MATH 416/516 ASSIGNMENT 4 SOLUTIONS

- Problem 8.3: Matlab

```
function [ mu ss ] = muss( X )
% compute mean and variance for X
  n = length(X); mu = X(1); ss = 0;
  for j = 2 : n, mo = mu; mu = mu + ( X(j) - mu )/j;
    ss = ( 1 - 1/(j-1) )*ss + j*( mu - mo )^2;
  end
% end muss
```

Test compares with Matlab *mean, var*:

```
x = rand(1,20); [ mu ss] = muss(x);
disp([mu ss]), disp([mean(x) var(x)])
      0.50467     0.049129
      0.50467     0.049129
```

- Problem 8.4: Matlab

```
mu = randn; j = 1; ss = 0;
while sqrt(ss/j) > .1 | j<100, j = j + 1; mo = mu;
  mu = mu+ (randn-mu)/j;
  ss=(1-1/(j-1))*ss + j*( mu - mo )^2;
end
disp([ j mu ss])
  100      0.06048      0.95353
```

a) standard normals have mean 0 and variance 1; for $\sqrt{Var(X)/n} < .1$, you expect $n \approx= 100$; b-e) the program stopped at $n = 100$, with sample mean is $\approx .06$, sample Var $\approx .95$, consistent with standard normal samples.

- Problem 8.6: Matlab

```
mu = exp(rand^2); j = 1; ss = 0;
while sqrt(ss/j) > .01 | j<100, j = j + 1; mo = mu;
  mu = mu+ (exp(rand^2)-mu)/j;
  ss = (1-1/(j-1))*ss + j*( mu - mo )^2;
end
disp([ j mu ss])
      2306       1.4699      0.23055
```

- Problem 8.7: $Var(X) = 9.0667$, and you need $\sqrt{Var(X)/N} < .1$, so $N \approx 9/(.1)^2 = 900$.

- Problem 8.8: Matlab

```
for i = 1:1000, n = 0; s = 0;
  while s < 1, s = s + rand; n = n + 1; end; N(i)=n;
end
disp([mean(N) var(N) 2*std(N)/sqrt(1000)])
        2.702       0.76196      0.055207
```

95% confidence interval is $\approx [2.65, 2.76]$.

- Problem 8.9*:
  a) $P\{M > n\} = 1/n!$ because this is the probability that the first n U_is are in order, and this is one possibility in $n!$ possible permutations.
  b) $E[M] = \sum_{n=0}^{\infty} P\{M > n\} = \sum_{n=0}^{\infty} 1/n! = e$.
  c) Matlab

```
for i=1:1000, n=2; uo = rand; un = rand;
  while uo < un, uo = un; un = rand; n = n + 1; end; N(i)=n;
end
disp([mean(N) 2*std(N)/sqrt(1000)])
        2.731     0.055259
```

  d) 95% confidence interval is $\approx [2.676, 2.786]$.

- Problem 8.12: for the given data the mean is 122.85, with $S = 16.8$.
  For 99% confidence level and error ± 0.5, we want $2.58S/\sqrt{N} \approx .5$.
  So $N \approx ((2)2.58S)^2 = 7514$, and we need approximately 7500 more samples.

- Problem 8.13*:
  a) sample $n-$vectors $\mathbf{X}$ with components taken uniformly from the data; for each $\mathbf{X}$, use RV I (1 or 0) to indicate if inequality is satisfied, then use $p \approx E[I]$.
  b) Matlab results

```
X = [56 101 78 67 93 87 64 72 80 69]; mu = mean(X);
R = 100; XV = X(ceil(10*rand(10,R)));
p = mean( abs( mean(XV) - mu ) < 5 ); disp(p)
         0.74
```

- Problem 8.14: $S^2(x_1, x_2) = ((x_1 - \frac{x_1+x_2}{2})^2 + (x_2 - \frac{x_1+x_2}{2})^2)/1 = (x_1 - x_2)^2/2$.
  There four choices for (x_1, x_2), with $S^2(1,1) = S^2(3,3) = 0$, $S^2(1,3) = S^2(3,1) = 2$.
  Then $mean(S^2) = 1$, so $Var(S^2) = ((0-1)^2 + (0-1)^2 + (2-1)^2 + (2-1)^2)/4 = 1$.

- Problem 8.15: this is like the previous problem, except you cannot compute 15^{15} possibilities, so just sample the possibilities from the data vector X, compute Var for each sample and then compute the var of the var's. Matlab

```
n = 15; X = [5 4 9 6 21 17 11 20 7 10 21 15 13 16 8];
R = 100; G = var( X(ceil(n*rand(n,R))) ); disp(var(G))
       55.964
```

- Problem 9.1: an antithetic estimator for θ is $X = (e^{U^2} + e^{(1-U)^2})/2 = e^{U^2}(1 + e^{1-2U})/2$, which is the estimator proposed by the text.
  Now, e^{x^2} is monotonic so the antithetic estimator will reduce variance.
  Matlab test results confirm this:

```
N = 1000; U = rand(1,N); X = exp(U.^2).*( 1+exp(1-2*U) )/2;
disp([mean(X) var(X)]) % Antithetic results
       1.4623     0.027588
X = ( exp(U.^2) + exp(rand(1,N).^2) )/2;
disp([mean(X) var(X)]) % Raw simulation results
       1.4745      0.11884
```

  Variance for the antithetic is smaller by factor of ≈ 5.

