

Hands-On 5 Variance Reduction Methods

• **Problem 5.1** (Monte Carlo with Antithetic Variables).

We want to calculate the integral $I = \int_{\mathbb{R}^d} \Psi(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$ where $f(\mathbf{x})$ is the joint PDF of a random vector $\mathbf{X} = (X_1, \dots, X_d)$ of independent random variables. Assume that:

- the function $\Psi : \mathbb{R}^d \rightarrow \mathbb{R}$ is monotone in each of its arguments;
- each variable X_j , $j = 1, \dots, d$ is such that $X_j^* = 2\mathbb{E}[X_j] - X_j$ has the same distribution as X_j (i.e. the probability density of X_j is symmetric with respect at the average value $\mathbb{E}[X_j]$).

The variable X_j^* is called the *antithetic* variable of X_j . It can then be proved that

$$\text{Cov}(\Psi(X_1, \dots, X_d), \Psi(X_1^*, \dots, X_d^*)) \leq 0.$$

We introduce the random variables

$$Y = \Psi(\mathbf{X}), \quad Z = \frac{\Psi(\mathbf{X}) + \Psi(\mathbf{X}^*)}{2}.$$

- Show that $\mathbb{E}[Y] = \mathbb{E}[Z] = I$ and that $\text{Var}[Z] \leq \text{Var}[Y]/2$. The previous result suggests introducing the following Monte Carlo method with variance reduction based on antithetic variables,

$$I \approx I_{AV} = \frac{1}{N} \sum_{i=1}^N Z^{(i)}$$

where $Z^{(i)}$ are n values sampled from the variable Z .

- We want to calculate the volume of an octant of the unit sphere in dimension 3:

$$\Omega = \{(x_1, x_2, x_3) \in [0, 1]^3 : x_1^2 + x_2^2 + x_3^2 \leq 1\}.$$

- Show that the function $f(\mathbf{x}) = \mathbf{1}_{\{\mathbf{x} \in \Omega\}}$ is monotone in all its arguments.
- Calculate the volume $|\Omega|$ using the Monte Carlo method with antithetic variables, adaptively determining the number of samples needed to have a relative error less than 0.05 with a confidence level of .95.
- Compare the number of samples found in the previous point with those needed for the crude Monte Carlo method to meet the same tolerance requirements.
- Repeat points 2(b) and 2(c) in the case we want to calculate the volume of the sphere in dimension $d = 8$, an issue already considered in **Problem 4.2, Hands On 4**.

- For the calculation of the integral $I = \int_{\mathbb{R}} e^{x-x^2/2} dx$, already considered in **Problem 4.1, Hands On 4**, compare the performance of the crude Monte Carlo method with the Monte Carlo method that makes use of antithetic variables.

- By assumption, the random vector $\mathbf{X}^* = 2\mathbb{E}[\mathbf{X}] - \mathbf{X}$ has the same distribution of \mathbf{X} . Hence,

$$\mathbb{E}[Z] = \mathbb{E}\left[\frac{\Psi(\mathbf{X}) + \Psi(\mathbf{X}^*)}{2}\right] = \frac{\mathbb{E}[\Psi(\mathbf{X})] + \mathbb{E}[\Psi(\mathbf{X}^*)]}{2} = \mathbb{E}[\Psi(\mathbf{X})] = \mathbb{E}[Y].$$

Moreover,

$$\begin{aligned} \text{Var}[Z] &= \text{Var}\left[\frac{\Psi(\mathbf{X}) + \Psi(\mathbf{X}^*)}{2}\right] = \frac{1}{4} (\text{Var}[\Psi(\mathbf{X})] + \text{Var}[\Psi(\mathbf{X}^*)] + 2\text{Cov}(\Psi(\mathbf{X}), \Psi(\mathbf{X}^*))) \\ &\leq \frac{\text{Var}[\Psi(\mathbf{X})]}{2} = \frac{\text{Var}[Y]}{2}. \end{aligned}$$

Observe that if we had defined Z starting from two independent samples: $Z = (Y_1 + Y_2)/2$ with independent Y_1 and Y_2 distributed as Y , we would have $\text{Var}[Z] = \text{Var}[Y]/2$. Under the hypotheses indicated, the fact of using dependent variables (the variable Y and its antithetic $Y^* = 2\mathbb{E}[Y] - Y$) allows then to reduce the variance.

