UNIVERSITEIT•STELLENBOSCH•UNIVERSITY
jou kennisvennoot • your knowledge partner

Simulation Part 1 / Simulasie Deel 1

LSCM344 - August 2022
Lecturer - HW Freiboth

Slides adapted from those originally compiled for the module by Dr Neil Jacobs (2019)

Toustaanteorie voorwaarde / Queueing theory condition

Tussenaankomstye en dienstye **moet** eksponensieel verdeel wees.

Indien die voorwaarde nie aan voldoen kan word nie kan simulasie van hulp wees.

Interarrival times and service times **must** be exponential distributed.

If the condition can not be met simulation can be of assistance.

Wat is simulasie? / What is simulation?

Definisie

Simulasie is 'n tegniek wat 'n regte-wêreld stelsel se verandering met die verloop van tyd naboots. Dit poog om die kenmerke, voorkoms en eienskappe van die regte wêreld stelsel te dupliseer.

Definition

Simulation is a technique which imitates a real world system's change with the elapse of time. It attempts to duplicate the features, appearance and characteristics of the real world system.

Waarom simuleer? / Why simulate?

- Bestudering van die interaksies tussen elemente van 'n komplekse stelsel;
- Effek van veranderinge in inligting, organisasie en omgewing van 'n stelsel op stelselgedrag;
- Rigtinggewing vir verbeterings / opgraderings;
- Verifiëring van analitiese resultate;
- Voorbereiding vir moontlike uitkomste;
- Animasie help met visualisering;
- Opleiding.

- Study of the interactions of elements in a complex system;
- Effect of change in information, organization and environment of a system on system behaviour;
- Provide direction for enhancements / upgrades.
- Verification of analytical results.
- Preparation for possible outcomes.
- Animation helps with visualization.
- Training.

'n Hawebestuurder se vraag / A port manager's question

Gegewe tyd-rekords van aankomste- en inspeksie-tye van voertuie by die hek van die hawe - hoe gaan kongestie geraak word deur 'n opgradering van toerusting wat inspeksie-tye verkort.

Sien QueueSimEx1.xlsx

Given time records of arrivals and inspection times of vehicles at the gate of the port - how will congestion be affected by upgrading of equipment that shortens inspection times.

See QueueSimEx1.xlsx

Stappe van simulasie / Steps of simulation

1. Definieer die probleem.
2. Bepaal belangrike veranderlikes.
3. Bou die simulasie model.
4. Stel parameters vir die model om teen te toets.
5. Voer die simulasie uit.
6. Ontleed die resultate.
7. Besluit op 'n aksieplan.

1. Define the problem.
2. Determine the important variables.
3. Build the simulation model.
4. Set the parameters for the model to test against.
5. Execute the simulation.
6. Analyze the results.
7. Decide on a plan of action.

Voordele van simulasie / Advantages of simulation

- Relatief voor die hand liggend en aanpasbaar.
- Rekenaars vergemaklik modellering.
- Maak die ontleding van komplekse stelsels moontlik.
- Ontleed "wat as"-scenarios.
- Geen impak op die regte stelsel.
- Maak die studie van inter-aksie van stelsel komponente moontlik.
- Tyd-verkorting.
- Maak die insluiting van verdelings moontlik wat nie met analitiese metodes moontlik is

- Relatively straightforward and flexible.
- Computers simplify modelling.
- Enables the analysis of complex systems.
- Analyzes "what if"-scenarios.
- No impact on the real system.
- Enables the study of interaction of system components.
- Time-compression.
- Enables the inclusion of probability distributions which are not possible with analytical methods.

Uitdagings met simulasie / Challenges with simulation

- Goeie modelle van komplekse stelsels kan baie duur wees.
- Simulasie genereer nie optimale oplossings soos wat met sommige analitiese tegnieke gevind kan word nie.
- Toestande en beperkinge van scenarios moet deur besluitnemers gestel word vir implementering in 'n simulasie, wat moeilik kan wees weens verskeidenheid van kombinasies en permutasies van opsies.
- Modelle is geneig om uniek te wees en resultate derhalwe nie in die algemeen oordraagbaar na

- Good models of complex systems can be very expensive.
- Simulation does not generate optimal solutions as might be found with some analytical techniques.
- Conditions and constraints of scenarios must be set my decisionmakers for implementation in a simulation, which could be difficult due to combinations and permutations of options.
- Models tend to be unique and results therefore generally not transferreable to different

Monte Carlo simulasie / simulation

As 'n stelsel kansgedrag openbaar kan die Monte Carlo simulasie benadering moontlik toegepas word. Voorbeelde van kansveranderlikes in die regte wêreld wat gesimuleer kan word:

1. Vraag na voorraad per periode bv. per week.
2. Leityd vir bestellings om aan te kom.
3. Passasiers wat nie opdaag nie.
4. Tyd tussen aankomste by 'n diensfasiliteit.
5. Aantal drywers afwesig per skof.
6. Aantal houers gelaai per uur.

