STATS 300A Theory of Statistics
Stanford University, Fall 2015

Problem Set 4

Due: Thursday, October 22, 2015

Instructions:

- You may appeal to any result proved in class or proved in the course textbooks.
- Any request to "find" requires proof that all requested properties are satisfied.

Problem 1 (Minimum Risk Scale Equivariance)**.** Suppose that $X = (X_1, \ldots, X_n)$ comes from a scale family with density

$$f_\tau(x) = \frac{1}{\tau^n} f\left(\frac{x_1}{\tau}, \ldots, \frac{x_n}{\tau}\right)$$

for known f and unknown scale parameter $\tau > 0$ and that

$$Z = \left(\frac{X_1}{X_n}, \ldots, \frac{X_{n-1}}{X_n}, \frac{X_n}{|X_n|}\right).$$

(a) Show that the minimum risk scale equivariant estimator of τ under the loss

$$\gamma\left(\frac{d}{\tau^r}\right) = \frac{(d-\tau^r)^2}{\tau^{2r}}$$

is given by

$$\delta^*(X) = \frac{\delta_0(X)\mathbb{E}_1\left[\delta_0(X) \mid Z\right]}{\mathbb{E}_1\left[\delta_0^2(X) \mid Z\right]}.$$

(b) Show that a minimum risk scale equivariant estimator of τ under the loss

$$\gamma\left(\frac{d}{\tau^r}\right) = \frac{|d-\tau^r|}{\tau^r}$$

is given by

$$\delta^*(X) = \frac{\delta_0(X)}{w^*(Z)}.$$

with $w^*(Z)$ any *scale median* of $\delta_0(X)$ under the conditional distribution of X given Z with $\tau = 1$. That is, $w^*(z)$ satisfies

$$\mathbb{E}_1\left[\delta_0(X)\mathbb{I}\left(\delta_0(X) \geq w^*(Z)\right) \mid Z\right] = \mathbb{E}_1\left[\delta_0(X)\mathbb{I}\left(\delta_0(X) \leq w^*(Z)\right) \mid Z\right].$$

Hint: You might find it useful to prove the claim in Exercise 3.7a in Chapter 3 of TPE.

In part (b) you may assume that X has a continuous probability density function $f(\cdot;\theta)$ with respect to Lebesgue measure.

Problem 2 (Bayes Estimation)**.** Suppose that Θ follows a log-normal distribution with known hyperparameters $\mu_0 \in \mathbb{R}$ and $\sigma_0^2 > 0$ and that, given $\Theta = \theta$, $(X_1, \ldots, X_n)$ is an i.i.d. sample from $\text{Unif}(0, \theta)$.

(a) What is the posterior distribution of $\log(\Theta)$?

(b) Let δ_τ represent the Bayes estimator of θ under the loss

$$L(\theta, d) = \begin{cases} 0 & \text{if } \frac{1}{\tau} \leq \frac{\theta}{d} \leq \tau \\ 1 & \text{otherwise} \end{cases}$$

for fixed $\tau > 1$. Find a simple, closed-form expression for the limit of δ_τ as $\tau \to 1$.

Note: Part (b) concerns Bayes estimators of θ, not of $\log(\theta)$, but part (a) is still relevant.

Problem 3 (Conjugacy)**.** A family $\Pi = \{\pi_\kappa : \kappa \in K\}$ of prior probability densities indexed by the hyperparameter κ is said to be *conjugate* for a model $\mathcal{P} = \{f(\cdot|\theta) : \theta \in \Omega\}$ of likelihoods if, for each prior $\pi_\kappa \in \Pi$, the posterior $\pi_\kappa(\theta|x) \propto f(x|\theta)\pi_\kappa(\theta)$ is also in Π. That is, $\pi_\kappa(\theta|x) = \pi_{\kappa'}(\theta)$ for some index $\kappa' \in K$ depending on κ and x. Posterior analysis is greatly simplified when the mapping $(\kappa, x) \mapsto \kappa'$ has a known closed form. For each model below, find a conjugate prior family under which the posterior hyperparameter κ' is a simple, closed-form function of the data x and the prior hyperparameter κ:

(a) $(X_1, \ldots, X_n) \overset{\text{iid}}{\sim} \text{Gamma}(\alpha, \beta)$ with unknown shape α and rate β.

(b) $(X_1, \ldots, X_n) \overset{\text{iid}}{\sim} \text{Beta}(\alpha, \beta)$ with unknown shape parameters (α, β).

(c) The family defined by the linear observation model $Y_i \overset{\text{ind}}{\sim} \mathcal{N}(\beta_1 + \beta_2 x_i, \sigma^2)$ for $i \in \{1, \ldots, n\}$ where $x_i \in \mathbb{R}$ and $\sigma^2 > 0$ are known and $\beta_1, \beta_2 \in \mathbb{R}$ are unknown.

Problem 4 (Posterior Quantiles)**.** Consider a Bayesian inference setting in which the posterior mean $\mathbb{E}[\Theta \mid X = x]$ is finite for each x. Show that under the loss function

$$L(\theta, a) = \begin{cases} k_1|\theta - a| & \text{if } a \leq \theta \\ k_2|\theta - a| & \text{otherwise} \end{cases}$$

with $k_1, k_2 > 0$ constant and for p an appropriate function of k_1 and k_2, every p-th quantile of the posterior distribution is a Bayes estimator.

Problem 5 (Bayesian Prediction)**.**

(a) Let $\mathbf{Z} \in \mathbb{R}^p$ be a random vector and $Y \in \mathbb{R}$ be a random variable. Our goal is to learn a prediction rule $f : \mathbb{R}^p \to \mathbb{R}$ to best predict the value of Y given $\mathbf{Z}$. We will measure the goodness of our predictions using the risk function

$$\mathbb{E}\left[(Y - f(\mathbf{Z}))^2\right]. \tag{1}$$

Assuming that the joint distribution of $(\mathbf{Z}, Y)$ is known, find the optimal prediction rule f.

(b) Now, let $\Theta \in \Omega$ be a random variable with distribution Λ, and, given $\Theta = \theta$, let $X_1, \ldots, X_{n+1}$ be drawn i.i.d. from a density p_θ. Let $\mu(\theta)$ represent the mean under p_θ. First, find the Bayes estimate of $\mu(\theta)$ under squared error loss given only the first n observations $X_1, \ldots, X_n$. Next, find the function f of $(X_1, \ldots, X_n)$ that best predicts X_{n+1} under the average risk

$$\mathbb{E}\left[\mathbb{E}_\Theta\left[(X_{n+1} - f(X_1, \ldots, X_n))^2\right]\right].$$

Note the close relationship between optimal prediction and optimal estimation.