2. (a) Taking a point in the first octant of the unit sphere (for which, therefore, $\Psi(\mathbf{x}) = 1$), if we increase any of the coordinates of the point, the characteristic function $\Psi(\mathbf{x})$ is still equal to 1 if the point remains inside the sphere and becomes 0 if the point leaves the sphere.

Similarly, if the starting point is outside the sphere ($\Psi(\mathbf{x}) = 0$) and we increase any of its coordinates, the point certainly remains outside the sphere and $\Psi(\mathbf{x})$ is still equal to zero. We deduce that the characteristic function Ψ is monotone decreasing.

In Matlab

```
d = 3; vol_ex = 4 * pi / 24; % volume of the octant
e = 0.05; p = 0.95;
n = 100; n1 = n; % let's start with 100 samples
while n1 >= n
    n = n1;
    u = rand(d, n); % sampling of points in [0,1]^3
    r = sqrt(sum(u.^2));
    Y1 = r < 1;
    ua = 1-u; % antithetic variables
    r = sqrt(sum(ua.^2));
    Y2 = r < 1;
    Z = (Y1 + Y2) / 2;
    V = sum(Z) / n;
    gammaZ = std(Z) / mean(Z);
    n1 = ceil(erfinv(p) * sqrt(2) * gammaZ / e)^2;
end
varZ = std(Z)^2
varZ = 0.0360
n1 antithetic MC
n1 = 225
err = abs(V-vol_ex) / vol_ex
err = 0.0068
```

*we use a loop
because we need to
understand how
many samples are
required to get a
relative error less
than: 0.05*

*} we would use only
these lines for crude MC*

*} we add these lines
since we're considering
antithetic variables
(1-u is the antithetic var.)*

*→ If we don't take into account
the antithetic variables, i.e.:
 $Z = Y_1$;
to get the same level of accuracy
we need:*

n1 = 1521
crude MC

Note the very low number of samples required to meet the accuracy requirements.

- (b) The crude MC method consists in approximating the volume as a sample mean $|\Omega| \approx \sum_{i=1}^N Y^{(i)}/N$, where the variables $Y^{(i)}$ are independent and distributed like the variable $Y = \mathbf{1}_\Omega$. The variance of the variable Y is given by

$$\text{Var}[Y] = \mathbb{E}[Y](1 - \mathbb{E}[Y]) = \frac{\pi}{6}(1 - \frac{\pi}{6}) \approx 0.2494.$$

The variance of Y is therefore approximately six times greater than the variance of Z , calculated at the previous point. Since the number of samples needed to calculate the volume with a relative error less than ε with confidence level $1 - \alpha$, given by

$$N = \left(\frac{z_{1-\alpha/2} \sigma_Y}{\mathbb{E}[Y] \varepsilon} \right)^2 = \text{Var}[Y] \left(\frac{z_{1-\alpha/2}}{\mathbb{E}[Y] \varepsilon} \right)^2$$

is proportional to the variance of the variable whose sample mean is calculated, it can be deduced that the use of antithetical variables of point 2(a) allows to reduce the number of samples required by about a factor of 6. However, keep in mind that each sampling of variable Z involves two evaluations of the integrating function. Therefore, if we count the number of function evaluations, the gain is reduced by a factor of 3.

- (c) In dimension $d = 8$, it is observed that the variance of variable Z is $\text{Var}[Z] \approx 8 \cdot 10^{-3}$. while the variance of Y is $\text{Var}[Y] = 0.0156$. In this case, therefore, the variance is reduced only by a factor of 2. Taking into account that each sampling of Z costs approximately twice as much as a sampling of Y , it follows that the gain, in the use of antithetic variables, is minimal.
- (d) To calculate the integral $I = \int_{\mathbb{R}} e^{x-x^2/2} dx$ we can sample a variable $X \sim N(0, 1)$ and calculate the sample mean

