Instruction: Please show your work for full credit.

1. Let X denote the failure time of a computer monitor and suppose that X has the continuous density

$$p(x,\theta) = \begin{cases} 2\theta x \exp(-\theta x^2), & \text{if } x > 0, \\ 0, & \text{otherwise,} \end{cases}$$

where $\theta \in \Theta = (0, \infty)$. Let $X_1, \ldots, X_n$ be i.i.d. as X, and let $p(\boldsymbol{x}, \theta)$ denote the joint density of $X_1, \ldots, X_n$.

   (a) Find the distribution function $F(x)$ of X. Express θ in terms of $F(x)$. Suggest an empirical plug-in estimate of θ.

   (b) Express $p(\boldsymbol{x}, \theta)$ as a canonical exponential family $q(\boldsymbol{x}, \eta)$. Give the canonical parameter space.

   (c) Give the natural sufficient statistic $T(\boldsymbol{X})$ and find $E_\theta\{T(\boldsymbol{X})\}$ and $\text{var}_\theta\{T(\boldsymbol{X})\}$.

   (d) Find the MLEs of η and θ.

   (e) Give a conjugate prior for $p(\boldsymbol{x}, \theta)$. Include the parameter space for θ and for the parameters in the prior density.

2. Consider the model

$$Y = \beta_0 + \beta_1 X_1 + b(X_1) X_2 + \varepsilon,$$

where ε is independent of (X_1, X_2), $E(\varepsilon) = 0$, $E(\varepsilon^2) = \sigma^2$ and $b(\cdot)$ is a function from $\mathbb{R}$ to $\mathbb{R}$. This model is called a varying-coefficient model.

   (a) Show that if $\text{var}(X_2 \mid X_1 = x_1)$ and $\text{cov}(X_2, Y \mid X_1 = x_1)$ exist, then

$$b(x_1) = \frac{\text{cov}(X_2, Y \mid X_1 = x_1)}{\text{var}(X_2 \mid X_1 = x_1)}.$$

   Hint: $\text{cov}(X_2, Y \mid X_1 = x_1) = E(X_2 Y \mid X_1 = x_1) - E(X_2 \mid X_1 = x_1) E(Y \mid X_1 = x_1)$.

   (b) Suppose $b(X_1) = \beta_2 + \beta_3 X_1$, and suppose that $(X_{11}, X_{12}, Y_1), \ldots, (X_{n1}, X_{n2}, Y_n)$ are i.i.d. as (X_1, X_2, Y). Conditioning on $(X_{i1}, X_{i2}) = (x_{i1}, x_{i2})$, $i = 1, \ldots, n$, find a least-squares estimate $\widehat{\boldsymbol{\beta}}$ of $(\beta_0, \beta_1, \beta_2, \beta_3)^T$. What assumptions do you need for $\widehat{\boldsymbol{\beta}}$ to be unique?

   Hint: Ignore part (a). Also introduce the variable $X_3 = X_1 X_2$.

3. Let $X_1, \ldots, X_n$ be i.i.d. as X where $X \sim p_\theta$. Write $\boldsymbol{X} = (X_1, \ldots, X_n)^T$ and let $\widehat{\theta} = \widehat{\theta}(\boldsymbol{X})$ be an unbiased estimate of θ. Assume that $0 < \text{var}(\widehat{\theta}) < \infty$ and consider the loss function $l(\theta, a) = (\theta - a)^2$. An estimate of θ of the form $\widetilde{\theta}_c = c\widehat{\theta}$ with $0 < c < 1$ is called a shrinkage estimate.

(a) Show that the risk of $\widetilde{\theta}_c$ is minimized when

$$c = \frac{\theta^2}{\theta^2 + \mathrm{var}(\widehat{\theta})}.$$

Give $\inf_c R(\theta, \widetilde{\theta}_c)$.

Hint: $E\{(c\widehat{\theta} - \theta)^2\} = E\{(c\widehat{\theta} - c\theta)^2 + (c\theta - \theta)^2\}$.

(b) Show that $R(\theta, \widetilde{\theta}_c) < R(\theta, \widehat{\theta})$ if and only if

$$\frac{1-c}{1+c} < \frac{\mathrm{var}(\widehat{\theta})}{\theta^2}.$$