KTH Electrical Engineering

Exam in EG2080 Monte Carlo Methods in Engineering, 11 December 2012, 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.

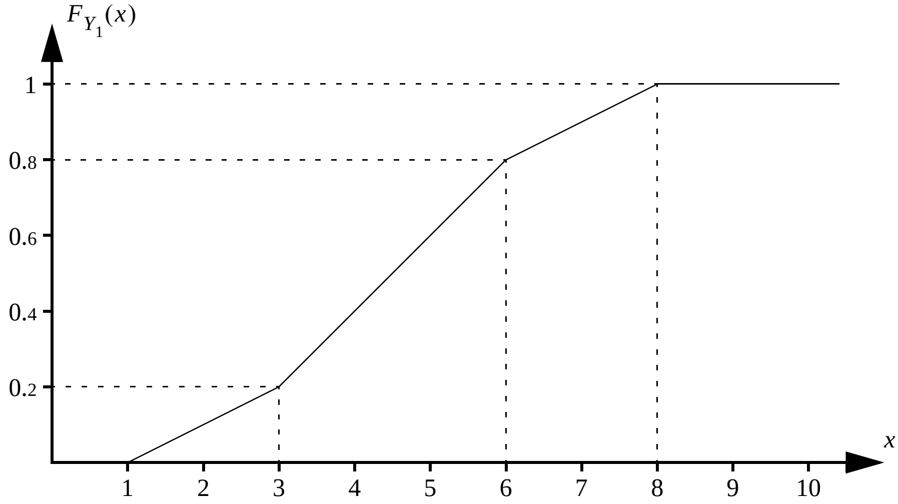
The exam can yield 40 points in total. The examinee is guaranteed to pass if the score is at least 33 points. There will be a possibility to complement for passing the exam if the result is at least 31 points.

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 two inputs. The first input, Y_1, is a continuous random variable, the probability distribution of which is shown in the figure above. The second input, Y_2, is a two-state random variable:

$$f_{Y_2} = \begin{cases} 0.7 & x = 6, \\ 0.3 & x = 12, \\ 0 & \text{all other } x. \end{cases}$$

This system is to be simulated using complementary random numbers (for Y_1) and dagger sampling (for Y_2). Generate five scenarios for this simulation using $U(0, 1)$-distributed random numbers from table 1.

Table 1 Random numbers.

0.815	0.127	0.632	0.278	0.958
0.906	0.913	0.098	0.547	0.965

Problem 2 (8 p)

Assume that the objective of a Monte Carlo simulation is to find an interval $m_X \pm 0.01 m_X$, which is a 95% confidence interval for the expectation value $E[X]$. The stopping criteria for the simulation is to test if the coefficient of variation is less than the relative tolerance level, ρ. Calculate an appropriate value for ρ.

Problem 3 (8 p)

Production Ltd. has two machines, one larger and one smaller. Data for these machines are shown in table 2. The production target, D (i.e., the number of units that Production Ltd. has to manufacture in a day), is varying randomly according to the probability distribution shown in the figure below.

Everyday, Production Ltd. tries to minimise their costs, which means that they produce as much as they can in the larger machine, then they use the smaller machine. If the capacity is insufficient, each unit that cannot be produced costs the company 200 ¤.

