



מרצה : פרופסור אבישי מנדלבאום

מתרגלת : ניצן כרמלי

תאריך הבחינה : 11.10.2012

מס' סטודנט _____

שם _____

הנדסת מערכות שירות – 096324

מועד ב' – סמסטר אביב תשע"ב 2012

פתרון

- המבחן נמשך שלוש שעות. באחריותך להחזירו במועד.
- המבחן בחומר סגור.
- מותר להשתמש במחשבון.
- המבחן כולל 33 עמודים, כולל עמוד זה ועמוד 33 (התפלגויות מהקורס בהסתברות).
- כאשר נדרשים חישובים, ניתן לעגלם באופן שאינו פוגע בהבנת התשובה.
- מלאו את הפרטים הנדרשים בראש העמוד, בכתב ברור בבקשה.
- על התשובות יש לענות במקומות המיועדים, שאמורים להספיק: מלל מיותר יגרע ניקוד.
- יש לתת הסברים או הוכחות רק אם התבקשתם במפורש לעשות זאת.

הרצאות\תרגולים\בחינות קודמות	מתוך 10 נקודות	שאלה 1
תרגילי בית	מתוך 10 נקודות	שאלה 2
יישום	מתוך 20 נקודות	שאלה 3
תיאוריה	מתוך 10 נקודות	שאלה 4

סה"כ _____ מתוך 50 נקודות

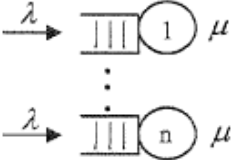
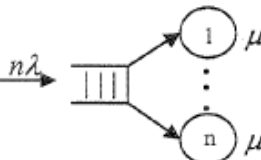
בהצלחה !

שאלה 1 - תרגילי כיתה/הרצאות/מבחנים קודמים (10 נקודות)

חלק 1 – Pooling Queues and Servers

להלן טבלה שהוצגה בכיתה המפרטת את ההבדלים בין 3 מערכות שונות של תורים:

1. n תורי M/M/1 עם הפרמטרים λ, μ - מערכת 1.
2. תור M/M/n עם הפרמטרים $\mu, n\lambda$ - מערכת 2.
3. תור M/M/1 עם הפרמטרים $n\mu, n\lambda$ - מערכת 3.

	$n \times M M 1$ λ, μ	$M M n$ $n\lambda, \mu$	$M M 1$ $n\lambda, n\mu$
	Multiple queues Multiple Servers	Single queue Multiple Servers	Single queue Single server
			
Tradeoff	Separate vs. single queue		Slow vs. fast server
Enabler	Process design		Technology
Gain	Load balancing $\forall t$ (avoid a long queue & idle servers)		Maximal capacity $\forall t$ $\leq n\mu$ vs. $n\mu$
Utilization	$\frac{\lambda}{\mu} = \rho$	A	B
$P\{W_q > 0\}$	ρ	2	ρ
$E[W_q W_q > 0]$	$\frac{1}{\mu} \cdot \frac{1}{1-\rho}$	1	$= \frac{1}{n\mu} \cdot \frac{1}{1-\rho}$
$E[W_q] =$ $E[W_q W_q > 0] \cdot P\{W_q > 0\}$	$\frac{1}{\mu} \cdot \frac{\rho}{1-\rho}$	4	3
$E[S]$	$\frac{1}{\mu}$	$= \frac{1}{\mu}$	$> \frac{1}{n\mu}$
$E[W_{sys}] =$ $= E[W_q] + E[S]$	$\frac{1}{\mu} \cdot \frac{1}{1-\rho}$	5	6

1.1.1 השלימו את הערכים החסרים המסומנים ב – A וב – B (הנצילות במערכת 1 ובמערכת 2 בהתאמה).

תשובה:

$$\rho = \frac{n\lambda}{(n)\mu} : A$$

$$\rho = \frac{n\lambda}{n\mu} : B$$

1.1.2 יחס (2) : מה היחס בין $P\{W_q > 0\} = \rho$ במערכת 1 ל- $E_{2,n}(\rho)$ במערכת

2. כלומר, האם מתקיים $E_{2,n}(\rho) \geq \rho$ או $E_{2,n}(\rho) \leq \rho$?

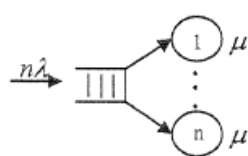
הסבירו את האינטואיציה מאחורי היחס הנ"ל.

תשובה:

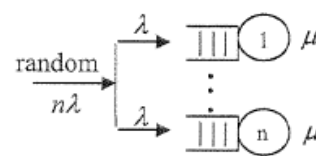
$$E_{2,n}(\rho) \leq \rho$$

הסבר : מערכת 1

מערכת 2



vs.



$$E_{2,n}(\rho)$$

≤

$$\rho$$

כשמופיע לקוח למערכת 2 ויש שרת פנוי הוא ייכנס באופן מיידי לשירות, ללא המתנה. כשמופיע לקוח למערכת 1 הוא למעשה נכנס באופן רנדומאלי לאחד מ- n התורים לשרתים השונים. ייתכן כי הלקוח נכנס לתור בו השרת עסוק ולכן יאלץ להמתין בעוד שיש שרת פנוי בתור אחר.

1.1.3 הסבירו מדוע מתקיים שוויון בין $E[W_q | W_q > 0]$ במערכת 2 לבין $E[W_q | W_q > 0]$ במערכת 3.