If a system contains random behaviour the Monte Carlo simulation approach can possibly be applied. Examples of random variables in the real world which can be simulated:

1. Inventory demand per time period eg. week.
2. Lead time for orders to arrive.
3. Passenger no-shows.
4. Times between arrivals at a service facility.
5. Number of drivers absent per shift.
6. Number of containers loaded per hour.

Diskrete kansveranderlikes vs Kontinue kansveranderlikes / Discrete random variables vs Continuous random variables.

Uit die voorgaande lys sal nommers 3, 5 en 6 altyd diskrete verdelings volg terwyl die ander meesal uit kontinue verdelings sal kom. Nommer 1 mag in sommige gevalle soos mielies uit 'n silo in kontinue veranderlike hoeveelhede aangevra kan word, teenoor sakke mielies wat eenheidsgewys benodig word.

Meer hieroor later.

From the previous list numbers 3, 5 and 6 will always follow discrete distributions while the others will mostly be from continuous distributions. Number 1 may in some instances be like maize from a silo which is demanded in continuous quantities versus bags of maize which are demanded unit wise.

More about this later.

Ses stappe van Monte Carlo simulasie / Six steps of Monte Carlo simulation

1. Bepaal waarskynlikheids-verdelings vir belangrike toevoer veranderlikes.
2. Bou 'n kumulatiewe waarskynlikheidsverdeling vir elke veranderlike in stap 1- as die KWV nie reeds bestaan nie.
3. Bepaal 'n interval vir kansgetalle vir elke kansveranderlike uitkoms, of in die geval van 'n kontinue veranderlike bepaal 'n inverse transform metode om die kansgetal in 'n kansveranderlike uitkoms te transformeer.

1. Establish probability distributions for important input variables.
2. Build a cumulative probability distribution for each variable in step 1 - if the CPD does not already exist.
3. Establish an interval for random numbers for each random variable outcome, or in the case of a continuous variable establish an inverse transformation method to translate the random number into a random variable outcome

Ses stappe van Monte Carlo simulasie / Six steps of Monte Carlo simulation

4. Genereer kansgetalle.
5. Genereer kansveranderlike uitkomste.
6. Simuleer 'n reeks steekproewe.

4. Generate random numbers.
5. Generate random variable outcomes.
6. Simulate a series of trials.

Monte Carlo simulasie / Monte Carlo simulation

Hierdie week sal ons fokus op stap 1: bepaal die waarskynlikheidsverdelings vir belangrike toevoerveranderlikes. Hiervoor sal ons die Chi-kwadraat passingstoets gebruik.

Daarna sal ons fokus op kansgetalle, voordat ons aanbeweeg na die gebruik van simulasie om toustaan- en voorraadprobleme te bestudeer.

This week we will focus on step 1: determining the probability distributions for important input variables. For this, we will use the Chi square *goodness-of-fit* test.

Thereafter we will focus on random numbers, before we move on to using simulation to study queueing and inventory problems.

Passingstoetse / Goodness-of-fit tests

- Toets of 'n versameling observasies 'n bepaalde verdeling volg.
- Lewer uitspraak oor 'n hipotese.
- Baseer die uitspraak op 'n statistiek uit die observasies.

Algemeen gebruikte passingstoetse is die Kolmogorov-Smirnov-toets and die χ^2-toets. Die K-S-toets word gebruik vir kontinue waarskynlikheidsverdelings terwyl die χ^2-toets vir beide kontinue en diskrete waarskynlikheids-verdelings gebruik kan word.

- Test whether a set of observations follows a certain distribution.
- Pronounces on a hypothesis
- Base the pronouncement on a statistic of the observations.

Commonly used GOF-tests are the Kolmogorov-Smirnov-test and χ^2-tests.The K-S-test is used for continuous probability distributions whereas the χ^2-test can be used for continuous and discrete probability distributions.

Essensie van passingstoetse / Essence of GOF-tests

Die aantal observasies in elke interval word getel en vergelyk met die verwagte aantal observasies in die intervalle en die tellings word verwerk tot 'n statistiek wat die mate van ooreenstemming weerspieël.

The number of observations in each interval is counted and compared with the number of expected observations in the intervals and the counts are translated to a statistic which reflects the extent of similarity.