$$I \approx \frac{1}{N} \sum_{i=1}^N Y^{(i)}$$

of random variables $Y^{(i)}$ independent and identically distributed as $Y = \sqrt{2\pi}e^X$. Similarly, if we want to use the technique of antithetic variables, we will calculate the integral as a sample mean

$$I \approx \frac{1}{N} \sum_{i=1}^N Z^{(i)}$$

where the random variables $Z^{(i)}$ are independent and distributed as

$$Z = \sqrt{2\pi}(e^X + e^{-X})/2.$$

In fact, observe that $-X$ is the antithetic variable to X being the probability density f_X symmetric with respect to zero. We estimate the variance reduction obtained with the antithetic variables method using 1000 samples of X :

```
n = 1000;
X = randn (1, n);

% crude Monte Carlo
Y = sqrt (2 * pi) * exp (X);
I = mean (Y)
I = 4.1567

varY = std (Y) ^ 2
varY = 26.8170

% antithetic variables
Ya = sqrt (2 * pi) * exp (-X);
Z = (Y + Ya) / 2;
I = mean (Z)
I = 4.1024

varZ = std (Z) ^ 2
varZ = 8.4339

varY / varZ = 3.1797
```

We see that the (estimated) variance of Z is more than three times smaller than the variance of Y . The method of the antithetic variables allows to considerably reduce the number of samples to be used (more than a factor of 3).

Problem 5.2 (Monte Carlo with Importance Sampling).

We want to calculate the integral

$$I = \int_{\Omega} \Psi(\mathbf{x}) f(\mathbf{x}) d\mathbf{x},$$

being $f(\mathbf{x})d\mathbf{x}$ a suitable probability measure. The crude Monte Carlo method may not be efficient if the integrand function has large values concentrated in small regions of the domain Ω . The importance sampling technique consists in introducing an auxiliary probability distribution $g(\mathbf{x})$, such that $g(\mathbf{x}) > 0$ in Ω , $\int_{\Omega} g(\mathbf{x}) d\mathbf{x} = 1$, and rewrite the integral as

$$I = \int_{\Omega} \left(\frac{\Psi(\mathbf{x}) f(\mathbf{x})}{g(\mathbf{x})} \right) g(\mathbf{x}) d\mathbf{x}.$$

The density g must have the property of making the regions in which $\Psi \cdot f$ is large “more probable”. Let \mathbf{X} be a random vector with distribution f and $Y = \Psi(\mathbf{X})f(\mathbf{X})/g(\mathbf{X})$. The integral can then be calculated as

$$I = \mathbb{E}_g[Y].$$

1. Resume the problem of calculating the volume of the unit sphere Ω in dimension 8 (see **Problem 4.2, Hands On 4**). Aiming at concentrating the points in the sphere as much as possible (and minimize those falling out), we take as auxiliary probability density a Gaussian density with zero mean and covariance matrix $\Sigma = \frac{1}{3}\mathbb{I}$:

$$g(\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^4} e^{-\sum_{i=1}^8 x_i^2 / 2\sigma^2}.$$

The volume $|\Omega|$ can be calculated as

$$|\Omega| = \int_{\mathbb{R}^8} \mathbf{1}_{\Omega}(\mathbf{x}) d\mathbf{x} = \int_{\mathbb{R}^8} \frac{\mathbf{1}_{\Omega}(\mathbf{x})}{g(\mathbf{x})} g(\mathbf{x}) d\mathbf{x}.$$

Use the importance sampling to compute $|\Omega|$ ensuring a relative error less than 0.05 with a confidence level of 0.95.

2. Compare the number of samples used in the previous point with those obtained in **Problem 4.2, Hands On 4**.

Modifying what we did in **Problem 4.2, Hands On 4**, we can use the following commands:

```
d=8; V_ex=pi^4/24;
s=1/3; % standard deviation Gaussian g(x)
tol=.05; alpha=.05;
n=100;
n1=n;

while n1>= n
    n=n1;
    x=s*randn(d,n);
    r2=sum(x.^2);
    Y=r2<1;
    Z=Y.*exp(r2/(2*s^2))*(2*pi*s^2)^4;
    I=mean(Z);
    varZ=var(Z);
    n1=ceil(varZ*(za/(tol*I))^2)
end

I = 4.0668
varZ = 27.9889

err=abs(I-V_ex)/V_ex
err = 0.0020
n = 2601
```

Note how the algorithm use less than 3000 samples, compared to the 10^5 samples used by the crude Monte Carlo method.