תשובה:

לפי PASTA, לקוח שמגיע למערכת 2 ורואה את כל השרתים עסוקים ו- i לקוחות בתור, או לקוח שמגיע למערכת 3 ורואה את השרת עסוק ו- i לקוחות בתור ייאלץ להמתין עד ש- $i+1$ לקוחות יקבלו שירות. זמן השירות של השרת המהיר במערכת 3 מתפלג $\exp(n\mu)$, זמן השירות של n השרתים יחד במערכת 2 גם הוא מתפלג $\exp(n\mu)$. לכן, בהינתן המתנה, זמן ההמתנה הן במערכת 2 והן במערכת 3 הינו סכום גיאומטרי עם פרמטר $(1-\rho)$ (בהינתן המתנה, מספר הלקוחות במערכת 2 ובמערכת 3 מתפלג $Geom(1-\rho)$ של מ"מ $\exp(n\mu)$, ולפיכך, בשתי

$$E[W_q | W_q > 0] = \frac{1}{n\mu} \cdot \frac{1}{1-\rho}$$

1.1.4 יחס (4) : מה היחס בין $E[W_q]$ במערכת 1 ל- $E[W_q]$ במערכת 2 (כלומר איזה סימן

מופיע ב- (4) , $\geq$, $\leq$ או $=$). הסבירו מדוע מתקבל יחס זה.

תשובה:

$$E[W_q]_1 = \frac{1}{\mu} \frac{\rho}{1-\rho} \geq E[W_q]_2 = \frac{1}{\mu} \frac{E_{2,n}}{n(1-\rho)}$$

באופן דומה להסבר בסעיף הקודם, בהינתן המתנה, מספר הלקוחות במערכת 1 ובמערכת 2 מתפלג $Geom(1-\rho)$, וזמן השירות מתפלג $\exp(\mu)$ ו- $\exp(n\mu)$ בהתאמה. לכן:

$$E[W_q | W_q > 0]_1 = \frac{1}{\mu} \frac{1}{1-\rho} \geq E[W_q | W_q > 0]_2 = \frac{1}{\mu} \frac{1}{n(1-\rho)}$$

בסעיף 1.1.2 הראינו כי $E_{2,n}(\rho) \leq \rho$ כאשר ρ זוהי ההסתברות להמתנה במערכת 1 ו-

$E_{2,n}(\rho)$ זוהי ההסתברות להמתנה במערכת 2.

כמו כן, ע"פ נוסחת ההסתברות השלמה $P\{W_q > 0\}$ $E[W_q] = E[W_q | W_q > 0]P\{W_q > 0\}$

ברור כי אם מתקיים $E[W_q | W_q > 0]_1 \geq E[W_q | W_q > 0]_2$ וכן $P\{W_q > 0\}_1 \geq P\{W_q > 0\}_2$ אזי

$$E[W_q]_1 \geq E[W_q]_2$$

1.1.5 יחס (6) : מה היחס בין $E[W_{sys}]$ במערכת 2 ל- $E[W_{sys}]$ במערכת 3 (כלומר איזה סימן

מופיע ב- (6) , $\geq$, $\leq$ או $=$). הסבירו מדוע מתקבל יחס זה.

תשובה:

ראשית נוכיח כי מתקיים בקשר הבסיסי: $1 \leq \frac{1-E_{2,n}}{1-\rho} \leq n$

הוכחה:

$$1 - E_{2,n} = P(\text{at least one server idle}) =$$

$$= P\left(\bigcup_{i=1}^n \{\text{server } i \text{ idle}\}\right) \begin{cases} \leq \sum_{i=1}^n P\{\text{server } i \text{ idle}\} = n(1-\rho) \\ \geq P\{\text{server } i \text{ idle}\} = 1-\rho \end{cases}$$

הסבר: ההסתברות לקבל שירות מידי במערכת 2 הינה ההסתברות שלפחות שרת אחד מבין n

השרתים פנוי. הסתברות זו שווה להסתברות על איחוד המאורעות $\bigcup_{i=1}^n \{\text{server } i \text{ is idle}\}$.

מנוסחת ההכלה וההפרדה, הסתברות האיחוד הנ"ל קטנה או שווה לסכום ההסתברויות

ההסתברות ששרת i פנוי הינה $(1-\rho)$. וגדולה כמובן מהסתברות המאורע הבודד $\{\text{server } i \text{ is idle}\}$. כאשר

ההסתברות ששרת i פנוי הינה $(1-\rho)$.

$$1 - E_{2,n} \leq n(1-\rho) \Rightarrow \frac{1-E_{2,n}}{n(1-\rho)} \leq 1 \Rightarrow \frac{1}{n(1-\rho)} \leq 1 + \frac{E_{2,n}}{n(1-\rho)} : \text{ כעת מקבלים:}$$

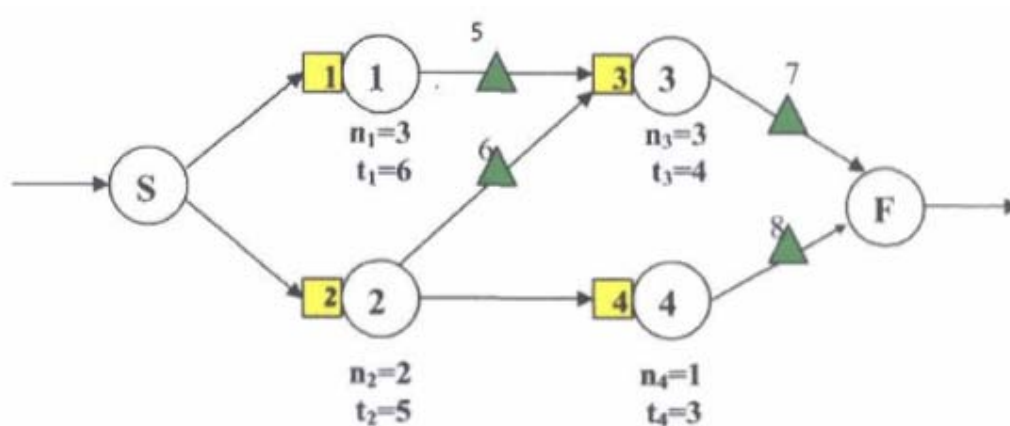
חלק 2 - רשתות דינמיות סטוכסטיות

נתונה רשת PERT/CPM דינמית סטוכסטית, כמתואר בגרף 1 וכפי שנלמד בכיתה: t_i הוא

ממוצע זמן השירות בתחנה i , ו- n_i הוא מספר השרתים בתחנה i , $i = 1, 2, 3, 4$.

