**Solution 1** The log likelihood contribution from a single observation,

$$\ell(\alpha, \lambda) = \alpha \log \lambda + (\alpha - 1)\log y - \lambda y - \log \Gamma(\alpha), \quad \alpha, \gamma > 0,$$

has first derivatives $\ell_\alpha = \log \lambda + \log y - \Psi(\alpha)$, $\ell_\lambda = \alpha/\lambda - y$ and second derivatives $\ell_{\alpha\alpha} = -\Psi'(\alpha)$, $\ell_{\alpha\lambda} = 1/\lambda$, $\ell_{\lambda\lambda} = -\alpha/\lambda^2$, so the observed information matrix equals the expected information matrix,

$$\imath(\alpha, \lambda) = n\imath_1(\alpha, \lambda) = n \begin{pmatrix} \Psi'(\alpha) & -1/\lambda \\ -1/\lambda & \alpha/\lambda^2 \end{pmatrix}.$$

(a) The score statistic is $\ell_\alpha$ for the sample divided by the square root of its asymptotic variance $n\Psi'(\alpha)$, all evaluated at $\alpha = 1$; this gives

$$\frac{\sum_{j=1}^n \log(\lambda Y_j) - n\Psi(1)}{\{n\Psi'(1)\}^{1/2}} \mathrel{\dot\sim} \mathcal{N}(0, 1).$$

(b) The score statistic is $\ell_\alpha^{\mathrm{T}} \jmath^{\alpha\alpha} \ell_\alpha$, evaluated at the maximum likelihood estimator when $\alpha = 1$, i.e., at $(\alpha, \widehat{\lambda}_\alpha)|_{\alpha=1} = (1, 1/\overline{Y})$, and with

$$\jmath^{\alpha\alpha} = (\widehat{\jmath}_{\alpha\alpha} - \widehat{\jmath}_{\alpha\lambda}\widehat{\jmath}_{\lambda\lambda}^{-1}\widehat{\jmath}_{\lambda\alpha})^{-1} = (\widehat{\imath}_{\alpha\alpha} - \widehat{\imath}_{\alpha\lambda}\widehat{\imath}_{\lambda\lambda}^{-1}\widehat{\imath}_{\lambda\alpha})^{-1} = \frac{1}{n\{\Psi'(\alpha) - 1/\alpha\}}$$

evaluated at $\alpha = 1$. Hence the score statistic is

$$\left\{\sum_{j=1}^n \log(Y_j/\overline{Y}) - n\Psi(1)\right\}^2 / \left[n\{\Psi'(1) - 1\}\right] \mathrel{\dot\sim} \chi_1^2;$$

in fact here, since the interest parameter is scalar, we might use the approximation

$$\frac{\sum_{j=1}^n \log(Y_j/\overline{Y}) - n\Psi(1)}{[n\{\Psi'(1) - 1\}]^{1/2}} \mathrel{\dot\sim} \mathcal{N}(0, 1).$$

(c) (i) The log likelihood function in terms of $\mu = \alpha/\lambda$ is obtained by setting $\lambda = \alpha/\mu$, and is

$$\ell^*(\alpha, \mu) \equiv \alpha \log \alpha - \alpha \log \mu + (\alpha - 1)\log y - \alpha y/\mu - \log \Gamma(\alpha), \quad \alpha, \mu > 0,$$

giving $\ell^*_{\alpha\mu} = y/\mu^2 - 1/\mu$, which has expectation zero; hence $\mu$ and $\alpha$ are orthogonal. This could also be found by using the Jacobian $J$ for the transformation $(\alpha, \mu) \mapsto (\alpha, \lambda)$ to compute $\imath^*(\alpha, \mu) = J\imath(\alpha, \lambda)J$. (ii) The original log likelihood function can be written as $\ell(\alpha, \lambda) \equiv \alpha \log y - \lambda y + \alpha \log \lambda - \log \Gamma(\alpha) - \log y$, so the complementary mean parameter for $\alpha$, $\mathrm{E}(Y) = \alpha/\lambda = \mu$, is orthogonal to $\alpha$ by Example 57.

In the orthogonal parametrisation $\ell^*_\alpha = 1 + \log(y/\mu) - y/\mu + \log \alpha - \Psi(\alpha)$ and $\ell^*_{\alpha\alpha} = 1/\alpha - \Psi'(\alpha)$. Now $\widehat{\mu}_\alpha = \overline{Y}$, and owing to the orthogonality $\jmath^{\alpha\alpha} = \widehat{\jmath}_{\alpha\alpha}^{-1} = 1/[n\{\Psi'(\alpha) - 1/\alpha\}]$, which gives $1/[n\{\Psi'(1) - 1\}]$ when $\alpha = 1$. The numerator of the score statistic is $\sum_{j=1}^n \{1 + \log(Y_j/\mu) - Y_j/\mu + \log \alpha - \Psi(\alpha)\}$ evaluated at $\alpha = 1$ and $\mu = \widehat{\mu}_1 = \overline{Y}$, and the upshot is that the statistic is the same as in (b).

