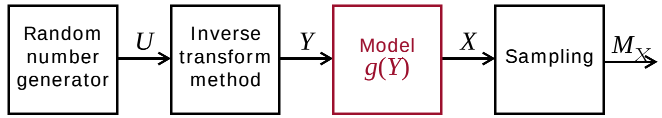
EG2080 Monte Carlo Methods
in Engineering

Theme 5

CONTROL VARIATES & CORRELATED SAMPLING

INTRODUCTION

- Only one model is simulated in simple sampling.
- Sometimes it is possible to increase the accuracy by sampling several models at the same time.



CONTROL VARIATES

Assume a population with the following properties:

- Each unit has a value, x_i, the probability distribution of which is to be investigated using Monte Carlo methods.
- Each unit also has another value, z_i, which is referred to as a control variate.
- The expectation value of the control variate is known:

$$E[Z] = \frac{1}{N} \sum_{i=1}^{N} z_i = \mu_Z.$$

CONTROL VARIATES

- Study $M_{(X-Z)} + \mu_Z$, where

$$m_{(X-Z)} = \frac{1}{n} \sum_{i=1}^{n} (x_i - z_i).$$

- Expectation value:

$$E[M_{(X-Z)} + \mu_Z] = E[M_{(X-Z)}] + \mu_Z =$$
$$= E[X - Z] + \mu_Z = E[X] - E[Z] + \mu_Z = \mu_X, \quad (9)$$

i.e., $M_{(X-Z)} + \mu_Z$ is an estimate of $E[X]$.

CONTROL VARIATES

- Variance:

$$Var[M_{(X-Z)} + \mu_Z] = Var[M_{(X-Z)}] =$$
$$= \frac{Var[X-Z]}{n} =$$
$$= \frac{1}{n}(Var[X] + Var[Z] - 2Cov[X, Z]). \qquad (10)$$

- Compare to simple sampling of X, where

$$Var[M_X] = \frac{Var[X]}{n}.$$

KTH Electrical Engineering

CONTROL VARIATES

- If we can find a control variate such that

$$Var[X] + Var[Z] - 2Cov[X, Z] < \mathrm{Var}[X]$$

then it would be more efficient to sample the difference $(X-Z)$ instead of the absolute values of X.

- In our simulation problem, the control variate can be generated by another model, i.e., $X = g(Y)$ and $Z = \tilde{g}(Y)$.

KTH Electrical Engineering

SIMPLIFIED MODEL

The model used to generate control variates, $\tilde{g}(Y)$, should be a simplified model of the system to be simulated, $g(Y)$.

- It must be possible to determine the expectation value of the outputs from the simplified model, i.e., $\mu_Z = E[\tilde{g}(Y)]$.
- The simplified model must preserve the main characteristics of the original model so that $X = g(Y)$ and $Z = \tilde{g}(Y)$ are sufficiently positively correlated.

KTH Electrical Engineering

SIMPLIFIED MODEL

- How can we find a simplified model?
- Obviously, we should try to remove those properties of $g(Y)$ which prevents us from computing the integral

$$E[g(Y)] = \int_{\mathcal{Y}} f_Y(y)g(y)dy. \qquad (11)$$

- It may not always be possible to find a good simplified model!

KTH Electrical Engineering

MODIFIED INPUTS

- The number of dimensions in the integral (11) is reduced if some elements of the input vector Y can be neglected in the simplified model.
- It is also possible to introduce extra inputs $Y_E = h(Y_1, \ldots, Y_J)$ which can replace the inputs $Y_1, \ldots, Y_J$ in the simplified model.

Cf. examples 17 and 18.

KTH Electrical Engineering

EXPLICIT FUNCTION

- It is easier to solve the integral (11) if the simplified model can be expressed as an explicit function.
- If $g(Y)$ is an implicit function, try to let $\tilde{g}(Y)$ be an explicit function.

KTH Electrical Engineering

EXAMPLE 22 - Simplified model of the Ice-Cream Company

Suggest a simplified model for the system in example 15.

Solution:

- Start by listing the main characteristics of the model. Which characteristics are essential to get an approximative value of the net income and the risk of missing a delivery?

KTH Electrical Engineering

EXAMPLE 22 - Simplified model of the Ice-Cream Company

- ☐ There are different flavours.
- ☐ Ingredients are not always available.
- ☐ The machines have limited capacity.
- ☐ Operation cost of machines.
- ☐ The difference between the normal price paid by the restaurants and the ingredient cost is 2 €/litre.
- ☐ Replacement ice-cream is sold for half price.

KTH Electrical Engineering

EXAMPLE 22 - Simplified model of the Ice-Cream Company

☐ There are different flavours.

☐ Ingredients are not always available.