פרויקטים חדשים מופיעים למערכת לפי תהליך פואסון עם ממוצע זמן בין מופעי של 3.5 ימים.

נניח שהרשת במצב יציב.



Graph Number 1

1.2.1 לרשת 8 תורים הממוספרים מ-1 עד 8. מי התורים הנובעים מפערי סינכרון ומי הנובעים

ממשאבים מוגבלים?

תשובה:

תורי סינכרון: 5, 6, 7, 8

תורי משאבים: 1, 2, 3, 4

1.2.2 מהו המספר המינימלי של שרתים הנדרשים לטפל בפעילות 3 כדי שנצילות כל אחד לא

תעלה על 33%?

תשובה:

מספר מינימלי: $\lceil 4 / (3.5 * n_3) \rceil \leq 0.33$. התוצאה היא 4 שרתים.

1.2.3 נסמן ב- T_{DS} את משך ביצוע פרויקט שמגיע למערכת במצב יציב. הסבירו במדויק ככל שתוכלו, אבל בשניים-שלושה משפטים לכל היותר, מדוע $ET_{DS} < 10$.

תשובה:

ההסבר כולל שני חלקים. מכיוון שהרשת סטוכסטית ועל סמך מבנה הקדימויות, אנחנו כבר יודעים ש- $ET_{DS} \leq 10$ (ראו הוכחה בהמשך). בנוסף, הדינאמיות של המערכת (ריבוי הפרויקטים) גורמת לתורי משאבים שמאריכים עוד יותר את משך ביצוע הפרויקט.

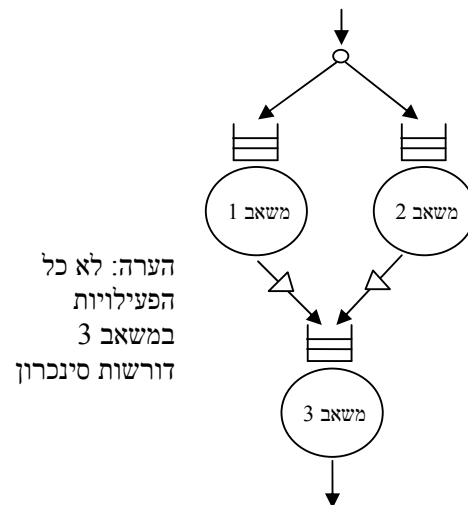
$$E\left(\max_j \sum_{i \in j} T_i\right) \geq \max_j E\left(\sum_{i \in j} T_i\right) = \max_j \sum_{i \in j} ET_i = 10$$

ברשת.

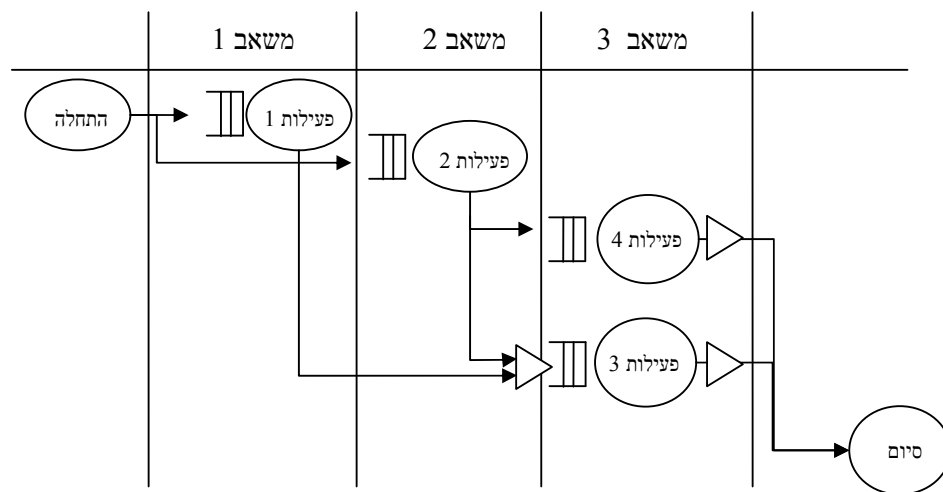
חלק זה של העמוד נשאר ריק

1.2.4 מנהל החברה הכשיר את השרתים שמבצעים את פעילות 3 ו-4 לבצע את שתי הפעילויות. צייר תרשים משאבים ותרשים משולב משאבים ופעילויות לפרויקט, תחת תרחיש חדש זה.

תרשים משאבים:



תרשים משולב פעילויות-משאבים:



שאלה 2 – תרגילי בית (10 נקודות)

מודלים אמפיריים ו-SEESTAT

להלן נתון חלק מדו"ח call by call data של בנק אמריקאי מיום ה-25 למאי 2001 (טבלה 1). בנוסף, מצורפת טבלה הכוללת נתוני הגעות מצטברות ועזיבות מצטברות מהמערכת, מהתור ומהשירות (טבלה 2). שעות פעילות המוקד ביום זה הן 06:00-23:00. הדוחות מתייחסים לסוג שירות "Summit".

לנוחיותכם, הדוחות מצורפים גם בשני העמודים האחרונים של הבחינה, כדי שתוכלו לתלוש עמודים אלו ולהעזר בהם למענה על השאלות הבאות.

