CSI 771 Problem Set 2

2.2a We will generate a standard normal deviate, that is, a realization of a random variable X with distribution $\mathrm{N}(0,1)$. If the desired distribution is $\mathrm{N}(\mu,\sigma^2)$, we can get it by the transformation $\sigma X + \mu$. In the usual notation, we have

$$p_X(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2), \quad -\infty < x < \infty,$$

and

$$g_Y(y) = \frac{1}{2\theta} \exp(-|y|/\theta) \quad -\infty < y < \infty.$$

We first note that we can generate from the nonnegative half (so that the densities are multiplied by 2), and then randomly change signs with probability 1/2.

We seek the minimum value of c so that $0 \le c g_Y(x) - p_X(x)$ for $0 \le x$; that is,

$$c\frac{1}{\theta}\exp(-x/\theta) \ge \sqrt{\frac{2}{\pi}}\exp(-x^2/2).$$

So

$$\begin{aligned} c &\ge \theta\sqrt{\frac{2}{\pi}}\exp(-x^2/2 + x/\theta) \\ &= \theta\sqrt{\frac{2}{\pi}}\exp\left(-\left(\frac{x}{\sqrt{2}} - \frac{1}{\sqrt{2}\theta}\right)^2 + \frac{1}{2\theta^2}\right) \\ &= \theta\sqrt{\frac{2}{\pi}}\exp\left(\frac{1}{2\theta^2}\right)\exp\left(-\left(\frac{x}{\sqrt{2}} - \frac{1}{\sqrt{2}\theta}\right)^2\right). \end{aligned}$$

Since the maximum of the last factor is 1 (when $x = 1/\theta$), we choose $c = \theta\sqrt{2/\pi}\exp\left(1/(2\theta^2)\right)$.

Now, for a given θ and the value of c above, the steps are

1. Generate y from an exponential with parameter θ. (Take $y = -\theta\log(u_1)$, where u_1 is from $\mathrm{U}(0,1)$.)
2. Generate u_2 from $\mathrm{U}(0,1)$. (It is best to use a different random number stream for u_2.)
3. If $u_2 > p_X(y)/(c g_Y(y))$, where c, $p_X(\cdot)$, and $g_Y(\cdot)$ are as above, reject y and go to step 1.
4. Generate u_3 from $\mathrm{U}(0,1)$. (It is best to use a different random number stream for u_3.)
5. If $u_3 > 0.5$, set $y = -y$.
6. Deliver y as the normal deviate.

The value of c depends on θ. The best value of θ is either 1 (because there is no multiplication in step 1) or the value that minimizes the probability of rejection. This is the integral over $(0,\infty)$ of $c g_Y(y) - p_X(y)$. Choosing $\theta = 1$ yields a relatively simple procedure, and it's probably what I would do.

2.2b A majorizing function must be at least as large as the given density over the full range. In this case, using the usual notation, we have

$$p_X(x) = \frac{1}{2\theta}\exp(-|y-\tau|/\theta) \quad -\infty < y\infty,$$

and

$$g_Y(y) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-(x-\mu)^2/(2\sigma^2)), \quad -\infty < x\infty.$$

The problem is that there is no constant c such that $p_X(x) \le c g_Y(x)$ as $x \to \infty$.

2.2c The bivariate product double exponential density centered at $(0,0)$ is just

$$g_{Y_1Y_2}(y_1,y_2) = \frac{1}{4\theta^2}\exp(-(|y_1|+|y_2|)/\theta) \quad -\infty < y_1, y_2 < \infty.$$

That is the distribution with the product density is just the distribution of two independently identically distributed double exponential. We can get that by generating y_1 and y_2 both as in step 1 above. If the target density had different variances, we might use different values of θ in the bivariate double exponential. In any problem like this, there is no end to the possible variations for improving the efficient. Remember that the basic problem is to keep the generation of the candidate deviates simple while trying to minimize the probability of rejecting the deviates.

