KTH Electrical Engineering

Exam in EG2080 Monte Carlo Methods in Engineering, 13 December 2010, 9:00–13:00, the seminar room

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:

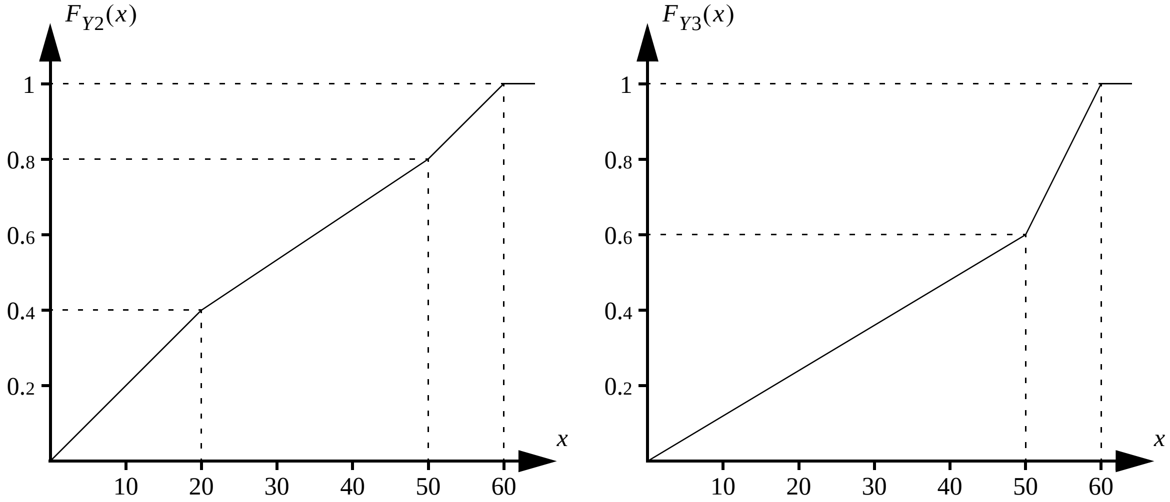
- Calculator without information relevant to the course.
- Formulae sheet.

Problem 1 (8 p)

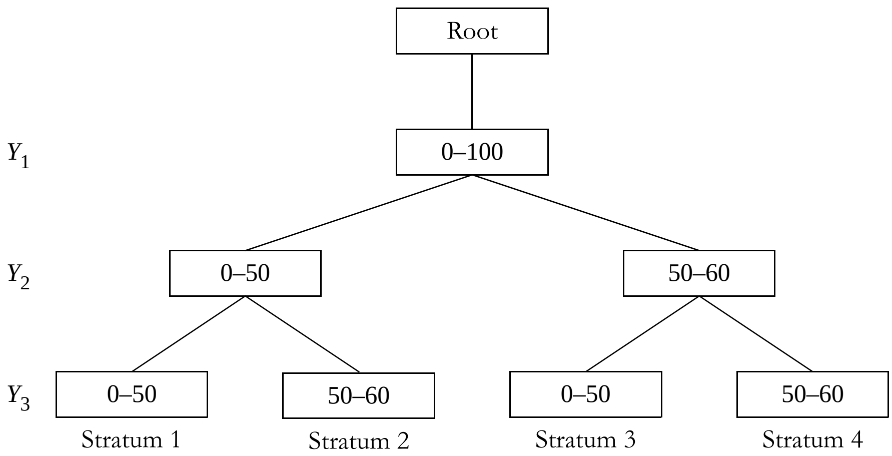
Consider a system with three independent input variables. Y_1 is a two-state random variable with the probability density function

$$f_{Y2}(x) = \begin{cases} 0.4 & x = 0, \\ 0.6 & x = 100, \\ 0 & \text{all other } x. \end{cases}$$

Y_2 and Y_3 have the distribution functions $F_{Y2}(x)$ and $F_{Y3}(x)$ according to the figure below.



This system is simulated using stratified sampling and dagger sampling. Strata have been defined according to the strata tree below. Generate two scenarios for stratum 1 using the random numbers 0.81, 0.75, 0.15, 0.92 and 0.64.



Problem 2 (8 p)

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

$$\sum_{i=1}^{1\,000} x_i = 51\,200, \quad \sum_{i=1}^{1\,000} x_i^2 = 2\,630\,000.$$

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

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

c) (4 p) Calculate a 95% confidence interval for the expectation value of X. It can be assumed that m_X is normally distributed.

Problem 3 (8 p)

A system has two inputs and one output. Table 1 shows the results from a Monte Carlo simulation of this system using importance sampling. What are the estimates of $E[X]$ and $Var[X]$ respectively?.

Table 1 Results from the Monte Carlo simulation in problem 3.

Scenario, i	Input 1			Input 2			Output value, x_i
	Value, $y_{1,i}$	Density function, $f_{Y1}(y_{1,i})$	Importance sampling function, $f_{Z1}(y_{1,i})$	Value, $y_{2,i}$	Density function, $f_{Y2}(y_{2,i})$	Importance sampling function, $f_{Z2}(y_{2,i})$	
1	90	1/100	1/40	99	1/100	1/20	379
2	20	1/100	1/160	91	1/100	1/20	202
3	81	1/100	1/40	12	1/100	1/180	173
4	87	1/100	1/40	13	1/100	1/180	197
5	96	1/100	1/40	23	1/100	1/180	215

Problem 4 (8 p)

Table 2 shows the results of the first 1 000 samples from a Monte Carlo simulation of a model with two outputs. Suggest an appropriate sample allocation for the next 1 000 scenarios to be generated.