To compare the variances between the two methods, let us remark that

$$|\Omega| = \mathbb{E}[Y], \quad \text{with } Y = 2^8 \cdot \mathbf{1}_{\mathbf{X} \in \Omega}, \quad \mathbf{X} \sim \mathcal{U}([-1, 1]^d).$$

being equal to 2^8 the volume of the hypercube $\Sigma = \prod_{i=1}^8 [-1, 1]$, containing Ω . Hence,

$$\text{Var}[Y] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = 2^8 |\Omega| - |\Omega|^2 = 1024.6$$

which is 40 times higher than the variance $\text{Var}[Z]$.

Problem 5.3 (Monte Carlo with Control Variates).

(to do)

Let Y be a random variable whose expected value is to be calculated $I = \mathbb{E}[Y]$. As known, the number of samples to be used in order to be the average estimator sample yields results with a given accuracy proportional to the variance of Y . Let us now consider an auxiliary random variable (or *control variate*) Z , depending on Y , of which we know the expected value $\mathbb{E}[Z]$. We can then build the random variable

$$Z_c = Y + c(Z - \mathbb{E}[Z])$$

with c a constant to be determined. Clearly, $\mathbb{E}[Z_c] = \mathbb{E}[Y]$ for any $c \in \mathbb{R}$.

1. Determine the value c_{opt} which minimizes the variance of Z_c , as well as the value of $\text{Var}[Z_{c_{opt}}]$. Verify that $\text{Var}[Z_{c_{opt}}] \leq \text{Var}[Y]$ is always valid, regardless of the choice of the control variate Z . This result suggests calculating the quantity I as the sample mean of the variable $Z_{c_{opt}}$ instead of Y . The constant c_{opt} can be estimated on the basis of a first sampling of values of Y and \hat{Y} .
2. Let X_1, X_2, \dots, X_5 be five Bernoulli random variables with parameter $p = 0.9$ (i.e. $P(X_i = 1) = p$ and $P(X_i = 0) = 1 - p$). We want to calculate the average value $I = \mathbb{E}[Y]$ of the random variable

$$Y = \min\{X_1, X_2, \dots, X_5\}.$$

Note that the exact value is $I = p^5$ and $\text{Var}[Y] = I(1 - I)$.

Take as control variate $Z = \sum_{i=1}^5 X_i$, whose mean is $\mathbb{E}[Z] = 5p$. Estimate the value of the constant c from 1000 samples of X_1, X_2, \dots, X_5 .

3. Estimate the variance reduction obtained by calculating the quantity I as the sample mean of Z_c instead of Y , using the value of the constant c obtained at the previous step.
4. Write an adaptive algorithm for the computation of I that has the following steps: given an initial choice n_0 of samples,
 - (a) generate n_0 samples of Y and Z and estimate the optimal constant c . Therefore, set $n = n_0$;
 - (b) generate n samples of Y and Z ;
 - (c) generate the quantity I as the sample mean of Z_c ;
 - (d) estimate the variation coefficient of Z_c and the number $\tilde{n}(\varepsilon, \alpha)$ of samples needed to have a relative error less than ε with confidence level $1 - \alpha$;
 - (e) if $\tilde{n}(\varepsilon, \alpha) > n$ set $n = \tilde{n}(\varepsilon, \alpha)$, estimate the optimal constant c starting from the sampled values of Y and \hat{Y} and go back to point (a).

Execute the algorithm with $\varepsilon = 0.01$ and $\alpha = 0.05$. Compare the performance of this algorithm with that of the crude Monte Carlo method.