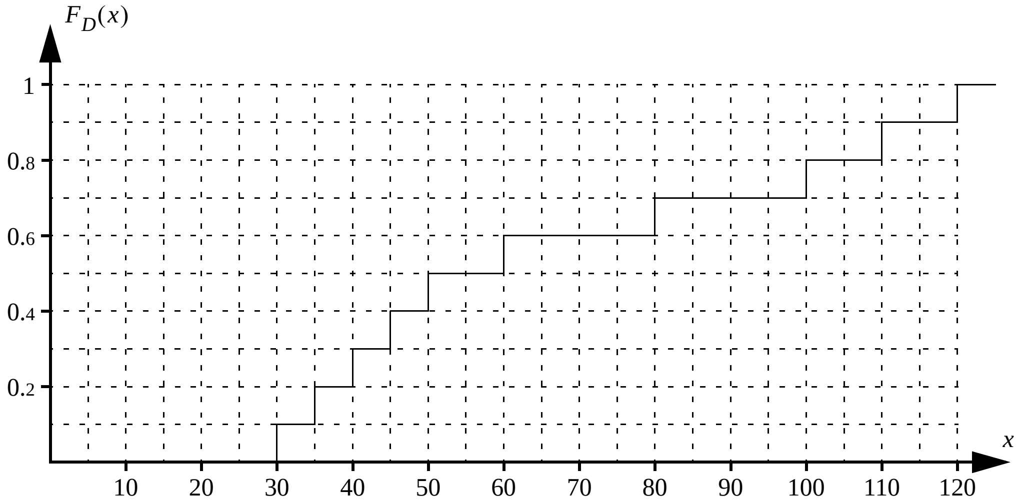


Table 2 Data for the machines of Production Ltd.

Machine	Capacity, C [units/day]	Production cost, β [¤/unit]	Availability, p[a] [%]
Small	50	80	98
Large	100	60	98

a. This is the probability that the machine can be used during a specific day. It can be assumed that the machines are either available or unavailable for each day, i.e., machines never fail during a day.

The operation of Production Ltd. is to be simulated using Monte Carlo techniques. In order to improve the accuracy, importance sampling is applied on the availability of the machines, i.e., the available capacity is randomised using the importance sampling functions

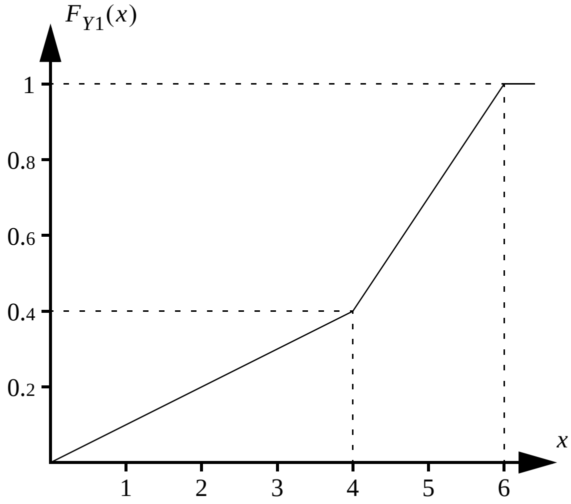
$$f_{C_{\text{small}}} = \begin{cases} 0.3 & x = 0, \\ 0.7 & x = 50, \\ 0 & \text{all other } x, \end{cases} \quad \text{and } f_{C_{\text{large}}} = \begin{cases} 0.3 & x = 0, \\ 0.7 & x = 100, \\ 0 & \text{all other } x, \end{cases} \quad \text{respectively.}$$

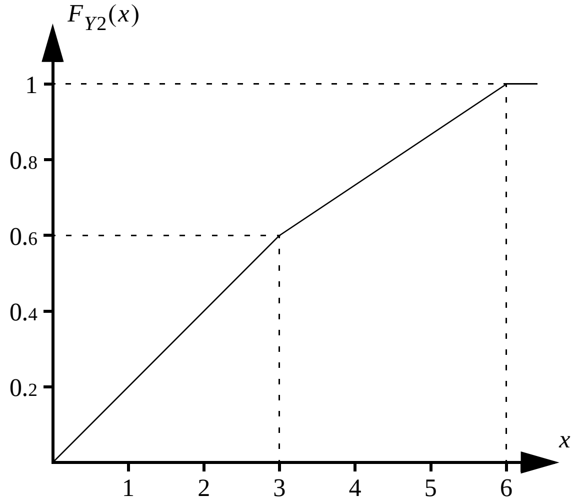
Table 3 shows five scenarios for this Monte Carlo simulation. What is the estimated daily production cost based on these five scenarios?