Die χ^2-toets / The χ^2-test

1. Lys die data.
2. Bepaal die teoretiese verdeling.
3. Stel die hipotese op waarvoor getoets gaan word.
4. Bepaal die kritieke χ^2-waarde waarby die H_0-hipotese verwerp gaan word, noem dit $\chi^2_{k-r-1,\alpha}$.
5. Bepaal die χ^2-statistiek van die data, χ^2_{data}.
6. Vergelyk χ^2 met $\chi^2_{k-r-1,\alpha}$ en verwerp H_0 as $\chi^2 > \chi^2_{k-r-1,\alpha}$

1. List the data.
2. Determine the theoretical distribution.
3. State the hypothesis which will be tested.
4. Determine the critical χ^2-value at which the null hypothesis will be rejected, call it $\chi^2_{k-r-1,\alpha}$.
5. Determine the χ^2-statistic of the data, χ^2_{data}.
6. Compare χ^2 with $\chi^2_{k-r-1,\alpha}$ and reject H_0 if $\chi^2 > \chi^2_{k-r-1,\alpha}$

Die χ^2-verdeling / the χ^2-distribution

Die χ^2-verdeling verskaf die waarskynlikheid van die som van ν-vierkante uit 'n standaard normaal verdeling, waar ν na verwys word as die grade van vryheid van die verdeling, $\nu = k - r - 1$.

The χ^2-distribution provides the probability of the sum of ν-squares from a standard normal distribution, where ν is referred as the number of degrees of freedom of the distribution, $\nu = k - r - 1$.

Die gemiddeld van die χ^2-verdeling is gelyk aan die grade van vryheid van die verdeling ν en die variansie daarvan is gelyk aan 'n getal twee keer ν.

The mean of a χ^2-distribution is equal to the degrees of freedom parameter ν and its variance is equal to a number twice of ν.

Bepaling van $\chi^2_{k-r-1,\alpha}$ / Determining $\chi^2_{k-r-1,\alpha}$

- k is die aantal intervalle.
- r is die aantal parameters wat beraam is uit die data.
- α is die beduidendheidsvlak.

As die data gebruik word om 'n parameter te beraam soos $\hat{\lambda}$ vir die eksponensiële verdeling, dan is $r = 1$.

α dui die besluitnemer se toleransie vir risiko van 'n Tipe I fout. In die afwesigheid van 'n spesifieke voorskrif neem aan dat $\alpha = 0.05$.

- k is the number of intervals.
- r is the number of parameters estimated from the data.
- α is the level of significance.

If the data is used to estimate a parameter such as $\hat{\lambda}$ for the exponential distribution, then $r = 1$.

α reflects the decisionmaker's tolerance for risk of a Type I error. In the absence of a specific prescription assume $\alpha = 0.05$.

Bepaling van χ^2_{data} / Determining χ^2_{data}

$$\chi^2_{data} = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \tag{1}$$

waar k sodanig bepaal is dat $e_i \geq 5$.

where k is determined such that $e_i \geq 5$.

Daar is geen vaste reël waarvolgens k bepaal word nie dus kan Sturge se reël, waarmee die aantal intervalle in histogramme ook bepaal word, voorgestel as 'n vertrekpunt.

There is no fixed rule according to which k is determined, therefore Sturge's rule, also used to determine the bin in histograms, is proposed as a departure point.

$$k = \lceil log_2 n \rceil + 1 \tag{2}$$

waar n die aantal waarnemings is.

where n is the number of observations.

Parameter van die eksponensiële verdeling $\hat{\lambda}$ / Parameter of the exponential distribution $\hat{\lambda}$

Gegewe 'n reeks van n tussen aankomstye $\{t_1, t_2, \ldots, t_n\}$, dan is

Given a series of n interarrival times $\{t_1, t_2, \ldots, t_n\}$, then

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i} \tag{3}$$

Intervalle / Intervals

Laat u_i die bogrens wees vir elke interval $i < k$ dan is

Let u_i be the upper limit of each interval $i < k$ then

$$u_i = -\frac{1}{\hat{\lambda}} \ln \left(1 - \frac{i}{k} \right) \tag{4}$$

Kansgetal generering / Generating random numbers

Diskrete verdeling voorbeeld / Discrete distribution example

Gebruik Tabel 13.4 op bladsy 485 van Render en Stair, om 'n reeks van 20 waardes te simuleer van die aantal drywers wat op 'n dag afwesig is van die werk. Die aantal afwesig volg die diskrete verdeling in die onderstaande tabel.

Use Table 13.4 on page 485 of Render and Stair to simulate a series of 20 values of the number of drivers who are absent from work on a day. The number absent follows the discrete distribution in the table below.

Aantal afwesig Number absent	Waarskynlikheid Probability
0	0.5
1	0.25
2	0.1
3	0.1
4	0.05

Kansgetal generators / Random number generators

Simulasie verg kansgetalle wat $U[0,1]$-verdeel is.

Tabelle soos wat in die vorige voorbeeld gebruik is, is van beperkte nut wanneer groot hoeveelhede kansgetalle gegenereer moet word. Daarom bestaan daar 'n behoefte aan 'n metode wat kansgetalle kan genereer wat

1. rekenaarmatig vinnig is,
2. min rekenaar geheue verg,
3. identies nadoenbaar is,
4. genoegsaam verspreid is en
5. 'n lang siklus het.

