BAN403

Simulation of Business Processes

Assignment 1
14.02.23 - 10.03.23

NHH

Candidate Numbers:

25, 87, 96

Task 1

a) To replicate the results presented in the paper, we first need to understand the results produced. The paper consists of a simulation design, data analysis, and a conclusion. If we manage to replicate the simulation design successfully, we are then able to reproduce the data analysis used, thereby achieving a similar conclusion. Using R, we will now attempt to replicate the simulation design, commenting on assumptions made along the way.

Simulation design

The first steps are to simulate the class size, response rate and random numbers presenting the evaluation score. According to Step 3 the response rate is simulated randomly, here an assumption is made that the response rate and class size will follow the normal distribution of figure 1 in the paper (He & Freeman, 2020). Another assumption is made on what basis the set of random numbers are generated. We first classified the simulated classes using the boundaries for class and response level. Each classification was then split into the corresponding subsets using a ratio based on class count in table 3 in the paper (He & Freeman, 2020). For each class every student was given a random number, which followed the normal distribution from the given subsets parameters and was rounded to integer numbers with 1-5 as the ceiling effects. A random sample for each class was then pulled using the response rate. In step 4 we encounter the next challenge when calculating the 80% confidence intervals of the relevant statistic from table. Since we do not have data on the exact class size of the historical data, we do not know the sample size to calculate the confidence interval. We will therefore proceed with the assumption that the confidence interval test is performed incorrectly, therefore using a more relaxed test with count of classes as the sample size. From *Table 1* we observe that the simulation fails to replicate the kurtosis of the historical dataset and the skewness to some extent.

Testing criteria	Successful cases
Initial simulation	16 387
Evaluation score	12 048
Evaluation score, SD	9 143
Evaluation score, SD, Skewness	3 185
Evaluation score, SD, Skewness, Kurtosis	34

Table 1- Simulation test criteria

Since we do not have historical data for the individual classes, we cannot replicate the regression or the scatter plots from the data analysis part of the paper. We can however make a regression for the simulated classes to see if we pick up similar trends. From *Table 2* we see that both Response rate and Class size are significant predictors of for the evaluation score, but the response rate has a greater impact. This is the same trend and conclusion as we see in the paper.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.9606	0.0207	190.537	~ 0
Response rate	0.2408	0.0270	8.912	~ 0
Class size	-0.0015	0.0003	-3.831	0.00012

Table 2 - Regression on simulated classes

Reasoning for results

Since we had to make several assumptions and alterations, in addition to only having a small margin of the simulations passing all the tests, can we conclude that it is not possible to replicate the results from the paper with the given information. The main reason for the simulation being unsuccessful is insufficient data factors. This is due to the simulations only being based on a single university, in addition to not providing the raw class data. If the raw class data was provided, it would be possible to perform correct confidence interval testing and base the simulations on more sufficient data. It would also enable us to get a better understanding and incorporation of the kurtosis variable in the simulation. Lastly having the historical class data would allow deviation from the simulation recipe to alternative methods like machine- and deep learning. Simulations are inherently subject to random variation. Therefore, replicating the simulation successfully would likely indicate data mining, i.e., the model fits a specific data set rather than reality. The model would probably have low accuracy with a different data set of teaching evaluations.

b) The paper provides valuable insights into teaching evaluations, but it needs to be adjusted to fit the characteristics of NHH. One significant difference is the class size. NHH has many bachelor classes with up to 550 students, there is then a big jump in class size down to the smaller master classes. Another difference is that NHH has a much more homogeneous population compared to the public university used in the paper. NHH being a pure economics school while the sample university most likely has a great variety of educational programs. In addition to NHH having a much smaller sample size with only 3600 students (NHH, 2023).

