Solution 1

(a) In case (i) the original likelihood is $\lambda e^{-\lambda y_1} \times \lambda\psi e^{\lambda\psi y_2}$, so the integrated likelihood is

$$\int_0^\infty \psi\lambda^2 \exp\{-\lambda(y_1+\psi y_2)\}\,\mathrm{d}\lambda = \frac{\psi}{(y_1+\psi y_2)^3}, \int_0^\infty u^2 e^{-u}\,\mathrm{d}u = \frac{2\psi}{(y_1+\psi y_2)^3}, \quad \psi > 0,$$

where we set $u = \lambda(y_1+\psi y_2)$. In case (ii) the original likelihood is $\lambda^{-1}e^{-y_1/\lambda} \times (\psi/\lambda)e^{\psi y_2/\lambda}$, giving

$$\int_0^\infty \psi\lambda^{-2} \exp\{-(y_1+\psi y_2)/\lambda\}\,\mathrm{d}\lambda = \frac{\psi}{y_1+\psi y_2}\int_0^\infty u^2 e^{-u}\,u^{-2}\mathrm{d}u = \frac{\psi}{y_1+\psi y_2}, \quad \psi > 0,$$

where we set $u = (y_1+\psi y_2)/\lambda$. Hence the result depends on the nuisance parametrisation, which is clearly unsatisfactory.

(b) In case (i) we now obtain

$$\int L(\psi,\lambda)\pi(\psi,\lambda)\,\mathrm{d}\lambda = \int_0^\infty \psi\lambda^2\exp\{-\lambda(y_1+\psi y_2)\}\pi(\psi,\lambda)\,\mathrm{d}\lambda,$$

but the prior corresponding to $(\psi, \mu = 1/\lambda)$ is $\pi^*(\psi,\mu) = \pi(\psi, 1/\mu)\mu^{-2}$, giving

$$\int_0^\infty L(\psi, 1/\mu)\pi^*(\psi,\mu)\,\mathrm{d}\mu = \int_0^\infty \psi\mu^{-2}\exp\{-(y_1+\psi y_2)/\mu\}\pi(\psi,1/\mu)\mu^{-2}\,\mathrm{d}\mu,$$

and this reduces to the result for (i) when we change the variable of integration to $\lambda = 1/\mu$.

Solution 2

(a) The likelihood is $\prod_{j=1}^n p_j$, and this equals zero if any of the $p_j = 0$, so we should take $p_j > 0$ for each j. Moreover if we had an optimal solution with $\sum p_j < 1$, we could increase the likelihood just by increasing (say) p_1 until $\sum p_j = 1$, so we should take $\sum p_j = 1$. Hence we can maximise $\sum \log p_j$ subject to $\sum p_j = 1$, and we can do this using Lagrange multipliers, by maximising

$$\sum_{j=1}^n \log p_j + \lambda\left(\sum_{j=1}^n p_j - 1\right),$$

differentiation of which with respect to p_j gives $p_j^{-1} + \lambda = 0$ for all j. As the p_j are equal and sum to unity, they must all equal n^{-1}. The second derivative is negative, so the point is a maximum.

(b) Now we maximise the (slightly eccentrically expressed) Lagrangian

$$\sum_{j=1}^n \log p_j - n\lambda^{\mathrm{T}}\left(\sum_{j=1}^n c_j(\theta)p_j - 0\right) - \mu\left(\sum_{j=1}^n p_j - 1\right),$$

with λ of dimension $d \times 1$ and μ scalar. Differentiation with respect to λ and μ gives the constraints, and differentiation with respect to p_j gives

$$p_j^{-1} - n\lambda^{\mathrm{T}}c_j(\theta) - \mu = 0 \quad \Longrightarrow \quad 1 = np_j\lambda^{\mathrm{T}}c_j(\theta) + \mu p_j,$$

addition of which over j gives $\mu = n$, and consequently $p_j^{-1} = n\{1 + \lambda^{\mathrm{T}}c_j(\theta)\}$, where λ is chosen to solve

$$\sum_{j=1}^n c_j(\theta)p_j = \sum_{j=1}^n \frac{c_j(\theta)}{n\{1+\lambda^{\mathrm{T}}c_j(\theta\}} = 0,$$

as required.

