Department of Industrial Engineering
University of Stellenbosch

Simulasie 442 : Simulation 442
2025

MEMORANDUM

Tutoriaal 5 *Tutorial 5*	Punt: 89 *Mark:*	Ingeedatum: **22-08-2025** (10:00) B3003 *Due date:*

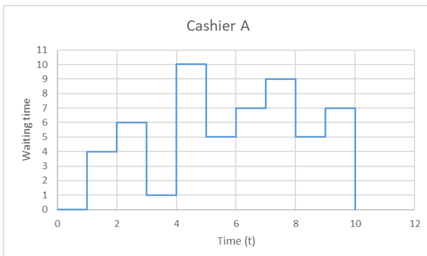
Instruksies:	Formatteer alle syfers sinvol. U mag in groepe van **twee** of minder werk om die vrae te beantwoord. Handig slegs een hardekopie van u antwoordstel in. Gebruik Excel, R of Matlab vir u berekenings. Die data vir hierdie tutoriaal is beskikbaar in die lêer Tut05_2025_RawData.xlsx. **Hierdie tutoriaal is verpligtend.** **Indien u nalaat om die vereistes betyds** **na te kom, sal u die module sak.**
Instructions:	*Format all numbers sensibly.* *You may work in groups of* ***two*** *or less when* *answering the questions.* *Submit one hardcopy only.* *Use Excel, R or Matlab for your calculations.* The data for this tutorial is available in the file Tut05_2025_RawData.xlsx. ***This tutorial is compulsory.*** ***You will fail the module if you do not*** ***comply with the requirements, on time.***

PLEASE NOTE:

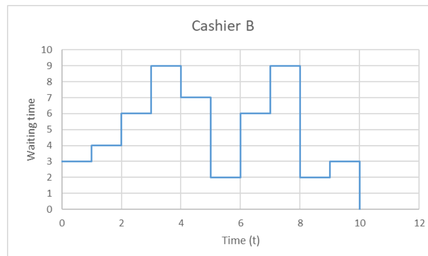
- student h* can be different to memo but must always be a sensible number less than h
- n* must always be rounded up to a whole number

Question 1 [8]

(a) Consider a scenario where two cashiers are working at a grocery store. Customers arrive at the store and are served by the cashiers. The average waiting time is influenced by how the cashiers operate and the customers' arrival patterns. Which cashier has a higher average waiting time? Provide a clear explanation supporting your answer. [5]

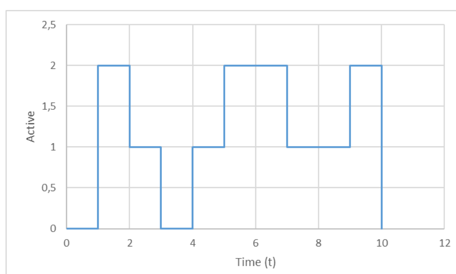


(a) Cashier A



(b) Cashier B

(b) Two resources are operated over a time period from t=0 to t=10. The operational level is defined as follows: when neither resource is operating, the level is 0; when one resource is active, the level is 1; when both resources are active simultaneously, the level is 2. Based on this information, analyze the combined utilisation of the two resources. [3]



Refer to pages 61 and 62 of the textbook.

- $Cashier A = (4 + 6 + 1 + 10 + 5 + 7 + 9 + 5 + 7)/10 = 54/10 = 5.4$ ✓✓
- $Cashier B = (3 + 4 + 6 + 9 + 7 + 2 + 6 + 9 + 2 + 3)/10 = 51/10 = 5.1$ ✓✓
- (a) Cashier A has a higher average waiting time ✓
- (b) $= (2 + 1 + 1 + 2 + 2 + 1 + 1 + 2)/(10 * 2) = 12/20 = 0.6$ ✓✓✓

Question 2 [19]

Refer to the data provided in the **Sim442_Tut_05_RawData** Excel file on sheet **Question 2**. Use a confidence level of 95% to answer the questions

1. Determine the confidence interval (CI) half-widths using both the normal and t-distribution. [11]
2. Determine n^* using the half-widths of both the normal and t-distribution. [4]
3. Explain the difference in the size of the CIs. [2]
4. Explain why a smaller h^* is better than selecting a larger h^*. [2]

QUESTION 2 [19]