טבלה 1

call_id	q_start	q_exit	ser_start	ser_exit	outcome	wait_time	service_time	Agent
510013309	06:23:00	06:23:07	06:23:08	06:25:01	Agent	7	113	23015
510013355	06:23:19	06:24:30	06:24:31	06:24:51	Agent	71	20	23007
510013559	06:24:52	06:24:52	06:24:52	06:26:57	Agent	0	125	23007
510013814	06:26:32	06:26:33	06:26:33	06:27:27	Agent	1	54	23015
510013830	06:26:36	06:26:38	00:00:00	00:00:00	Abandon	2	0	0
510013871	06:26:47	06:26:57	06:26:58	06:28:45	Agent	10	107	23007
510014006	06:27:53	06:27:53	06:27:53	06:29:46	Agent	0	113	23015
510014071	06:28:15	06:28:45	06:28:46	06:29:04	Agent	30	18	23007
510014116	06:28:35	06:29:04	06:29:05	06:30:31	Agent	29	86	23007
510014132	06:28:43	06:29:46	06:29:47	06:31:27	Agent	63	100	23015
510014196	06:29:05	06:30:31	06:30:32	06:34:32	Agent	86	240	23007
510014425	06:30:38	06:31:27	06:31:27	06:34:31	Agent	49	184	23015
510014578	06:31:32	06:34:31	06:34:32	06:35:54	Agent	179	82	23015
510014674	06:32:06	06:34:32	06:34:33	06:37:04	Agent	146	151	23007
510014702	06:32:18	06:33:00	00:00:00	00:00:00	Abandon	42	0	0
510014712	06:32:26	06:35:54	06:35:55	06:37:37	Agent	208	102	23015
510014750	06:32:37	06:36:08	06:36:09	06:47:58	Agent	211	709	23041
510014927	06:33:46	06:37:04	06:37:05	06:43:49	Agent	198	404	23007
510015175	06:35:19	06:36:04	00:00:00	00:00:00	Abandon	45	0	0
510015368	06:36:28	06:37:37	06:37:38	06:39:53	Agent	69	135	23015
510015393	06:36:35	06:37:54	00:00:00	00:00:00	Abandon	79	0	0
510015405	06:36:39	06:39:53	06:39:54	06:51:04	Agent	194	670	23015
510015441	06:36:55	06:43:49	06:43:50	06:45:58	Agent	414	128	23007
510015539	06:37:41	06:45:58	06:45:59	06:47:44	Agent	497	105	23007
510015752	06:38:54	06:39:14	00:00:00	00:00:00	Abandon	20	0	0
510015802	06:39:13	06:40:31	00:00:00	00:00:00	Abandon	78	0	0
510016057	06:40:42	06:40:59	00:00:00	00:00:00	Abandon	17	0	0
510016208	06:41:36	06:47:44	06:47:45	06:51:29	Agent	368	224	23007
510016446	06:42:56	06:47:58	06:47:59	06:48:31	Agent	302	32	23041
510016552	06:43:35	06:45:36	00:00:00	00:00:00	Abandon	121	0	0
510016619	06:44:02	06:48:31	06:48:32	06:50:56	Agent	269	144	23041
510016625	06:44:05	06:46:13	00:00:00	00:00:00	Abandon	128	0	0
510016700	06:44:32	06:50:56	06:50:57	06:56:20	Agent	384	323	23041
510016708	06:44:36	06:51:04	06:51:05	06:53:01	Agent	388	116	23015
510016925	06:45:53	06:51:29	06:51:30	06:53:14	Agent	336	104	23007
510017302	06:48:04	06:53:01	06:53:02	06:54:12	Agent	297	70	23015
510017389	06:48:31	06:49:30	00:00:00	00:00:00	Abandon	59	0	0
510017423	06:48:42	06:53:14	06:53:14	06:54:56	Agent	272	102	23007
510017606	06:49:55	06:54:12	06:54:12	06:55:49	Agent	257	97	23015
510017636	06:50:03	06:51:57	00:00:00	00:00:00	Abandon	114	0	0
510017717	06:50:38	06:54:56	06:54:57	06:56:50	Agent	258	113	23007

טבלה 2

time	Arr_sys	Dep_sys	Arr_q	Dep_q	Arr_ser	Dep_ser
06:15	6	6	6	6	5	5
06:16	6	6	6	6	5	5
06:17	6	6	6	6	5	5
06:18	6	6	6	6	5	5
06:19	7	6	7	7	6	5
06:20	7	6	7	7	6	5
06:21	7	6	7	7	6	5
06:22	8	6	8	8	7	5
06:23	9	6	9	9	8	5
06:24	11	7	11	10	9	6
06:25	12	9	12	12	11	8
06:26	12	10	12	12	11	9
06:27	15	12	15	15	13	10
06:28	16	13	16	16	14	11
06:29	19	14	19	17	15	12
06:30	20	16	20	19	17	14
06:31	21	17	21	20	18	15
06:32	22	18	22	21	19	16
06:33	26	18	26	21	19	16
06:34	27	19	27	22	19	16
06:35	27	21	27	24	21	18
06:36	28	22	28	25	22	19
06:37	32	24	32	27	23	20
06:38	33	27	33	30	25	22
06:39	34	27	34	30	25	22
06:40	35	29	35	32	26	23
06:41	36	31	36	34	26	23
06:42	37	31	37	34	26	23
06:43	38	31	38	34	26	23
06:44	39	32	39	35	27	24
06:45	43	32	43	35	27	24
06:46	44	34	44	37	28	25
06:47	44	35	44	38	28	25
06:48	44	37	44	40	30	27
06:49	47	38	47	41	31	28
06:50	48	39	48	42	31	28
06:51	51	40	51	43	32	29
06:52	53	43	53	46	34	31
06:53	54	43	54	46	34	31
06:54	54	45	54	48	36	33
06:55	54	47	54	50	38	35
06:56	56	49	56	52	39	36
06:57	59	52	59	55	41	38
06:58	59	53	59	56	42	39
06:59	60	54	60	57	43	40
07:00	60	55	60	59	44	40
07:01	62	57	62	62	46	41
07:02	64	57	64	64	48	41
07:03	70	62	70	68	51	45
07:04	70	63	70	70	53	46
07:05	71	65	71	71	54	48
07:06	74	67	74	74	57	50
07:07	79	69	79	78	61	52
07:08	80	72	80	80	63	55

2.1 כיצד תחשבו את ממוצע זמן ההמתנה של לקוחות שנכנסו למוקד בין 06:28 ל-07:01. הסבירו ופרטו את תשובתכם, תוך שימוש בשמות העמודות בטבלאות שבנתונין אתם משתמשים. אין צורך בחישובים.