To make the simulation applicable to NHH, variables such as class sizes, response rates, and evaluation scores need to be adjusted to reflect NHH's context. NHH's current course evaluations are more comprehensive compared to that of the paper. The current evaluations use multiple evaluation questions to provide a more comprehensive view of teaching performance, this adds more dimensions to the case and require a more complex analysis. NHH should also consider other factors when assessing teaching quality, such as student engagement, faculty development programs, and institutional support for teaching in order to capture an evaluation of the school as a whole.

Regardless of all these differences, the essence and conclusion of the paper still mostly applies to NHH's case. As the paper states "Large class sizes help to achieve accurate teaching evaluations" (He & Freeman, 2020). The class size and a consensus amongst the more homogeneous students on courses, might allow NHH to have response rates even lower that the relaxed response rate of 50%, while still maintaining accurate teaching evaluations. For the smaller master and elective classes however, we would recommend NHH following the conclusion of the paper by focusing on maintaining high response rates to achieve accurate teaching evaluations. This can be achieved by implementing benefits for completing course evaluations. To conclude, the simple answer to the question is yes, it is applicable, but to a limited extent and with several adjustments,

Task 2

Question 1: The system only has the possibility to support 3 cars making the different states look like Figure 1**.** Since the system also has to be in balance and there are 8 different states the steady-state probability is derived as.

$50x_1 = 40x_4$	$50x_4 + 40x_4 = 40x_7 + 30x_2$	$50x_7 + 40x_7 = 30x_5$
$50x_2 + 30x_2 = 50x_1 + 40x_5$	$50x_5 + 40x_5 + 30x_5 = 50x_4 + 40x_8 + 30x_3$	$40x_8 = 50x_7 + 30x_6$
$30x_3 = 50x_2 + 40x_6$	$40x_6 + 30x_6 = 50x_5$	$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1$

Figure 2 - Steady state equations

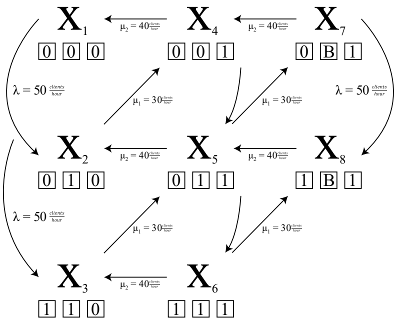

Figure 1 - Basis of equations

This gives us the results for each steady-state probability:

$$x_1 = 5\%, x_2 = 11\%, x_3 = 32\%, x_4 = 6\%, x_5 = 15\%, x_6 = 11\%, x_7 = 5\%, x_8 = 15\%$$

Question 2: The average number of clients in the system is found by using the formula for finding the expected number of clients in the system. Throughput (λ) is given as the average inflow which is equal to the average outflow ($\lambda = R_i = R_o$).

$$L = \sum_{n=0}^{\infty} nP_n = 1{,}585$$

But since the given average inflow is higher than the possible process rate of the microphone station and pickup station the effective average λ is considerably lower. The microphone station processes average $60 \div 30 = 2$ minutes every order and the pickup station delivers the food with an average $60 \div 40 = 1.5$ minutes for every order. The average rate of processing orders in the system is $60 \div (2 + 1.5) = 17.143$ orders per hour. Using Little's law, $L = \lambda \cdot W$, we can calculate the expected time spent in the system and in the queue. Solving for W, we get $W = L \div \lambda = 0.0925$ hours. Converting to minutes, we get $0.0925 \cdot 60 = 5.547$ minutes as the expected time that a customer spends in the system.

Question 3: Whether or not the system is stable is determined by the ratio between the flowrate and the processing rate: $\rho = \lambda \div \mu$. This ratio is defined as ρ, with a ratio higher than 1 describing an unstable system. The system in which we are describing in this task has a bottleneck of $\mu = 30$ which gives a value of $\rho = 1.6\overline{6}$. In theory, this is an unstable system if we do not consider the fact that the queue to the order system has a limit of 1 which disposes of additional potential customers. When taking this threshold into account the system is considered stable as the queue cannot grow beyond 1 under any circumstance (Laguna & Marklund, 2013), even with a positive ρ.

