STAT 581
Homework #8
11/25/2013

Solutions for Hoff Exercises 8.4, 8.7, and 8.8

8.4 Let $\mathcal{P}$ be the class of all probability distributions on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with continuous CDFs. Let $X_1, \ldots, X_n$ be an iid sample from $P \in \mathcal{P}$, and consider estimation of the median via the loss $L(P, d) = (P(-\infty, d] - 1/2))^2$. Show that the estimation problem is invariant under the group $\mathcal{G}$ of transformations of the form $g_h(x_1, \ldots, x_n) = (h(x_1), \ldots, h(x_n))$, where h is a continuous, strictly increasing function on $\mathbb{R}$. Identify the induced transformations $\bar{\mathcal{G}}$ and $\tilde{\mathcal{G}}$.

As setup, we first note that each distribution in the class $\mathcal{P}$ can be uniquely identified by the CDF F associated with X_1. Let $\mathcal{F}$ be the set of all continuous CDFs. We will use the notation P_F to refer to the model $P \in \mathcal{P}$ where X_1 is assumed to have CDF $F \in \mathcal{F}$. In this way, we can think of $\mathcal{F}$ as our parameter space, even though $\mathcal{F}$ isn't a space with a finite number of dimensions.

Now we proceed in several steps:

a) For an arbitrary g_h, identify the distribution of $g_h \boldsymbol{X}$, where $X_1, \ldots, X_n \sim_{iid} P_F$. (This will give us a handle on $\bar{g}_h$.)

b) Check invariance of the parameter space, i.e. $\bar{g}_h \mathcal{F} = \mathcal{F}$.

c) Identify the induced transformation on decision rules to attain invariance of loss.

---

a) Say $X_1, \ldots, X_n \sim_{iid} P_F$. Then $X_1, \ldots, X_n$ have joint cdf given by

$$\Pr(X_1 \le x_1, \ldots, X_n \le x_n) = \prod_i F(x_i).$$

What about the transformed variables $g_h \boldsymbol{X}$? Let $X_i' = h(X_i)$. Then

$$\begin{aligned}
\Pr(X_1' \le x_1, \ldots, X_n' \le x_n) &= \Pr(h(X_1) \le x_1, \ldots, h(X_n) \le x_n) \\
&= \Pr(X_1 \le h^{-1}(x_1), \ldots, X_n \le h^{-1}(x_n)) \\
&= \Pr(X_1 \le h^{-1}(x_1)) \cdots \Pr(X_n \le h^{-1}(x_n)) \\
&= \prod_i F(h^{-1}(x_i))
\end{aligned}$$

That is, $X_1', \ldots, X_n' \sim_{iid} P_{F \circ h^{-1}}$. That is, $\bar{g}_h$ is given by $\bar{g}_h : F \mapsto F \circ h^{-1}$.

The induced group of transformations is given by $\bar{\mathcal{G}} = \{\bar{g}_h : F \mapsto F \circ h^{-1}\}$.

b) Firstly, we note that since h is continuous and strictly increasing, its inverse exists and is also continuous. Thus, $F \circ h^{-1}$ is also a continuous cdf, so applying any transformation g_h to $\boldsymbol{X}$ won't cause us to leave the model.

Secondly, fix a transformation g_h. We need to show that $\bar{g}_h$ doesn't reduce $\mathcal{F}$. That is, for each $F \in \mathcal{F}$, there is an $F' \in \mathcal{F}$ such that $\boldsymbol{X} \sim P_{F'} \Rightarrow g_h\boldsymbol{X} \sim P_F$. Fixing an F and taking $F' = F \circ h$, we see that $\bar{g}_h F' = F' \circ h^{-1} = F \circ h \circ h^{-1} = F$, so $\bar{g}_h$ doesn't reduce $\mathcal{F}$.

Together, these show invariance of the parameter space under the group of parameter transformations $\bar{\mathcal{G}}$.

c) Finally, we look at loss. For invariance of loss, we want to find a definition of $\tilde{g}_h$ that satisfies $L(P_F, d) = L(P_{\bar{g}_h F}, \tilde{g}_h d)$. Working backwards, one option is

$$\begin{aligned}
& L(P_F, d) = L(P_{\bar{g}_h F}, \tilde{g}_h d) \\
\iff & (F(d) - 1/2)^2 = (F(h^{-1}(\tilde{g}_h(d))) - 1/2)^2 \\
\Leftarrow & F(d) = F(h^{-1}(\tilde{g}_h(d))) \\
\Leftarrow & d = h^{-1}(\tilde{g}_h(d)) \\
\Leftarrow & \tilde{g}_h(d) = h(d).
\end{aligned}$$

That is, taking $\tilde{g}_h : d \mapsto h(d)$ gives us invariance of loss. The induced group of transformations is then $\tilde{\mathcal{G}} = \{\tilde{g}_h : d \mapsto h(d)\}$.