Table 2 Results from the Monte Carlo simulation in problem 4.

Stratum, h	Stratum weight, ω_h	Number of samples, n_h	Estimated stratum mean, m_{Xh}		Estimated stratum variance, s_{Xh}^2	
			Output 1	Output 2	Output 1	Output 2
1	0.5	200	500	0	3 600	0
2	0.3	400	800	12	10 000	1
3	0.2	400	1 350	45	90 000	6.25

Problem 5 (8 p)

Table 3 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 9. 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 3 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	0.5	1 600	1 800	1 500	1 650
2	0.3	2 400	2 200	1 900	1 800
3	0.2	3 000	3 200	2 200	2 350

Problem 1

The values of the first input are generated using the dagger transform. The dagger cycle length is $S = \text{floor}(1/p) = \{p = 0.4\} = 2 \Rightarrow$

$$F^{\ddagger}_{Y1,j}(x) = \begin{cases} 0 & \text{if } 0.4(j-1) \le x < 0.4j, \\ 100 & \text{otherwise,} \end{cases} \quad j = 1, 2.$$

If we use the first random number then we get

$$y_{1,j} = F^{\ddagger}_{Y1,j}(0.81) = \begin{cases} 100 & j = 1, \\ 100 & j = 2. \end{cases}$$

The second input should be in the range between 0 and 50 according to the strate tree. This means that we need to transform uniformly distributed random numbers between $\underline{u}_{2h} = F_{Y2}(0) = 0$ and $\bar{u}_{2h} = F_{Y2}(50) = 0.8$. After scaling the next two random numbers to the specified interval we get

$$y_{2,1} = F^{-1}_{Y2}(0.8 \cdot 0.75) = 35,$$

$$y_{2,2} = F^{-1}_{Y2}(0.8 \cdot 0.15) = 6.$$

For Y_2 we need to transform uniformly distributed random numbers between $\underline{u}_{2h} = F_{Y3}(0) = 0$ and $\bar{u}_{2h} = F_{Y2}(50) = 0.6$. After scaling the next two random numbers to the specified interval we get

$$y_{3,1} = F^{-1}_{Y2}(0.6 \cdot 0.92) = 46,$$

$$y_{3,2} = F^{-1}_{Y2}(0.8 \cdot 0.64) = 32.$$

The resulting scenarios are shown in the table below.

Scenario	Input 1	Input 2	Input 3
1	100	35	46
2	100	6	32

Problem 2

a) $m_X = \frac{1}{1\,000} \sum_{i=1}^{1\,000} x_i = 51.20.$

b) $s_X^2 = \frac{1}{1\,000}\left(\sum_{i=1}^{1\,000} (x_i - m_X)^2\right) = \frac{1}{1\,000}\left(\sum_{i=1}^{1\,000} x_i^2\right) - m_X^2 = 8.56.$

c) $\delta = \frac{t_\alpha \cdot s_X}{\sqrt{n}} \approx \{t_{0.95} = 1.9600\} \approx 0.022 \Rightarrow 51.20 \pm 0.18$ is a 95% confidence interval for $E[X]$.

Problem 3

The weight factors for the five scenarios are given by

$$w_i = \frac{f_{Y1}(y_{1,i})}{f_{Z1}(y_{1,i})} \cdot \frac{f_{Y2}(y_{2,i})}{f_{Z2}(y_{2,i})} = \begin{cases} 0.08 & i = 1, \\ 0.32 & i = 2, \\ 0.72 & i = 3, \\ 0.72 & i = 4, \\ 0.72 & i = 5. \end{cases}$$

The estimated expectation value is then given by

$$m_X = \frac{1}{5} \sum_{i=1}^{n} w_i x_i \approx 103$$

and the estimated variance is

$$s_X^2 = \frac{1}{5}\left(\sum_{i=1}^{n} w_i x_i^2\right) - m_X^2 \approx 10\,808.$$

Problem 4

According to the Neyman allocation, the target distribution of 2 000 samples is

$$n_h = n \frac{\omega_h s_{Xh}}{\sum_{k=1}^{L} \omega_k s_{Xk}} = \begin{cases} 500 & h = 1, \\ 500 & h = 2, \\ 1\,000 & h = 3, \end{cases}$$

for the first output, and

$$n_h = n \frac{\omega_h s_{Xh}}{\sum_{k=1}^{L} \omega_k s_{Xk}} = \begin{cases} 0 & h = 1, \\ 750 & h = 2, \\ 1\,250 & h = 3, \end{cases}$$

for the second. Using the mean of these two distributions as a mean and taking into account the scenarios that already have been generated, we should choose to generate 50 scenarios in stratum 1, 225 scenarios in stratum 2, and 725 scenarios in stratum 3.

Problem 5

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

$$m_{(X-Z)h} = \frac{1}{200} \sum_{i=1}^{h} ((x_{h,i} + x_{h,i}^*) - (z_{h,i} + z_{h,i}^*)) = \begin{cases} 1.25 & h = 1, \\ 4.5 & h = 2, \\ 8.25 & h = 3. \end{cases}$$

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

$$m_{(X-Z)} = \sum_{h=1}^{3} \omega_h m_{(X-Z)h} = 3.625.$$

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

$$m_X = m_{(X-Z)} + \mu_Z = 12.625.$$