**Solution 2**

The log likelihood for a sample $y_1, \ldots, y_n$ from the $\mathcal{N}(\mu, \sigma^2)$ distribution is

$$\ell(\mu, \sigma^2) \equiv -\frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\sum_{j=1}^n (y_j - \mu)^2 = -\frac{n}{2}\log \sigma^2 - \frac{1}{2\sigma^2}\left\{\sum_{j=1}^n (y_j - \overline{y})^2 + n(\overline{y} - \mu)^2\right\}.$$

(a) With $\mu = \theta$ and $\sigma^2 = 1$ the formula above reduces to

$$\ell(\theta) \equiv -\frac{n}{2}(\overline{y} - \theta)^2 \equiv n\overline{y}\theta - n\theta^2/2, \quad \theta \in \mathbb{R}.$$

and $\widehat{\theta} = \overline{y}$, so

$$r(\theta) = \mathrm{sign}(\widehat{\theta} - \theta)\sqrt{2\{\ell(\widehat{\theta}) - \ell(\theta)\}} = \mathrm{sign}(\overline{y} - \theta)\sqrt{n(\overline{y} - \theta)^2} = \sqrt{n}(\overline{y} - \theta).$$

The log likelihood has second derivative $-n$, so $\widehat{\jmath} = n$, and therefore $t(\theta) = \widehat{\jmath}^{1/2}(\widehat{\theta} - \theta) = r(\theta)$. Hence $\log\{q(\theta)/r(\theta)\} = 0$, so $r^*(\theta) = r(\theta)$, i.e., inferences based on either will be the same. This is not surprising, because $r(\theta) \sim \mathcal{N}(0, 1)$ exactly.

(b) For the second model $\mu = 0$ and $\sigma^2 = 1/\theta$, giving log likelihood

$$\ell(\theta) = \frac{n}{2}\log\theta - \frac{\theta}{2}\sum_{j=1}^{n} y_j^2 = \frac{n}{2}(\log\theta - \theta\overline{s}), \quad \theta > 0,$$

say. This has second derivative $-n/(2\theta^2)$ and gives $\widehat{\theta} = 1/\overline{s}$, so $\widehat{\jmath} = n\overline{s}^2/2$. Hence

$$r(\theta) = \mathrm{sign}(\widehat{\theta} - \theta)\sqrt{2\{\ell(\widehat{\theta}) - \ell(\theta)\}} = \mathrm{sign}(1 - \overline{s}\theta)\sqrt{n\left\{\overline{s}\theta - \log(\overline{s}\theta) - 1\right\}},$$

because $\overline{s} > 0$, and $q(\theta) = t(\theta) = \sqrt{n/2}(1 - \overline{s}\theta)$. As $\overline{s}$ is an average of $n$ variables each with mean $1/\theta$ and variance $2/\theta^2$, the mean and variance of $q(\theta)$ are 0 and 1, as we would expect.

This is the same as in Example 45 but with $n$ replaced by $n/2$, so the accuracy will be as high as it was there (i.e., $r^*(\theta)$ gives nearly perfect inferences, when $n = 2$ here, or $n = 1$ in Example 45).

**Solution 3**

(a) The modified likelihood root depends on $r(\psi)$ and $q(\psi)$, and we need only consider how $q(\psi)$ is affected by this change, which does not affect the information matrices. But

$$\varphi(\widehat{\theta}) - \varphi(\widehat{\theta}_\psi) \mapsto B\{\varphi(\widehat{\theta}) - \varphi(\widehat{\theta}_\psi)\}, \quad \varphi_\lambda(\widehat{\theta}_\psi) \mapsto B\varphi_\lambda(\widehat{\theta}_\psi), \quad \varphi_\theta(\widehat{\theta}_\psi) \mapsto B\varphi_\theta(\widehat{\theta}_\psi),$$

so

$$\frac{|\ \varphi(\widehat{\theta}) - \varphi(\widehat{\theta}_\psi)\ \ \varphi_\lambda(\widehat{\theta}_\psi)\ |}{|\ \varphi_\theta(\widehat{\theta})\ |} \mapsto \frac{|\ B\varphi(\widehat{\theta}) - B\varphi(\widehat{\theta}_\psi)\ \ B\varphi_\lambda(\widehat{\theta}_\psi)\ |}{|\ B\varphi_\theta(\widehat{\theta})\ |} = \frac{|B|\times |\ \varphi(\widehat{\theta}) - \varphi(\widehat{\theta}_\psi)\ \ \varphi_\lambda(\widehat{\theta}_\psi)\ |}{|B|\times |\ \varphi_\theta(\widehat{\theta})\ |},$$

which leaves $q(\psi)$ unchanged, because $|B|$ cancels from top and bottom.

