Solution 1

(a) Clearly $Y_{(n)} \le y$ iff all the $Y_j \le y$, and since the Y_j are independent, $\mathrm{P}(Y_{(n)} \le y) = F(y)^n$, so the density function is $nf(y)F(y)^{n-1}$.

Likewise $Y_{(1)} > y$ iff all the $Y_j > y$, so $\mathrm{P}(Y_{(1)} > y) = \{1 - F(y)\}^n$, giving $\mathrm{P}(Y_{(1)} \le y) = 1 - \{1 - F(y)\}^n$ and density function $nf(y)\{1 - F(y)\}^{n-1}$.

(b) The probability density for a particular permutation of $Y_1, \ldots, Y_n$ to equal $y_1, \ldots, y_n$ is $\prod_{j=1}^n f(y_j)$. But since there are $n!$ such permutations, all with the same density, the density values for all $n!$ permutations must be added to give the joint density for the order statistics, i.e.,

$$f_{Y_{(1)},\ldots,Y_{(n)}}(y_1, \ldots, y_n) = n! \prod_{j=1}^n f(y_j), \quad y_1 < \cdots < y_n.$$

We obtain the marginal density of $Y_{(n)}$ by integration over y_1, then y_2, ..., up to y_{n-1}. The first step gives that the joint density of $Y_{(2)}, \ldots, Y_{(n)}$ is

$$\int_{-\infty}^{y_2} n! \prod_{j=1}^n f(y_j)\,\mathrm{d}y_1 = n! F(y_2) \prod_{j=2}^n f(y_j),$$

and then integration over $y_2 \in (-\infty, y_3)$ gives $n!F(y_3)^2 \prod_{j=3}^n f(y_j)/2!$, and then similar integrations over $y_3, \ldots, y_{n-1}$ lead to

$$n!F(y_n)^{n-1} \prod_{j=n}^n f(y_j)/(n-1)! = nf(y_n)F(y_n)^{n-1}.$$

The computation for the minimum is similar, integrating over $y_n \in (y_{n-1}, \infty)$, then over $y_{n-1} \in (y_{n-2}, \infty)$, etc. and resulting in $nf(y_1)\{1 - F(y_1)\}^{n-1}$.

(c) The uniform density on $(0, a)$ is $1/a$ for $y \in (0, a)$, which gives $n!/a^n$ for $0 < y_1 < \cdots < y_n < a$ when inserted into (b).

Solution 2

(a) As $\min(Y_1, \ldots, Y_r) > x$ if and only if $Y_1 > x, \ldots, Y_r > x$, we have

$$\mathrm{P}\{\min(Y_1, \ldots, Y_r) \le x\} = 1 - \mathrm{P}\{\min(Y_1, \ldots, Y_r) > x\} = 1 - \mathrm{P}(Y_1 > x)^r = 1 - \exp(-r\lambda x), \quad x > 0,$$

and for $x, y > 0$, $\mathrm{P}(Y - x > y \mid Y > x)$ equals

$$\frac{\mathrm{P}(Y - x > y, Y > x)}{\mathrm{P}(Y > x)} = \frac{\mathrm{P}(Y > y + x)}{\mathrm{P}(Y > x)} = \exp\{-\lambda(x + y)\}/\exp(-\lambda x) = \exp(-\lambda y),$$

as required.

(b) As $\mathrm{P}(E_j/\lambda \le x) = \mathrm{P}(E_j \le \lambda x) = 1 - \exp(-\lambda x) = \mathrm{P}(Y_j \le x)$, we have $Y_j \stackrel{\mathrm{D}}{=} E_j/\lambda$. We argue as follows:

- $Y_{(1)}$ is the smallest of n independent exponential variables, so it is exponential with parameter $n\lambda$ and therefore we can write $Y_{(1)} \stackrel{\mathrm{D}}{=} E_1/(n\lambda)$;
- the remaining $n - 1$ variables have the lack of memory property, so given that $Y_{(1)} = x$ the remaining $Y_j - x$ have exponential distributions with parameter λ. Thus $Y_{(2)} - Y_{(1)}$ is the minimum of $n - 1$ exponential variables, i.e., $Y_{(2)} - Y_{(1)} \stackrel{\mathrm{D}}{=} E_2/\{(n-1)\lambda\}$;

- iterating the argument by successively conditioning on $Y_{(2)}, \dots, Y_{(n-1)}$ and obtaining the distributions of $Y_{(3)} - Y_{(2)}, \dots, Y_{(n)} - Y_{(n-1)}$ gives the stated representation.