תשובה:

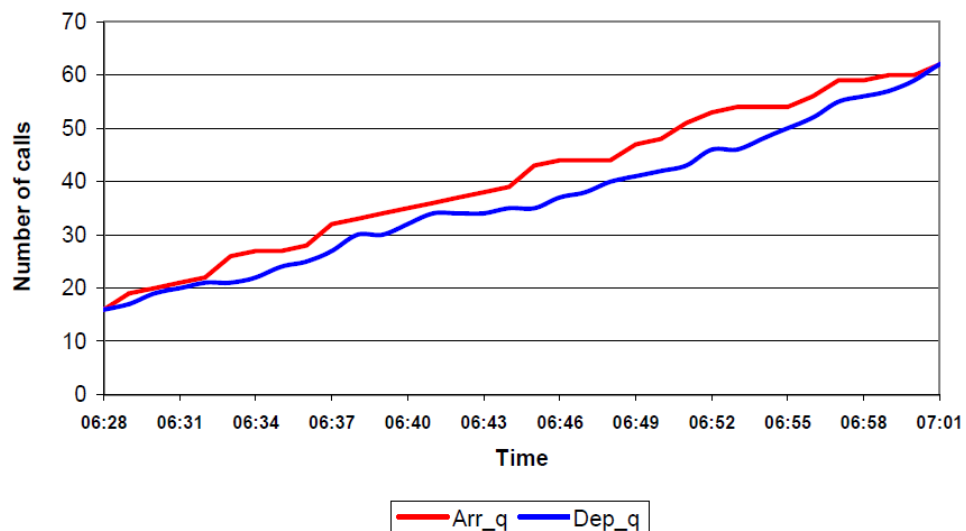
יש לחשב ממוצע על פני עמודת wait_time (טבלה 1) עבור שיחות נכנסות בין 06:28 ל-07:01. הממוצע יחושב ע"י סכימת כל הערכים המופיעים בעמודת wait_time (כולל זמני המתנה השווים ל-0) ומתאימים לשיחות שנכנסו לתור בין 06:28 ל-07:01 (ע"פ עמודת ה-q_start), וחלוקת הסכום הנ"ל בסה"כ השיחות שנכנסו באינטרוול זה.

2.2 כיצד ניתן בעזרת שימוש במודל נוזלים לאמוד את ממוצע זמן ההמתנה של לקוח הנכנס למוקד בין 06:28 ל-07:01. הסבירו תחילה מה עליכם להניח כדי שמודל הנוזלים יהיה רלוונטי, תארו את תהליך האמידה והסבירו האם האמד שתקבלו שונה מהאמד שמחושב לפי סעיף 2.1, או שווה לו. **רמז:** השתמשו בנתונים המופיעים בטבלה 2.

תשובה:

כאשר אנו משתמשים במודל הנוזלים אנו למעשה מסתכלים על סך ההגעות המצטברות למערכת (במקרה זה המערכת היא התור) אל מול סך העזיבות המצטברות מהמערכת (תור). בגרף מטה זמן ההמתנה הכולל של לקוחות בתור הינו השטח בין שתי העקומות – ההגעות המצטברות והעזיבות המצטברות. על מנת לקבל את ממוצע זמן ההמתנה עלינו לחלק שטח זה במספר הלקוחות שנכנסו למערכת באינטרוול הזמן הרצוי.

25 May- Arrivals and departures from queue.



לשם שימוש במודל הנוזלים מניחים כי לקוחות מקבלים שירות לפי מדיניות שירות First Come First Served.

תיאור תהליך האמידה: מכיוון שאנו מניחים מדיניות FCFS, נתאים ללקוח ה- i שנכנס לתור את זמן היציאה ה- i מהתור, נחשב את זמני ההמתנה תחת הנחה זו, ואת ממוצעם. נשים לב כי בתחילת האינטרוול (6: 28) ובסופו (7: 01) התור ריק (מטבלה 2: Arr_q=Dep_q) כלומר, בשעה 6: 28 התור היה ריק וכל הלקוחות שהגיעו החל משעה זו עזבו את התור עד 7: 01. לכן, השטח בין שתי העקומות הינו סך זמן המתנה של כל הלקוחות, והוא אינו תלוי במדיניות התור, ומכאן ששני האמדים יהיו זהים בערכם.

2.3 נשים לב כי הנתונים אינם כוללים את מספר המוקדנים הנותנים שירות. הניחו כי רמת האיוש נשארת קבועה לאורך כל שעה והציעו גישה לאמידת מספר המוקדנים תוך שימוש בנתונים הקיימים. הסבירו את תשובתכם (אין צורך בחישובים).