- Problem 9.2: for antithetic variables use the RV $X = (e^{(U_1+U_2)^2} + e^{(2-U_1-U_2)^2)})/2$; this will be better than simple MC, because the integrand is monotone increasing in both variables.

- Problem 9.3: a) generate $X_i = -\ln(1-U_i)$ and use RV I (1 or 0) to indicate
  if $\sum_{i=1}^{5} iX_i > 21.6$; then average over I values to estimate θ.
  b) for antithetic simulation just replace the $1-U_i$'s by U_i's and average the two results..
  c) Matlab
  ```
  N = 1000; U = rand(5,N);
  X = -log(1-U); I = [1 2 3 4 5]*X > 21.6;
  disp( [mean(I) var(I) 2*std(I)/sqrt(N)] )
          0.167      0.13925     0.023601
  N = N/2; U = rand(5,N);
  X = -log(1-U); I = [1 2 3 4 5]*X > 21.6;
  Y = -log(U);   J = [1 2 3 4 5]*Y > 21.6; J = (I+J)/2;
  disp( [mean(J) var(J) 2*std(J)/sqrt(N)] )
          0.184     0.058261     0.021589
  ```
  Antithetic is slightly more efficient.

- Problem 9.5: a) antithetic variables simulator uses $X = (Z^3e^Z - Z^3e^{-Z})/2$ with $Z \sim N(0,1)$.
  b) Matlab test results for 95% confidence .05 error :
  ```
  f = @(z)( z^3*exp(z) - z^3*exp(-z) )/2;
  N = 1; mu = f(randn); ss = 0;
  while N < 10 | 2*sqrt(ss/N) > .05, N = N + 1;
    mo = mu; mu = mu + ( f(randn) - mu )/N;
    ss = ( 1 - 1/(N-1) )*ss + N*( mu - mo )^2;
  end
  disp( [N mu 2*sqrt(ss/N)] )
     3.3523e+06       6.6124          0.05
  ```
  It looks like more than 3 million samples are needed.
  Note: exact value for $E[Z^3e^Z] = 4\sqrt{e} \approx 6.59489$.

- Problem 9.7: minimizing $f(c) = Var(X) + c^2Var(Y) + 2cCov(X,Y)$, set
  $f'(c) = 2cVar(Y) + 2Cov(X,Y) = 0$, so optimal $c^* = -Cov(X,Y)/Var(Y)$.
  Checking $f''(c^*) = 2Var(Y) > 0$; $f(c^*)$ is a minimum.

- Problem 9.8: $f(c^*) = Var(X) + \frac{Cov(X,Y)^2}{Var(Y)} - 2\frac{Cov(X,Y)^2}{Var(Y)} = \frac{Var(X)Var(Y)-Cov(X,Y)^2}{Var(Y)}$.

- Problem 9.9*: Matlab

```
for i = 1 : 1000; un = rand; n = 1; sm = un; I = [1 1 1 1];
  while sum(I) > 0
    n = n + 1; uo = un; un = rand; sm = sm + un;
    if uo < un  & I(1), N = n; I(1) = 0; end
    if uo > un  & I(2), M = n; I(2) = 0; end
    if sm > 1   & I(3), S = n; I(3) = 0; end
    if sm < n-1 & I(4), T = n; I(4) = 0; end
  end, A(i) = ( N + M + S + T )/4;
end, disp([ mean(A) var(A) 2*std(A)/sqrt(1000) ] )
      2.7228      0.066386      0.016296
```

  Variance $\approx .066$ is much smaller than .76 from problem 8.8 (raw simulation).

- Problem 9.11*: let $f(\alpha) = Var(\alpha X + (1-\alpha)W)$, so
  $f(\alpha) = \alpha^2 Var(X) + (1-\alpha)^2 Var(W) + 2\alpha(1-\alpha)Cov(X,W)$.
  $f'(\alpha) = 2\alpha Var(X) - 2(1-\alpha)Var(W) + 2(1-2\alpha)Cov(X,W) = 0$, when
  $\alpha(Var(X) + Var(W) - 2Cov(X,W)) = Var(W) - Cov(X,W)$; solving for α gives α^* in (8.3).
  Note $f''(\alpha^*) = 2(Var(X)+Var(W)-2Cov(X,W)) = 2Var(X-W) >= 0$, so $f(\alpha*)$ is minimum.
  Substitution (plus algebra) gives $f(\alpha^*) = \frac{Var(X)Var(W)-Cov(X,W)^2}{Var(X)+Var(W)-2Cov(X,W)}$.

- Problem 9.12: a) a simple choice for a CV is $Y = e^U$, with $\mu_Y = e-1$.
  b-c) Matlab

```
N = 100; U = rand(1,N); X = exp(U.^2); Y = exp(U);
Xb = mean(X); Yb = mean(Y); muY = exp(1)-1;
cs = -sum( (X-Xb).*(Y-Yb) )/sum( (Y-Yb).^2 );
Z = X + cs*( Y - muY );      % Control Variate
disp([mean(Z) var(Z) 2*std(Z)/sqrt(N), cs])
       1.4667     0.011072     0.021045     -0.93296
W = ( X + exp((1-U).^2) )/2; % Antithetic Variate
disp([mean(W) var(W) 2*std(W)/sqrt(N)])
       1.4628     0.025681     0.032051
```

  d) the control variate method is better.