MATH 557 - MID-TERM 2008 - SOLUTIONS

1. (a) Note first that by standard expansion into a quartic polynomial

$$\left(\frac{x-\theta}{\sigma}\right)^4 = w_0(\theta,\sigma) + \sum_{j=1}^{k} w_j(\theta,\sigma)x^j = w_0(\theta,\sigma) + \sum_{j=1}^{k} w_j(\theta,\sigma)t_j(x)$$

say, where $w_j(\theta,\sigma)$ are constant functions of θ and σ. Thus

$$f_{X|\theta,\sigma}(x|\theta,\sigma) = h(x)c(\theta,\sigma)\exp\left\{\sum_{j=1}^{k} w_j(\theta,\sigma)t_j(x)\right\}$$

where

$$h(x) = 1 \qquad c(\theta,\sigma) = \exp\{w_0(\theta,\sigma) - \kappa(\theta,\sigma)\} \qquad t_j(x) = x^j,\ j = 1,\ldots,4.$$

and hence the distribution is an Exponential Family distribution.

6 Marks

(b) By inspection, and using the Neyman factorization theorem in this Exponential family setting, we have

$$\underset{\sim}{T}(\underset{\sim}{X}) = \left(T_1(\underset{\sim}{X}), T_2(\underset{\sim}{X}), T_3(\underset{\sim}{X}), T_4(\underset{\sim}{X})\right)^{\mathsf{T}} \qquad T_j(\underset{\sim}{X}) = \sum_{i=1}^{n} t_j(X_i) = \sum_{i=1}^{n} X_i^j \qquad j = 1,\ldots,4$$

is a sufficient statistic.

6 Marks

2. (a) We have

$$\begin{aligned} f_{\underset{\sim}{X}|\theta}(\underset{\sim}{x}|\theta) &= \left(\frac{1}{2\pi\theta^2}\right)^{n/2} \exp\left\{-\frac{1}{2\theta^2}\sum_{i=1}^{n}(x_i-\theta)^2\right\} \\ &= \left(\frac{1}{2\pi\theta^2}\right)^{n/2} \exp\left\{-\frac{1}{2\theta^2}\sum_{i=1}^{n}x_i^2 + \frac{1}{\theta}\sum_{i=1}^{n}x_i - \frac{n}{2}\right\} \\ &= \left(\frac{1}{2\pi\theta^2}\right)^{n/2} e^{-n/2} \exp\left\{-\frac{1}{2\theta^2}T_2(\underset{\sim}{x}) + \frac{1}{\theta}T_1(\underset{\sim}{x})\right\} \end{aligned}$$

so that $\underset{\sim}{T}(\underset{\sim}{X}) = \left(T_1(\underset{\sim}{X}), T_2(\underset{\sim}{X})\right)^{\mathsf{T}}$ where

$$T_1(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i \qquad T_2(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i^2$$

is a sufficient statistic. Note that $T_1(\underset{\sim}{X})$ and $T_2(\underset{\sim}{X})$ are linearly independent, and that for two vectors $\underset{\sim}{x}, \underset{\sim}{y}$

$$\frac{f_{\underset{\sim}{X}|\theta}(\underset{\sim}{x}|\theta)}{f_{\underset{\sim}{X}|\theta}(\underset{\sim}{y}|\theta)} = \exp\left\{-\frac{1}{2\theta^2}(T_2(\underset{\sim}{x}) - T_2(\underset{\sim}{y}) + \frac{1}{\theta}(T_1(\underset{\sim}{x}) - T_1(\underset{\sim}{y}))\right\}$$

does not depend on θ iff $T_1(\underset{\sim}{x}) = T_1(\underset{\sim}{y})$ and $T_2(\underset{\sim}{x}) = T_2(\underset{\sim}{y})$. Thus $\underset{\sim}{T}$ is minimal sufficient.

6 Marks

(b) For each i, we have from the formula sheet properties of Normal distributions that

$$E_{f_{X_i|\theta}}[X_i^2] = 2\theta^2$$

so that

$$E_{f_{T_2|\theta}}[T_2] = 2n\theta^2$$

Also from mgf results.

$$T_1(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i \sim Normal(n\theta, n\theta^2)$$

Thus if $S(\underset{\sim}{X}) = \{T_1(\underset{\sim}{X})\}^2$.

$$E_{f_{S|\theta}}[S] = n\theta^2 + (n\theta)^2 = n(n+1)\theta^2.$$

Therefore the function

$$g(t_1, t_2) = \frac{t_1^2}{n(n+1)} - \frac{t_2}{2n}$$

is such that $E_{f_{T_1,T_2|\theta}}[g(T_1, T_2)] = 0$, and hence $\underset{\sim}{T}(\underset{\sim}{X})$ is not complete.

6 Marks

3. (a) In the Poisson model, the likelihood is

$$L(\lambda|\underset{\sim}{x}) \propto \lambda^{\sum\limits_{i=1}^{n} x_i} e^{-n\lambda}$$

so therefore the conjugate prior is $Gamma(\alpha, \beta)$

$$\pi_\lambda(\lambda) \propto \lambda^{\alpha-1} e^{-\beta\lambda}$$

yielding posterior

$$\pi_{\lambda|\underset{\sim}{x}}(\lambda|\underset{\sim}{x}) \propto \lambda^{\alpha + \sum\limits_{i=1}^{n} x_i - 1} e^{-(\beta+n)\lambda} \equiv Gamma(\alpha + \sum_{i=1}^{n} x_i, \beta + n).$$

From lectures, the estimate under squared-error loss is the posterior mean, that is, from the formula sheet

$$\widehat{\lambda}_B(\underset{\sim}{x}) = \frac{\alpha + \sum\limits_{i=1}^{n} x_i}{\beta + n} = \frac{n}{\beta + n}\,\overline{x}_n + \frac{\beta}{\beta + n}\left(\frac{\alpha}{\beta}\right) = w_n \overline{x}_n + (1 - w_n) m$$

10 Marks

(b) If $T(\underset{\sim}{X}) = a\overline{X}_n + b$, then

$$\begin{aligned} R_T(\lambda) &= \int_{\mathcal{X}} (a\overline{X}_n + b - \lambda)^2 f_{\underset{\sim}{X}|\lambda}(\underset{\sim}{x}|\lambda)\, d\underset{\sim}{x} = \int_{\mathcal{X}} ((a\overline{X}_n - a\lambda) + (b + (a-1)\lambda))^2 f_{\underset{\sim}{X}|\lambda}(\underset{\sim}{x}|\lambda)\, d\underset{\sim}{x} \\ &\geq \int_{\mathcal{X}} (a\overline{X}_n - a\lambda)^2 f_{\underset{\sim}{X}|\lambda}(\underset{\sim}{x}|\lambda)\, d\underset{\sim}{x} \\ &> \int_{\mathcal{X}} (\overline{X}_n - \lambda)^2 f_{\underset{\sim}{X}|\lambda}(\underset{\sim}{x}|\lambda)\, d\underset{\sim}{x} = R_{T_0}(\lambda) \end{aligned}$$

as $a > 1$, where

$$T_0(\underset{\sim}{X}) = \overline{X}_n.$$

This follows by expanding the integrand in the second integral of line 1, noting that the integral of the cross term is zero, and that the integral of the $(b + (a-1)\lambda))^2$ term is non-negative. Hence T is inadmissible.

6 Marks