Question 4: From the simulation we find that the system is stable since the *MicrophoneQueue* does not exceed the capacity of 1, therefore not congesting the queueing system. The simulation has been simulated for 400 hours. During the time 8304 clients has been processed by the system. Making the $\lambda = 8304 \div 400 = 20{,}74$ (clients per hour) which is higher than what was calculated before. Reasons for this is the random seeds given has in 400 hours been faster, and in more processed time simulated it will be tangent towards what has been calculated. For the time processed the simulation shows that a client spends on average $0{,}089 * 60 = 5{,}34$ minutes to go through the drive-through. The difference here can also be explained as the throughput. Lastly for the average clients in the system. Here it is expected a higher rate than calculated over. Which we find by using Little's law $20{,}74 * 0{,}089 = 1{,}846$.

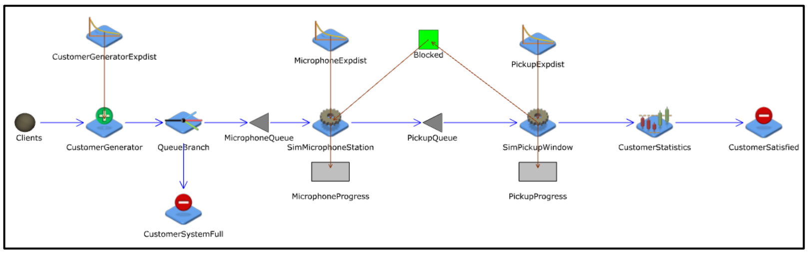

Figure 3- JaamSim drive-through model, which can be found in "Assig1-task2-4.cfg"

Question 5: Whilst changing the service time distribution from an exponential to a normal distribution is possible, it changes the assumptions about how the system behaves. By changing the expected time interval to a normal distribution, you in turn assume that the service time has a fixed mean and standard deviation, rather than being dependent on the rate of arrival of events. In addition to this, the normal distribution is defined for all real numbers, which is not a realistic time interval for a given food preparation time as the service time cannot be negative. This leads us to conclude that the food preparation time is unable to follow a true normal distribution.

Task 3

Vaccination Center Simulation

a) In this assignment the main focus is building a discrete event simulation model of a vaccination center and using the resulting insights in order to inform Bergen's decision making when deciding how to vaccinate their population in an epidemic scenario. The goal of the model is to create an accurate representation of the flow of people in a vaccination center during a pandemic scenario, in order to get a sense of the staffing, time frame and capacity of vaccinating a given population.

Building the model

We used the data from two papers (Aaby et al., 2005 & 2006) that conducted a physical exercise to measure the processing times of different center layouts. We refer to these papers as the source papers. We also adopted a similar center design as the source papers to avoid making categorical assumptions about the layout effects. We chose the data from the physical exercise over the CDC regulations because they are more relevant for a governing body during a pandemic. We will state any assumptions as they arise in our model description.