(b) If $h$ is normal, then $\log h(u) \equiv -u^2/2$, so $(\log h)'\{(y_j^{\mathrm{o}} - \eta)/\tau\} = -(y_j^{\mathrm{o}} - \eta)/\tau^2$, and in the notation of the example (recall that $e_j^{\mathrm{o}} = (y_j^{\mathrm{o}} - \widehat{\eta}^{\mathrm{o}})/\widehat{\tau}^{\mathrm{o}}$) this gives

$$\begin{aligned}
\varphi(\theta) &= \left(\sum_{j=1}^{n}(\eta - y_j^{\mathrm{o}})/\tau^2, \sum_{j=1}^{n}(\eta - y_j^{\mathrm{o}})/\tau^2 \times e_j^{\mathrm{o}}\right)^{\mathrm{T}} \\
&= (n\eta/\tau^2 - a/\tau^2, b\eta/\tau^2 - c/\tau^2)^{\mathrm{T}} \\
&= \begin{pmatrix} n & a \\ b & c \end{pmatrix}\begin{pmatrix} \eta/\tau^2 \\ 1/\tau^2 \end{pmatrix} \\
&= B\begin{pmatrix} \eta/\tau^2 \\ 1/\tau^2 \end{pmatrix}
\end{aligned}$$

for $a = -\sum y_j^{\mathrm{o}}$, $b = \sum e_j^{\mathrm{o}}$ and $c = -\sum y_j^{\mathrm{o}} e_j^{\mathrm{o}}$ that depend only on the data $y^{\mathrm{o}}$ and are easily computed, and $B$ is non-singular with probability one.

In view of (a) we can therefore take $\varphi(\theta) = (\eta/\tau^2, 1/\tau^2)^{\mathrm{T}}$ for computing $q(\theta)$. This invariance can also greatly simplify computations in other examples.

**Solution 4** The log likelihood in the non-orthogonal parametrization is

$$\ell^*(\psi,\gamma) = \log\gamma - \gamma y_1 + \log(\gamma\psi) - \gamma\psi y_2 = 2\log\gamma + \log\psi - \gamma(y_1+\psi y_2), \quad \gamma,\psi>0,$$

so its observed and Fisher information matrices have elements $-\ell^*_{\gamma\gamma} = 2/\gamma^2$, $-\ell^*_{\gamma\psi} = y_2$ and $-\ell^*_{\psi\psi} = 1/\psi^2$, and $\imath^*_{\gamma\gamma} = 2/\gamma^2$, $\imath^*_{\gamma\psi} = 1/(\gamma\psi)$ and $\imath^*_{\psi\psi} = 1/\psi^2$. Hence the partial differential equation giving the orthogonal parametrization is

$$\frac{\partial\gamma}{\partial\psi} = -\imath^{*-1}_{\gamma\gamma}(\psi,\gamma)\imath^*_{\gamma\psi}(\psi,\gamma) = -\frac{\gamma^2}{2}\frac{1}{\gamma\psi} = -\frac{\gamma}{2\psi}, \quad \gamma,\psi>0,$$

as required.

To check that $\lambda = \gamma\psi^{1/2}$ is orthogonal to $\psi$, we write $\gamma = \lambda\psi^{-1/2}$ so the log likelihood becomes

$$\ell(\psi,\lambda) = 2\log\lambda - \lambda\psi^{-1/2}(y_1+\psi y_2),$$

and note that $\ell_{\lambda\psi} = (y_1-\psi y_2)/(2\psi^{3/2})$ has expectation $\{\gamma^{-1} - \psi/(\gamma\psi)\}/(2\psi^{3/2}) = 0$. Hence $\lambda$ is orthogonal to $\psi$, as required.

The question only asks you to check that the given solution provides an orthogonal transformation, so the material below is included only to show how the PDE

$$2\psi\frac{\partial\gamma}{\partial\psi} = -\gamma$$

would be solved if the solution had not been given in the question. Now (e.g., Theorem 2, page 50 of Sneddon, *Elements of Partial Differential Equations*, 1957),

> "The general solution of the linear partial differential equation
>
> $$P\frac{\partial z}{\partial x} + Q\frac{\partial z}{\partial y} = R$$
>
> is $F(u,v)=0$, where $F$ is an arbitrary function and $u(x,y,z)=c_1$ and $v(x,y,z)=c_2$ form a solution of the equations
> $$\frac{\mathrm{d}x}{P} = \frac{\mathrm{d}y}{Q} = \frac{\mathrm{d}z}{R}."$$

In the present setting $z=\gamma$, $x=\psi$, $P=2\psi$, $Q=0$ and $R=-\gamma$, so we need to solve

$$\frac{\mathrm{d}\psi}{2\psi} + \frac{\mathrm{d}\gamma}{\gamma} = 0 \quad\Longrightarrow\quad \tfrac{1}{2}\log\psi + \log\gamma = c \quad\Longrightarrow\quad \gamma\psi^{1/2} = c,$$

and thus according to the theorem, the general solution is any function of $\gamma\psi^{1/2}$, such as $\lambda(\psi,\gamma) = \gamma\psi^{1/2}$.