Midterm Project

The purpose of the midterm project is to use the theoretical techniques learned in class to derive some large sample and finite sample properties of estimators, and then to use computational methods to check the accuracy of the large sample properties and also to compare competitive estimators. This project will be considered as your midterm examination. The project must be handed in by 2:30 p.m., Monday, November 6.

You should each do your own work for the project. Be sure to cite any references that you use.

1. Consider the logistic location family where \(X_1, \ldots, X_n\) are i.i.d. \(F(x - \theta)\), where \(-\infty < \theta < \infty\), and \(F(x) = \left[1 + \exp(-x)\right]^{-1}\). We wish to compare the following estimators of \(\theta\):
   (i) The MLE of \(\theta\).
   (ii) The one-step Newton-Raphson estimator of \(\theta\).
   (iii) The one-step Fisher-scoring estimator of \(\theta\).
   (iv) The sample median \(Y_5\)
   (v) The sample mean.
   (vi) An L-estimator of the form \(W = aY_p + (1 - 2a)Y_5 + aY_{1-p}\), where \(a = 0.3\), \(p = 0.25\), and \(Y_p\) is the sample \(p\)th quantile.

2. Obtain the limiting distributions of these estimators and obtain their A.R.E.s. You may use the following theorem concerning the asymptotic distribution of sample quantiles:

   **Theorem:** Suppose that \(X_1, \ldots, X_n\) are i.i.d. from a distribution with density \(g(x)\) and c.d.f. \(G(x)\). Let \(Y_{p_1}, \ldots, Y_{p_k}\) be the sample \(p_1, \ldots, p_k\) quantiles, respectively, where \(p_1 < p_2 < \ldots < p_k\). Let \(x_{p_j}\) satisfy \(G(x_{p_j}) = p_j\). Suppose that \(g(x_{p_j}) > 0\) for each \(j\). Then
   \[
   \sqrt{n} \begin{pmatrix} Y_{p_1} - x_{p_1} \\ \vdots \\ Y_{p_k} - x_{p_k} \end{pmatrix} \xrightarrow{d} N_k(0, \Psi)
   \]
   where \(\Psi\) has \((i, j)\) entry
   \[
   \psi_{i,j} = \frac{p_{\min\{i,j\}} - p_ip_j}{g(x_{p_i})g(x_{p_j})}.
   \]
3. Carry out a simulation study to compare the estimators. You need to design an experiment with various sample sizes. For each sample size, you should generate a large number of samples and compute all the estimates for each sample.

4. Then summarize your results by reporting estimated bias, variance, and mean squared error for each estimator for each sample size.

5. Make some conclusions about your simulation study. Some questions that you should consider include:
   • How accurate are the results reported in step 4?
   • Do the simulation results agree with the theoretical results?
   • For what sample sizes are the asymptotic results appropriate?
   • Which estimator is best? If none is best, under what conditions are different estimators better?
   You may want to use statistical methods in helping you to arrive at your conclusions.

6. A TeXed or typed report should be prepared. The derivations for the theoretical results can be attached as an appendix. A possible outline for your report follows:
   a. Introduction
   b. Theoretical results
   c. Description of the simulation
   d. Results of the simulation presented in tables or figures
   e. Summary and conclusions.