Table 3 Scenarios for the simulation of Production Ltd.

Scenario, i	Available capacity in the small machine, C_{small} [units/day]	Available capacity in the large machine, C_{large} [units/day]	Production target, D [units/day]
1	0	0	80
2	50	100	30
3	50	100	110
4	50	100	120
5	50	100	80

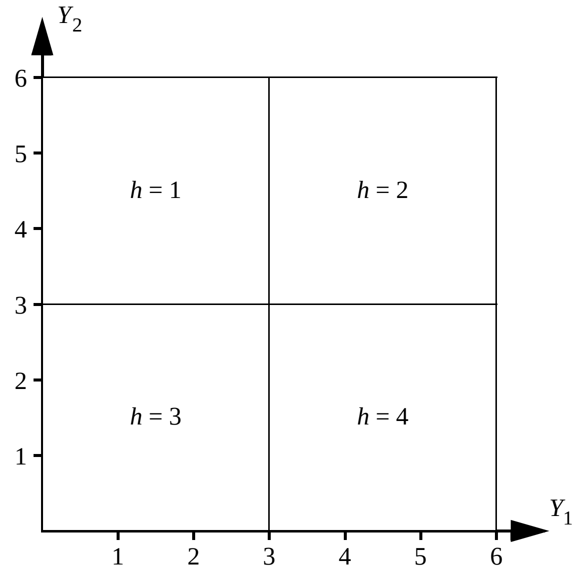
Problem 4 (8 p)





Y_1 and Y_2 are two independent random variables with distribution functions $F_{Y1}(x)$ and $F_{Y2}(x)$ according to the figure above.

a) (3 p) The system $X = g(Y_1, Y_2)$ is simulated using stratified sampling. Strata have been defined according to the figure below. Calculate the stratum weights.



b) (5 p) Table 4 shows ten scenarios in a Monte Carlo simulation of the model $g(Y_1, Y_2)$. Calculate the estimated expectation value according to these samples.

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

Scenario	1	2	3	4	5	6	7	8	9	10
Y_1	0.5	1.2	1.6	2.8	3.4	4.4	4.5	5.2	5.5	5.9
Y_2	1.3	4.8	2.2	5.2	5.1	3.9	2.4	4.3	0.6	3.6
X	2.1	5.7	3.3	8.5	14.7	13.6	7.3	16.9	5.7	15.6

Problem 5 (8 p)

Consider a Monte Carlo simulation which is using stratified sampling, complementary random numbers (on one input) and importance sampling (on the remaining inputs). The results of the simulation are shown in table 5. What is the estimated expectation value of this system?

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

Stratum, h	Stratum weight, ω_h	Samples per stratum, n_h	Results original scenarios, $\sum_{i=1}^{n_h/2} \frac{f_Y(y_{h,i})}{f_Z(y_{h,i})} g(y_{h,i})$	Results from complementary scenariosl, $\sum_{i=1}^{n_h/2} \frac{f_Y(y_{h,i}^*)}{f_Z(y_{h,i}^*)} g(y_{h,i}^*)$
1	0.5	200	1 650	1 550
2	0.3	400	3 800	4 600
3	0.2	400	7 400	6 200

Problem 1

For the first input, we get five scenarios by generating three random numbers according to F_{Y_1} and then using two complementary random numbers:

$$y_{1,1} = F_{Y_1}^{-1}(0.815) = 6.15,$$

$$y_{1,2} = F_{Y_1}^{-1}(1 - 0.815) = 2.85,$$

$$y_{1,3} = F_{Y_1}^{-1}(0.127) = 2.27,$$

$$y_{1,4} = F_{Y_1}^{-1}(1 - 0.127) = 6.73,$$

$$y_{1,5} = F_{Y_1}^{-1}(0.632) = 3.216.$$

The second input has a dagger cycle length of three, which means that we need to generate two dagger cycles in order to get five random values (the last value of the second cycle is however discarded):

