Problem Set 1

1. Consider a continuous random variable Y^* with density $f_{Y^*}(\cdot)$, and a censored version of Y^*: $Y = \min\{Y^*, C\}$ for a known constant C. Assume that Y^* has a density function that is everywhere positive.

   (a) Find the c that solves $\min_{c\in R} E(\rho_\tau(Y-c))$, where $\rho_\tau(u) = (\tau - 1(u<0))u$ for a given number $\tau \in (0,1)$. You should consider both the case where $\Pr(Y < C) > \tau$ and the case where $\Pr(Y < C) \leq \tau$. Is the minimizer unique?

   (b) Now suppose that $Y^* = X'\beta_0 + \varepsilon$ for a finite dimensional vector of independent variables X, and a parameter value $\beta_0 \in B$ for a parameter space $B \subseteq R^{d_x}$. Suppose that the $100\tau\%$ conditional quantile of ε given X is zero: $\text{quantile}_\tau(\varepsilon|X) = 0$. Assume that the conditional density $f_{\varepsilon|X}(\cdot|x)$ is everywhere positive for all x on the support of X. Consider the criterion function:

   $$\hat{Q}_n(\beta) = n^{-1}\sum_{i=1}^{n}\rho_\tau(Y_i - \min\{X_i'\beta, C\}),$$

   where $\{(Y_i, X_i')'\}_{i=1}^n$ is an i.i.d. sample of $(Y, X')'$. Assume that the parameter space B is a compact subset of R^{d_x}. Also assume that $E|\varepsilon| < \infty$ and $E\|X\| < \infty$. Show the uniform convergence of $\hat{Q}_n(\beta)$ as $n \to \infty$, and find the limiting criterion function $Q(\beta)$.

   (c) Show that β_0 minimizes $Q(\beta)$. (Hint: use part (a)).

   (d) Suppose that the matrix $E[XX'1(X'\beta_0 < C)]$ has full rank. Show that β_0 is the unique minimizer of $Q(\beta)$.

2. (Monte Carlo Simulation) Consider the following random coefficient binary choice model:

   $$Y_i = 1\{\alpha + \beta_i X_i + \varepsilon_i > 0\},$$

   where we assume that $\varepsilon \sim N(0,1)$, $\beta_i \sim N(\mu, \sigma^2)$, β_i, X_i and ε_i are mutually independent. In this model, $\theta = (\alpha, \mu, \sigma^2)'$ is the vector of unknown parameters.

   With an i.i.d. data set $\{(Y_i, X_i)\}_{i=1}^n$, this model could be estimated by maximum likelihood, and the log-likelihood function is

$$L_n(\theta) = \sum_{i=1}^{n} \log(p(X_i, \theta)^{Y_i}(1 - p(X_i, \theta))^{1-Y_i}), \tag{1}$$

where $p(X_i, \theta) = \int_{-\infty}^{\infty} \sigma^{-1}\Phi(\alpha + bX_i)\phi((b-\mu)/\sigma)db$, where $\Phi(\cdot)$ is the standard normal CDF and $\phi(\cdot)$ is the standard normal PDF. The complication is that $p(X_i, \theta)$ does not have an analytical form and have to be obtained by numerical integration for every i and every θ value.

One numerical integration method is by Monte Carlo simulation, that is, to use independent draws $\{Z_s\}_{s=1}^S$ from $N(0, 1)$ for a large number S to form an approximation for $p(X_i, \theta)$:

$$p^S(X_i, \theta) = \frac{1}{S}\sum_{s=1}^{S} \Phi(\alpha + (\sigma Z_s + \mu)X_i).$$

This leads to the simulated log-likelihood function:

$$L_n^S(\theta) = \sum_{i=1}^{n} \log(p^S(X_i, \theta)^{Y_i}(1 - p^S(X_i, \theta))^{1-Y_i}).$$

Then we can define the simulated maximum likelihood estimator

$$\hat{\theta}^S = \arg\max L_n^S(\theta).$$

Note that in the simulated log-likelihood above, $\{Z_s\}_{s=1}^S$ is hold the same for every i. This is called same-draw simulated log-likelihood. An alternative is to use independent draws for different i's. That is, we draw n independent samples $\{Z_{i,s}\}_{s=1}^S$: $i = 1, \ldots, n$ from $N(0, 1)$, and let

$$p^{S2}(X_i, \theta) = \frac{1}{S}\sum_{s=1}^{S} \Phi(\alpha + (\sigma Z_{i,s} + \mu)X_i).$$

Let $L_n^{S2}(\theta)$ and $\hat{\theta}^{S2}$ be defined analogously.

With these simulated MLEs in mind, conduct the following Monte Carlo experiments:

(a) Let the true value of θ be $(0, 1, 1)$ and let $X_i \sim N(0, 1)$. Let $n = 500$ and

$S = 50$. Generate M independent samples $\{(Y_i^m, X_i^m)\}_{i=1}^n : m = 1, \ldots, M$ for $M = 1000$. For each m, use the sample $\{(Y_i^m, X_i^m)\}_{i=1}^n$ to calculate the two simulated maximum likelihood estiamtors, and denote them by $\hat{\theta}_m^S$ and $\hat{\theta}_m^{S2}$, respectively.

Caculate the bias, the standard deviation, and the root-mean-squared error (=square root of (bias^2+standard deviation^2)) of each element of $\hat{\theta}^S$ and of $\hat{\theta}^{S2}$ using the simulated samples $\{\hat{\theta}_m^S\}_{m=1}^M$ and $\{\hat{\theta}_m^{S2}\}_{m=1}^M$, and report them in a well-labeled table.

(b) Increase S to 100 and re-do part (a).

(c) Decease S to 25 and re-do part (a).

(d) Note that even though $\{Z_s\}_{s=1}^S$ (or $\{Z_{i,s}\}_{s=1}^S$) is drawn from $N(0,1)$, the empirical distribution of the draws typically does not have mean 0 and variance 1 because S is finite (and relatively small in parts (a)-(c)). It is reasonable to believe that such finite sample discrepancy may lead to errors in the estimators $\hat{\theta}^S$ and $\hat{\theta}^{S2}$. Now define the normalized version of $\{Z_s\}_{s=1}^S$

$$\tilde{Z}_s = (Z_s - \bar{Z}_S)/\hat{\sigma}_Z,$$

where $\bar{Z}_S = \frac{1}{S}\sum_{s=1}^S Z_s$ and $(\hat{\sigma}_Z)^2 = \frac{1}{S}\sum_{s=1}^S (Z_s - \bar{Z}_S)^2$. Define $\{\tilde{Z}_{i,s}\}_{s=1}^S$ analogously based on $\{Z_{i,s}\}_{s=1}^S$ for each $i = 1, \ldots, n$.

Now re-do parts (a)-(c) using $\{\tilde{Z}_s\}_{s=1}^S$ and $\{\tilde{Z}_{i,s}\}_{s=1}^S : i = 1, \ldots, n$ in place of $\{Z_s\}_{s=1}^S$ and $\{Z_{i,s}\}_{s=1}^S : i = 1, \ldots, n$, respectively.

(e) Discuss your findings.