1(a)

$$\mathbb{P}(X=x;\lambda)=\frac{\lambda^x e^{-\lambda}}{x!}=\exp\{x\log\lambda-\lambda-\log x!\}$$

The probability mass function is of the exponential family form with

natural parameter $\theta=\log\lambda$

natural statistic $s(x)=x$

cumulant transform $k(\theta)=e^\theta$

Hence the cumulant generating function of X is

$$K_X(t)=k(\theta+t)-k(\theta)=e^{\theta+t}-e^\theta=e^\theta(e^t-1)=\lambda(e^t-1)$$

1(b) By the properties of the cumulant generating function

$$K_Y(t)=K_{\Sigma X_i}(t)\overset{\{X_i\}\text{ independent}}{=}\sum_{i=1}^n K_{X_i}(t)=(e^t-1)\sum_{i=1}^n\lambda_i$$

Comparing the above expression with $K_X(t)$ in 1(a),

$$Y\sim\text{Poisson}\left(\sum_{i=1}^n\lambda_i\right)$$

1(c) Now, $X_i\overset{iid}{\sim}P(\lambda)$, $i=1,\dots,n$.

By the properties of the cumulant generating function

$$K_Z(t)=K_{\Sigma X_i/\sqrt{n\lambda}-\sqrt{n\lambda}}(t)=-\sqrt{n\lambda}\,t+K_{\Sigma X_i}(t/\sqrt{n\lambda})$$

$$=-\sqrt{n\lambda}\,t-n\lambda+n\lambda\exp\{t/\sqrt{n\lambda}\}$$

Expanding $\exp\{t/\sqrt{n\lambda}\}$ around $t=0$ gives

$$K_Z(t)=\cancel{-\sqrt{n\lambda}\,t}\;\cancel{-n\lambda}+\{\cancel{n\lambda}+\cancel{n\lambda\frac{t}{\sqrt{n\lambda}}}+n\lambda\frac{t^2}{2n\lambda}+O(n^{-1/2})\}$$

So $K_Z(t)=\frac{t^2}{2}+O(n^{-1/2})$.

As $n\to\infty$ $K_Z(t)\longrightarrow\frac{t^2}{2}$ which is the cumulant generating function of a N(0,1) random variable.

2(a) Omitting any quantities that are constant wrt λ

$$\ell(\lambda) = \log\lambda \sum X_i - n\lambda, \text{ and so } U(\lambda) = \frac{d\ell(\lambda)}{d\lambda} = \frac{\sum X_i}{\lambda} - n$$

Hence, $U(\lambda) = 0 \implies \hat{\lambda} = \frac{\sum X_i}{n}$

We need to check that $\ell(\lambda)$ is maximized at $\hat{\lambda}$.

$$\frac{d^2\ell(\lambda)}{d\lambda^2} = -\frac{\sum X_i}{\lambda^2} \text{ and so } \left.\frac{d^2\ell(\lambda)}{d\lambda^2}\right|_{\lambda=\hat{\lambda}} = -\frac{n^2}{\sum X_i} < 0 \; [X_i > 0]$$

So $\hat{\lambda}$ is the maximum likelihood estimator.

The expected information on λ is

$$i(\lambda) = \mathbb{E}_\lambda\left\{-\frac{d^2\ell(\lambda)}{d\lambda}\right\} = \frac{\sum \mathbb{E}(X_i)}{\lambda^2} = \frac{n\lambda}{\lambda^2} = \frac{n}{\lambda}$$

2(b) $\hat{\theta} = \log\hat{\lambda} = \log\frac{\sum X_i}{n}$

$$\tilde{i}(\theta) = \left\{\frac{d\lambda}{d\theta}\right\}^2 i(e^\theta) = e^{2\theta}\frac{n}{e^\theta} = ne^\theta$$

$$\tilde{U}(\theta) = \frac{d\lambda}{d\theta} U(e^\theta) = \cancel{e^\theta}\frac{\sum X_i}{\cancel{e^\theta}} - ne^\theta = \sum X_i - ne^\theta$$

2(c)

Because θ is the natural parameter $\frac{d^2\tilde{\ell}(\theta)}{d\theta^2}$ does not depend on the data. Furthermore, by the Bartlett identity $\frac{d^2\tilde{\ell}(\theta)}{d\theta^2} = -\tilde{i}(\theta) = -ne^\theta$ and so

$\frac{d^3\tilde{\ell}(\theta)}{d\theta^3} = -ne^\theta$. Hence,

$$b(\theta) = \frac{\mathbb{E}_\theta\left\{\frac{d^3\tilde{\ell}(\theta)}{d\theta^3}\right\} + \cancelto{0}{2\mathbb{E}_\theta\left\{\frac{d\tilde{\ell}(\theta)}{d\theta}\frac{d^2\tilde{\ell}(\theta)}{d\theta^2}\right\}}}{2\{\tilde{i}(\theta)\}^2} = -\frac{ne^\theta}{2(ne^\theta)^2} = -\frac{1}{2ne^\theta}$$

$$\hat{\theta}_{bc} = \hat{\theta} - b(\hat{\theta}) = \log\frac{\sum X_i}{n} + \frac{1}{2\sum X_i}$$

