



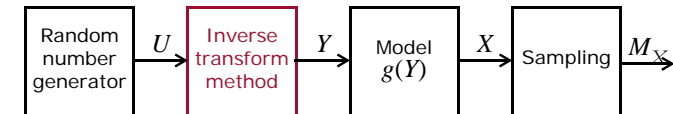
Theme 4

COMPLEMENTARY RANDOM NUMBERS & DAGGER SAMPLING

1

INTRODUCTION

- All observations in simple sampling are independent of each other.
- Sometimes it is possible to increase the probability of a good selection of samples if the samples are not independent.



2

MEAN OF TWO ESTIMATES

- Consider two estimates M_{X1} and M_{X2} of the same expectation value μ_X , i.e.,

$$E[M_{X1}] = E[M_{X2}] = \mu_X.$$

- Study the mean of these estimates, i.e.,

$$M_X = \frac{M_{X1} + M_{X2}}{2}.$$



3

MEAN OF TWO ESTIMATES

- Expectation value:

$$\begin{aligned} E[M_X] &= E\left[\frac{M_{X1} + M_{X2}}{2}\right] = \\ &= \frac{1}{2}(E[M_{X1}] + E[M_{X2}]) = \frac{1}{2}(\mu_X + \mu_X) \\ &= \mu_X, \end{aligned} \tag{5}$$

i.e., the combined estimate M_X is also an estimate of μ_X .



4

MEAN OF TWO ESTIMATES

- Variance:

$$\begin{aligned} \text{Var}[M_X] &= \text{Var}\left[\frac{M_{X1} + M_{X2}}{2}\right] = \\ &= \frac{1}{4}\text{Var}[M_{X1}] + \frac{1}{4}\text{Var}[M_{X2}] + \\ &\quad + \frac{1}{4} \cdot 2\text{Cov}(M_{X1}, M_{X2}). \end{aligned} \quad (6)$$



KTH Electrical Engineering

5

MEAN OF TWO ESTIMATES

- If M_{X1} and M_{X2} both are independent estimates obtained with simple sampling and the number of samples is the same in both simulations, then we get $\text{Var}[M_{X1}] = \text{Var}[M_{X2}]$ and $\text{Cov}[M_{X1}, M_{X2}] = 0$. Hence, (6) yields

$$\begin{aligned} \text{Var}[M_X] &= \text{Var}\left[\frac{M_{X1} + M_{X2}}{2}\right] = \\ &= \frac{1}{4}(\text{Var}[M_{X1}] + \text{Var}[M_{X1}]) = \frac{\text{Var}[M_{X1}]}{2}. \end{aligned} \quad (7)$$



KTH Electrical Engineering

6

MEAN OF TWO ESTIMATES

- The variance obtained in (7) is equivalent to one run of simple sampling using $2n$ samples; according to theorem 11 we should get half the variance when the number of samples is doubled.
- However, if the M_{X1} and M_{X2} are not independent but negatively correlated then the covariance term in (6) will make $\text{Var}[M_X]$ smaller than the corresponding variance for simple sampling using the same number of samples.



KTH Electrical Engineering

7

COMPLEMENTARY RANDOM NUMBERS

- A negative correlation between estimates can be created using complementary random numbers.
- The complementary random number of a $U(0, 1)$ -distributed random variable is $U^* = 1 - U$.
 - Correlation coefficient: $\rho_{U, U^*} = -1$.



KTH Electrical Engineering

8

COMPLEMENTARY RANDOM NUMBERS



KTH Electrical Engineering

- If a random value is calculated using the inverse transform method, i.e., $Y = F_Y^{-1}(U)$, then the complementary random number is given by $Y^* = F_Y^{-1}(1 - U)$.
 - The complementary random number is calculated in the same way also for normally distributed random numbers generated according to the approximative inverse transform method.
 - For symmetrical distributions we get that if $Y = \mu_Y + \delta$ then $Y^* = \mu_Y - \delta$.
 - Correlation coefficient: $\rho_{Y, Y^*} \geq \rho_{U, U^*}$.

