Solution 1

(a) Since $G(y) = \exp\{-\Lambda(y)\}$, $G(y)^T = \exp\{-T\Lambda(y)\}$, so we need to consider $T\Lambda(y)$. This equals

$$\begin{aligned} T\left(1+\xi\frac{y-\eta}{\tau}\right)_+^{-1/\xi} &= \left\{1+\left(T^{-\xi}-1\right)+\xi\frac{y-\eta}{\tau T^{\xi}}\right\}_+^{-1/\xi} \\ &= \left\{1+\xi\frac{y-\eta+\tau T^{\xi}(T^{-\xi}-1)/\xi}{\tau T^{\xi}}\right\}_+^{-1/\xi} \\ &= \left\{1+\xi\frac{y-\eta-\tau(T^{\xi}-1)/\xi}{\tau T^{\xi}}\right\}_+^{-1/\xi} \\ &= \left(1+\xi_T\frac{y-\eta_T}{\tau_T}\right)_+^{-1/\xi_T}, \end{aligned}$$

where $\eta_T = \eta + \tau(T^{\xi}-1)/\xi$, $\tau_T = \tau T^{\xi}$ and $\xi_T = \xi$. This proves the result.

(b) We can write

$$G(y;\eta,\tau,\xi) = \exp\left[-\exp\left\{-\frac{1}{\xi}\log\left(1+\xi\frac{y-\eta}{\tau}\right)_+\right\}\right].$$

and since $\log(1+a) = a - a^2/2 + \cdots$ as $a \to 0$, we have

$$\lim_{\xi\to 0} -\frac{1}{\xi}\log\left(1+\xi\frac{y-\eta}{\tau}\right)_+ = -\frac{(y-\eta)}{\tau},$$

as $1+\xi(y-\eta)/\tau > 0$ for small enough ξ and any $(y-\eta)/\tau$. The function $\exp\{-\exp(-x)\}$ is continuous for all x, so

$$\lim_{\xi\to 0} G(y;\eta,\tau,\xi) = \exp\left[-\exp\left\{-(y-\eta)/\tau\right\}\right].$$

Furthermore, $\eta_T = \eta + \tau(T^{\xi}-1)/\xi \to \eta + \tau\log T$, $\tau_T \to \tau$ and $\xi_T \to 0$ as $\xi \to 0$.

(c) As $T \to \infty$, we have $\eta_T \to \begin{cases} +\infty, & \xi > 0 \\ +\infty, & \xi = 0 \\ \eta - \tau/\xi, & \xi < 0 \end{cases}$, $\tau_T \to \begin{cases} +\infty, & \xi > 0 \\ \tau, & \xi = 0 \\ 0, & \xi < 0 \end{cases}$ and $\xi_T \to \xi$.

Increasing T corresponds to taking maxima over a larger block of variables, and the maximum of 100 random variables is always higher than the maximum over only 10 of them. Intuitively, we therefore expect the distribution to shift to the right as we increase T, so the behaviour of η_T makes sense. The behaviour of τ_T is less intuitive, but when $T \to \infty$ and $\xi > 0$ we see that τ_T increases, i.e., the GEV becomes more dispersed, and when $\xi < 0$, $\tau_T \to 0$. i.e., the limiting distribution becomes less dispersed, because the largest values bunch up near the finite upper support point.

(d) The support of $Y \sim G$ corresponds to the set of values $S = \{y : 0 < G(y) < 1\}$.

When $\xi > 0$, $G(y) > 0 \iff \left(1+\xi\frac{y-\eta}{\tau}\right) > 0 \iff y > \eta - \tau/\xi$, so $S = (\eta - \tau/\xi, +\infty)$.

When $\xi < 0$, $G(y) < 1 \iff \left(1+\xi\frac{y-\eta}{\tau}\right) > 0 \iff y < \eta - \tau/\xi$, so $S = (-\infty, \eta - \tau/\xi)$.

When $\xi = 0$, $0 < G(y) < 1$ for all values of $y \in \mathbb{R}$, so $S = \mathbb{R}$.