3(a)

$$f_X(x;\mu,\lambda) = \frac{\sqrt{\lambda}}{\sqrt{2\pi}} x^{-3/2} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}$$

$$= \exp\left\{-\frac{\lambda x^2}{2\mu^2 x} + \frac{2\lambda x\mu}{2\mu^2 x} - \frac{\lambda\mu^2}{2\mu^2 x} - \frac{3}{2}\log x + \log\frac{\sqrt{\lambda}}{\sqrt{2\pi}}\right\}$$

$$= \exp\left\{-\frac{\lambda}{2\mu^2}x - \frac{\lambda}{2}\frac{1}{x} + \frac{\lambda}{\mu} + \frac{1}{2}\log\lambda - \frac{3}{2}\log x - \log\sqrt{2\pi}\right\}$$

So $f_X(x;\mu,\lambda)$ is of the exponential family form with

natural parameters $\theta_1 = -\frac{\lambda}{2\mu^2}$, $\theta_2 = -\frac{\lambda}{2}$

natural statistics $s_1(x) = x$, $s_2(x) = 1/x$

cumulant transform $k(\theta_1,\theta_2) = -\frac{\lambda}{\mu} - \frac{1}{2}\log\lambda$

$$= -\sqrt{4\theta_1\theta_2} - \frac{1}{2}\log(-2\theta_2)$$

3(b)

By the properties of exponential families

$$\mathrm{Cov}(X, 1/X) = \mathrm{Cov}(S_1(X), S_2(X)) = \frac{\partial^2 k(\theta_1,\theta_2)}{\partial\theta_1\partial\theta_2} =$$

$$= -\frac{1}{2\sqrt{\theta_1\theta_2}}$$

Substituting ~~for~~ $\theta_1 = -\frac{\lambda}{2\mu^2}$ and $\theta_2 = -\frac{\lambda}{2}$ gives

$$\mathrm{Cov}(X, 1/X) = -\frac{\mu}{\lambda}$$

4(a) Denote $\lambda(\psi,\xi)$ the parameter λ as a function of ψ, ξ

Then the log-likelihood on ψ and ξ is

$$\ell(\psi,\xi) = s_1(Y)\theta_1(\psi) + s_2(Y)\theta_2(\lambda(\psi,\xi)) - k(\theta_1(\psi), \theta_2(\lambda(\psi,\xi))) + c(y)$$

For ψ and ξ to be orthogonal $i_{\psi\xi}(\psi,\xi) = 0$, where $i_{\psi\xi}$ is the component of the Fisher information matrix that corresponds ~~to~~ ψ and ξ.

The score function on ξ is

$$\ell_\xi(\psi,\xi) = \frac{\partial \ell(\psi,\xi)}{\partial \xi} = s_2(Y)\frac{\partial \theta_2(\lambda(\psi,\xi))}{\partial \xi} - \frac{\partial k(\theta_1(\psi), \theta_2(\lambda(\psi,\xi)))}{\partial \xi}$$

and so

$$\ell_{\psi\xi}(\psi,\xi) = \frac{\partial^2 \ell(\psi,\xi)}{\partial\psi\,\partial\xi} = s_2(Y)\frac{\partial^2 \theta_2(\lambda(\psi,\xi))}{\partial\psi\,\partial\xi} - \frac{\partial^2 k(\theta_1(\psi), \theta_2(\lambda(\psi,\xi)))}{\partial\psi\,\partial\xi}$$

Hence,

$$i_{\psi\xi}(\psi,\xi) = \mathbb{E}\left(-\ell_{\psi\xi}(\psi,\xi)\right) =$$

$$= -\xi\frac{\partial^2 \theta_2(\lambda(\psi,\xi))}{\partial\psi\,\partial\xi} + \frac{\partial^2 k(\theta_1(\psi), \theta_2(\lambda(\psi,\xi)))}{\partial\psi\,\partial\xi}$$

The above expression can be shown to be 0, by noting that $\mathbb{E}(\ell_\xi(\psi,\xi)) = 0$ and calculating $\frac{\partial}{\partial\psi}\mathbb{E}(\ell_\xi(\psi,\xi))$.

4(b) $f_Y(y,\varphi,n) = \exp\{-\frac{1}{\varphi}y + n\log y - n\log\varphi - \log\Gamma(n) - \log y\}$

So this is of the exponential family form with natural parameters $\theta_1 = n$, $\theta_2 = -1/\varphi$, natural statistics $s_1(y) = \log y$, $s_2(y) = y$.