The bivariate normal density with means μ_1 and μ_2, variances σ_1^2 and σ_2^2, and correlation ρ is

$$p_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp(-((x_1-\mu_1)^2/\sigma_1 - 2\rho(x_1-\mu_1)/\sigma_1(x_2-\mu_2)/\sigma_2 + (x_2-\mu_2)^2/\sigma_2)/(2(1-\rho^2))).$$

In this exercise $\mu_1 = \mu_2 = 0$ and $\sigma_1^2 = \sigma_2^2 = 1$.

This is similar to the problem above; however, because of the correlation, it is more difficult to split the distributions up into positive and negative parts. We go through similar steps over the full range, including completing the square twice, to determine a value of c. The smaller we make c so that $cg_{Y_1Y_2}(y_1,y_2)$ majorizes $p_{X_1,X_2}(x_1,x_2)$, the better, but a c larger than the minimum still works.

In step 1 above, we generate two exponential deviates (y_1,y_2), and then immediately do the equivalent of step 5 on each of them. Then we do step 2 and 3, and if they are accepted in step 3, the pair is delivered as the bivariate normal.

An alternative approach is to generate x_1 as a $\mathrm{N}(0,1)$ and then to generate x_2 conditionally as a $\mathrm{N}(\rho x_1,\ 1-\rho^2)$. In R, for example to get n pairs, after initializing n, rho, and rhosq, we could use

```
x1<-rnorm(n)
x2<-(1-rhosq)*rnorm(n)+rho*x1
```

We can compute the sample statistics by

```
mean(x1)
mean(x2)
var(x1)
var(x2)
cor(x1,x2)
```

There is also a multivariate normal random number generator in R, but I wanted you to work it without using that one.

2.4 Initial n, the number of deviates to generate, and x0, the starting value. Then,

```
x<-rep(NaN,n)
x[1]<-x0
yi<-1
for (i in 1:n){
    u1 <- runif(1)
    u2 <- runif(1)
    y <- rnorm(1)
    yip1 <- ifelse(u1>0.5, -log(u2)+y, log(u2)+y)
    lr2 <- yi^2-yip1^2
    if(lr2>0)
       yi<-yip1
    else{
       u<-runif(1)
```

```
            if(lr2>2*log(u))
               yi<-yip1
            else
               yi<-x[i-1]
         }
         x[i]<-yi
      }
```

There are various goodness-of-fit tests, such as chi-squared tests and K-S tests, that could be used. In addition to formal tests, it is instructive to look at q-q plots. A q-q plot with a normal reference distribution is provided by the R function qqnorm( ).

2.7a Notice that the variance is not specified; it is a nuisance parameter so far as the test of the mean is concerned. The "standard test statistic" has a distribution that is independent of the variance, so it is the one to use.

The Monte Carlo test procedure involves generating m samples each of size n from the hypothesized distribution, that is, from a normal with mean $\mu = 0$, as the hypothesis states. Let x_{ij} be the i^{th} observation in the j^{th} sample, and let $\bar{x_j}$ be the mean of the j^{th} sample. For each of these m samples we compute the test statistic,

$$t_j^* = \frac{\bar{x_j}}{\sqrt{\sum_i (x_{ij} - \bar{x_j})^2/(n(n-1))}}$$

Now rank all m t_j^*'s.

Compute the same quantity for the given sample, t_o, and its empirical quantile within the t_j^*'s. Because this is a two-sided test, we count the number of t_j^*'s greater than t_o in absolute value. If r is this number, the empirical p-value is $(r+1)/(m+1)$.

2.7b This test can be studied analytically; it is equivalent to a t test, except for the larger variance due to the randomness from the Monte Carlo sampling.

To study the power empirically, we need to set a significance level so that for each Monte Carlo sample, we can say whether or not we reject the null hypothesis.

One way to study the power would be to generate a single set of m samples and generate reference samples with various values of μ, say $\mu = 0.01, 0.05, 0.10, \ldots$ and for each of the reference samples conduct the Monte Carlo test as above. At each value of μ we determine whether or not we reject. Then we repeat this whole procedure k times. Out of the k trials, the proportion of rejections at each value of μ is an estimate of the power of the test at that value.