Homework 4: Due 02/19/2018

1. (5 points) Consider the Bernoulli example we used to motivate the invariant estimation. Let $X_1, \cdots, X_n \overset{\text{iid}}{\sim} Ber(p), p \in (0,1)$. Often only the total count is recorded, $Y = \sum_{i=1}^n X_i \sim Bin(n,p), p \in (0,1)$. Let $\mathcal{P} = \{Bin(n,p) : p \in (0,1)\}$, and $g(x) = n - x$. Please clearly define the following:

   (a) the sample space, $\mathcal{X}$ (for Y)

   (b) parameter space, Θ

   (c) the group of transformations on the sample space, G

   (d) the induced group of transformations on the parameter space, $\bar{G}$

   (e) the orbits of $\bar{G}$. (Is the group transitive?)

2. (15 points) Although invariance seems a very natural, maybe the most convincing, principle in estimation problem, it may run into difficulties. For example, an estimation problem might be invariant under two different groups, and each group leads to a different MRIE.
   Let $(X_1, X_2) \sim N_2(0, \Sigma)$ and $(Y_1, Y_2) \sim N_2(0, \theta\Sigma), \theta > 0$. Suppose (X_1, X_2) and (Y_1, Y_2) are independent. Consider the problem of estimating θ.

   (a) Let $g_1(x_1, x_2, y_1, y_2) = (x_1^*, x_x^*, y_1^*, y_2^*)$, where

   $$\begin{aligned} x_1^* &= c_1 x_1 + c_2 x_2 & \qquad y_1^* &= d(c_1 y_1 + c_2 y_2) \\ x_2^* &= b x_2 & y_2^* &= d b y_2 \end{aligned}$$

   Denote $G_1 = \{g_1 : c_1, c_2, b, d > 0\}$. (A subgroup of scale transformations in $\mathbb{R}^2$.) Show that the model is invariant under G_1 (the distribution family is invariant).

   (b) Show that the loss function is invariant under G_1 iff $L(\theta, a) = L(\frac{a}{\theta})$.

   (c) Show that an estimator δ is invariant under G_1 iff $\delta(x_1, x_2, y_1, y_2) = k\frac{y_2^2}{x_2^2}$, for some value of k (a.e.).

   (d) Show that an estimator δ^* is an MRIE under G_1, iff, $\delta^* = k^* \frac{y_2^2}{x_2^2}$, where k^* minimizes

   $$E_\theta \left[L(k\frac{Y_2^2}{\theta X_2^2}) \right] = E_1 \left[L(k\frac{Y_2^2}{X_2^2}) \right]$$

   (e) Now consider G_2 by switching x_1^* and x_2^* and switching y_1^* and y_2^*. What is an MRIE under G_2? Please comment on how it compares with an MRIE under G_1.

3. (5 points) Recall that we have defined risk-unbiasness in location invariant problem. An estimator δ is called risk-unbiased if

   $$E_\theta \left[L(\theta, \delta(X)) \right] \le E_\theta \left[L(\theta^*, \delta(X)) \right], \quad \forall \theta^* \in \Theta.$$

In another word, on the average, δ is at least as close to the true value (θ or a function of it) as it is to any false value (θ^* or a function of it). Furthermore, a group, G is called **commutative** iff $g_1 g_2 = g_2 g_1, \forall g_1, g_2 \in G$. Please prove that an MRIE is risk-unbiased if $\bar{G}$ is transitive and $\tilde{G}$ is commutative.

4. (15 points) Consider the location scale family generated by a known density f_{01} on $\mathbb{R}^p$.

$$\mathcal{P} = \{f(\boldsymbol{x}|\mu, \sigma) : f(\boldsymbol{x}|\mu, \sigma) = \frac{1}{\sigma^p} f_{01}(\frac{\boldsymbol{x} - \mu\mathbf{1}}{\sigma}), \mu \in \mathbb{R}, \sigma \in \mathbb{R}^+\}.$$

Consider the transformations $G = \{g : g(x) = cx + b, c > 0, b \in \mathbb{R}\}$. The model is invariant under G.

(a) Show that the right-invariant measure on

$$\bar{G} = \{\bar{g} : \bar{g}(\mu, \sigma) = (c\mu + b, c\sigma)c > 0, b \in \mathbb{R}\}$$

satisifies

$$\mu_r(A \times B) = \int_A \int_B \frac{1}{c} \; db \; dc,$$

so that μ_r has density $1/c$ with respect to Lebesgue measure on $\mathbb{R}\times\mathbb{R}^+$.Furthermore, the induced right-invariant prior on $\Theta = \{(\mu, \sigma)\} \in \mathbb{R} \times \mathbb{R}^+$, π_r is same as μ_r.

(b) Let $X_1, X_2, \cdots, X_n \overset{\text{iid}}{\sim} N(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+$.
(That is, f_{01} is the density of standard normal distribution, and $p = n$.)
Derive the posterior distribution of (μ, σ^2) under π_r. Describe the distribution name and the parameter of $\mu|\bar{x}, \sigma^2$ and $1/\sigma^2|\bar{x}$.
Remark: the marginal posterior of $\mu|\bar{x}$ can be obtained by integrating the joint pdf over σ^2, which is

$$\frac{\mu - \bar{x}}{s/\sqrt{n}}\Big|(X_1, \cdots, X_n) = \frac{\mu - \bar{x}}{s/\sqrt{n}}\Big|(\bar{x}, s) \sim t_{n-1}$$

(c) Derive the MRIE for σ^2 under loss function

$$L((\mu, \sigma^2), a) = (\sigma^2 - a)^2/\sigma^4.$$

Compare it with the result in Example 4.14 on Page 257 of Shao 2003.

(d) Note that in the previous part, the loss function assigns much heavier penalities to overestimation than to underestimation since

$$\lim_{a\to\infty} L((\mu, \sigma^2), a) = \infty, \quad \lim_{a\to 0} L((\mu, \sigma^2), a) = 1.$$

Alternatively, Stein's loss,

$$L((\mu, \sigma^2), a) = (a/\sigma^2) - log(a/\sigma^2) - 1,$$

is more evenhanded. Derive the MRIE under Stein's loss and compare it with the previous MRIE.