☑ The machines have limited capacity.

☐ Operation cost of machines.

☑ The difference between the normal price paid by the restaurants and the ingredient cost is 2 €/litre.

☐ Replacement ice-cream is sold for half price.

EXAMPLE 22 - Simplified model of the Ice-Cream Company

- The conclusion is that we should use a model which only considers the total demand for ice-cream, the net income of selling 1 litre of ice-cream for the normal price and the total capacity of the machines.
- However, it is easier to calculate μ_Z for the net income, if we also ignore the total capacity of the machines, because then we get

$$\mu_{\tilde{NI}} = E\left[2 \sum_{f=1}^{6} D_f\right] = 1\,050.$$

EXAMPLE 22 - Simplified model of the Ice-Cream Company

- The total capacity of the machines is 800 litres of ice-cream per day.
- The capacity is insufficient if the demand is maximal for all flavours.

  The maximal value for D_{tot} is $3\cdot100 + 3\cdot200 = 900$.
- The capacity is also insufficient if the demand for one flavour is average while the demand for the other flavours is maximal.

  The difference between maximal and average demand is either 50 or 75 litres, i.e., in these cases we get D_{tot} equal to 850 or 825 litres.

EXAMPLE 22 - Simplified model of the Ice-Cream Company

- The capacity is sufficient if the demand for one flavour is low or if the demand for two flavours are average while the demand for the other flavours is maximal.

  The difference between maximal and low demand is either 100 or 150 litres, and the difference between maximal demand and average demand for two flavours is either 100, 125 or 150 litres, i.e., in these cases we get D_{tot} equal to 850 or 825 litres.

EXAMPLE 22 - Simplified model of the Ice-Cream Company

- For the simplified model, we get the following probability of missing a delivery:
- Now we can calculate the expectation values of the simplified model:

$$\mu_{\tilde{MD}} = 0.25^6 + 6 \cdot 0.25^5 \cdot 0.5 = 0.32\%.$$

EXAMPLE 22 - Simplified model of the Ice-Cream Company

- Finally, the simplified model can be summarised with the following two formulae:

$$\tilde{NI} = 2D_{tot},$$

$$\tilde{MD} = \begin{cases} 0 & \text{if } D_{tot} \leq 800, \\ 1 & \text{if } D_{tot} > 800. \end{cases}$$

EXPECTATION VALUE OF THE CONTROL VARIATE

- It does not matter if we sample $X - Z + \mu_Z$ instead of $X - Z$ when we are estimating $E[X]$:

$$m_{(X - Z + \mu_Z)} = \frac{1}{n}\sum_{i=1}^{n}(x_i - z_i + \mu_Z) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - z_i) + \mu_Z = m_{(X-Z)} + \mu_Z. \qquad (12)$$

EXPECTATION VALUE OF THE CONTROL VARIATE

- The estimated variance is not affected neither:

$$s^2_{(X - Z + \mu_X)} = \frac{1}{n}\sum_{i=1}^{n}((x_i - z_i + \mu_Z) - m_{(X - Z + \mu_Z)})^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - z_i + \mu_Z - (m_{(X-Z)} + \mu_Z))^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - z_i - m_{(X-Z)})^2 = s^2_{(X-Z)}. \qquad (13)$$

EXPECTATION VALUE OF THE CONTROL VARIATE

- However, the coefficient of variation is affected:

$$a_{(X-Z+\mu_Z)} = \frac{s_{(X-Z+\mu_Z)}}{m_{(X-Z+\mu_Z)}\sqrt{n}} = \frac{s_{(X-Z)}}{m_X\sqrt{n}}, \quad (14)$$

