KTH Electrical Engineering

Exam in EG2080 Monte Carlo Methods in Engineering, 12 January 2015, 14:00–18:00, B23

Instructions

The answer to each problem must begin on a new sheet, but answers to different parts of the same problem (a, b, c, etc.) can be written on the same sheet. The fields *Namn* (Name), *Blad nr* (Sheet number) and *Uppgift nr* (Problem number) must be filled out on every sheet.

Solutions should include sufficient detail that the argument and calculations can be easily followed.

The exam can yield 40 points in total. The examinee is guaranteed to pass if the score is at least 33 points. There will be a possibility to complement for passing the exam if the result is at least 31 points.

Allowed aids

The following aids are allowed in this exam:

- Calculator without information relevant to the course.
- Formulae sheet.

Problem 1 (8 p)

Assume that six values for the random variable Y should be generated, and that Y is a two-state variable with the probability density function

$$f_Y(x) = \begin{cases} 0.15 & x = 0, \\ 0.85 & x = 1, \\ 0 & \text{all other } x. \end{cases}$$

What is the probability that all six random values are equal to one (i.e, $y_i = 1$, $i = 1, \ldots, 6$) in the following cases?

a) (2 p) The random values are generated independently using the inverse transform method (i.e., simple sampling).

b) (3 p) The random values are generated using complementary random numbers.

c) (3 p) The random values are generated using dagger sampling.

Problem 2 (8 p)

The following results have been obtained from a Monte Carlo simulation using simple sampling:

$$\sum_{i=1}^{100} x_i = 20\,000, \quad \sum_{i=1}^{100} x_i^2 = 4\,000\,000.$$

a) (2 p) Calculate the estimated expectation value of X.

b) (2 p) Calculate the estimated variance of X.

c) (4 p) Would it be appropriate to have a stopping rule for the simulation of X, which stops the simulation if the coefficient of variation is less then 0.01? Do not forget to motivate your answer!

Problem 3 (8 p)

Consider a system with three inputs,

$$X = g(Y_1, Y_2, Y_3) = \begin{cases} 0 & \text{if } Y_1(Y_2 + Y_3) > 0, \\ 1 & \text{if } Y_1(Y_2 + Y_3) = 0. \end{cases}$$

The frequency function of each input is

$$f_Y(x) = \begin{cases} 0.01 & x = 0, \\ 0.99 & x = 1, \\ 0 & \text{all other } x. \end{cases}$$

The system is simulated using the importance sampling function

$$f_Z(x) = \begin{cases} 0.4 & x = 0, \\ 0.6 & x = 1, \\ 0 & \text{all other } x, \end{cases}$$

for each component. Table 1 shows five scenarios for this system.

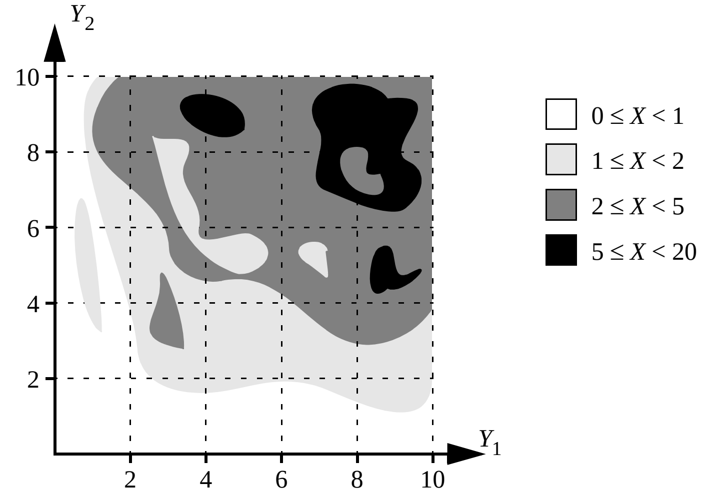
Table 1 Five scenarios for the Monte Carlo simulation in problem 3.

Scenario	Y_1	Y_2	Y_3
1	1	1	0
2	1	1	0
3	0	1	1
4	1	0	1
5	1	1	1

Calculate the estimated expectation value of X.

Problem 4 (8 p)

A system has two independent inputs, which both are uniformly distributed between 0 and 10. The output of the system, $X = g(Y_1, Y_2)$ is illustrated in the figure below, where the different shadings indicate possible ranges of the output value.



a) (5 p) The expectation value of X should be estimated using stratified sampling. Use a strata tree to define appropriate strata.

b) (3 p) Calculate the stratum weights for the strata that were defined in part a.

Problem 5 (8 p)

Table 2 shows the results from a Monte Carlo simulation. Complementary random numbers are applied to one of the inputs of the systems, i.e., for each scenario y_i that is generated, we also get a complementary scenario y_i*. A simplified model is used to generate control variates, and the expectation value of the simplified model has been calculated to 19. Moreover, stratified sampling has been used. 100 original and 100 complementary scenarios have been generated for each stratum. Calculate the estimate of $E[X]$.

Table 2 Results from the Monte Carlo simulation in problem 5.