(c) We saw in (a) that ℓ_{E} is maximised when $p_j \equiv 1/n$, and in this case $0 = \sum p_j(y_j - \theta)$ yields $\widehat{\theta} = \overline{y}$. The equation $0 = \sum p_j(y_j - \theta)$ and constraint $\sum p_j = 1$ imply that $\sum y_j p_j = \sum y_j/\{1 + \lambda(y_j - \theta\} = \sum \theta p_j = \theta$, so the y_j are reweighted so that their weighted average equals θ; this is only possible in the convex hull $(\min y_j, \max y_j)$ of the data.

Solution 3

(a) Under the null hypothesis we can write $Y_j \stackrel{\mathrm{D}}{=} E_j/\lambda$, where $E_1, \ldots, E_n \stackrel{\mathrm{iid}}{\sim} \exp(1)$, so the test statistic

$$T = \sum_{j=1}^{n} \log(Y_j/\overline{Y}) \stackrel{\mathrm{D}}{=} \sum_{j=1}^{n} \log(E_j/\overline{E}),$$

which does not depend on λ. Hence the statistic is invariant to λ. This could be simulated by generating $E_1, \ldots, E_n \stackrel{\mathrm{iid}}{\sim} \exp(1)$ and hence computing a null distribution for T.

(b) Under the null hypothesis the data are exponential, so the minimal sufficient statistic is $S = Y_1 + \cdots + Y_n$, which has a gamma (n, λ) distribution. Hence the conditional density of the data given S is

$$\frac{\lambda^n \exp\{-\lambda(y_1 + \cdots + y_n)\}}{\lambda^n s^{n-1} \exp(-\lambda s)/\Gamma(n)} = \frac{\Gamma(n)}{s^{n-1}}, \quad 0 < y_1, \ldots, y_n < s, \sum y_j = s.$$

This is the uniform distribution on an n-dimensional simplex, and of course it does not depend on λ. The statistic T depends only on the $Y_j/\overline{Y} = nY_j/S$, which is invariant to s, so the same simulation algorithm as in (a) will work.

Solution 4

(a) For $x \in (0,1)$ and because the events $P \le u$ and $P \ge 1-u$ are disjoint for $u < 1/2$ we have

$$\mathrm{P}(Q \le x) = \mathrm{P}\{\min(P, 1-P) \le x/2\} = \mathrm{P}(P \le x/2) + \mathrm{P}(P \ge 1 - x/2) = x/2 + x/2 = x,$$

as required.

(b) These are computed using the R code

```
> ppois(1:7,lambda=2,lower.tail=FALSE)
[1] 0.593994150 0.323323584 0.142876540 0.052653017 0.016563608 0.004533806 0.001096719
> ppois(0:2,lambda=2)
[1] 0.1353353 0.4060058 0.6766764
```

The values are fairly limited in both cases, though the limitations for (ii) are not surprising.

For confidence intervals we would solve the equation $\mathrm{P}(Y \ge 2; \mu) = \alpha$ for some specific values of α, so the discreteness is not a major issue.

Solution 5 The Poisson distribution is an exponential family with canonical statistic y and canonical parameter $\varphi = \log \psi$, and the test with a critical region $\mathcal{Y}_1 = \{y, y+1, \ldots\}$ is therefore the most powerful critical region of size

$$\alpha = \mathrm{P}_0(Y \in \mathcal{Y}_1) = \sum_{x=y}^{\infty} \lambda_0^x e^{-\lambda_0}/x!$$

against any alternative $\lambda > \lambda_0$. Hence (whether or not they knew it) the test used by the physicists could not have been improved (provided of course that the underlying Poisson model is reasonable).