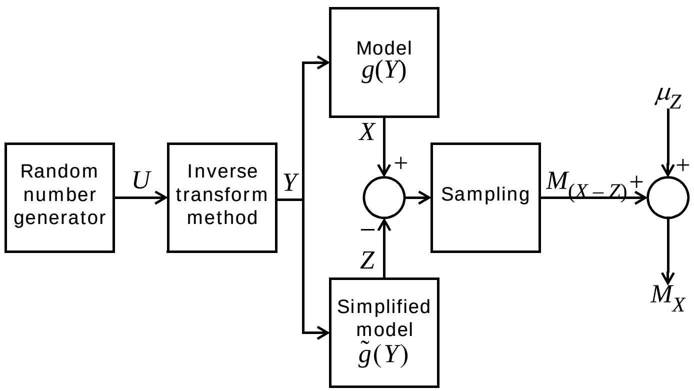
and

$$a_{(X-Z)} = \frac{s_{(X-Z)}}{m_{(X-Z)}\sqrt{n}}. \quad (15)$$

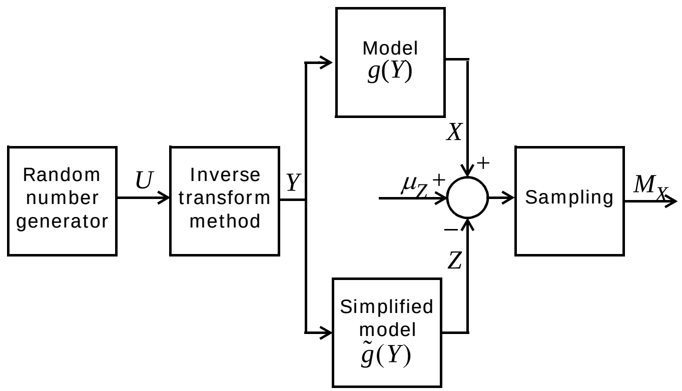
EXPECTATION VALUE OF THE CONTROL VARIATE

- Sampling $X - Z + \mu_X$ results in a coefficient of variation which is directly comparable to simple sampling.
- Sampling $X - Z$ saves n additions during the sampling process.

SIMULATION PROCEDURE



SIMULATION PROCEDURE



SIMULATION PROCEDURE

- Step 1. Generate the first batch of scenarios, y_i, $i = 1, \ldots, n_b$.
- Step 2. Calculate $x_i = g(y_i)$ and $z_i = \tilde{g}(y)$, $i = 1, \ldots, n_b$.
- Step 3. Calculate the sums $\sum_{i=1}^{n} (x_i - z_i)$ and $\sum_{i=1}^{n} (x_i - z_i)^2$ or store all samples $x_i - z_i$ (depending on the objective of the simulation).

SIMULATION PROCEDURE

- Step 4. Test stopping rule. If not fulfilled, repeat step 1 and 3 for the next batch.
- Step 5. Calculate estimates and present results.

EXAMPLE 23 - Effectiveness of ICC simulation

Table 8 Results of ICC simulation in example 23.

Simulation method	Expected net income [€/day]			Risk for missed delivery [%]			Average simulation time [h:min:s]
	Min	Av.	Max	Min	Av.	Max	
Enumeration		874			0.32		0:34:54
Simple samp.	829	872	909	0.00	0.27	1.00	0:02:25
Compl. r.n.	842	875	914	0.00	0.27	1.00	0:02:25
Dagger samp.	856	877	900	0.00	0.27	1.00	0:02:27
Control variate	852	871	890	0.32	0.32	0.32	0:02:25

CORRELATED SAMPLING

- We want to compare two expectation values, μ_{X1} and μ_{X2}.
- Study $M_{(X1-X2)} = M_{X1} - M_{X2}$.
- Expectation value:

$$E[M_{(X1-X2)}] = E[M_{X1} - M_{X2}] = \\ = E[M_{X1}] - E[M_{X2}] = \mu_{X1} - \mu_{X2}, \qquad (16)$$

i.e., $M_{(X1-X2)}$ is an estimate of the difference between the systems.

KTH Electrical Engineering

CORRELATED SAMPLING

- Variance:

$$Var[M_{(X1-X2)}] = Var[M_{X1} - M_{X2}] =$$
$$= Var[M_{X1}] + Var[M_{X2}] - 2Cov[M_{X1}, M_{X2}]. \quad (17)$$

- If M_{X1} and M_{X2} are estimated in separate simulations, we will have $Cov[M_{X1}, M_{X2}] = 0$.
- The variance will be lower if M_{X1} and M_{X2} are positively correlated.

KTH Electrical Engineering

CORRELATED SAMPLING

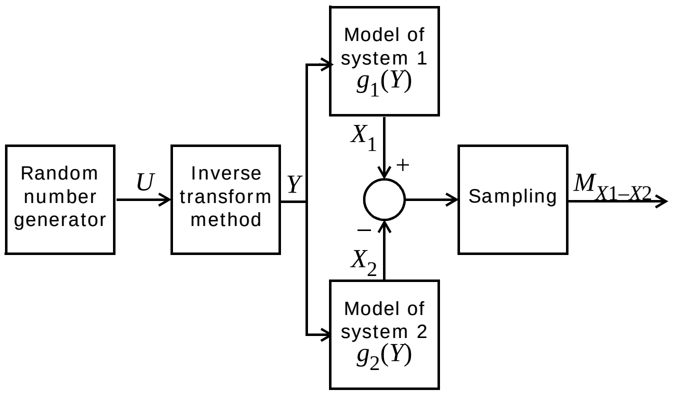
- If the probability distributions for the inputs to the models g_1 and g_2 are the same, then we can use the same scenarios y_i, $i = 1, \ldots, n$, for both models:

$$m_{X1} = \frac{1}{n}\sum_{i=1}^{n} g_1(y_i) \text{ and } m_{X2} = \frac{1}{n}\sum_{i=1}^{n} g_2(y_i).$$

If the models have some common properties then we would have $Cov[M_{X1}, M_{X2}] > 0$.