Observations			
36	**n =**	20	**[1]**
11	**Avg =**	38,65	**[2]**
49	**Std dev =**	23,69	**[2]**
34	**(α=0.05) t =**	2,09	**[1]**
69	**(α=0.05) z =**	1,96	**[1]**
53			
38	**h_t =**	11,09	**[2]**
70	**n* =**	615	**[2]**
45			
6	**h_z =**	10,38	**[2]**
59	**n* =**	539	**[2]**
75			
4			
26	**h_t *=**	2	
30	**h_z *=**	2	
17			
19			
60			
4			
68			

3 The z-value is less than the t-value for small n, so the t-value gives a wider (more conservative) CI half-width. ✓ As $n \to \infty$, the t-distribution approaches the normal. ✓

4 A smaller h^* requires a larger n^*, which means more simulation replications ✓ and therefore higher precision in estimated results. ✓

Question 3 [24]

Sam is a dedicated marathon runner who trains regularly on various routes around Stellenbosch. She decided to conduct her final year project on her training performance by creating a simulation model of her daily running times, based on various factors such as weather, route difficulty, and fatigue levels. The simulation model's output was captured in the **Sim442_Tut_05_RawData** Excel file on sheet **Question 3**. Show all calculations.

(a) Define the confidence interval for this parameter at a confidence level of $\alpha = 1\%$ AND $\alpha = 5\%$ based on:
(i) the normal distribution, and
(ii) the t-distribution. [14]

(b) Interpret the confidence intervals for $\alpha = 1\%$ [1]

(c) Interpret the confidence intervals for $\alpha = 5\%$ [1]

(d) With reference to your results in **(a)**, explain what the implication of using the normal distribution compared to the t-distribution is, when used to determine the confidence intervals for a small number of samples. [2]

(e) With reference to your results in **(a)**, explain the difference in the confidence interval width due to an α value of 0.01 compared to an α value of 0.05. [2]

(f) Calculate the total number of simulation run replications required to reduce the normal distribution-based confidence interval to a suitable width for:
(i) $\alpha = 1\%$
(ii) $\alpha = 5\%$ [2]

(g) Explain the difference in the number of required simulation run replications. [2]

t-distribution (a = 1%)		
alpha =	1%	
n =	30	
$\bar{X}$ =	49,63	
$S_{\bar{X}} =$	9,626	
$t_{n-1;1-a/2}$ =	2,756	[1] (a)
h =	4,844	[1] (a)
h* =	2	
n* =	176	
CI Lower =	44,789	[1] (a)
CI Upper =	54,477	

Normal distribution (a = 1%)		
alpha =	1%	
n =	30	
$\bar{X}$ =	49,633	
$S_{\bar{X}} =$	9,626	
z =	2,576	[1] (a)
h =	4,527	[1] (a)
h* =	2	
n* =	154	[1] (f)
CI Lower =	45,107	[1] (a)
CI Upper =	54,160	

t-distribution (a = 5%)		
alpha =	5%	
n =	30	
$\bar{X}$ =	49,633	
$S_{\bar{X}} =$	9,63	
$t_{n-1;1-a/2}$ =	2,045	[1] (a)
h =	3,594	[1] (a)
h* =	2	
n* =	97	
CI Lower =	46,039	[1] (a)
CI Upper =	53,228	

Normal distribution (a = 5%)		
alpha =	5%	
n =	30	
$\bar{X}$ =	49,633	
$S_{\bar{X}} =$	9,63	
z =	1,960	[1] (a)
h =	3,444	[1] (a)
h* =	2	
n* =	89	[1] (f)
CI Lower =	46,189	[1] (a)
CI Upper =	53,078	

General		
Mean =	49,633	[1] (a)
Std Dev =	9,626	[1] (a)
n =	30	

Figure 2: Question 3 (a) & (f)

(a)

(b) The calculated confidence interval may be one out of 99 out of 100, which contain the true population mean of manufacturing duration.✓

(c) The calculated confidence interval may be one out of 95 out of 100, that contain the true population mean of manufacturing duration.✓

(d) The confidence interval of the t-distribution is wider ✓than that of the normal distribution, thus the probability of coverage is higher ✓.

(e) The confidence interval for $\alpha = 1\%$ is wider than that of $\alpha = 5\%$ since a larger proportion of intervals are required to cover the true mean ✓. Thus 1% yields a more "conservative" interval width ✓.