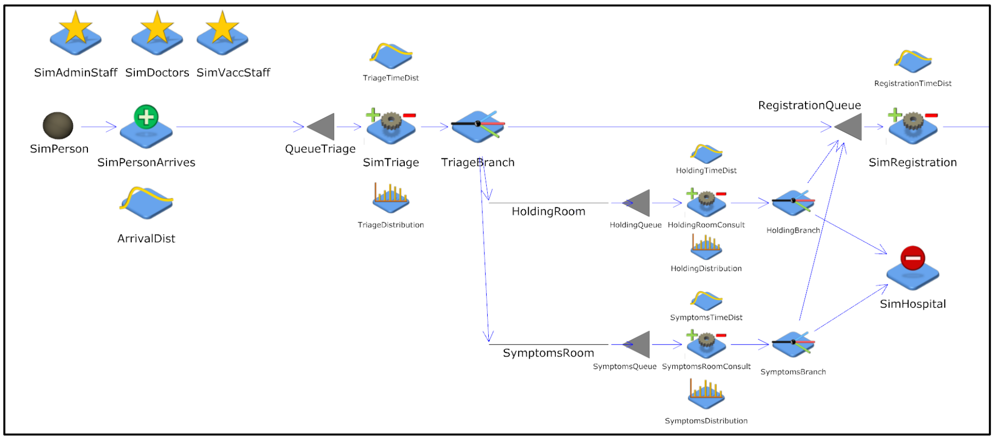

Figure 4 - Part I of the JaamSim model, which can be found in "Assig1-task3.cfg"

One factor which is important to capture in this model is the arrival rate of the people that need to be vaccinated. We have assumed that the government has created a ranked system of when to tell people to come into the center and that the center won't accept drop-in vaccinations. This is because the papers emphasized that when the center is close to full capacity variability in arrival is a key cause of congestion. On the other hand, it is unrealistic to assume constant inter arrival times, which led to us to assume that the inter arrival time is gamma distributed with a scale factor of 4 (SCV = 0.25) and a mean time between events of $60 \div 250 = 0.24$ as we assume the flow of people to be 250 per hour. Under this constant arrival scheme, some people will remain in the system when the simulation ends. We assume that these people are not admitted to the clinic unless they can be vaccinated before closing time. This assumption does not affect the outputs of the model, such as the mean time in the system and the number of people processed.

We have divided the staff of the center into three competence areas: Administration, Vaccination and Doctors. Various servers in the model make use of these resources and the parallel capacity of the various servers are based upon the baseline clinic staffing in the source papers. Every server or station has a queue and a gamma-distributed time interval distribution which represents the time it takes to finish a mandatory task given a person's routing. This distribution is most often used as a model for service time. All the gamma-distributions are assumed to have a scale factor of 1 which makes them equivalent to an exponential distribution. The mean of these distributions is gathered from the exercise measurements from the source papers. We assume that these figures are representative of the simulation model we have created. The specific mean processing times we have used as our framework can be found in Table 3.

As this center includes divergent person pathing based on examinations and symptoms, we have included branching in this model to represent this. Any given path is determined by the measured routing probability in the exercise in the source papers with a discrete probability distribution. This is a simplification in the simulation model as the routing probabilities will most likely be different based upon the severity of the pandemic and other factors which could feasibly increase variation in the model.

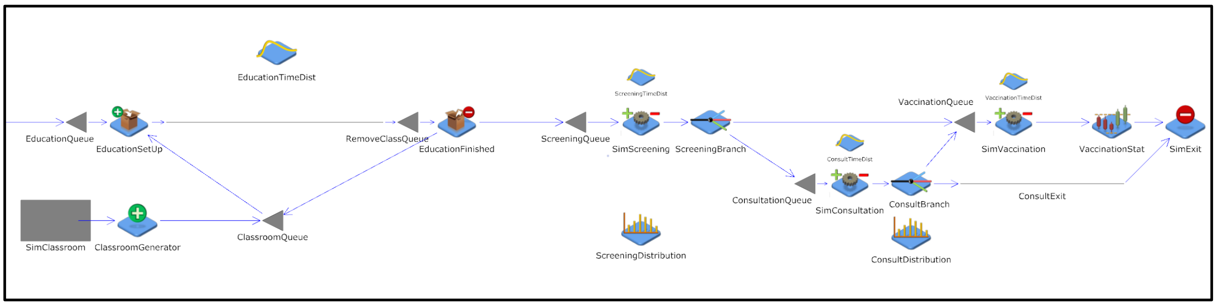

Figure 5 - Part II of the JaamSim model, which can be found in "Assig1-task3.cfg"