Altogether, from parts a, b, and c, we have invariance of the estimation problem under $\mathcal{G}$. ∎

8.7 Consider a scale model $\mathcal{P} = \{p_\theta(x) = p_1(x/\theta)/\theta : \theta > 0\}$, where p_1 is a known probability density on $\mathbb{R}^+$.
8.7.a Show that the model is invariant under the group $\mathcal{G} = \{g : x \mapsto cx, c > 0\}$, and identify the induced group $\bar{\mathcal{G}}$ on the parameter space.

Firstly, if $X \sim P_\theta$, what's the distribution of $g_c X = cX$? Appealing to the known formula for scale transformations, since X has pdf $p_1(x/\theta)/\theta$, cX has pdf $p_1(x/c\theta)/c\theta$. That is, $cX \sim P_{c\theta}$, suggesting $\bar{g}_c(\theta) = c\theta$. Or, in group form, $\bar{\mathcal{G}} = \{\bar{g}_c : \theta \mapsto c\theta, c > 0\}$.

For model invariance, we need to check that $\bar{g}_c\Theta = \Theta$. Abusing notation somewhat, we're checking that $c\mathbb{R}^+ = \mathbb{R}^+$, which is indeed true so long as $c > 0$. ∎

8.7.b Show that the problem of estimating θ is invariant under $\mathcal{G}$ for loss functions of the form $L(\theta, d) = f(d/\theta)$, where f attains its minimum value of zero at $d = \theta$. Identify the induced group $\tilde{\mathcal{G}}$ on the decision space.

For loss invariance, we want to find $\tilde{g}_c$ such that $L(\theta, d) = L(\bar{g}_c\theta, \tilde{g}_c d)$.

$$\begin{aligned}
&L(\theta, d) = L(\bar{g}_c\theta, \tilde{g}_c d) \\
\iff &f(d/\theta) = f(\tilde{g}_c d / c\theta) \\
\Leftarrow &\frac{d}{\theta} = \frac{\tilde{g}_c d}{c\theta} \\
\iff &\tilde{g}_c d = cd.
\end{aligned}$$

That is, one option for invariance of loss is to take $\tilde{g}_c d = cd$. Depending on the function f, it may not be the only option, but it's certainly sufficient for invariance of loss. Thus, the induced group we'll consider on the decision space is $\tilde{\mathcal{G}} = \{\tilde{g}_c : d \mapsto cd, c > 0\}$. ■

8.7.c Characterize the class of equivariant estimators and the UMREE in terms of f and p_1.

For equivariance, we require that $\delta(g_c(x)) = \tilde{g}_c(\delta(x))$. That is, $\delta(cx) = c\delta(x)$. Using our usual trick of considering $c = 1/x$, we see that it is necessary and sufficient to have $\delta_a(X) = aX$ for some $a \in \mathbb{R}$.

As we've seen in the course notes, this particular $\bar{\mathcal{G}}$ acts transitively on the parameter space, so the risk of any equivariant estimator is constant across the parameter space. Thus, to find the UMREE, it suffices to find the estimator which minimizes risk at a convenient θ, say $\theta = 1$. Then

$$\begin{aligned}
R(1, aX) &= \mathrm{E}[L(1, aX)] \\
&= \mathrm{E}[f(aX)|\theta = 1] \\
&= \int f(ax)p_1(x)\,dx
\end{aligned}$$

If we let a^* be the value of a which minimizes the above integral, then the UMREE is given by a^*X. ■

8.7.d In the case $f(r) = (1-r)^2$, find the UMREE in terms of p_1.