תשובה:

נציע שתי שיטות לאמידת רמת האיוש במוקד:

1. בכל שעה נמצא את מספר הלקוחות המקסימלי שנמצאים בשירות בו זמנית, ע"י חיסור עמודת Dep_ser מעמודת Arr_ser בטבלה 2.
2. מספר קודי המוקדנים השונים הנותנים שירות במהלך השעה, מתוך עמודת Agent בטבלה 1 (לא ניקח בחשבון את הערך 0)

למרכז שרות US Bank פונים לקוחות מכמה סוגים, ביניהם לקוחות המוגדרים Retail ו- Premier. לקוחות Premier נחשבים לקוחות VIP המקבלים עדיפות גבוהה בנגישות לשרות. למרות העובדה כי השרות עצמו, הניתן לשני סוגי הלקוחות, זהה, הוגדרו שתי קבוצות נציגי שירות נפרדים: "מוקדני Retail" ו"מוקדני Premier", כאשר מוקדני Retail אינם מוסמכים לתת שירות ללקוחות Premier אך מוקדני Premier יכולים לתת שירות ללקוחות Retail.

מנהל השיווק החליט להשתמש בתוכנת SEESat על מנת לבחון את רמת השירות הניתנת לשני סוגי הלקוחות אלו. הוא טוען שבחודש יולי 2001 רמת השירות לא היתה טובה, וזאת על סמך שני הגרפים המופיעים בעמוד הבא:

- גרף 1 - Average waiting time for Retail and Premier calls (July-01).
- גרף 2 - % immediately answered for Retail and Premier calls (July-01).

Figure 1: Average waiting time for Retail and Premier calls (July-01)

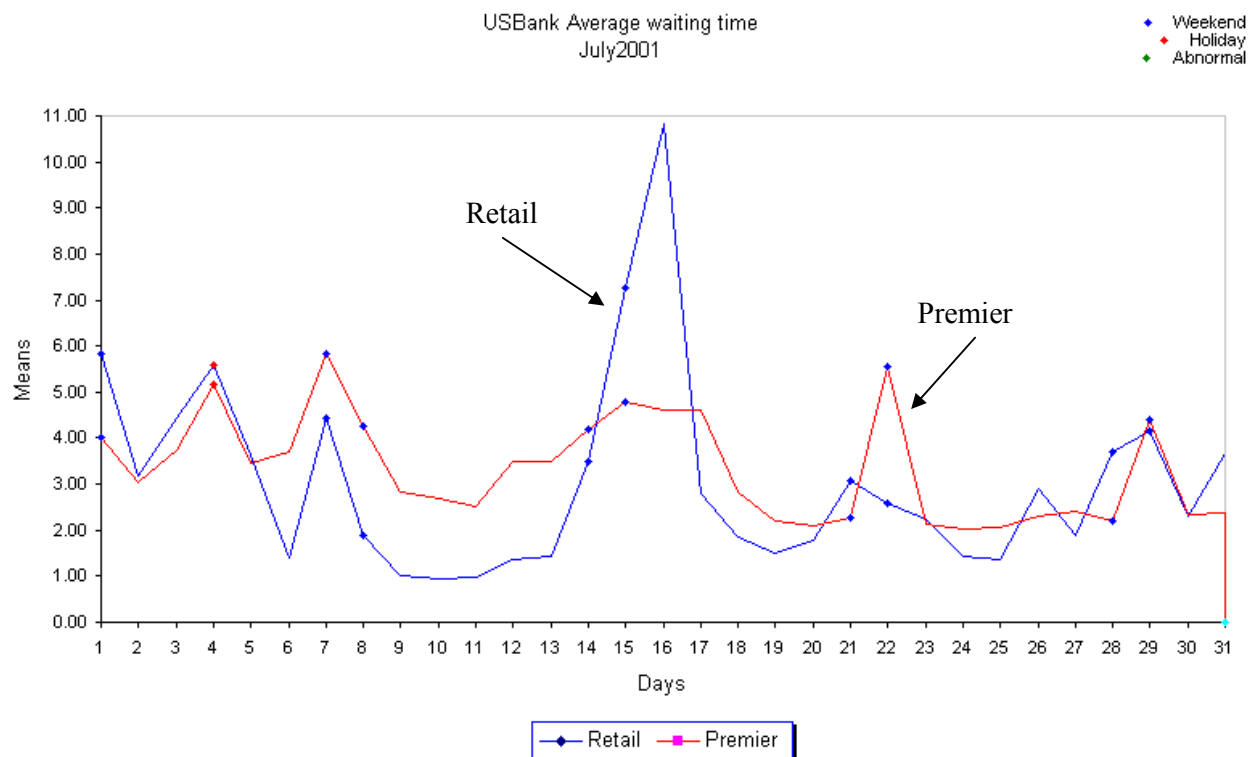
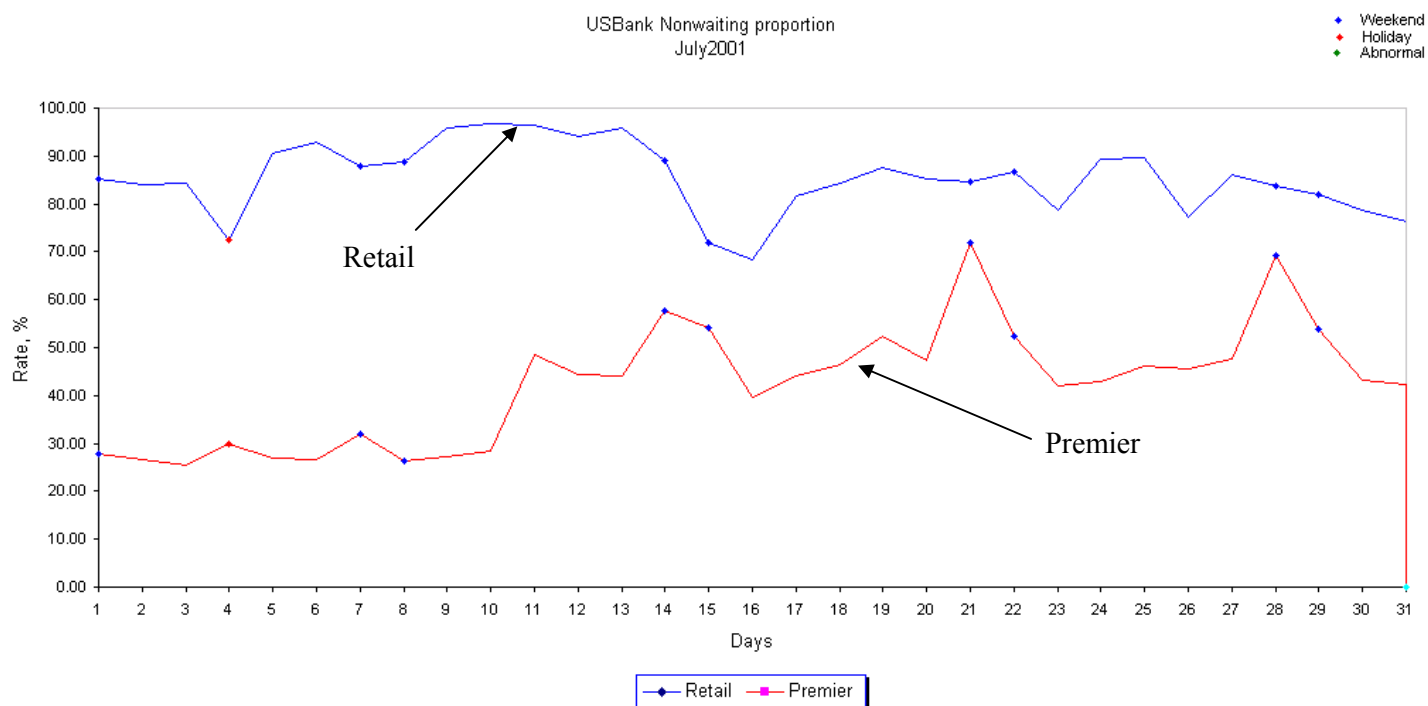