9

COMPLEMENTARY RANDOM NUMBERS



KTH Electrical Engineering

- If a random value is calculated by analysing a scenario, i.e., $X = g(Y)$, then the complementary random number is given by $X^* = g(Y^*)$.
 - Correlation coefficient: $\rho_{X, X^*} \geq \rho_{Y, Y^*}$.

10

ESTIMATES USING COMPLEMENTARY RANDOM NUMBERS



KTH Electrical Engineering

- Let M_{X1} be an estimate based on x_1, \dots, x_n .
- Let M_{X2} be an estimate based on the corresponding complementary random numbers, x_1^*, \dots, x_n^* .
- Obviously M_{X1} and M_{X2} are negatively correlated.

11

ESTIMATES USING COMPLEMENTARY RANDOM NUMBERS



KTH Electrical Engineering

- In practice, there is no need to differentiate between the origin of the samples; the final estimate is simply

$$m_X = \frac{1}{2n} \sum_{i=1}^n (x_i + x_i^*).$$

12

MULTIPLE INPUTS



KTH Electrical Engineering

- Generally, the input vector Y has more than one element.
- The complementary random numbers of a vector can be created by combining the original and complementary random numbers of each element in the vector.
It can be noted that each scenario y_i will have more than one complementary scenario.
- It is possible to apply complementary random numbers only to a selection of the elements in Y .

13

EXAMPLE 16 - Complementary random vector



KTH Electrical Engineering

The input vector Y has the three elements A, B, C . The outcomes a_i, b_i, c_i have been randomised for each element, and the corresponding complementary random numbers, a_i^*, b_i^*, c_i^* , have been calculated.

Enumerate the original and complementary scenarios if complementary random numbers are applied to the following elements:

- A, B, C .
- A, B .

14

EXAMPLE 16 - Complementary random vector



KTH Electrical Engineering

Solution: a) Randomise three values, a, b, c and calculate their complementary random numbers. The results are combined as follows:

$[a, b, c], [a, b, c^*], [a, b^*, c], [a, b^*, c^*],$
 $[a^*, b, c], [a^*, b, c^*], [a^*, b^*, c], [a^*, b^*, c^*].$

b) Randomise six values, a, b, c_1, c_2, c_3, c_4 and calculate the random numbers a^* and b^* . The results are combined as follows:

$[a, b, c_1], [a, b^*, c_2], [a^*, b, c_3], [a^*, b^*, c_4].$

15

EFFECTIVENESS



KTH Electrical Engineering

- The variance reduction effect is depending on the correlation between $x_i = g(y_i)$ and $x_i^* = g(y_i^*)$. This correlation depends on the correlation between Y and X .
- All inputs may not have a significant correlation to the outputs \Rightarrow no use in applying complementary random numbers to these inputs.
- Sometimes a stronger correlation can be obtained by modifying the input probability distribution.

16

MODIFIED INPUTS



KTH Electrical Engineering

- It might be possible to find a subset of the inputs, Y_i , $i = 1, \dots, J$ such that there is a strong correlation between $Y_E = h(Y_1, \dots, Y_J)$ (where h is an arbitrary function) and X .
- When it can be worthwhile to introduce Y_E as an extra input variable to the system, and apply complementary random numbers to Y_E instead of Y_i , $i = 1, \dots, J$.

17

EXAMPLE 17 - Modifying discrete random variables



KTH Electrical Engineering

- In example 15, the net income is correlated to the demand.
- The correlation between the demand for an individual flavour and the net income is not as strong as the correlation between the total demand and the net income.
- Hence, we can introduce the input variable

$$D_{tot} = \sum_{f \in \mathcal{F}} D_f.$$

18

EXAMPLE 17 - Modifying discrete random variables



KTH Electrical Engineering

- We now have to find a joint probability distribution for D_{tot} and D_f , $f = 1, \dots, 6$.
- The demand for each flavour has three possible states. The total number of combinations is then $3^6 = 729$.
- Enumerate all possible states, calculate the probability for each state and sort the states according to increasing values of D_{tot} .

