KTH Electrical Engineering

Complementary test in EG2080 Monte Carlo Methods in Engineering, 19 January 2012, 9:00–11:00, the seminar room

Instructions

Only the problems indicated in your exam have to be answered (the score of the remaining problems is kept from the exam).

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 maximal score of the complementary test in is 40 points including the points that are kept from the exam. You are guaranteed to pass if you get at least 33 points.

Allowed aids

The following aids are allowed in this complementary test:

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

Problem 1 (8 p)

Y_1 and Y_2 are two correlated random variables. The joint density function is given by

$$f_Y(x_1, x_2) = \begin{cases} 0.008 & \text{if } 0 \le x_1 \le 5,\ 0 \le x_2 \le 10, \\ 0.008 & \text{if } 5 \le x_1 \le 10,\ 0 \le x_2 \le 5, \\ 0.016 & \text{if } 5 \le x_1 \le 10,\ 5 < x_2 \le 10, \\ 0 & \text{all other } (x_1, x_2). \end{cases}$$

Generate random values of Y_1 and Y_2 using the random numbers 0.22 and 0.64 from a $U(0, 1)$-distribution.

Problem 2 (8 p)

The following samples have been obtained in a Monte Carlo simulation using simple sampling:

300, 26, 93, 67, 222, 144, 61, 2, 172, 59.

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

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

c) (4 p) Calculate a 99% confidence interval for the expectation value of X. It can be assumed that m_X is normally distributed.

Problem 3 (8 p)

Consider a simplified model

$$Z = \tilde{g}(Y) = \begin{cases} Y & \text{if } Y \le 10, \\ 2Y & \text{if } Y > 10, \end{cases}$$

of the system $X = g(Y)$, where the input Y has the density function

$$f_Y(x) = \begin{cases} 0 & x < 0, \\ 0.05 & 0 \le x \le 20, \\ 0 & 20 < x. \end{cases}$$

Use the simplified model to suggest an importance sampling function for this system.

Problem 4 (8 p)

A system $X = g(Y)$ is simulated using stratified sampling. Three strata have been defined and the properties of these strata are shown in table 1. Moreover, the true mean and variance of the entire population are $\mu_X = 0.019$ and $\sigma_X^2 = 0.0186$ respectively.

Assume that during the simulation, n scenarios will be generated, and samples can be distributed exactly according to the Neyman allocation. Is this stratification then more efficient than using simple sampling with the same number of samples?

Table 1 Properties of the strata in problem 4.

Stratum, h	Weight, ω_h	Mean, μ_{Xh}	Variance, σ_{Xh}^2
1	0.9	0	0
2	0.09	0.1	0.09
3	0.01	1	0

Problem 5 (8 p)

Monte Carlo simulation has been used to compare the two systems $X_1 = g_1(Y)$ and $X_2 = g_2(Y)$. Some results of the simulation are shown in table 2. Calculate the estimated difference between $E[X_1]$ and $E[X_2]$. If the table does not contain all the information you need to solve the problem, then you may assume an arbitrary value for the missing information.

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

Stratum, h	Stratum weight, ω_h	Samples per stratum, n_h	Results from system 1 $\sum_{i=1}^{n_h} g_1(y_{h,i})$	Results from system 2l, $\sum_{i=1}^{n_h} g_2(y_{h,i})$
1	0.5	200	1 600	1 500
2	0.3	400	4 800	3 800
3	0.2	400	8 000	6 000

Suggested solution for complementary test in EG2080 Monte Carlo Methods in Engineering, 19 January, 2012

Problem 1

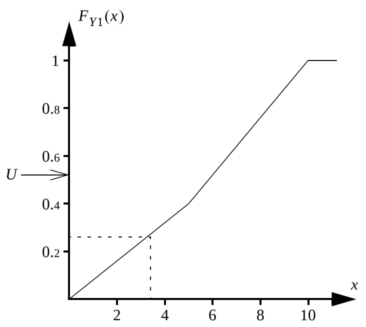
Let us start by randomising a value for the first input. To apply the inverse transform method, we need the distribution function $F_{Y1}(x_1)$, which can be found by integrating the density function $f_{Y1}(x_1)$. To find the density function of a single variable in a bivariate distribution, we need to compute

$$f_{Y1}(x_1) = \int_{-\infty}^{\infty} f_Y(x_1, x_2)dx_2.$$

However, in this case the bivariate distribution is rather straightforward, and it is easy to see that all states of Y_1 in the interval (0, 5) are equally probable and the same goes for all states in the interval (5, 10). We can also see that the total probability that $Y_1 < 5$ is 40%, because

$$P(Y_1 < 5) = \int_{-\infty}^{5} \int_{-\infty}^{\infty} f_Y(x_1, x_2)dx_2dx_1 = 0.008 \cdot 5 \cdot 10 = 0.4.$$

