Statistics 221 – Assignment 3

Due: Friday, April 9, 2004

1. Lange 20.12

   Devise an acceptance-rejection method for generating beta deviates based on the inequality $x^{\alpha-1}(1-x)^{\beta-1} \leq x^{\alpha-1} + (1-x)^{\beta-1}$.

2. Let

$$\theta = \int_0^1 \frac{e^x - 1}{e - 1} dx$$

   For each of the following Monte Carlo methods specified below, find an estimator of θ and compare its efficiency with that of the crude Monte Carlo estimator.

   (a) Crude Monte Carlo (e.g. sample $X \sim U(0,1)$)

   (b) Importance sampling

   (c) Control variates

   (d) Antithetic variates

3. Lange 21.8

   The method of control variates can be used to estimate the moments of the sample moments of the sample median $X_{(n)}$ from a random sample of size $2n-1$ from a symmetric distribution. Because we expect the difference of $X_{(n)} - \bar{X}$ between the sample median and the sample mean to be small, the moments of $\bar{X}$ serve as a first approximation to the moments of $X_{(n)}$. Put this insight into practice by writing a Monte Carlo program to compute $Var(X_{(n)})$ for a sample from the standard normal distribution.

   Plus: Using your code, find the variance for $n = 10$ and $n = 20$ using $m = 500$ imputations in each case.

4. Consider calculating the p-value for a test statistic Z which has a $N(0,1)$ distribution asymptotically. Assume that we are interested in a one-sided hypothesis so that the p-value of interest is $P[Z \geq c]$, where c is the observed value of the test statistic. One approach to calculating the p-value is to generate $z_1, \ldots, z_n$ under H_0 and to estimate the p-value with

$$\hat{p} = \frac{1}{n} \sum_{i=1}^{n} I(z_i \geq c),$$

   where $I(z_i \geq c)$ is the indicator function of whether $z_i \geq c$. An alternative to this estimator is to use importance sampling. Instead of sampling realizations of the test statistic under H_0, sample under a member of H_A, as this can give a more precise estimate, particularly when c is large.

Let us examine which points in the alternative could be used for the importance sampling estimate. Assume that $Z \sim N(0,1)$ under H_0 and you want to sample from $N(b,1)$. This gives an importance sampling estimate of the p-value, $\tilde{p}_b$.

Find a condition (and bounds if possible) on the choice of b which minimizes $Var(\tilde{p}_b)$.

In addition, what is the efficiency of the optimum choice for b relative to $b = 0$ and calculate this for $c = 2$, 3, and 4.

Hint: Let $S(x) = 1 - \Phi(x) = P[Z \geq x]$ be the survivor function for the standard normal distribution. Then for $x > 0$, one possible set of bounds on the Mills' ratio is

$$\frac{1}{x}\left(1 - \frac{1}{x^2}\right) \leq \frac{S(x)}{\phi(x)} \leq \frac{1}{x}.$$