1. We have

$$\text{Var}[Z_c] = \text{Var}[Y] + c^2 \text{Var}[Z] + 2c \text{Cov}(Y, Z)$$

so that, aiming at minimizing the variance of Z_c ,

$$\frac{d}{dc} \text{Var}[Z_c] = 2c \text{Var}[Z] + 2 \text{Cov}(Y, Z) = 0 \quad \Rightarrow \quad c_{opt} = -\frac{\text{Cov}(Y, Z)}{\text{Var}[Z]}.$$

As a consequence,

$$\text{Var}[Z_{c_{opt}}] = \text{Var}[Y] - \frac{\text{Cov}(Y, Z)^2}{\text{Var}[Z]} \leq \text{Var}[Y];$$

in particular, $\text{Var}[Z_{c_{opt}}] = \text{Var}[Y]$ if and only if Y and Z are uncorrelated.

2. We can estimate the optimal constant c_{opt} by the following Matlab commands:

```
p = 0.9; m = 5;
n = 1000;
U = rand(m, n); % uniform in [0,1]
X = U>1-p; % Bernoulli Be (p)
Y = min(X); % variable Z = min{X1,X2,...X5}
Z = sum(X); % control variate
Cmat = cov(Y,Z); % covariance matrix
c = - Cmat(1,2) / Cmat(2,2) % optimal constant c_opt
c = -0.6695
```

3. The variance of $Z_{c_{opt}} = Y + c_{opt}(Z - \mathbb{E}[Z])$ can be computed using the formula found at the previous point, or by direct calculation, constructing $Z_{c_{opt}}$ from the sampled values of Y and Z , and calculating the sample variance. The two values coincide.

```
varY=Cmat(1,1) % variance of Y
varY = 0.2434

varZc = varY - Cmat(1,2)^2 / Cmat(2,2) % variance of Z (formula)
varZc = 0.0468

meanZ = m*p;
Zc = Y + c*(Z - meanZ);
varZc1 = var(Zc) % variance of Zc (direct calculus) varZ1 =
VarZc1 = 0.0468

rid_var = 1 - varZc/varY % variance reduction
rid_var = 0.8076
```

We can remark a variance reduction of about 80%, thus showing the effectivity of the control variate method.

4. Our goal is to sample the variable $Z_c = Y + c(Z - \mathbb{E}[Z])$ and calculate its average sample. We denote by $\hat{\mu}_{cv}$ the sample mean estimator of Z_c , calculated from a sample $(Y^{(i)}, Z^{(i)})$ of independent random variables distributed as the pair (Y, Z) :

$$\hat{\mu}_{cv} = \frac{1}{N} \sum_{i=1}^N \left(Y^{(i)} + c_{opt}(Z^{(i)} - \mathbb{E}[Z]) \right) = \hat{\mu}_Y + c_{opt}(\hat{\mu}_Z - \mathbb{E}[Z]).$$

If the value of the optimal constant c_{opt} is known a priori or it is calculated independently from the sample $(Y^{(i)}, Z^{(i)})$, with a suitable estimator, we have

$$\mathbb{E}[\hat{\mu}_{cv}] = \mathbb{E}[\hat{\mu}_Y] + \mathbb{E}[c_{opt}] \mathbb{E}[(\hat{\mu}_Z - \mathbb{E}[Z])] = \mathbb{E}[Y]$$

whence $\hat{\mu}_{cv}$ is an unbiased estimator for the mean value of Y . However, this is no longer true, in general, if the constant c_{opt} is estimated starting from the same values $(Y^{(i)}, Z^{(i)})$ used in the $\hat{\mu}_{cv}$ estimator. It is therefore a good practice to estimate c_{opt} and $\hat{\mu}_{cv}$ using independent samples $(Y^{(i)}, Z^{(i)})$ in order to have an unbiased estimator. The proposed algorithm achieves this goal.

```
tol = 0.01; alpha = 0.05;
za = sqrt(2)*erfinv(1-alpha);
n = 100; % we start from 100 samples and estimate c_opt
U = rand(m,n);
X = U > 1 - p;
Y = min(X);
Z = sum(X);
Cmat = cov(Y,Z);
c = -Cmat(1,2) / Cmat(2,2)
c = -0.0688
```

```

n1 = n;      % adaptive algorithm
while n1 >= n
    n = n1;
    U = rand(m,n);
    X = U > 1-p;
    Y = min(X);
    Z = sum(X);
    Zc = Y + c* (Z-meanZ);
    I = mean(Zc);
    varZc = var(Zc);
    n1 = ceil(varZc*(za/(tol*I))^2)
    Cmat = cov(Y,Z);
    c = -Cmat(1,2)/Cmat(2,2);
end

I = 0.5881
n = 5156