Stratum, h	Stratum weight, ω_h	Results from detailed model		Results from simplified model	
		Original scenarios, $\sum_{i=1}^{100} x_{h,i}$	Complementary scenarios, $\sum_{i=1}^{100} x_{h,i}^*$	Original scenarios, $\sum_{i=1}^{100} z_{h,i}$	Complementary scenarios, $\sum_{i=1}^{100} z_{h,i}^*$
1	0.7	1 750	2 250	1 500	1 900
2	0.2	2 900	2 100	1 700	2 300
3	0.1	6 000	5 000	3 000	3 800

Problem 1

a) In each trial there is an 85% probability that we get the value 1; hence, for six independent trials the probability that the result is equal to one every time is $0.85^6 \approx 37.7\%$.

b) This time we will only use three pseudorandom numbers, and each time there will be a 70% probability that we get the value 1 for both the original and complementary random number; hence, the probability that the result is equal to one every time is $0.7^3 = 34.3\%$.

c) The dagger cycle length will be equal to six; hence, we will be generating exactly one dagger cycle and the probability that we get only the result one is equal to the probability that the pseudorandom number falls within the rest interval, i.e., 10%.

Problem 2

a) $m_X = \frac{1}{100}\sum_{i=1}^{100} x_i = 200.$

b) $s_X^2 = \frac{1}{100}\sum_{i=1}^{100}(x_i - m_X)^2 = \frac{1}{100}\sum_{i=1}^{100} x_i^2 - m_X^2 = 0.$

c) The estimated variance is equal to zero, which means that all the 100 collected samples produced the same result. There are two possible explanations: either X is not random (and then there would not be any point in sampling X) or X is duogeneous and only conformist units have been sampled so far. In the latter case it is not appropriate to stop the simulation until some diverging units have been sampled. As the coefficient of variation will be equal to zero when the estimated variance is zero—because $a_X = s_X/(m_X\sqrt{n})$—we need some additional criteria in the stopping rule, for example that $s_X > 0$.

Problem 3

First we calculate the weight, w_i, and result x_i, of each scenario:

$$w_1 = \frac{f_Y(y_{1,1})}{f_Z(y_{1,1})}\cdot\frac{f_Y(y_{2,1})}{f_Z(y_{2,1})}\cdot\frac{f_Y(y_{3,1})}{f_Z(y_{3,1})} = \frac{0.99}{0.6}\cdot\frac{0.99}{0.6}\cdot\frac{0.01}{0.4} \approx 0.0681, \qquad x_1 = 0,$$

$$w_2 = \frac{0.99}{0.6}\cdot\frac{0.99}{0.6}\cdot\frac{0.01}{0.4} \approx 0.0681, \qquad x_2 = 0,$$

$$w_3 = \frac{0.01}{0.4}\cdot\frac{0.99}{0.6}\cdot\frac{0.99}{0.6} \approx 0.0681, \qquad x_3 = 1,$$

$$w_4 = \frac{0.99}{0.6}\cdot\frac{0.01}{0.4}\cdot\frac{0.99}{0.6} \approx 0.0681, \qquad x_4 = 0,$$

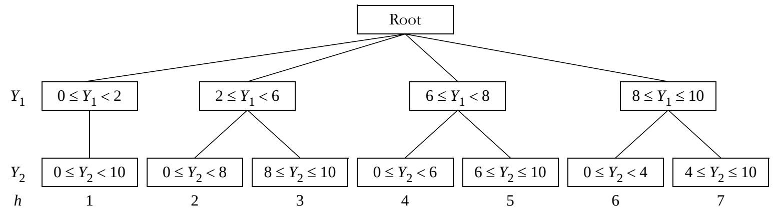
$$w_5 = \frac{0.99}{0.6}\cdot\frac{0.99}{0.6}\cdot\frac{0.99}{0.6} \approx 4.4921, \qquad x_5 = 0.$$

Then, $E[X]$ is estimated by

$$m_X = \frac{1}{5}\sum_{i=1}^{5} w_i x_i \approx 0.0136.$$

Problem 4

a) The idea of stratified sampling is to define strata which are as homogeneous as possible. In this case, the largest differences are between the black areas where the output varies between 5 and 20 (in comparison to between 0 and 5 for the lighter shaded areas). In this solution, we will therefore focus on finding the black areas. One possible strata tree according to this principle is shown below.



b) Both inputs are uniformly distributed; hence, the stratum weight will be equal to the size of stratum divided by the size of the rectangle which comprises the sample space. Thus, we get the following stratum weights:

$\omega_1 = 2 \cdot 10/100 = 0.2$, $\omega_2 = 4 \cdot 8/100 = 0.32$, $\omega_3 = 4 \cdot 2/100 = 0.08$, $\omega_4 = 2 \cdot 6/100 = 0.12$,
$\omega_5 = 2 \cdot 4/100 = 0.08$, $\omega_6 = 2 \cdot 4/100 = 0.08$, $\omega_7 = 2 \cdot 6/100 = 0.12$.

Problem 5

We start by calculating the estimated difference between the detailed model and the control variate for each stratum:

$$m_{(X-Z)h} = \frac{1}{200}\sum_{i=1}^{h}\left((x_{h,i} + x_{h,i}^*) - (z_{h,i} + z_{h,i}^*)\right) = \begin{cases} 3 & h = 1, \\ 5 & h = 2, \\ 21 & h = 3. \end{cases}$$

Then we calculate the estimated difference between the detailed model and the control variate for the entire population:

$$m_{(X-Z)} = \sum_{h=1}^{3}\omega_h m_{(X-Z)h} = 5.2.$$

Finally, we add the expectation value of the control variate to get the estimated expectation value of the detailed model:

$$m_X = m_{(X-Z)} + \mu_Z = 24.2.$$