19

EXAMPLE 17 - Modifying discrete random variables



KTH Electrical Engineering

- The demand for each flavour can now be randomised with one value from the pseudo-random number generator, U .
- The complementary demand for each flavour is calculated using $U^* = 1 - U$.
- The total demand based on U will be negatively correlated to the total demand based on U^* .

20

EXAMPLE 18 - Modifying normal distributions



KTH Electrical Engineering

- Gismo Inc. is selling their products to K retailers, and the demand of each retailer, D_k , is $N(\mu_k, \sigma_k)$ -distributed.
- The costs of Gismo Inc. has a stronger correlation to the total demand compared to the demand of an individual retailer.
- Introduce the input variable $D_{tot} = D_1 + \dots + D_K$.

21

EXAMPLE 18 - Modifying normal distributions



KTH Electrical Engineering

- The probability distribution of D_{tot} is calculated using the addition theorems for normally distributed random variables.
- To generate the original scenario, randomise D_{tot} as well as temporary values of D_1, \dots, D_K . Then scale the individual retailer demands to match $D_{tot'}$ i.e., let

$$D_k = \frac{D_{tot}}{\sum_j D_j} D_k', \quad \text{where } D_k' \text{ is the temporary value.}$$

22

EXAMPLE 18 - Modifying normal distributions



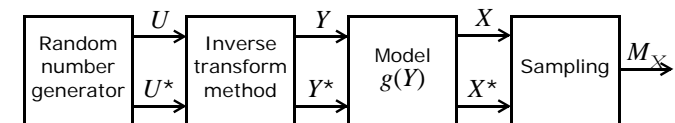
KTH Electrical Engineering

- To generate the complementary scenario, calculate the complementary random value of D_{tot} and randomise new temporary values of D_1, \dots, D_K . Then scale the individual retailer demands to match D_{tot}^* .

23

SIMULATION PROCEDURE

- Complementary random numbers



KTH Electrical Engineering

- Step 1.** Generate a random number $u_{k,i}$ and calculate $y_i = F_{Yk}^{-1}(u_{i,k})$. If applicable to this scenario parameter, also calculate $y_i^* = F_{Yk}^{-1}(1 - u_{i,k})$.
- Step 2.** Repeat step 1 for all scenario parameters. Create all complementary scenarios and calculate $x_i = g(y_i)$.

24

SIMULATION PROCEDURE

- Complementary random numbers

- **Step 3.** Repeat step 1 and 2 until enough scenarios have been analysed for this batch.
- **Step 4.** Update the sums

$$\sum_{i=1}^n x_i \text{ and } \sum_{i=1}^n x_i^2$$

or store all samples x_i (depending on the objective of the simulation).



KTH Electrical Engineering

25

SIMULATION PROCEDURE

- Complementary random numbers

- **Step 5.** Test stopping rule. If not fulfilled, repeat step 1–4 for the next batch.
- **Step 6.** Calculate estimates and present results.



KTH Electrical Engineering

26

EXAMPLE 19 - Effectiveness of ICC simulation

Consider the same system as in example 15. Compare the true expectation values to the probability distribution of estimates from a complementary random number simulation using 200 samples.



KTH Electrical Engineering

27

EXAMPLE 19 - Effectiveness of ICC simulation

Table 6 Results of ICC simulation in example 19.

| 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 sampling | 829 | 872 | 909 | 0.00 | 0.27 | 1.00 | 0:02:25 |
| Complementary random numbers | 842 | 875 | 914 | 0.00 | 0.27 | 1.00 | 0:02:25 |



KTH Electrical Engineering

28

DAGGER SAMPLING



KTH Electrical Engineering

- We have seen that a negative correlation between the outputs from different scenarios can have a variance reducing effect.
- When the input is a two-state probability distribution, we can create a negative correlation by using **dagger sampling** instead of complementary random numbers.

