Math 350 - Homework 10 - Solutions

1. *In certain situations a random variable X, whose mean is known, is simulated so as to obtain an estimate of $P\{X \leq a\}$ for a given constant a. The raw simulation estimator from a single run is I, where*

$$I = \begin{cases} 1 & \textit{if } X \leq a \\ 0 & \textit{if } X > a. \end{cases}$$

Because I and X are clearly negatively correlated, a natural attempt to reduce the variance is to use X as a control—and so use an estimator of the form $I + c(X - E[X])$.

(a) *Determine the percentage of variance reduction over the raw estimator I that is possible (by using the best c) if X were uniform on $(0, 1)$.*

(b) *Repeat (a) if X were exponential with mean 1.*

(c) *Explain why we knew that I and X were negatively correlated.*

The following remarks are useful. I will assume that X takes values in $[0, \infty)$. (This is the case for the random variables of part (a) and (b).) $E[I] = P(X \leq a) = F(a)$, where $F(x) = P(X \leq x)$ is the cumulative distribution function of X. Also recall that, for any function $H(x)$ such that the expected value of $H(X)$ makes sense to take (i.e., $H(x)$ is integrable with respect to the probability distribution of X), we have $E[H(X)] = \int xF'(x)\, dx$. In particular, if $I(x)$ is the function which is 1 if $x \leq a$ and 0 otherwise, then

$$E[XI] = \int_0^a xF'(x)\, dx.$$

Therefore, the covariance of X and I can be written as

$$\mathrm{Cov}(X, I) = E[XI] - E[X]E[I] = \int_0^a xF'(x)\, dx - F(a) \int_0^L xF'(x)\, dx$$

where L (possibly ∞) is the upper endpoint of the interval of values of X. We also have:

$$\mathrm{Var}(X) = E[X^2] - E[X]^2 = \int x^2 F'(x)\, dx - \left(\int xF'(x)\, dx \right)^2$$

and

$$\mathrm{Var}(I) = E[I^2] - E[I]^2 = E[I](1 - E[I]) = F(a)(1 - F(a)).$$

Finally, recall that the fraction of variance reduction obtained by using I_{c^*} is

$$\frac{\mathrm{Var}(I_{c^*})}{\mathrm{Var(I)}} = 1 - \mathrm{Corr}^2(X, I)$$

where $\mathrm{Corr}^2(X, I) = \mathrm{Var}^2(X, I)/(\mathrm{Var}(X)\mathrm{Var}(I))$.

Part (a). If X is uniform over $(0,1)$, then $F(x) = x$, $F'(x) = 1$ and

$$\mathrm{Cov}(X, I) = \int_0^a x\,dx - a\int_0^1 x\,dx = \frac{a^2 - a}{2}.$$

$$\mathrm{Var}(X) = \int_0^1 x^2\,dx - (x\,dx)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

$$\mathrm{Var}(I) = a(1-a).$$

Therefore, $\mathrm{Corr}^2(X, I) = 3a(1-a)$ and

$$\frac{\mathrm{Var}(I_{c^*})}{\mathrm{Var}(I)} = 1 - 3a(1-a).$$

Part (b). If X is exponential with mean $\lambda = 1$, then $F'(x) = e^{-x}$, $F(x) = 1 - e^{-x}$. In this case

$$\mathrm{Cov}(X, I) = \int_0^a xe^{-x}\,dx - (1 - e^{-a})\int_0^\infty xe^{-x}\,dx = -ae^{-a} - e^{-a} + 1 - 1 + e^{-a} = -ae^{-a}.$$

$$\mathrm{Var}(X) = \frac{1}{\lambda^2} = 1.$$

$$\mathrm{Var}(I) = F(a)(1 - F(a)) = e^{-a}(1 - e^{-a}).$$

Thus

$$\frac{\mathrm{Var}(I_{c^*})}{\mathrm{Var}(I)} = 1 - \frac{a^2}{e^a - 1}.$$

Part (c). It is intuitively clear that if X is large then I is more likely to be 0, and if X is small, then I is more likely to be 1. I give now a sketch of proof that $\mathrm{Cov}(X, I) < 0$. (See if you can fill in the details.) I assume that the range of values of X is $[0, L)$, where L may be finite or ∞.