```

The algorithm uses about 5000 samples to meet the required accuracy requirements. With the crude Monte Carlo method, we expect to use a number of samples about 5 times higher, based on the estimates of point 3. This can be easily checked with the commands

```

n=100; n1=n;
while n1>=n
    n=n1;
    U=rand(m,n);
    X=U>1-p;
    Y=min(X);
    I=mean(Y);
    n1=ceil(var(Y)*(za/(tol*I))^2);
end

I = 0.5918
n = 26498

```

Problem 5.4 (Monte Carlo with Control Variates).

Consider a system made by 5 components as reported in Figure 1.

Each component of the system is active with probability $p = 0.6$; it can therefore be described by a Bernoulli random variable $s_i, i = 1, \dots, 5$ with parameter p . The system is considered to be working if there is at least one active path that connects the entrance (point A) to the exit (point B). We want to calculate the probability P that the system is in operation. We introduce the Bernoulli variable

$$Y = \max\{s_1 s_4, s_2 s_5, s_1 s_3 s_5, s_2 s_3 s_4\}.$$

The probability sought is given by $P = \mathbb{E}[Y]$.

1. Calculate the probability P by means of the crude Monte Carlo method. In particular, adaptively determine the number of samples to be used to have a relative error less than 0.01 with a confidence level of 0.95.
2. Now consider the control variate

$$\hat{Y} = \sum_{i=1}^5 s_i,$$

whose mean value is $\mu = 5p$. Compute the probability P by means of the control variate technique, adaptively determining the optimal constant c_{opt} and the number of samples necessary to satisfy the same accuracy requirements as at the previous point. Quantify the obtained variance reduction.

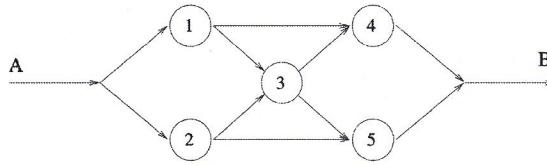


Figure 1: A system made by five components.

In the `c.m` script both the crude MC method and the control variate method are implemented:

```

p = 0.6; m = 5;
tol = 0.01; alpha = 0.05;
za = sqrt(2) * erfinv(1-alpha);

% S: m x n matrix containing on each row n samples of s_i, i=1,...,5
fY='max([S(1,:).*S(4,:); S(2,:).*S(5,:); S(1,:).*S(3,:).*S(5,:); S(2,:).*S(3,:).*S(4,:)])';

% crude monte Carlo
disp('**** crude Monte Carlo ****')

n = 100;
n1 = n;

while n1 >= n
    n=n1;
    U=rand(m,n); % uniform in [0,1]
    S=U>1-p; % sampling of m Be(p) variables
    Y=eval(fY); % evaluation of the function Y
    I=mean(Y);
    varY=var(Y);
    n1=ceil(varY*(za/(tol*I))^2);
end

% control variate
disp('**** control variate ****')

n = 100;
U = rand(m,n); % uniform in [0,1]
S = U>1-p; % sampling of m Be(p) variables
Y = eval(fY); % evaluation of the function Y
Z = sum(S); % control variate
Cmat = cov(Y,Z);
c = -Cmat(1,2) / Cmat(2,2); % estimation of the optimal constatnt

n1 = n;

while n1 >= n
    n = n1;
    U = rand(m,n);
    S = U>1-p;
    Y = eval(fY);
    Z = sum(S);
    Zc = Y + c*(Z-m*p);
    I = mean(Zc);
    varZc = var(Zc);
    n1 = ceil(varZc*(za/(tol*I))^2);
    Cmat = cov(Y,Z);
    c = -Cmat(1,2) / Cmat(2,2);
end

```

Running it we obtain:

```
>> HandsOn5_4

        crude Monte Carlo
I      = 0.6569
varY  = 0.2254
n      = 22789

control variates
I      = 0.6633
varZc = 0.1097
c     = -0.3098
n      = 9627
```

The probability P calculated in both cases is equal to $P = 0.66$ (the exact value is $P = 0.6595$). Furthermore, the variance of Z_c is about half of that of Y , as confirmed by the fact that the adaptive algorithm with control variable uses about half of the samples used by the crude Monte Carlo method.