29

DAGGER SAMPLING



KTH Electrical Engineering

- Consider a two-state input variable Y with the frequency function

$$f_Y(x) = \begin{cases} 1-p & x = \text{A}, \\ p & x = \text{B}, \\ 0 & \text{all other } x, \end{cases} \quad (8)$$

where $p < 0.5$.

- In dagger sampling, the inverse transform method is replaced by a dagger transform function.

30

DAGGER TRANSFORM



KTH Electrical Engineering

Definition 13: Assume that Y is distributed according to (8) and let S be the largest integer such that $S \leq 1/p$.^{*} Then, Y_i , $i = 1, \dots, S$, is randomised according to

$$F_{Y_i}^\dagger(U) = \begin{cases} \text{B} & \text{if } (i-1)p \leq U < ip, \\ \text{A} & \text{otherwise,} \end{cases}$$

where U is a random value from a $U(0, 1)$ -distribution.

^{*} The value of S is referred to as the *dagger cycle length*.

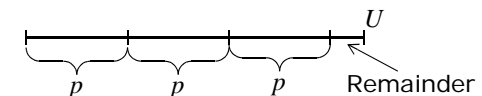
31

DAGGER TRANSFORM



KTH Electrical Engineering

- The dagger transform can be illustrated by dividing an interval between 0 and 1 in S sections, where each section has the width p .
- There will also be a remainder interval if $S \cdot p < 1$.



- Each value of U generates S values of Y .
- $y_k = \text{B}$ if U is found in the k :th section, and all other $y_i = \text{A}$, $i = 1, \dots, S$, $i \neq k$.

32

EXAMPLE 20 - Dagger transform

The random variable Y is either equal to 2 (70% probability) or 5 (30% probability).

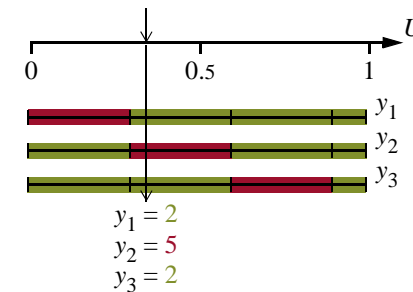
a) Randomise three values of Y using dagger sampling and the random value $U = 0.34$ from a $U(0, 1)$ -distribution.

b) Randomise three values of Y using dagger sampling and the random value $U = 0.94$ from a $U(0, 1)$ -distribution.

EXAMPLE 20 - Dagger transform

Solution: Here $p = 0.3 \Rightarrow S = 3$. Moreover, we have $A = 2$ and $B = 5$.

a)



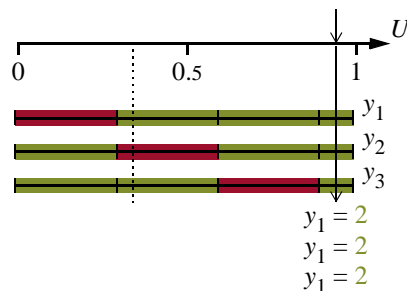
33

34

EXAMPLE 20 - Dagger transform

Solution (cont.)

b)



35

CORRELATION OF DAGGER SAMPLES

- The outputs X_1, \dots, X_S from the scenarios generated in a dagger cycle will be negatively correlated if the inputs Y_1, \dots, Y_S are negatively correlated (but the correlation might be weaker).
- Are the input values $Y_i = F_{Y_i}^\dagger(U)$, $i = 1, \dots, S$ negatively correlated, i.e., what is the value of

$$\text{Cov}[Y_i, Y_j] = E[Y_i Y_j] - E[Y_i]E[Y_j]?$$

36

CORRELATION OF DAGGER SAMPLES



KTH Electrical Engineering

- There are two possible values for the product $Y_i Y_j$: AB or AA . There are two sections there either Y_i or Y_j will be transformed to B ; hence, the probability for this event is $2p$. Thus, we get

$$E[Y_i Y_j] = 2pAB + (1 - 2p)AA.$$



37

CORRELATION OF DAGGER SAMPLES



KTH Electrical Engineering

- Moreover, we have

$$E[Y_i] = E[Y_j] = (1 - p)A + pB.$$

- The covariance can now be calculated as

$$\begin{aligned} \text{Cov}[Y_i, Y_j] &= E[Y_i Y_j] - E[Y_i]E[Y_j] = \\ &= 2pAB + (1 - 2p)AA - ((1 - p)A + pB)^2 = \\ &= \dots = -p^2(A + B)^2 < 0. \end{aligned}$$