Every person has to either exit the center or go through registration when arriving. After registration there is an education server which has some special properties. There is a set number of available classrooms which must be filled up before starting an educational video. According to the source papers, determining the classroom size is not significant in predicting average time in the clinic, so we assume that it is 30. We assume that filling up and emptying a classroom takes 30 seconds, which is represented in the *EducationSetUp* and *EducationFinished* objects in the simulation. A classroom must be emptied before a new class can come into the room. This is the staff distributions and mean processing times in this model: (Aaby, et. al, 2005)

Station	Staff type	Number of staff	Mean processing times	Station Capacity	Percentage of Patient Served	Constraint on clinic capacity
Triage	Administration	2	16 s	450	100.0	450
Registration	Administration	9	7 s	4 628	97.3	4756
Education	Classrooms	8	22 min 7 s	651	97.3	669
Screening	Administration	16	1 min 43 s	559	97.3	574
Consultation	Doctor	7	3 min 42 s	113	25.5	443
Vaccination	Vaccination	16	3 min 16 s	294	95.8	307
Symptoms	Doctor	1	1 min 12 s	50	4.8	1041
Contact	Doctor	1	3 min 48 s	15	3.2	468

Table 3 - Server resource allocation

Model output

As the key performance indicators of this model is the measured capacity and average time spent in the system, we ran the model in eight-hour intervals, with ten repetitions in order to get a sense of both the mean processing time and capacity and the degree of which it varied across the days. We based the 8 hour shifts in a normal working day in Norway. Utilization is calculated from the mean units in use and the number of staff.

Performance metric (per 8 hours)	Mean (10 reps)	Standard deviation (10 reps)
Time in center (mean)	0.8709 hours	0.0257 hours
Time in center (st.dev)	0.1137 hours	0.0086 hours
Max time in center	1.4442 hours	0.0498 hours
Min time in center	0.5968 hours	0.0355 hours
Vaccinated People	1702 people	25 people
Exit without vaccination	296 people	21 people

Table 4 - Simulation Performance

Resource	Administration	Doctors	Vaccination
Number of staff	27	9	16
Average units in use (mean, st.dev)	(7.92, 6.43)	(4.25, 2.48)	(12.74, 5.34)
Units in use (min, max)	(0, 22)	(0, 9)	(0, 16)
Utilization (mean/n)	29,3 %	47,2 %	79,6 %

Table 5 - Resource utilization

Application of the simulation model

The purpose of this section is to utilize the key performance indicators obtained from the simulation model to assist Bergen in enhancing its readiness for potential pandemics. The indicators derived from the simulation model will be analyzed to provide insights into the performance of the vaccination center and to identify areas for improvement. By doing so, this report aims to provide valuable recommendations for Bergen to enhance its pandemic preparedness, based on a data-driven approach. To ascertain the necessary capacity of Bergen for vaccinating its population, a fundamental aspect to consider is whether the requirement is to determine the number of individuals that can be vaccinated within a specific timeframe, or the number of days required to vaccinate a given population. This report assumes that the objective of Bergen is to achieve a negative reproduction number of the virus.

Drawing upon epidemiological reproduction theory, we can infer that Bergen seeks to devise a plan that identifies the required number of operational vaccination stations and shifts to control the spread of a virus, contingent on the virus's true reproduction number (R_0). If we assume that a virus with a similar reproduction number as Covid-19 were to hit Bergen, we find that 54.75% of the population, or 157 097 people would need to get vaccinated in order for the reproduction rate R of the virus to get below one, if we assume that vaccinated people are not susceptible to the virus[1]. According to the output from the simulation the center is able to vaccinate 1702 $\pm 1.96 \frac{25}{\sqrt{10}}$ (95% confidence) people on average per shift. By extrapolating this measure, we can determine that in order to reach the desired figure of 157 097 people vaccinated Bergen would need ~ 93 shifts to vaccinate a sufficient amount of people to reduce the R_0 number.

