Solution 1 Clearly

$$\sum_j (Y_j - \overline{Y})^2 = \sum_j (Y_j^2 - 2Y_j\overline{Y} + \overline{Y}^2) = \sum_j Y_j^2 - 2\overline{Y}\sum_j Y_j + n\overline{Y}^2 = \sum_j Y_j^2 - 2n\overline{Y}^2 + n\overline{Y}^2 = \sum_j Y_j^2 - n\overline{Y}^2$$

as required.

(a) The variance of a sum of independent variables is the sum of the variances, so

$$\mathrm{var}(\overline{Y}) = n^{-2}\mathrm{var}\left(\sum_j Y_j\right) = n^{-2}\sum_j \mathrm{var}(Y_j) = \sigma^2/n.$$

Moreover $\overline{Y} - \mu$ does not depend on j, so

$$\sum_{j=1}^n \{Y_j - \mu - (\overline{Y} - \mu)\}^2 = \sum_j (Y_j - \mu)^2 - 2(\overline{Y} - \mu)\sum_j (Y_j - \mu) + \sum_j (\overline{Y} - \mu)^2 = \sum_j (Y_j - \mu)^2 - n(\overline{Y} - \mu)^2,$$

and this has expectation $\sum_j \mathrm{var}(Y_j) - n\mathrm{var}(\overline{Y}) = n\sigma^2 - n\sigma^2/n = (n-1)\sigma^2$, as required.

(b) We have

$$\sum_{j,k=1}^n (Y_j - Y_k)^2 = \sum_{j,k} (Y_j^2 + Y_k^2 - 2Y_jY_k) = 2n\sum_j Y_j^2 - 2n^2\overline{Y}^2 = 2n\left(\sum_j Y_j^2 - n\overline{Y}^2\right) = 2n(n-1)S^2,$$

as required.

Solution 2

(a) If $\widehat{\theta}$ is unbiased, then we must have $\mathrm{E}(\sum_j a_jT_j) = \sum_j a_j\theta = \theta$ for any possible θ, so $\sum a_j = 1$. Now $\mathrm{var}(\sum_j a_jT_j) = \sum_j a_j^2 v_j$, and we seek to minimise this subject to $\sum a_j = 1$. The corresponding Lagrangian is

$$\sum_j a_j^2 v_j + \lambda\left(\sum a_j - 1\right),$$

and differentiation with respect to a_j and to λ gives

$$2v_ja_j + \lambda = 0, \quad j = 1, \ldots, n, \quad \sum_j a_j = 1,$$

resulting in $a_j = v_j^{-1}/\sum_i v_i^{-1}$ and $\mathrm{var}(\widehat{\theta}) = (\sum_j v_j^{-1})^{-1}$.

(b) If the $T_j \stackrel{\text{ind}}{\sim} \mathcal{N}(\theta, \sigma^2 v_j)$ then the log likelihood function is

$$\ell(\theta, \sigma^2) = -\tfrac{1}{2}\sum_{j=1}^n \left\{\log(2\pi\sigma^2 v_j) + (t_j - \theta)^2/(\sigma^2 v_j)\right\} \equiv -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum (t_j - \theta)^2/v_j,$$

where the second expression ignores additive constants. Differentiation gives

$$\ell_\theta = \frac{1}{\sigma^2}\sum (t_j - \theta)/v_j, \quad \ell_{\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum (t_j - \theta)^2/v_j,$$

in an obvious notation, and it is easy to check that $\widehat{\theta}$ is the sole solution of the first equation whatever the value of σ^2, and that $\widehat{\sigma}^2 = n^{-1}\sum (t_j - \widehat{\theta})^2/v_j$; the corresponding unbiased estimator uses the denominator $n - 1$.

(c) In this case we maximise not over $\sigma^2 \in (0,\infty)$ but over $\sigma^2 \geq 1$, so the estimator becomes $\tilde\sigma^2 = \max(\widehat\sigma^2, 1)$.

Solution 3

The probability of infection is $\mathrm{P}(Y > 0) = 1 - e^{-\mu}$, so the likelihood is

$$(1-e^{-\mu})^r (e^{-\mu})^{m-r} = (e^{\mu}-1)^r (e^{-\mu})^m, \quad \mu > 0.$$

The log likelihood $\ell(\mu) = r\log(e^\mu - 1) - m\mu$ has first and second derivatives

$$\ell'(\mu) = \frac{re^\mu}{e^\mu - 1} - m, \quad \ell''(\mu) = -r\frac{e^\mu}{(e^\mu-1)^2},$$

so $e^{-\widehat\mu} = (m-r)/m$, giving $\widehat\mu = \log\{m/(m-r)\}$, and observed information $J(\widehat\mu) = r^2(m-r)/m^2$, while the expected information is $I(\mu) = \mathrm{E}(R)e^\mu/(e^\mu-1)^2 = m/(e^\mu - 1)$, because $\mathrm{E}(R) = m\mathrm{P}(Y=1) = m(1-e^{-\mu})$.