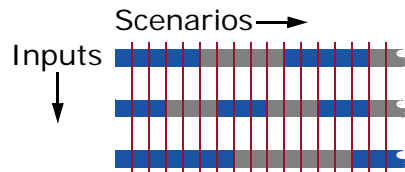
- The input values are negatively correlated and we can therefore expect a negative correlation in the outputs too.

38

MULTIPLE INPUTS

Assume that a computer simulation has several two-state input variables. Which dagger cycle length should be chosen?

- Alternative 1:** Independent dagger cycles.

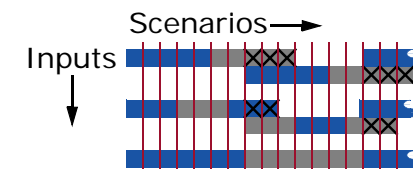


KTH Electrical Engineering

39

MULTIPLE INPUTS

- Alternative 2:** Reset all dagger cycles at the end of the longest cycle.

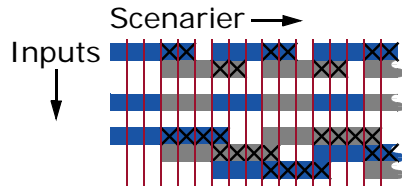


KTH Electrical Engineering

40

MULTIPLE INPUTS

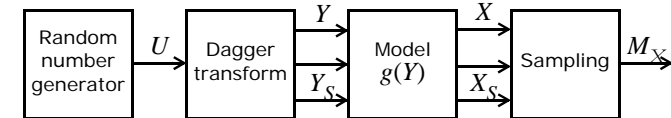
- **Alternative 3:** Reset all dagger cycles at the end of the shortest cycle.



41

SIMULATION PROCEDURE

- Dagger sampling



- **Step 1.** Generate a random number $u_{k,i}$ and calculate $y_i = F_{Yi}^\dagger(u_{i,k})$ for one dagger cycle.
- **Step 2.** Repeat step 1 for all scenario parameters. Combine the dagger cycles and calculate $x_i = g(y_i)$.

42

SIMULATION PROCEDURE

- Dagger sampling

- **Step 3.** Repeat step 1 and 2 until enough scenarios have been analysed for this batch.
- **Step 4.** Update the sums

$$\sum_{i=1}^n x_i \text{ and } \sum_{i=1}^n x_i^2$$

or store all samples x_i (depending on the objective of the simulation).



43

SIMULATION PROCEDURE

- Dagger sampling

- **Step 5.** Test stopping rule. If not fulfilled, repeat step 1–4 for the next batch.
- **Step 6.** Calculate estimates and present results.



44

EXAMPLE 21 - Effectiveness of ICC simulation

Consider the same system as in example 15. Compare the true expectation values to the probability distribution of estimates from a dagger sampling simulation using 200 samples.



KTH Electrical Engineering

45

EXAMPLE 21 - Effectiveness of ICC simulation

Table 7 Results of ICC simulation in example 21.

| 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 sampling | 856 | 877 | 900 | 0.00 | 0.27 | 1.00 | 0:02:27 |



KTH Electrical Engineering

46

VARIANCE REDUCTION

- To achieve a variance reduction, we need to have some information of the simulated system.
- For **complementary random numbers**, we need to identify the inputs where a negative correlation between scenarios will result in negative correlations for one or more outputs.
- For **dagger sampling**, we need to identify two-state inputs where a negative correlation between the scenarios will result in negative correlations for one or more outputs.



KTH Electrical Engineering

47