(c) A standard exponential variable has mean and variance both equal to 1, so

$$\mathrm{E}(Y_{(r)}) = \frac{1}{\lambda}\sum_{j=1}^{r}\frac{1}{n+1-j}, \quad \mathrm{cov}(Y_{(r)}, Y_{(s)}) = \frac{1}{\lambda^2}\sum_{j=1}^{m}\frac{1}{(n+1-j)^2}, \quad r, s, \in \{1, \dots, n\},$$

with $m = \min(s, r)$ and the second formula giving the variance when $r = s$. These are useful in assessing QQplots, since they give the expectation and variance of (individual) order statistics.

Solution 3

(a) We saw in the lectures that the joint density of the times of the n events is

$$e^{-\mu(0,t_0)}\prod_{j=1}^{n}\dot{\mu}(t_j), \quad 0 < t_1 < \cdots < t_n < t_0,$$

and on setting $\dot{\mu}(t) = \lambda$ this reduces to $\lambda^n e^{-\lambda t_0}$. If we successively integrate over t_1, then t_2 etc. we get the integral in the question, which reduces to $(\lambda t_0)^n e^{-\lambda t_0}/n!$, i.e., the density function of a Poisson variable with mean λt_0.

(b) The log likelihood based on the times has derivatives $\ell'(\lambda) = n/\lambda - t_0$ and $\ell''(\lambda) = -n/\lambda^2$ and the sole solution to the equation $\ell'(\lambda) = 0$ is $\widehat{\lambda} = n/t_0$ (the empirical rate of events per unit of time). The second derivative is negative (unless $n = 0$, in which case $\ell(\lambda)$ is maximised at $\lambda = 0$), so $\widehat{\lambda}$ gives a maximum. The expected information is $\mathrm{E}\{-\ell''(\lambda)\} = \mathrm{E}\{N(t_0)\}/\lambda^2 = \lambda t_0/\lambda^2 = t_0/\lambda$, where we have replaced the observed n by the corresponding random variable $N(t_0)$ to take the expectation. The information increases linearly with the length of the observation period, which is a surrogate for the sample size.

The log likelihood based on $N(t_0)$ would be

$$\log \mathrm{P}\{N(t_0) = n\} = n\log\lambda + n\log t_0 - \lambda t_0 - \log n!, \quad \lambda > 0,$$

which equals $\ell(\lambda)$ apart from additive constants, so the MLE and expected information will be the same. It is clear from $\ell(\lambda)$ that under this model the number of events is a sufficient statistic (*statistique exhaustive* — check this if unsure), so under this particular model the times are irrelevant for inference; what matters is the number of events.

(c) This conditional density is

$$\lambda^n e^{-\lambda t_0} \Big/ \left\{(\lambda t_0)^n e^{-\lambda t_0}/n!\right\} = \frac{n!}{t_0^n}, \quad 0 < t_1 < \cdots < t_n < t_0.$$

Hence the conditional distribution of the event times, given that there are n events, is that of the order statistics of a uniform sample on $(0, t_0)$.

(d) The plot clearly shows that the data are under the diagonal line that would correspond to a uniform sample, and this is confirmed by the Kolmogorov–Smirnov test, which has a tiny P-value. Hence there is strong evidence against the model. The connection with (c) is that the rescaled data t_j/t_0 should be a uniform sample on $U(0, 1)$ if the model is correct, and clearly this is not the case.