Given the determined resource allocation, we may proceed to utilize the available data to determine the optimal number of vaccination centers in Bergen. Although Bergen is a large city in Norway, it does not have unlimited access to large open buildings in which effective vaccination can take place. We can determine the average amount of people in the building at the same time by using Little's Law. With a cycle time of 0.8709 and an average throughput of $1998/8 = 249.5$, we find that average WIP of the center is $\sim 217 + staff$. In a pandemic there might also be restrictions on required space between people, (4 m^2 if required distance is 1 meter). The allocated space must therefore be at least 868 m^2. If we add in the numerous other operations at the center and parking etc. only schools and sports avenues remain as good candidates for this vaccination center design.

Pandemics increase medical demand and strain healthcare workers. Therefore, vaccination efforts should avoid disrupting hospital functions and work absences. Table 5 shows that vaccination workers have the highest average utilization rate. Therefore, we recommend that staff downsizing should focus on reducing administrative and doctor personnel.

The optimal number of centers will therefore depend on the trade-off between the socio-economic costs of diverting staff from hospitals and reducing workforce participation in the form of less hospital capacity and more time off work for the vaccinated people, and the benefits of faster vaccination. These costs are not considered in our model. We suggest that the number of centers should vary over time, with more centers available at the start of the vaccination program and fewer centers needed as R gets below 1 and the pandemic is somewhat under control.

---

[1] 286 930 people in Bergen times the needed non-susceptible rate with an R_0 of 2.21 equals 157 097 people. See Appendix 1 for specific calculations. (Bergen Kommune, 2023) (Locatelli, Trächsel, & Rousson, 2021)

Appendix:

Appendix 1:
Calculating the needed non- susceptible share of the population in order to get a reproductive number less than or equal to 1: (Thomas & Freiburger, 2020)

s: *Share of population that is susceptible to the virus*
R_0: *True reproduction number of the disease*
R: *Actual reproduction number in the population*

The actual reproduction number is the product of the *true* reproduction number of any given virus times the share of the population that is susceptible to catching the virus:

$$R = sR_0$$

In order for $R \leq 1$ there needs to be a given value for s which together with R_0 gives us a negative reproduction rate:

$$sR_0 < 1$$

Therefore, we need to find the share of the population that is *not* susceptible to the disease, which is given as $1 - s$:

$$1 - s < 1 - \frac{1}{R_0}$$

This formula gives us the vaccination level in the population that is needed in order to contain the virus, given that people who take the vaccine is non-susceptible to the virus (this is a simplification).

Bibliography

Bergen Kommune. (2023, March 9). *Fakta om Bergen*. Hentet fra Bergen Kommune: https://www.bergen.kommune.no/omkommunen/fakta-om-bergen/befolkning/folketall-per-1-januar-2022

He, J., & Freeman, L. A. (2020). *Can we trust teaching evaluations when response rates are not.* Dearborn: University of Michigan-Dearborn.

Laguna, M., & Marklund, J. (2013). *Business process modeling, simulation and design.* New York: CRC Press.

Locatelli, I., Trächsel, B., & Rousson, V. (2021). Estimating the basic reproduction number for COVID-19 in Western Europe. *PLoS ONE*, 16(3): e0248731.

NHH. (2023, March 10). Hentet fra https://www.nhh.no/om-nhh/tal-og-fakta/#:~:text=NHH%20har%20ein%20internasjonal%20profil,435%20tilsette.

Thomas, R., & Freiburger, M. (2020, April 2). *"R nought" and herd immunity*. Hentet fra Plus Math: https://plus.maths.org/content/maths-minute-r0-and-herd-immunity

Aaby, K., Cook, T., Herrmann, J., Jordan, C., & Wood, K. (2005). *Simulating a Mass Vaccination Clinic.*

Aaby, K., Cook, T., Herrmann, J., Jordan, C., & Wood, K. (2006). Montgomery County's Public Health Service Uses Operations Research to Plan Emergency Mass Dispensing and Vaccination Clinics. *Interfaces*, 36(6): 569-579.