Problem Set # 1

Econ 722

Spring 2018

Instructions: Answer each of the following. Questions 4 and 5 require numerical calculations. For these, please submit full source code, commented and cleanly formatted, along with your answers. All solutions must be submitted electronically on canvas by 1:30pm on Thursday, March 22nd. Late problem sets will not be accepted: it is much better to turn in partial solutions rather than nothing at all. You may discuss these problems with your classmates, but if you work together please list the names of the students with whom you have collaborated at the top of your solutions.

1. Let Θ be a discrete set and π be a prior distribution that gives strictly positive probability to each element of Θ. Show that if $\widehat{\theta}$ is a Bayes rule with respect to π, it is admissible.

2. Derive the KL divergence from a $N(\mu_0, \sigma_0^2)$ distribution to a $N(\mu_1, \sigma_1^2)$ distribution.

3. Suppose we observe a random sample $\{(\mathbf{x}_t, y_t)\}_{t=1}^T$ from some population and decide to predict y from $\mathbf{x}$ using the following linear model:

$$y_t = \mathbf{x}_t'\beta + \varepsilon_t$$

   Let $\widehat{\beta}$ denote the ordinary least squares estimator of β based on $\{(\mathbf{x}_t, y_t)\}_{t=1}^T$. Now suppose that we observe a *second* random sample $\{(\widetilde{\mathbf{x}}_t, \widetilde{y}_t)\}_{t=1}^T$ from the sample population that is *independent* of the first. Show that

$$\mathbb{E}\left[\frac{1}{T}\sum_{t=1}^T (y_t - \mathbf{x}_t'\widehat{\beta})^2\right] \leq \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^T (\widetilde{y}_t - \widetilde{\mathbf{x}}_t'\widehat{\beta})^2\right]$$

   In other words, show that the in-sample squared prediction error is an overly optimistic estimator of the out-of-sample squared prediction error.

4. Let $X \sim N(\theta, 1)$ where $\theta \in \Theta = [-m, m]$ $m > 0$. Further define:

$$\pi(\theta) = \begin{cases} 1/2, & \theta = -m \\ 0, & -m < \theta < m \\ 1/2, & \theta = m \end{cases}$$

   (a) Show that $\widehat{\theta}(X) = m \tanh(mX)$ is the Bayes rule with respect to π under squared error loss. Recall that $\tanh(z) = (e^z - e^{-z})/(e^z + e^{-z})$.

   (b) Write code to calculate the risk function $R(\theta, \widehat{\theta})$ and Bayes risk $r(\pi, \widehat{\theta})$ numerically. I suggest numerical integration rather than a simulation-based approximation.

   (c) For each value of $m \in \{0.5, 0.75, 1, 1.25\}$ plot the following:

      (i) X vs. $\widehat{\theta}(X)$ for $X \in [-3, 3]$, with the 45-degree line indicated in red.

      (ii) θ vs. $R(\theta, \widehat{\theta})$ for $\theta \in [-m, m]$ with $r(\pi, \widehat{\theta})$ indicated as a red horizontal line.

   (d) Explain your findings from part (c) above. How does $\widehat{\theta}$ compare to the MLE? For which, if any, of the values of $m \in \{0.5, 0.75, 1, 1.25\}$ is $\widehat{\theta}$ minimax?

5. Let $X \sim N(\theta, I)$ where θ is a p-vector for $p \geq 3$ and I is the $(p \times p)$ identity matrix. For this problem, we showed in class that the maximum likelihood estimator for θ, $\widehat{\theta} = X$ is inadmissible, as it is dominated by the James-Stein estimator:

$$\widehat{\theta}^{JS} = \widehat{\theta}\left(1 - \frac{p-2}{\widehat{\theta}'\widehat{\theta}}\right)$$

   I also argued, without proof, that the James-Stein estimator is itself inadmissible, as it is dominated by the so-called "positive-part" James-Stein estimator, namely

$$\widetilde{\theta}^{JS} = \widehat{\theta}\left[\max\left\{1 - \frac{p-2}{\widehat{\theta}'\widehat{\theta}}\,,\, 0\right\}\right]$$

   This estimator takes its name from the fact that, unlike the plain-vanilla James-Stein estimator, it can never shrink "past" zero and hence cannot have a different sign than the MLE. Design and carry out a simulation experiment comparing the risk of $\widehat{\theta}$, $\widehat{\theta}^{JS}$, and $\widetilde{\theta}^{JS}$ under squared error loss. Your results should be based on 10,000 simulation replications over a range of values for $p \geq 3$ and different configurations of the true mean vector θ. Write a brief summary of your results, accompanied by tables and or figures, as needed.