Now,

$\mathbb{E}(s_2(Y)) = \mathbb{E}(Y) = n\varphi$. Hence, by 4(a), $n\varphi$ is orthogonal to n.

5(a)

The log-likelihood on μ, σ^2 (omitting any quantities that do not depend on μ and σ^2) is

$$\ell(\mu,\sigma^2) = -\frac{1}{2} n \log \sigma^2 - \frac{\sum(x_i-\mu)^2}{2\sigma^2}$$

The maximum likelihood estimate of σ^2 for fixed μ is

$$\hat{\sigma}_\mu^2 = \frac{1}{n}\sum(x_i-\mu)^2$$

[if you are not confident using the above result directly, you may verify it by solving $\partial\ell(\mu,\sigma^2)/\partial\sigma^2 = 0$]

Hence the profile likelihood on μ is

$$\ell_p(\mu) = -\frac{1}{2} n \log \hat{\sigma}_\mu^2 - \frac{n}{2}$$

$$= -\frac{1}{2} n \log \frac{\sum(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2}{n} - \frac{1}{2}n$$

$$= -\frac{1}{2} n \log\left(\hat{\sigma}^2 + (\mu-\hat{\mu})^2\right) - \frac{1}{2}n$$

5(b) The approximate $100(1-\alpha)\%$ confidence interval based on $\ell_p(\mu)$ is $\{\mu : 2(\ell_p(\hat{\mu}) - \ell_p(\mu)) \le \chi^2_{1-\alpha}\} =$

$$\left\{\mu : n\log\left\{1 + \frac{(\mu-\hat{\mu})^2}{\hat{\sigma}^2}\right\} \le \chi^2_{1-\alpha}\right\} =$$

$$\left\{\mu : 1 + \frac{(\mu-\hat{\mu})^2}{\hat{\sigma}^2} \le \exp\left\{\frac{\chi^2_{1-\alpha}}{n}\right\}\right\} =$$

$$\vdots$$

$$\left\{\mu : \hat{\mu} - \hat{\sigma}\sqrt{\exp\left\{\frac{\chi^2_{1-\alpha}}{n}\right\} - 1} \le \mu \le \hat{\mu} + \hat{\sigma}\sqrt{\exp\left\{\frac{\chi^2_{1-\alpha}}{n}\right\} - 1}\right\}$$

6(a)

The likelihood ratio statistic for testing $H_0: \mu = \mu_0$ is

$$w(\mu_0) = +2\left\{ \ell(\hat{\mu}, \hat{\sigma}^2) - \ell(\mu_0, \hat{\sigma}^2_{\mu_0}) \right\}$$

$$= 2\left\{ \ell_p(\hat{\mu}) - \ell_p(\mu_0) \right\}$$

[see exercise 5]

$$= n \log\left\{ 1 + \frac{(\mu_0 - \hat{\mu})^2}{\hat{\sigma}^2} \right\}$$

where $\hat{\mu} = \overline{x}$ and $\hat{\sigma}^2 = \Sigma(x_i - \overline{x})^2/n$

But
$$\frac{(\mu_0 - \hat{\mu})^2}{\hat{\sigma}^2} = \frac{(\overline{x} - \mu_0)^2}{\frac{\Sigma(x_i - \overline{x})^2}{n}} = \frac{n(\overline{x} - \mu_0)^2}{\frac{(n-1)\Sigma(x_i - \overline{x})^2}{(n-1)}} = \frac{t^2}{n-1}$$

Using the expansion $\log(1+x) = x - \frac{x^2}{2} + \ldots$

$$w(\mu_0) = n \log\left\{1 + \frac{t^2}{n-1}\right\} = \frac{n}{n-1}t^2 - \frac{n}{2(n-1)^2}t^4 + O_p(n^{-2})$$

and so

$$\mathbb{E}_{\mu_0}(w(\mu_0)) = \frac{n}{n-1}\mathbb{E}(t^2) - \frac{n}{2(n-1)^2}\mathbb{E}(t^4) + O(n^{-2})$$

R1
$$= \frac{n}{\cancel{n-1}}\,\frac{\cancel{n-1}}{n-3} - \frac{n}{2\cancel{(n-1)^2}}\,\frac{3\cancel{(n-1)^2}}{(n-3)(n-5)} + O(n^{-2})$$

$$= \frac{1}{1 - 3/n} - \frac{3}{2n}\,\frac{1}{(1 - 3/n)(1 - 5/n)} + O(n^{-2})$$

R2
$$= 1 + \frac{3}{n} - \frac{3}{2n}\left[1 + O(n^{-1})\right] + O(n^{-2})$$

$$= 1 + \frac{3}{2n}.$$

A Bartlett corrected version of $w(\mu_0)$ is

$$\tilde{w}(\mu_0) = \frac{w(\mu_0)}{1 + \frac{3}{2n}} \doteq \frac{2n\,w(\mu_0)}{2n+3}$$