y_2)\}\Lambda(y_2)\,\mathrm{d}y_2 = \prod_{j=3}^{r-1}\{-\dot\Lambda(y_j)\}\times\left[-\Lambda(u)^2/2!\right]_{y_3}^{\infty} = \prod_{j=3}^{r-1}\{-\dot\Lambda(y_j)\}\times\Lambda(y_3)^2/2!,$$

and repeating the argument leads to the entire integral in (**??**) being $\Lambda(y_r)^{r-1}/(r-1)!$, giving

$$f(y_r) = \exp\{-\Lambda(y_r)\}\{-\dot\Lambda(y_r)\}\times\Lambda(y_r)^{r-1}/(r-1)!, \quad y_r\in\mathbb{R},$$

as required.

(c) Given the joint density from (a) and the density of y_{r+1} from (b) we compute

$$\begin{aligned}
f(y_1,\dots,y_r\mid y_{r+1}=u) &= \frac{f(y_1,\dots,y_{r+1})}{f(y_{r+1})}\\
&= \frac{\prod_{i=1}^{r+1}\{-\dot\Lambda(y_i)\}\exp\{-\Lambda(y_{r+1})\}}{\{-\dot\Lambda(y_{r+1})\}\dfrac{\Lambda(y_{r+1})^r}{r!}\exp\{-\Lambda(y_{r+1})\}}\\
&= r!\prod_{i=1}^{r}\frac{\{-\dot\Lambda(y_i)\}}{\Lambda(u)},\quad y_1>\cdots>y_r>u.
\end{aligned}$$

Let $X_1,\dots,X_r$ be i.i.d. random variables with distribution function $H(y)=1-\Lambda(y)/\Lambda(u)$ for $y>u$. Then by independence

$$h(y_1,\dots,y_r) = \prod_{i=1}^{r}h(y_i) = \prod_{i=1}^{r}\frac{\{-\dot\Lambda(y_i)\}}{\Lambda(u)}.$$

Note that the distribution $H(y)=1-\Lambda(y)/\Lambda(u)$ corresponds to a generalized Pareto distribution, because $\Lambda(y)=\{1+\xi(y-\eta)/\tau\}_+^{-1/\xi}$, so

$$H(y) = \begin{cases}1-(1+\xi x/\sigma_u)_+^{-1/\xi}, & \xi\neq 0,\\ 1-\exp(-x/\sigma_u), & \xi=0,\end{cases}$$

where $\sigma_u=\tau+\xi(u-\eta)$.

Solution 2

(a) In the POT case with $\xi\neq 0$ and following the note to slide 83, 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 3

(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 ...