$$y_{2,1} = F^{\ddagger}_{Y_2,1}(0.278) = 12,$$

$$y_{2,2} = F^{\ddagger}_{Y_2,2}(0.278) = 6,$$

$$y_{2,3} = F^{\ddagger}_{Y_2,3}(0.278) = 6,$$

$$y_{2,4} = F^{\ddagger}_{Y_2,1}(0.958) = 6,$$

$$y_{2,5} = F^{\ddagger}_{Y_2,2}(0.958) = 6.$$

Problem 2

$m_X \pm \delta$, where $\delta = \frac{t_{0.95} \cdot s_X}{\sqrt{n}}$, is a 95% confidence interval. If $\delta \le 0.01 m_X$ then we get $\frac{t_{0.95} \cdot s_X}{\sqrt{n}} \le 0.01 m_X$; hence $a_X = \frac{s_X}{m_X \sqrt{n}} \le \frac{0.01}{t_{0.95}}$. Consequently, $\rho = 0.01/1.9600 \approx 0.005$ is an appropriate relative tolerance.

Problem 3

Let the capacity of the smaller machine be input 1 and the capacity of the larger machine be input 2. In order to compute the estimate, we need to calculate the input density function $f_{Y_j}(y_j)$, $j = 1, 2$, the importance sampling function $f_{Z_j}(y_j)$, $j = 1, 2$, and the output x for the five scenarios:

Scenario	$f_{Y_1}(y_1)$	$f_{Z_1}(y_1)$	$f_{Y_2}(y_2)$	$f_{Z_2}(y_2)$	$w = \frac{f_{Y_1}(y_1) f_{Y_2}(y_2)}{f_{Z_1}(y_1) f_{Z_2}(y_2)}$	x
1	0.02	0.3	0.02	0.3	0.0044	16 000
2	0.98	0.7	0.98	0.7	1.96	1 800
3	0.98	0.7	0.98	0.7	1.96	6 800
4	0.98	0.7	0.98	0.7	1.96	7 600
5	0.98	0.7	0.98	0.7	1.96	4 800

The estimated expected operation cost of Production Ltd. is then given by

$$m_X = \frac{1}{5} \sum_{i=1}^{5} w_i x_i \approx 8\,246 \text{ ¤/day}.$$

Problem 4

a) $\omega_1 = P(Y_1 \le 3) \cdot P(Y_2 \ge 3) = F_{Y1}(3) \cdot (1 - F_{Y2}(3)) = 0.12$.
Similarly, $\omega_2 = (1 - F_{Y1}(3)) \cdot (1 - F_{Y2}(3)) = 0.28$, $\omega_3 = F_{Y1}(3) \cdot F_{Y2}(3) = 0.18$ and $\omega_4 = (1 - F_{Y1}(3)) \cdot F_{Y2}(3) = 0.42$.

b) We start by sorting out which scenario belongs to which stratum and then we calculate the estimated expectation value for each stratum. The result is as follows:

Stratum, h	Scenarios, i	Estimated expectation value, $m_{Xh} = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{h,i}$
1	2, 4	7.1
2	5, 6, 8, 10	15.2
3	1, 3	2.7
4	7, 9	6.5

The estimate for the entire population is obtained by

$$m_X = \sum_{h=1}^{4} \omega_h m_{Xh} = 8.324.$$

Problem 5

We start by calculating the estimated expectation value for each stratum:

$$m_{X\,h} = \frac{1}{n_h} \sum_{i=1}^{n_h/2} \left(\frac{f_Y(y_{h,i})}{f_Z(y_{h,i})} g(y_{h,i}) + \frac{f_Y(y_{h,i}^*)}{f_Z(y_{h,i}^*)} g(y_{h,i}^*) \right) = \begin{cases} 16 & h = 1, \\ 21 & h = 2, \\ 34 & h = 3. \end{cases}$$

Then we calculate the estimate for the entire population:

$$m_X = \sum_{h=1}^{3} \omega_h m_{Xh} = 21.1.$$