Plugging in for f above, we now want to find a which minimizes

$$\int (1 - ax)^2 p_1(x)\,dx.$$

Equivalently, we're minimizing

$$\int 1 p_1(x)\,dx - 2a \int x p_1(x)\,dx + a^2 \int x^2 p_1(x)\,dx.$$

This is a quadratic function in a which is minimized by at

$$a^* = \frac{\int x p_1(x)\,dx}{\int x^2 p_1(x)\,dx} = \frac{\mathrm{E}[X|\theta = 1]}{E[X^2|\theta = 1]}.$$

Thus, the UMREE is given by

$$\delta_{a^*} = \frac{\int x p_1(x)\, dx}{\int x^2 p_1(x)\, dx} X = \frac{\mathrm{E}[X|\theta = 1]}{E[X^2|\theta = 1]} X.$$

■

> **8.7.e** For the case $f(r) = (1-r)^2$, show that the risk of any equivariant estimator can be expressed as a posterior risk under a (possibly improper) prior distribution. Show that the UMREE can be viewed as a Bayes estimator under this prior.

What's the risk of an equivariant estimator?

$$\begin{aligned}
R(\theta, aX) &= R(1, aX) \\
&= \int_0^\infty f(au) p_1(u)\, du \\
&= \int_0^\infty f(ud_x/x) p_1(u)\, du \text{ (where } d_x = ax) \\
&= \int_\infty^0 f(d_x/\theta) p_1(x/\theta) \frac{-x}{\theta^2}\, d\theta \text{ (change of variables : } \theta = x/u) \\
&= \int_0^\infty f(d_x/\theta) p_1(x/\theta) \frac{x}{\theta^2}\, d\theta,
\end{aligned}$$

which looks like a posterior expected loss if we can just find a posterior distribution of the form

$$\pi(\theta|x) = p_1(x/\theta) \frac{x}{\theta^2}.$$

Recalling that $\pi(\theta|x) \propto_\theta p(x|\theta)\pi(\theta)$, and that $p(x|\theta) = p_1(x/\theta)/\theta$, it looks like we should be okay by taking a prior $\pi(\theta) \propto_\theta \frac{x}{\theta}$, or equivalently, $\pi(\theta) \propto_\theta \frac{1}{\theta}$. We can then confirm by computing

$$\pi(\theta|x) = \frac{p(x|\theta)\pi(\theta)}{\int p(x|\theta)\pi(\theta)}$$

that this prior does indeed lead to the desired posterior.

Then, under this improper prior, the Bayes estimator is the estimator for each value of x which minimizes the posterior risk. But we've already shown that the posterior risk is equivalent to the usual risk $R(\theta, aX)$, so minimizing with respect to a will give us the same estimator that we saw as the UMREE in part (d). ■

> **8.8** Obtain the UMREE for the univariate location problem $X \sim p_0(x - \theta), \theta \in \mathbb{R}$, for the group $\mathcal{G} = \{f : x \mapsto x + c, c \in \mathbb{R}\}$ and absolute loss $L(\theta, d) = |\theta - d|$.

From the lecture notes, we already have $\bar{g}_c\theta = \theta + c$ and $\tilde{g}_c d = d + c$ for this problem. Then an equivariant estimator should have $\delta(g_c(x)) = \tilde{g}_c(\delta(x))$. That is, $\delta(x + c) =$

$\delta(x)+c$. With our usual tricks (taking $c=-x$), we find that all equivariant estimators must be of the form $\delta_a(x) = x + a$ for some $a \in \mathbb{R}$.

All that remains is to find the value of a which minimizes risk. Noting that $\bar{\mathcal{G}}$ acts transitively on $\Theta = \mathbb{R}$, it's enough to minimize $R(0, \delta_a)$ with respect to a. What is this risk?

$$\begin{aligned}
R(0, \delta_a(X)) &= \mathrm{E}[L(0, \delta_a(X)) | \theta = 0] \\
&= \mathrm{E}[|X + a| | \theta = 0] \\
&= \int |x + a| p_0(x)\, dx \\
&= \int_{-\infty}^{-a} -(x+a) p_0(x)\, dx + \int_{-a}^{\infty} (x+a) p_0(x)\, dx.
\end{aligned}$$

Applying Leibniz's rule and differentiating with respect to a gives

$$\frac{\partial}{\partial a} R(0, \delta_a(X)) = \int_{-a}^{\infty} p_0(x)\, dx - \int_{-\infty}^{-a} p_0(x)\, dx,$$

which is 0 when $-a = \text{median}(p_0)$. (For thoroughness, you could also check that we're finding a minimum and not a maximum here.) Thus, the UMREE is $X - \text{median}(p_0)$. ■