$$\begin{aligned}
\mathrm{Cov}(X, I) &= E[XI] - E[X]E[I] \\
&= \int_0^a xF'(x)\,dx - F(a)\int_0^L xF'(x)\,dx \\
&= (1 - F(a))\int_0^a xF'(x)\,dx - F(a)\int_a^L xF'(x)\,dx \quad \text{(Splitting the integral from 0 to } L \text{ in two)} \\
&= -(1 - F(a))\int_0^a F(x)\,dx + aF(a)(1 - F(a)) - F(a)\int_a^L xF'(x)\,dx \quad \text{(Integration by parts)} \\
&= -(1 - F(a))\int_0^a F(x)\,dx + aF(a)(1 - F(a)) - \bar{x}F(a)(1 - F(a)) \\
&= -(1 - F(a))\left(\int_0^a F(x)\,dx + (\bar{x} - a)F(a)\right)
\end{aligned}$$

where $\bar{x}$ is the conditional expectation of X assuming that $X \geq a$. Since $\bar{x} > a$, $F(a) < 1$, and $F(x) \geq 0$, we conclude that the covariance is negative.

2. *Show that $Var(\alpha X + (1 - \alpha)W)$ is minimized by α being equal to the value given in Equation (8.3) and determine the resulting variance.*

Due to the general properties of variance and covariance, we can write:

$$\begin{aligned}\mathrm{Var}(\alpha X+(1-\alpha)W) &= \alpha^2\mathrm{Var}(X)+(1-\alpha)^2\mathrm{Var}(W)+2\alpha(1-\alpha)\mathrm{Cov}(X,W)\\ &= \mathrm{Var}(W)+2(\mathrm{Cov}(X,W)-\mathrm{Var}(W))\alpha+(\mathrm{Var}(X)+\mathrm{Var}(W)-2\mathrm{Cov}(X,W))\alpha^2.\end{aligned}$$

This is a quadratic function in α, which is bounded below since $\mathrm{Var}(\alpha X+(1-\alpha)W)\geq 0$. Therefore, the unique critical point is a point of minimum, and it is given by the unique root α^* of the derivative of the quadratic function:

$$2(\mathrm{Cov}(X,W)-\mathrm{Var}(W))+2\alpha^*(\mathrm{Var}(X)+\mathrm{Var}(W)-2\mathrm{Cov}(X,W))=0$$

or

$$\alpha* = \frac{\mathrm{Var}(W)-\mathrm{Cov}(X,W)}{\mathrm{Var}(X)-2\mathrm{Cov}(X,W)+\mathrm{Var}(W)} = \frac{\mathrm{Var}(W)-\mathrm{Cov}(X,W)}{\mathrm{Var}(X-W)}.$$

Replacing this value of α into the expression

$$\mathrm{Var}(W)+2(\mathrm{Cov}(X,W)-\mathrm{Var}(W))\alpha+(\mathrm{Var}(X)+\mathrm{Var}(W)-2\mathrm{Cov}(X,W))\alpha^2$$

and after some straightforward algebraic manipulation, we obtain

$$\mathrm{Var}(\alpha^*X+(1-\alpha^*)W)=\mathrm{Var}(W)-\frac{(\mathrm{Cov}(X,W)-\mathrm{Var}(W))^2}{\mathrm{Var}(X)-2\mathrm{Cov}(X,W)+\mathrm{Var}(W)}.$$

An alternative way to write this expression which is more explicitly symmetric is

$$\mathrm{Var}(\alpha^*X+(1-\alpha^*)W)=\frac{\mathrm{Var}(X)\mathrm{Var}(W)-\mathrm{Cov}^2(X,W)}{\mathrm{Var}(X-W)}.$$

3. *(a) Explain how control variables may be used to estimate θ in Exercise 1. (The latter is Problem 1 of homework assignment 9.)*

   (b) Do 100 *simulation runs, using the control given in (a), to estimate first c^* and then the variance of the estimator.*

   (c) Using the same data as in (b), determine the variance of the antithetic variable estimator.

   (d) Which of the two types of variance reduction techniques worked better in this example?

