Solution 1

(a) If $X_1 \neq X_2$ then $(X_1+X_2)/2 = (\theta+1+\theta-1)/2 = \theta$, so in this case $\mathrm{P}(\theta \in \mathcal{C}) = 1$. If $X_1 = X_2$ then $\mathrm{P}(X_1 - 1 = \theta) = \mathrm{P}(X_1 = \theta + 1) = 1/2$. Moreover $\mathrm{P}(X_1 = X_2) = \mathrm{P}(X_1 = X_2 = \theta - 1) + \mathrm{P}(X_1 = X_2 = \theta + 1) = (1/2)^2 + (1/2)^2 = 1/2$, by independence, so

$$\mathrm{P}(\theta \in \mathcal{C}) = \mathrm{P}(\theta \in \mathcal{C} \mid X_1 = X_2)\mathrm{P}(X_1 = X_2) + \mathrm{P}(\theta \in \mathcal{C} \mid X_1 \neq X_2)\mathrm{P}(X_1 \neq X_2) = 1/2 \times 1/2 + 1 \times 1/2 = 3/4,$$

as required. This does not seem sensible, since if $X_1 \neq X_2$ we know θ with certainty, i.e., our confidence set has coverage 100%.

(b) The sample space consists of the three lines $\mathcal{Y} = \{(x, y) : |x - y| = 0, \pm 1\}$. Two possible classes of reference sets are $\mathcal{S}_1(x) = \{(x, y) : |x - y| = 1\}$ and $\mathcal{S}_2(x) = \{(x, y) : x = y\}$. If $(x_1, x_2) \in \mathcal{S}_1(x)$ for some x, then $\mathcal{C}_1 = \{(x_1 + x_2)/2\}$ is a 100% confidence set, and if $(x_1, x_2) \in \mathcal{S}_2(x)$ for some x, then $\mathcal{C}_2 = \{x_1 + 1, x_1 - 1\}$ is a 100% confidence set.

Solution 2

(a) In this case

$$\ell(\theta) = -y/\theta - \log\theta, \quad \tilde{\ell}(\phi) = \log\phi - \phi y, \quad \phi = 1/\theta, \quad \theta > 0,$$

so $\partial\phi/\partial\theta = -1/\theta^2 = -\phi^2$, $\partial^2\phi/\partial\theta^2 = 2/\theta^3 = 2\phi^3$, and (in shorthand notation)

$$\ell'(\theta) = y/\theta^2 - 1/\theta, \quad \ell''(\theta) = -2y/\theta^3 + 1/\theta^2, \quad \tilde{\ell}'(\phi) = 1/\phi - y, \quad \tilde{\ell}''(\phi) = -1/\phi^2.$$

Note that $\widehat{\phi} = 1/y$ and $\widehat{\theta} = y$, so $\widehat{\phi} = 1/\widehat{\theta}$; checking the rest is tedious but easy.

(b) We need $\mathrm{E}(Y^r) = \mathrm{E}(e^{rX}) = M_X(r)$ for $r = 1, 2$, and using the given formula leads to

$$\mathrm{E}(Y) = \exp(\mu+\sigma^2/2) = \psi, \quad \mathrm{var}(Y) = \exp(2\mu+2\sigma^2) - \{\exp(\mu+\sigma^2/2)\}^2 = \exp(2\mu+\sigma^2)(e^{\sigma^2}-1) = \psi^2\lambda,$$

as required. Hence $\sigma^2 = \log(1 + \lambda)$ and $\mu = \log\psi - \frac{1}{2}\log(1 + \lambda)$.

The brute force approach to maximum likelihood estimation of ψ and λ would be to compute the density of Y, which is (check this if it is not obvious)

$$f_Y(y; \mu, \sigma^2) = f_X(x; \mu, \sigma^2)\left|\frac{\partial x}{\partial y}\right|\Bigg|_{x=\log y} = \frac{1}{y\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(\log y - \mu)^2}{2\sigma^2}\right\}, \quad y > 0, \mu \in \mathbb{R}, \sigma^2 > 0,$$

then write $\sigma^2 = \log(1 + \lambda)$ and $\mu = \log\psi - \frac{1}{2}\log(1 + \lambda)$, compute the corresponding log likelihood based on $y_1, \dots, y_n$ and hence obtain $\widehat{\psi}$ and $\widehat{\lambda}$. However it is much simpler to appeal to invariance: $\log Y_1, \dots, \log Y_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$, so

$$\widehat{\mu} = n^{-1}\sum_{j=1}^n \log y_j, \quad \widehat{\sigma}^2 = n^{-1}\sum_{j=1}^n (\log y_j - \widehat{\mu})^2,$$

and $\widehat{\psi} = \exp(\widehat{\mu} + \widehat{\sigma}^2/2)$ and $\widehat{\lambda} = \exp(\widehat{\sigma}^2) - 1$.

Solution 3

(a) The hazard function may be interpreted as $\lim_{h\to 0} h^{-1}\mathrm{P}\{Y \in [y, y + h) \mid Y > y\}$, i.e., the instantaneous probability of failure at time y conditional on survival to then, accounting for the term 'force of mortality'.

Since $f(y) = h(y)\mathcal{F}(y)$ the likelihood contribution is $f(t)^d\mathcal{F}(t)^{1-d} = h(y)^d\mathcal{F}(t)$, as stated.