Figure 2: % immediately answered for Retail and Premier calls (July-01)



2.4 הסבירו את טענתו של מנהל השיווק, תוך שימוש בגרפים הנ"ל.

תשובה:

המנהל משווה בין רמות השירות של שני סוגי הלקוחות בשני היבטים: ממוצע זמן המתנה וההסתברות לקבל שירות מיידי. הוא מגלה שלקוחות ה-Retail קיבלו רמת שירות טובה יותר מלקוחות ה-Premier, כיוון שבכל ימי החודש ההסתברות שלקוחות ה-VIP יקבלו שירות מיידי נמוכה יותר וכן ברוב הימים (ובייחוד בין ה-5 ל-14 ביולי) לקוחות ה-VIP מחכים יותר.

2.5 הציעו לפחות 2 פתרונות אפשריים לבעיה, ללא תוספת כ"א.

תשובה:

א. ניתן להעביר קבוצת מוקדנים מ-Retail ל-Premier (השירות הניתן לשתי קבוצות הלקוחות זהה).
ב. ניתן לאחד את שתי קבוצות המוקדנים ולהגדיר מדיניות ניתוב שתיתן עדיפות ללקוחות ה-VIP.
ג. ניתן לשנות את מדיניות הניתוב, למשל:
- לאפשר למוקדני ה-Retail לתת שירות גם ללקוחות ה-Premier, אם למשל יש יותר מ-X לקוחות Premier בתור.
- להגדיר את המדיניות הבאה: "אל תנתב שיחות Retail למוקדני Premier כאשר יש פחות מ-X מוקדני Premier פנויים."

2.6 תארו, במדויק ככל שתוכלו, כיצד ניתן לייצר את גרף 2 -

% immediately answered for Retail and Premier calls (July-01)

בעזרת תוכנת SEESat. (לשם תזכורת הפונקציות של התוכנה, צרפנו תצוגה של שני עמודים מממשק התוכנה).

תשובה:

1. נבחר במודל Time Series, באפשרות של Daily Totals.
2. נבחר במשתנה Nonwaiting proportion, ב-options באופציה Rate, בסוג שירות Retail ו-Premier.
3. נעבור לבחירת התאריכים ונסמן את יולי 2001, נבחר באופציית ה-Days for one month.

Variables
Options
Select Categories

Select by

Open

Variable	Variable Type	Category	Subcategory	
Number of agents	Counts	service	node	
Average service incoming calls	Means	service	node	
Incoming calls	Counts	service	node	
Business calls	Counts	service	node	
Served < 5	Counts	service	node	
Served 5 to 19	Counts	service	node	
Proportion served < 5	Rate	service	node	
Proportioned served 5 to 19	Rate	service	node	
Number agents < 5	Counts	service	node	
Proportion agents < 5	Rate	service	node	
VRU entries	Counts	service		
Announcement entries	Counts	service		
Message entries	Counts	service		
Direct entries	Counts	service		
VRU only	Counts	service		
VRU only proportion	Rate	service		
Announcement only	Counts	service		
Announcement only proportion	Rate	service		
Message only	Counts	service		
Message only proportion	Rate	service		
Arrivals to offered	Counts	service	node	
Handled	Counts	service	node	
Handled proportion	Rate	service	node	
Unhandled	Counts	service	node	
Unhandled proportion	Rate	service	node	
Short abandons	Counts	service	node	
Short abandons proportion	Rate	service	node	
Abandons	Counts	service	node	
Abandons proportion	Rate	service	node	
Other unhandled	Counts	service	node	
Other unhandled proportion	Rate	service	node	
Disconnected in service	Counts	service	node	
Disconnected in service proportion	Rate	service	node	
Continuation	Counts	service	node	
Continuation proportion	Rate	service	node	
Nonwaiting proportion	Rate	service	node	
Average waiting time given waiting	Means	service	node	
Average waiting time	Means	service	node	
Average customer service time	Means	service	node	
Arrivals to offered (all sub-calls)	Counts	service	node	
Handled (all sub-calls)	Counts	service	node	

Months
Days

Month	Year	
March	2001	
April	2001	
May	2001	
June	2001	
July	2001	
August	2001	
September	2001	
October	2001	
November	2001	
December	2001	
January	2002	
February	2002	
March	2002	
April	2002	
...	----	

Time series type
☒ Days for one month
☐ Aggregated days by months

2.7 בעת ביצוע ניתוחים שונים בתוכנת ה-SEESat, הציג מנהל המוקד גרפים המתארים את זמני ההמתנה של לקוחות Platinum ולקוחות EBO שהמתינו בתור, בימי חול בחודש מרץ 2002. הגרפים נראו למנהל המוקד מוזר במקצת. האם אתם יכולים להסביר מדוע הם נראים כך?