Problem 5.5 (Monte Carlo with Control Variates).

Let Σ be the unit circle in \mathbb{R}^2 centered at the origin of the axes and Υ the circle of radius 0.3 and center (0.8, 0). We want to calculate the area of the region $\Omega = \Sigma \setminus \Upsilon = \Sigma \cap \Upsilon^c$ (see Figure 2).

Let X be a point randomly chosen in $[-1, 1]^2$, with uniform distribution, and $Y = \mathbf{1}_{\{\mathbf{x} \in \Omega\}}$ the characteristic function of the domain Ω . We therefore have $|\Omega| = 4\mathbb{E}[Y]$.

1. Consider as control variate the variable $Z = \mathbf{1}_{\mathbf{x} \in \Sigma}$, for which $\mathbb{E}[Z] = \pi/4$. Denoting by $\lambda_\omega = \mathbb{E}[Y]$ and $\lambda_\sigma = \mathbb{E}[Z]$, quantify the maximum variance reduction obtainable with the control variate technique for the calculation of $|\Omega|$. Carry out the calculations analytically and express the result as a function of λ_σ and λ_ω .
2. Calculate the area $|\Omega|$ with the control variate technique, adaptively determining the optimal constant c_{opt} as well as the number of samples required to have a relative error less than 0.01 with a confidence level of 0.95.
3. Compare the number of samples found in the previous point with those required by the crude Monte Carlo method to meet the same accuracy requirements.

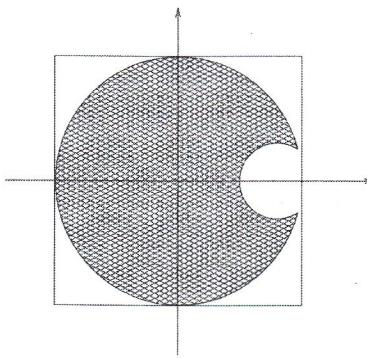


Figure 2: The area to be computed.

1. The variable Y has mean $\mathbb{E}[Y] = \lambda_\omega$ and variance $\text{Var}[Y] = \lambda_\omega(1 - \lambda_\omega)$. Analogously for Z , replacing λ_ω with λ_σ .

Let us now consider the variable

$$Z_c = Y + c(Z - \mathbb{E}[Z])$$

with $c = c_{opt}$; as a consequence,

$$\text{Var}[Z_{c_{opt}}] = \text{Var}[Y] - \frac{\text{Cov}(Y, Z)^2}{\text{Var}[Z]} \leq \text{Var}[Y].$$

Moreover, we have that $YZ = Y$, since the inclusion $\Omega \subset \Sigma$ holds, so that

$$\text{Cov}(Y, Z) = \mathbb{E}[YZ] - \mathbb{E}[Y]\mathbb{E}[Z] = \mathbb{E}[Y] - \mathbb{E}[Y]\mathbb{E}[Z] = \lambda_\omega(1 - \lambda_\sigma).$$

Hence,

$$\text{Var}[Z_{c_{opt}}] = \lambda_\omega(1 - \lambda_\omega) - \frac{\lambda_\omega^2(1 - \lambda_\sigma)^2}{\lambda_\sigma(1 - \lambda_\sigma)} = \lambda_\omega(1 - \lambda_\omega) \left(1 - \frac{\lambda_\omega(1 - \lambda_\sigma)}{\lambda_\sigma(1 - \lambda_\omega)}\right).$$

The maximum variance reduction we can obtain is thus

$$1 - \frac{\text{Var}[Z_{c_{opt}}]}{\text{Var}[Y]} = \frac{\lambda_\omega(1 - \lambda_\sigma)}{\lambda_\sigma(1 - \lambda_\omega)}.$$

From the previous formula, we deduce that if we are able to find a domain $\Sigma \supset \Omega$ such that $|\Sigma|$ is slightly larger than $|\Omega|$, we can achieve a remarkable variance reduction through the control variate technique.