(f) Refer to figure 2.
One ✓for suitable h* value and subsequent n* value for $\alpha = 1\%$, and another ✓for $\alpha = 5\%$

(g) More replications are needed to ensure a higher level of confidence (99%)✓. The confidence interval for $\alpha = 1\%$ is also wider to ensure 99% coverage✓.

Question 4 [38]

Keegan runs a small manufacturing business. He decided to record the time it takes to produce products in his factory over a monthly period. The **Sim442_Tut_05_RawData** Excel file on sheet **Question 4** contains the daily observations of the time it takes to produce a product.

(a) Determine the number of replications to be simulated to estimate the parameter X = Expected Time to Produce a product. [10]

(b) Determine the number of replications to be simulated to estimate the parameter Y = Expected number of products per day. [10]

(c) What is the minimum number of replications to be executed according to (a) and (b)? [2]

(d) Construct a histogram for the original observations (time to produce a product) from the dataset used. [6]

(e) Construct a histogram for the daily means of the time to produce a product from the dataset used. [6]

(f) Explain the shape of the histogram in (e). [4]

(a+b) See Excel

QUESTION 4(a) [10]

Daily average	3,691	4,379	2,891	3,032	4,148	4,199	2,376	4,052	3,473	2,395	3,601	3,769	4,181	5,329	2,202	2,791	3,313	3,900	4,184	4,789	3,297	2,819	3,363	3,295	3,470	3,375	3,759	4,042	3,474	3,422	[2]

Mean of means =	3,567	[1]
Std dev of means =	0,70266405	[1]
n=	30	
a =	0,05	[1]
t =	2,04522964	[1]
h =	0,26237907	[2]
CI Lower =	3,305	
CI Upper =	3,829	
h* =	0,2	[1]
n* =	52	[1]

QUESTION 4(b) [10]

Number of obs per day =	22	24	21	17	17	16	19	17	19	14	19	20	21	19	19	24	19	15	17	18	23	14	18	17	20	15	22	13	19	23	[2]

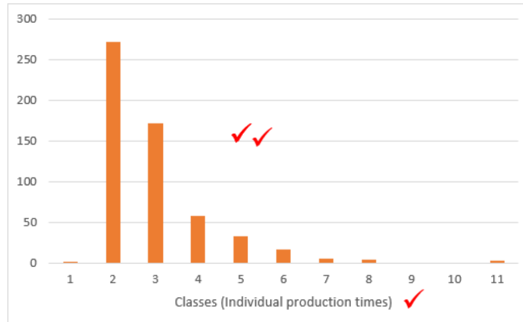
Mean of means =	18,700	[1]
Std dev of means =	2,97286579	[1]
n=	30	
a =	0,05	[1]
t =	2,04522964	[1]
h =	1,11008633	[2]
CI Lower =	17,590	
CI Upper =	19,810	
h* =	0,75	[1]
n* =	66	[1]

(c) Maximum of n** of (a) and (b) thus 66 replications. ✓✓
Depends on what the student calculated as n*.

QUESTION 4(d) [6]

Time to produce (original data)		
Min =	0,130	
Max =	24,852	[1]
Count =	561	
#Classes =	10	[1]
Class width =	2,4722032	[1]

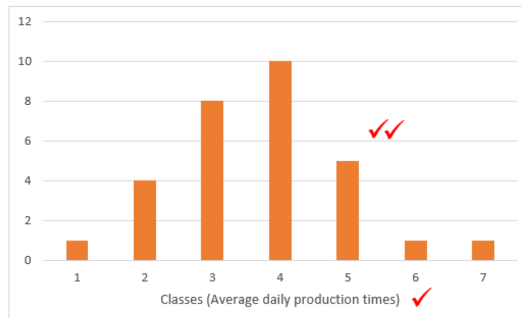
Classes	Obs
0,130	1
2,602	271
5,075	172
7,547	58
10,019	33
12,491	16
14,963	5
17,436	3
19,908	0
22,380	0
24,852	2



QUESTION 4(e) [6]

Time to produce (original data)		
Min =	2,202	
Max =	5,329	[1]
Count =	30	
#Classes =	5	[1]
Class width =	0,6253762	[1]

Classes	Obs
2,202	1
2,828	4
3,453	8
4,078	10
4,704	5
5,329	1
5,954	1



(f) The shape tends towards a normal distribution ✓✓, this is due to the central limit theorem ✓and the fact that the observations are means of the original observation times ✓.

Total: Cross-check: 89