(b) In this case the observed data are $(t_1, d_1), \ldots, (t_n, d_n)$, $h(y) = \lambda$ and $\mathcal{F}(y) = \exp(-\lambda y)$, so the log likelihood is

$$\ell(\lambda) = \sum_{j=1}^{n} \{d_j \log h(t_j) + \log \mathcal{F}(t_j)\} = \sum_{j=1}^{n} (d_j \log \lambda - \lambda t_j), \quad \lambda > 0.$$

Differentiation gives $\widehat{\lambda} = \sum d_j / \sum t_j$ and $\ell''(\lambda) = -\sum d_j/\lambda^2$ and as $D_j = I(Y_j \leq c)$ is an indicator variable, $\mathrm{E}(D_j) = \mathrm{P}(Y_j \leq c) = 1 - e^{-\lambda c}$, leading to the stated formula. This seems reasonable because a proportion $e^{-\lambda c}$ of the data are lost to censoring, and the lack of memory property means that if they exceed c we have no idea of their values and thus no information on λ from them.

(c) Treating c as a realisation of C means that we must compute $\mathrm{E}(e^{-\lambda C})$, which is the moment-generating function $M_C(t)$ of C evaluated at $t = -\lambda$, i.e.,

$$\imath(\lambda) = \mathrm{E}_C\{\imath(\lambda, C)\} = \frac{n}{\lambda^2}\left\{1 - \mathrm{E}\left(e^{-\lambda C}\right)\right\} = \frac{n}{\lambda^2}\left\{1 - \left(\frac{\lambda\alpha}{\lambda\alpha + \lambda}\right)^{\nu}\right\} = \frac{n}{\lambda^2}\left\{1 - (1 + 1/\alpha)^{-\nu}\right\}.$$

(i) When $\alpha \to 0$, $\mathrm{E}(C) \to \infty$, so no observations will be censored in the limit, and thus $\imath(\lambda)$ tends to the usual quantity without censoring.

(ii) When $\alpha \to \infty$, $\mathrm{E}(C) \to 0$, the censoring probability tends to unity, and thus $\imath(\lambda)$ tends to zero.

(iii) When $\alpha = \nu = 1$ then $(1 + 1/\alpha)^{-\nu} = 1/2$, and $\imath(\lambda)$ is half that for an uncensored sample.

(iv) When $\alpha, \nu \to \infty$ for fixed $\mu = \nu/\alpha$, $(1 + 1/\alpha)^{-\nu} = (1 + \mu/\nu)^{-\nu} \to e^{-\mu}$, and $\imath(\lambda)$ corresponds to censoring at a fixed time μ/λ; as $\mathrm{E}(C) = \nu/(\alpha\lambda) \to \mu/\lambda$ and $\mathrm{var}(C) = \nu/(\alpha\lambda)^2 \to 0$, $C \xrightarrow{P} \mu/\lambda$.

Solution 4

(a) If we assume that the times to death have common distribution F, then the probability of death by time c is $F(c)$, and the probability of being alive is thus $1 - F(c)$. Hence if d is the indicator that the individual is alive, then their likelihood contribution is $F(c)^{1-d}\{1 - F(c)\}^d$, which yields the given likelihood, if the outcomes are independent.

(b) Writing $p(\lambda) = \exp(-\lambda c)$ and with $s = \sum_j d_j$ survivors, the log likelihood can be written as

$$\ell(\lambda) = (n - s)\log\{1 - p(\lambda)\} + s \log p(\lambda), \quad \lambda > 0,$$

so $p(\widehat{\lambda}) = s/n$, which yields $\widehat{\lambda} = c^{-1}\log(n/s)$. For the Fisher information we note that $S \sim B\{n, p(\lambda)\}$, and then after a little work obtain

$$\mathrm{E}\left\{-\frac{\partial^2 \ell(\lambda)}{\partial \lambda^2}\right\} = -\mathrm{E}\left\{\left(\frac{\partial p}{\partial \lambda}\right)^2 \frac{\partial^2 \ell}{\partial p^2} + \frac{\partial^2 p}{\partial \lambda^2}\frac{\partial \ell}{\partial p}\right\} = \frac{nc^2 p(\lambda)}{1 - p(\lambda)},$$

because $\partial p(\lambda)/\partial \lambda = -cp(\lambda)$, $\mathrm{E}(\partial \ell/\partial p) = 0$ and $\mathrm{E}(S) = np(\lambda)$.

(c) In this case the likelihood contribution for an individual is $(\lambda e^{-\lambda y})^{1-d}(e^{-\lambda c})^d$, so with $s = \sum_j d_j$ the overall log likelihood is

$$\sum_{j=1}^{n}(1 - d_j)(\log \lambda - \lambda y_j) - \lambda c s, \quad \lambda > 0.$$

This has second derivative $-(n - s)/\lambda^2$, leading to Fisher information $\{n - \mathrm{E}(S)\}/\lambda^2 = n\{1 - p(\lambda)\}/\lambda^2$ so the asymptotic relative efficiency of using current status data is

$$\frac{nc^2 p(\lambda)}{1 - p(\lambda)} \div \frac{n\{1 - p(\lambda)\}}{\lambda^2} = \frac{\lambda^2 c^2 p(\lambda)}{\{1 - p(\lambda)\}^2} = \frac{p(\lambda)\{\log p(\lambda)\}^2}{\{1 - p(\lambda)\}^2}.$$

Perhaps surprisingly, this is fairly high: it equals $0.999, 0.961, 0.655$ when $p = 0.9, 0.5, 0.1$ respectively, so despite the strong censoring, relatively little information is lost overall. Unfortunately this is a feature of the exponential distribution, and not of the problem itself.