This results in the distribution function shown below. Applying the inverse transform method on the random number 0.22 results in the value $y_1 = 2.75$.



Now, we need the probability distribution of Y_2 given that the outcome of the first variable is equal to 2.75, i.e. we need to compute

$$f_{Y2|Y_1 = 2.75}(x_2) = \frac{f_Y(6, x_2)}{f_{Y1}(2.75)}.$$

Fortunately, we may observe that $f_Y(6, x_2)$ has a constant value in the interval (0, 10) and is equal to zero otherwise; hence, Y_2 must be uniformly distributed between 0 and 10. Applying the inverse transform method on the random number 0.64 then yields $y_1 = 6.4$.

Problem 2

a) $m_X = \frac{1}{10} \sum_{i=1}^{10} x_i = 114.6.$

b) $s_X^2 = \frac{1}{10} \sum_{i=1}^{10} (x_i - m_X)^2 \approx 7\,929.2.$

c) $\delta = \frac{t_\alpha \cdot s_X}{\sqrt{n}} \approx \{t_{0.99} = 2.5758\} \approx 72.5 \Rightarrow 114.6 \pm 72.5$ is a 99% confidence interval for $E[X]$.

Problem 3

Choose the importance sampling function

$$f_Z(\psi) = \frac{\tilde{g}(\psi)f_Y(\psi)}{\mu_Z},$$

where

$$\mu_Z = E[\tilde{g}(Y)] = \int_{-\infty}^{\infty} f_Y(x)\tilde{g}(x)dx = \int_0^{10} 0.05x dx + \int_{10}^{20} 0.05 \cdot 2x dx = 2.5 + 15 = 17.5.$$

Hence, we get

$$f_Z(\psi) = \begin{cases} 0 & \psi < 0, \\ \frac{0.05 \cdot \psi}{17.5} & 0 \le \psi \le 10, \\ \frac{0.05 \cdot 2\psi}{17.5} & 10 < \psi \le 20, \\ 0 & 20 < \psi. \end{cases}$$

Problem 4

The variance of the estimate for stratified sampling is given by

$$Var[M_X] = \sum_{h=1}^{L} \omega_h^2 Var[M_{Xh}].$$

However, in this case there is no uncertainty in strata 1 and 3 (since the standard deviations of these strata are equal to zero), i.e., $Var[M_{X1}] = Var[M_{X3}] = 0$, and will not contribute to the variance of the variance of the estimate for the entire population. Moreover, the Neyman allocation will not assign any scenarios to these two strata, which means that all the n samples will be collected from stratum 2. Thus, we get

$$Var[M_X] = \omega_2^2 Var[M_{X2}] = \omega_2^2 \frac{Var[X_2]}{n} = 0.09^2 \cdot 0.09/n = 0.000729/n.$$

The variance of the estimate when applying simple sampling is given by

$$Var[M_X] = \frac{\sigma_X^2}{n} = 0.0186/n.$$

Assuming that the differences in computation time are negligible, the variance of the estimate is much lower for stratified sampling; hence, this application of stratified sampling is efficient.

Problem 5

The information is sufficient. We start by calculating the estimated difference between the two systems for each stratum:

$$m_{(X1-X2)h} = \frac{1}{n_h} \sum_{i=1}^{h} (g_1(y_{h,i}) - g_2(y_{h,i})) = \begin{cases} 0.5 & h = 1, \\ 2.5 & h = 2, \\ 5 & h = 3. \end{cases}$$

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

$$m_{(X1-X2)} = \sum_{=1}^{3} \omega_h m_{(X1-X2)i} = 2.$$