(a) Let $X=e^{U^2}$, so that $\theta=E\left[e^{U^2}\right]$. One possible choice of control variable is $Y=U^2$. The expected value of Y is $E[Y]=\int_0^1 x^2\,dx=1/3$. So we can use the unbiased estimator of θ given by

$$X_{c^*}=e^{U^2}+c^*(U^2-1/3)$$

where $c^*=-\mathrm{Cov}(e^{U^2},U^2)/\mathrm{Var}(Y)$.

The following Matlab script can be used to answer the other questions (notice that the expected value of Y used is the exact one, equal to 1/3, not the estimated value):

```
m     = 100;
U     = rand(1,m);
Y     = U.^2;
Ybar  = 1/3;
X     = exp(Y);
```

```
Xbar  = sum(X)/m;
J     = ones(1,m);
A     = sum((X-Xbar*J).*(Y-Ybar*J));
B     = sum((Y-Ybar*J).^2);
C     = sum((X-Xbar*J).^2);
CovXY = A/(m-1);
VarY  = B/(m-1);
VarX  = C/(m-1);
c     = -A/B;
%Estimator:
Xc    = X + c*(Y-Ybar*J);
Xcbar = sum(Xc)/m;
%Variance of estimator:
VarXc = (VarX - CovXY^2/VarY)/m
%Antithetic estimator
Xa    = (exp(U.^2)+exp((1-U).^2))/2;
Xabar = sum(Xa)/m;
VarXa = sum((Xa-Xabar*J).^2)/(m-1)
```

(b) One run of the above program gave the following: the estimated value of c^* was -1.5950, and the variance of the estimator X_{c^*} was $\mathrm{Var}(X_{c^*}) = 4.5860 \times 10^{-5}$.

(c) The variance of the antithetic variable estimator (using the same U) was: $\mathrm{Var}(X_a) = 0.0306$, so the reduction in variance is

$$\frac{\mathrm{Var}(X_{c^*})}{\mathrm{Var}(X_a)} = 0.0015.$$

(d) It is clear from part (c) that it is better to use the control variable method. Confidence intervals using the control variable method is $\sqrt{0.0015} = 0.0387$ shorter than confidence intervals obtained by using the antithetic variable method.

4. *Show that in estimating $\theta = E[(1-U^2)^{1/2}]$ it is better to use U^2 rather than U as the control variate. To do this, use simulation to approximate the necessary covariances.*

One run of the below program gave the following values: the variance of the estimator X_1 of $\theta = \pi/4$ that uses the control variable U is $V_1 = 6.3796 \times 10^{-5}$ and the variance of the estimator that uses the control variable U^2 is $V_2 = 1.6246 \times 10^{-5}$. So it is more efficient to use U^2. The quotient is

$$\frac{V_2}{V_1} = 0.2547$$

and the ratio of standard deviations is approximately 0.5.

```
m      = 100;
U      = rand(1,m);
X      = (1-U.^2).^(1/2);
Xbar   = sum(X)/m;
Y1     = U;
Y2     = U.^2;
Y1bar  = 1/2;
```

```
Y2bar  = 1/3;
J      = ones(1,m);
A1     = sum((X-Xbar*J).*(Y1-Y1bar*J));
B1     = sum((Y1-Y1bar*J).^2);
C1     = sum((X-Xbar*J).^2);
CovXY1 = A1/(m-1);
VarY1  = B1/(m-1);
VarX   = C1/(m-1);
c1     = -A1/B1;
%Variance of first estimator, X1c:
X1c     = X + c1*(Y1-Y1bar*J);
X1cbar  = sum(X1c)/m;
VarX1c  = (VarX - CovXY1^2/VarY1)/m
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
A2     = sum((X-Xbar*J).*(Y2-Y2bar*J));
B2     = sum((Y2-Y2bar*J).^2);
C2     = sum((X-Xbar*J).^2);
CovXY2 = A2/(m-1);
VarY2  = B2/(m-1);
VarX   = C2/(m-1);
c2     = -A2/B2;
%Variance of second estimator, X2c:
X2c     = X + c2*(Y2-Y2bar*J);
X2cbar  = sum(X2c)/m;
VarX2c  = (VarX - CovXY2^2/VarY2)/m
```