Simulation requires random numbers which are $U[0,1]$ distributed.

Tables as used in the previous example is of limited use when large numbers of random numbers are to be generated. Therefore a need exist for a method which can generate random numbers that are

1. computationally fast,
2. requires little computer memory,
3. is sufficiently spread out,
4. is identically replicable and
5. has a long cycle.

Lineêr kongruensiële kansgetal generator / Linear congruential random number generator

Die Lineêr Kongruensiële Kansgetal Generator-metode, of LKG, is 'n metode wat, **as dit korrek toegepas word**, aan die vereistes voldoen.

Elke kansgetal R_i, $i \leq n$, in die kansgetal reeks van n kansgetalle, word bereken met die formule:

Linear Congruential Random Number Generator-method, or LCG, is a method which, **if it is applied correctly**, fulfills the requirements.

Every random number R_i, $i \leq n$, in the random number series of n random numbers , are calculated with the formula:

$$R_i = \frac{x_i}{m} \tag{5}$$

waar / where

$$\tag{6}$$

$$x_i = ax_{i-1} + c \quad \text{mod } m. \tag{7}$$

Lineêr kongruensiële kansgetal generator / Linear congruential random number generator

Komponente van

$$x_i = ax_{i-1} + c \mod m. \tag{8}$$

word soos volg genoem

- a is die konstante vermenigvuldiger,
- c is die inkrement,
- m is die modulus,
- x_0 is die saadgetal.

Components of

are called the following

- a is the constant multiplier,
- c is the increment,
- m is the modulus,
- x_0 is seed number.

Goeie parameters vir die LKG / Good parameters for the LCG

$$a = 7^5 \tag{9}$$
$$m = 2^{31} - 1 \tag{10}$$

dus / thus

$$\tag{11}$$
$$x_i = 7^5 x_{i-1} + c \mod (2^{31} - 1). \tag{12}$$

werk goed, met x_0 en c deur die modelleerder gespesifiseer.

works well, with x_0 and c specified by the modeller.

Kontinue waarskynlikheidsverdelings van belang / Continuous probability distributions of importance

Daar is drie kontinue waarskynlikheidsverdelings van belang in hierdie kursus. Hoewel die drie baie gebruik word in die praktyk is daar baie meer waarskynlikheidsverdelings om van te kies in modellering. Die drie is

- die uniforme verdeling,
- die eksponensiële verdeling en
- die normaal verdeling.

There are three continuous probability distributions of importance in this course. Although these three are used a lot in practice there are many more probability distributions to choose from in practice. The three are

- the uniform distribution,
- the exponential distribution and
- the normal distribution.

Uniforme verdeling / Uniform distribution

$$t \sim \mathrm{Uniform}\,(a, b)$$

Waarskynlikheidsdigtheidsfunksie:

Probability density function:

$$f(t) = \begin{cases} 0 & t < a, t > b \\ \frac{1}{b-a} & a \leq t \leq b \end{cases}$$

Kumulatiewe waarskynlikheidsdigtheidfunksie:

Cumulative probability density function:

$$F(t) = \begin{cases} 0 & t < a \\ \frac{x-a}{b-a} & a \leq t \leq b \\ 1 & t > b \end{cases}$$

Inverse transform:

Inverse transform:

$$t = a + (b-a)R_i$$

Eksponensiële verdeling / Exponential distribution

$$t \sim \mathrm{Expon}\,(\lambda)$$

Waarskynlikheidsdigtheidsfunksie:

Probability density function:

$$f(t) = \begin{cases} 0 & t < 0 \\ \lambda \mathrm{e}^{-\lambda x} & t \geq 0, \lambda > 0 \end{cases}$$

Kumulatiewe waarskynlikheidsdigtheidfunksie:

Cumulative probability density function:

$$F(t) = \begin{cases} 0 & t < 0 \\ 1 - \mathrm{e}^{-\lambda t} & t \geq 0, \lambda > 0 \end{cases}$$

Inverse transform:

Inverse transform:

$$t_i = -\frac{1}{\lambda} \ln\left(1 - R_i\right)$$

Normaal verdeling / Normal distribution

$$t \sim \mathrm{Normal}[\mu, \sigma]$$

Waarskynlikheidsdigtheidsfunksie:

Probability density function:

$$f(t) = \mathrm{e}^{\frac{-(t-\mu)^2}{2\sigma^2}}$$

Gebruik die Excel funksie "=norm.inv(R_i, μ, σ)" om kansveranderlike uitkomste te verkry.

Use the Excel function "=norm.inv(R_i, μ, σ)" to find random variable outcomes.