KTH Electrical Engineering

SIMULATION PROCEDURE



KTH Electrical Engineering

SIMULATION PROCEDURE

- Step 1. Generate the first batch of scenarios, y_i, $i = 1, \ldots, n_b$.
- Step 2. Calculate $x_{1,i} = g_1(y_i)$ and $x_{2,i} = g_2(y_i)$, $i = 1, \ldots, n_b$.
- Step 3. Calculate the sums $\sum_{i=1}^{n}(x_{1,i} - x_{2,i})$ and $\sum_{i=1}^{n}(x_{1,i} - x_{2,i})^2$ or store all samples $x_i - z_i$ (depending on the objective of the simulation).

KTH Electrical Engineering

SIMULATION PROCEDURE

- Step 4. Test stopping rule. If not fulfilled, repeat step 1 and 3 for the next batch.
- Step 5. Calculate estimates and present results.

KTH Electrical Engineering

CORRELATED SAMPLING

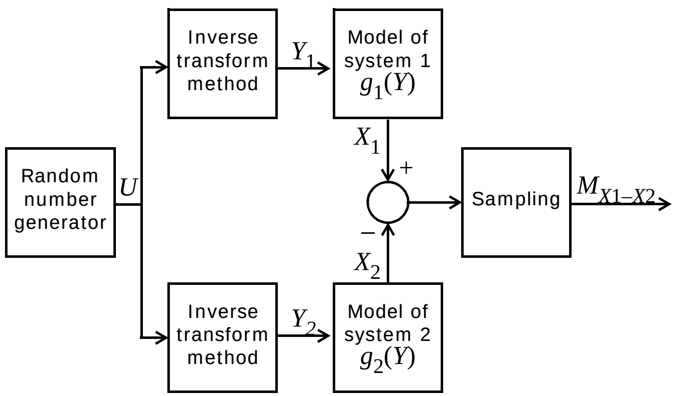
- If the models have the same inputs but different probability distributions then we can try to at least use the same values u_i, $i = 1, \ldots, n$, from the random number generator:

$$m_{X1} = \frac{1}{n}\sum_{i=1}^{n} g_1(F_{Y1}^{-1}(u_i)) \text{ and}$$

$$m_{X2} = \frac{1}{n}\sum_{i=1}^{n} g_2(F_{Y2}^{-1}(u_i)).$$

KTH Electrical Engineering

SIMULATION PROCEDURE



KTH Electrical Engineering

EXAMPLE 24 - Effectiveness of ICC simulation

- The ice-cream company is considering to introduce a new flavour: licorice ice-cream.
- The company does not expect any larger demand for this flavour: the probability that 100 litres will be ordered is 25%, the probability that 50 litres will be ordered is 25% and the probability of not receiving any orders at all is 50%.

EXAMPLE 24 - Effectiveness of ICC simulation

- It could be worthwhile to invest in a new machine if the new flavour is introduced, because otherwise it will not be possible to manufacture all flavours during the same day.

Compare the true changes in the expectation values to the results from simple sampling (using 200 scenarios per set-up) and correlated sampling (using the same 200 scenarios for each set-up).

EXAMPLE 24 - Effectiveness of ICC simulation

Table 9 Results of ICC simulation in example 24.

Simulation method	Expected increase in net income [€/day]						Average simulation time [h:min:s]
	New flavour			New flavour & new machine			
	Min	Av.	Max	Min	Av.	Max	
Enumeration		+51			+60		4:10:04
Simple samp.	+3	+53	+127	+23	+62	+92	0:07:13
Correlated sampling	+40	+51	+61	+48	+59	+72	0:07:12

EXAMPLE 24 - Effectiveness of ICC simulation

Table 10 Results of ICC simulation in example 24.

Simulation method	Increased risk for missed delivery [missed deliveries/100 days]						Average simulation time [h:min:s]
	New flavour			New flavour & new machine			
	Min	Av.	Max	Min	Av.	Max	
Enumeration		+1.38			–0.29		4:10:04
Simple samp.	–0.50	+1.10	+3.00	–1.00	–0.20	+1.00	0:07:13
Correlated sampling	+0.50	+1.50	+4.00	–1.00	–0.27	±0	0:07:12

VARIANCE REDUCTION

- To achieve a variance reduction, we need to have some information of the simulated system.
- For control variates, we need to create and analyse a simplified model of the system.
- Correlated sampling only requires that we know that we are comparing two slightly different versions of the same system.