It is also possible to provide a lower bound to the variance reduction we can achieve. Let $\Sigma' \subset \Omega$, $Y' = \mathbf{1}_{\Sigma'}$ and $\lambda_L = \mathbb{E}[Y'] < \lambda_\omega$. We then have

$$\text{variance reduction} = 1 - \frac{\text{Var}[Z_{c_{opt}}]}{\text{Var}[Y]} \geq \frac{\lambda_L}{\lambda_\sigma} \frac{(1 - \lambda_\sigma)}{(1 - \lambda_L)}.$$

In our case, we can take as Σ' the portion of the unit circle for which $x < 0.5$; this set is indeed contained in Ω . In this case, $\lambda_L = \pi/6 + \sqrt{3}/16$ (computing the integral...), $\lambda_\sigma = \pi/4$ and we can obtain that

$$\text{variance reduction} = 1 - \frac{\text{Var}[Z_{c_{opt}}]}{\text{Var}[Y]} \geq 0.47.$$

2. Let $x=2*\text{rand}(2)-1$ a random point in $[-1, 1]^2$. To know if the point falls in Ω , we can simply check if the point falls in Σ and not in Υ . This can be done as in the lines 2-7 appearing in the `while` cycle below.

```

tol=.01; alpha=.05;
za=sqrt(2)*erfinv(1-alpha);

% crude monte Carlo
disp('**** crude Monte Carlo ****')

n=100;
n1=n;

while n1>=n
    n=n1;
    x=2*rand(2,n)-1; % random point in [-1,1]^2
    r1=sqrt(sum(x.^2)); % distance from the origin
    Y1=r1<1; % Y1=1 if x falls in Sigma
    r2= sqrt((x(1,:)-.8).^2+x(2,:).^2); % distance from (.8,0)
    Y2=r2>.3; % Y2=1 if x does not fall in Upsilon
    Y=Y1.*Y2; % Y=1 if x belongs to Omega

    meanY=mean(Y);
    varY=var(Y);
    I=4*meanY;
    n1=ceil(varY*(za/(tol*meanY))^.2);
end

%% control variate
disp('**** control variate ****')

```

```

tol=.01; alpha=.05;
za=sqrt(2)*erfinv(1-alpha);

n=100;
x=2*rand(2,n)-1; % random point in [-1,1]^2
r1=sqrt(sum(x.^2)); % distance from the origin
Y1=r1<1; % Y1=1 if x falls in Sigma
r2= sqrt((x(1,:)-.8).^2+x(2,:).^2); % distance from (.8,0)
Y2=r2>.3; % Y2=1 if x does not fall in Upsilon
Y=Y1.*Y2; % Y=1 if x belongs to Omega
Z=Y1; % the control variate is Y1
Cmat=cov(Y,Y1);
c=-Cmat(1,2)/Cmat(2,2)

n1=n;
while n1>=n
    n=n1;
    x=2*rand(2,n)-1;
    r1=sqrt(sum(x.^2));
    Y1=r1<1;
    r2= sqrt((x(1,:)-.8).^2+x(2,:).^2);
    Y2=r2>.3;
    Y=Y1.*Y2;
    Z=Y1;
    Zc=Y+c*(Z-pi/4);
    meanZc=mean(Zc);
    varZc=var(Zc);
    I=4*meanZc;
    n1=ceil(varZc*(za/(tol*meanZc))^2);
    Cmat=cov(Y,Z);
    c=-Cmat(1,2)/Cmat(2,2);
end

% control variate Y1:
reduction=(1-varZc/varY)*100

```

Running this script, we obtain

```

>> HandOn5_5
      crude Monte Carlo
I= 2.8795
varY = 0.2017
n = 15107

      control variate

c = -0.9620
I = 2.8901
varZc = 0.0582
c = -0.9199
n = 4871

reduction = 71.1425

```

We remark that the value we sought is $2.9 \pm 1\%$, the variance of Z_c is about 4 times smaller than the variance of Y , with a variance reduction larger than 70%.