The asymptotic distribution of $\widehat\mu$ is $\mathcal{N}\{\mu, I(\mu)^{-1}\}$

Alternatively we write $\widehat\mu = g(r/m) = -\log(1-r/m)$, with $g(u) = -\log(1-u)$ and $g'(u) = 1/(1-u)$, note that $R \sim B(m, 1-e^{-\mu})$, so $R/m \dot\sim \mathcal{N}\{1-e^{-\mu}, e^{-\mu}(1-e^{-\mu})/m\}$, and apply the delta method, which gives that $g(R/m)$ is approximately normal with mean and variance

$$g(1-e^{-\mu}) = \mu, \quad g'(1-e^{-\mu})^2 \mathrm{var}(R/m) = (e^\mu - 1)/m.$$

The asymptotic variance is $(e^\mu-1)/m > 0$ for any $\mu > 0$, but the exact variance is infinite for any m, because

$$\mathrm{E}(\widehat\mu) = \sum_{r=0}^{m} \log\{m/(m-r)\}\mathrm{P}(R=r) = \log m - \sum_{r=0}^{m}\log(m-r)\mathrm{P}(R=r) = \infty,$$

as $\mathrm{P}(R=m) = (1-e^{-\mu})^m > 0$ for any μ and m.

Solution 4

(a) The minimum value of zero is attained when $\nabla_y \log f(y;\theta') - \nabla_y \log f(y;\theta) \equiv 0$, and this clearly occurs when $\theta' = \theta_g$, say. Now suppose that $\nabla_y \log f(y;\theta') - \nabla_y \log g(y) \equiv 0$, and integrate with respect to y to obtain

$$\log f(y;\theta') - \log g(y) = a(\theta')$$

where $a(\cdot)$ is an arbitrary function of θ alone. Hence $g(y)b(\theta') = f(y;\theta')$ for all y, which implies that as both f and g are densities, they must be identical.

(b) (i) In this case the log likelihood is $-\frac{1}{2}\log(2\pi\tau^2) - (y-\eta)^2/(2\tau^2)$, and the two derivatives are $-(y-\eta)/\tau^2$ and $-1/\tau^2$, so the objective function is

$$\mathrm{E}\left\{(Y-\eta)^2/\tau^4 - 2/\tau^2\right\} = \tau^{-4}\left\{\mathrm{E}\left\{(Y-\mu+\mu-\eta)^2\right\} - 2\tau^2\right\} = \tau^{-4}\left\{\sigma^2 + (\mu-\eta)^2 - 2\tau^2\right\},$$

which is clearly minimised by taking $\eta = \mu$ and then setting $\tau^2 = \sigma^2$. The empirical estimators minimise

$$\tau^{-4}\sum_{j=1}^{n}(Y_j - \eta)^2 - 2n/\tau^2 = \tau^{-4}\sum_{j=1}^{n}(Y_j - \overline{Y})^2 + \tau^{-4}n(\overline{Y}-\eta)^2 - 2n/\tau^2,$$

so $\tilde\eta = \overline{Y}$ and $\tilde\tau^2 = n^{-1}\sum_{j=1}^n (Y_j - \overline{Y})^2$; these are the maximum likelihood estimators.

(ii) Here the log likelihood is $\log\lambda' - y\lambda'$ for $\lambda' > 0$, so the two derivatives are $-\lambda'$ and 0 and thus the objective function is $(\lambda')^2$, which is minimised by taking $\tilde\lambda = 0$. This is not a sensible solution, and it arises because the argument from the original expression to that above fails (the integration by parts involves another, non-zero, term). The original expression would be $\mathrm{E}\{[-\lambda' - (-\lambda)\}^2] = (\lambda - \lambda')^2$, which is minimised by taking $\lambda' = \lambda$.

(c) On writing

$$\{\nabla_y \log f(y;\theta) - \nabla_y \log g(y)\}^2 = \{\nabla_y \log f(y;\theta)\}^2 - 2\nabla_y \log f(y;\theta)\nabla_y \log g(y) + \{\nabla_y \log g(y)\}^2,$$

we see that the population version of the estimator is

$$\theta_g = \mathrm{argmin}_\theta \int \{\nabla_y \log f(y;\theta)\}^2 w(y)g(y)\,\mathrm{d}y - 2\int \{\nabla_y \log f(y;\theta)\nabla_y \log g(y)\}\, w(y)g(y)\,\mathrm{d}y,$$

and the previous argument and integration by parts shows that the second integral here is

$$[w(y)\nabla_y \log f(y;\theta)g(y)]_{y_-}^{y_+} - \int \nabla_y\{w(y)\nabla_y \log f(y;\theta)\}g(y)\,\mathrm{d}y,$$

which leads to (1) if $w(y)$ ensures that the first term here equals zero. This is the case with $w(y) = y$, $g(y) = \lambda e^{-\lambda y}$ and $\nabla_y \log f(y;\theta) = -\lambda'$, and then as $\mathrm{E}(Y) = 1/\lambda$, (1) becomes $(\lambda')^2/\lambda - 2\lambda'$, minimisation of which with respect to λ' gives $\lambda' = \lambda$, as expected.