Solution 2

(a) $X \sim \mathrm{GEV}(0,1,\xi)$ has CDF $G_0(x) = \exp\{-\Lambda_0(x)\}$, say, where $\Lambda_0(x) = (1+\xi x)_+^{-1/\xi}$, and

$$\mathrm{P}(\eta + \tau X \le y) = \mathrm{P}\{X \le (y-\eta)/\tau\} = \exp\left[-\Lambda_0\{(y-\eta/\tau\}\right] = \exp\{-\Lambda(y)\} = \mathrm{P}(Y \le y),$$

and hence $Y \stackrel{\mathrm{D}}{=} \eta + \tau X$. This implies that $\mathrm{E}(Y) = \eta + \tau E(X)$ and $\mathrm{var}(Y) = \tau^2 \mathrm{var}(X)$.

(b) The PDF of X is

$$\frac{\mathrm{d}G_0(x)}{\mathrm{d}x} = \{-\dot{\Lambda}_0(x)\} \exp\{-\Lambda_0(x)\},$$

so

$$\mathrm{E}\left\{X^r G_0(X)^s\right\} = \int x^r \exp\{-s\Lambda_0(x)\}\{-\dot{\Lambda}_0(x)\} \exp\{-\Lambda_0(x)\}\, \mathrm{d}x.$$

If we write $z = (s+1)\Lambda_0(x)$, we need expressions for x and $\{-\dot{\Lambda}_0(x)\}$ in terms of z. The second is easy, because $\mathrm{d}z/\mathrm{d}x = (s+1)\{-\dot{\Lambda}_0(x)\}$, and

$$x = \xi^{-1}\left\{\left(\frac{z}{s+1}\right)^{-\xi} - 1\right\}.$$

Hence

$$\mathrm{E}\left\{X^r G(X)^s\right\} = \frac{1}{s+1}\int_0^\infty \xi^{-r}\left\{\left(\frac{z}{s+1}\right)^{-\xi} - 1\right\}^r e^{-z}\, \mathrm{d}z.$$

Setting $s = 0$ and $r = 1$, and provided $\xi < 1$ so the gamma function is finite, we have

$$\mathrm{E}(X) = \frac{1}{\xi}\int_0^\infty \left(z^{-\xi} - 1\right) e^{-z} \mathrm{d}z = \frac{1}{\xi}\left(\int_0^\infty z^{-\xi} e^{-z} \mathrm{d}z - 1\right) = \frac{\Gamma(1-\xi) - 1}{\xi}.$$

With $s = 0$ and $r = 2$, we have, provided $\xi < 1/2$,

$$\begin{aligned}
\mathrm{E}\left(X^2\right) = \frac{1}{\xi^2}\int_0^\infty \left(z^{-\xi} - 1\right)^2 e^{-z}\mathrm{d}z &= \frac{1}{\xi^2}\int_0^\infty \left(z^{-2\xi} - 2z^{-\xi} + 1\right) e^{-z}\mathrm{d}z \\
&= \frac{1}{\xi^2}\left(\int_0^\infty z^{-2\xi} e^{-z}\mathrm{d}z - 2\int_0^\infty z^{-\xi} e^{-z}\mathrm{d}z + 1\right) \\
&= \frac{\Gamma(1-2\xi) - 2\Gamma(1-\xi) + 1}{\xi^2}.
\end{aligned}$$

Note that $\mathrm{E}(X^2)$ exists if and only if $\xi < 1/2$, which is therefore also a condition for $\mathrm{var}(X)$ to be finite. In fact the computations above yield

$$\mathrm{var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2 = \frac{\Gamma(1-2\xi) - \Gamma(1-\xi)^2}{\xi^2}.$$

Solution 3

(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 $\mathrm{P}(Y_1 \le \infty) = \exp\{-\Lambda(\infty)\} = 1$. Using the density in (1) 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 (2) 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 $\sigma_u=\tau+\xi(u-\eta)$.