Figure 3: Average waiting time for Platinum calls (Week days, March 2002)

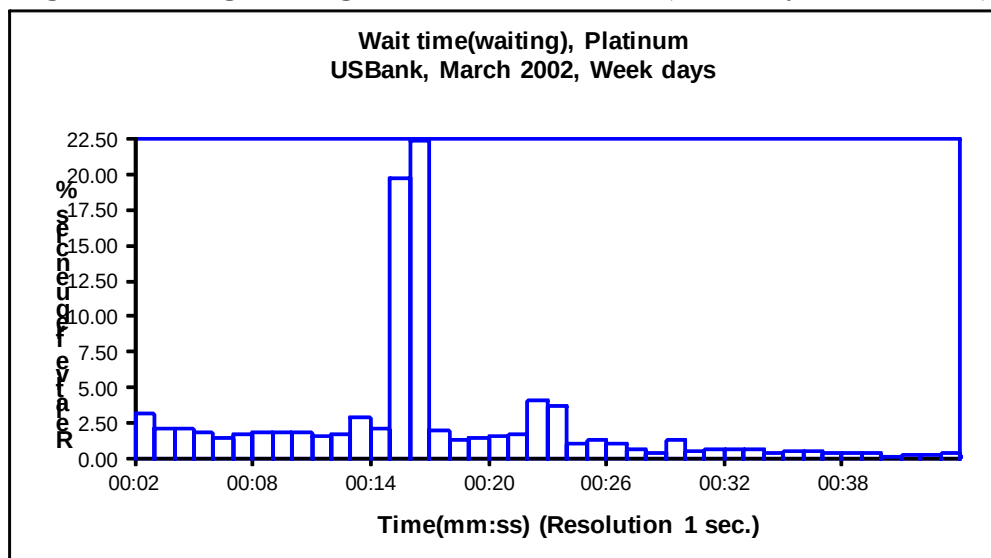
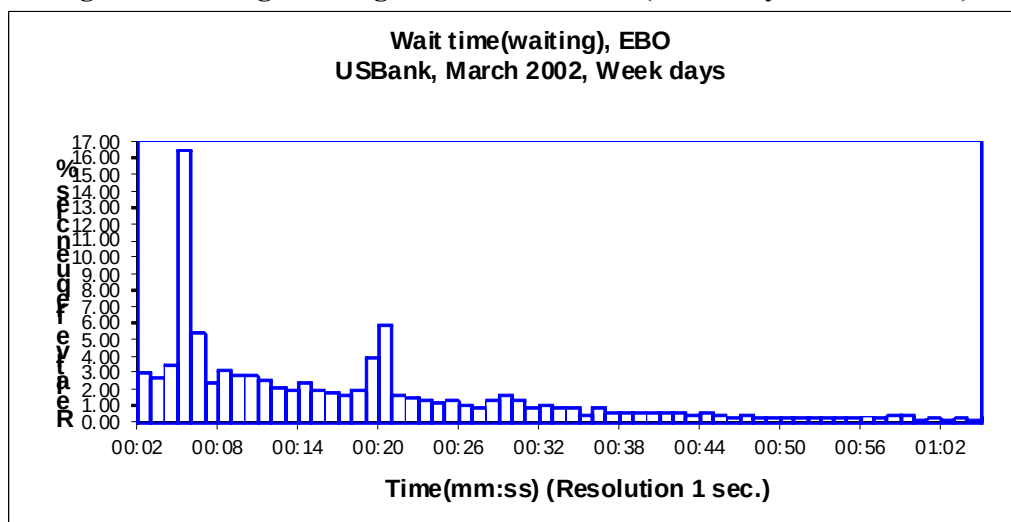


Figure 4: Average waiting time for EBO calls (Week days, March 2002)



תשובה:

בהיסטוגרמה של זמני ההמתנה של לקוחות Platinum שהמתינו בתור, ניתן לזהות peak המתרחש עבור זמן המתנה של 15 שניות. קפיצה זו נוצרת כתוצאה מפרוטוקול הניתוב של מערכת הטלפונים במוקד: אם לקוח מסוג זה ממתין יותר מ- 15 שניות, הוא מועבר לתור כללי, ואז סביר שיענה מיידית (ע"י מוקדנים פנויים מהמוקדים האחרים). מאותה סיבה ניתן לראות peak ב - 6 שניות בהיסטוגרמה של לקוחות EBO.

שאלה 3 – יישום (20 נקודות)

בתי משפט צבאיים

להלן כתבה שהתפרסמה בעיתון הארץ בדצמבר 2011.

אין צורך לקרוא את הכתבה, כל המידע הדרוש יופיע בשאלות עצמן.

מחקר של שני שופטים בדק את משך הזמן הממוצע של דיונים בתיקים

אורך הדיונים בבימ"ש צבאי: 43 דקות בממוצע

ברקו לקח ודהאן את נתוני שנת 2008 בבתי המשפט הצבאיים והתגלה כי 95.5% מ-4,000 התיקים שנדונו באותה השנה הסתיימו בהסדר או בהודאה, ללא התערבות כלל של בית המשפט הצבאי.

כפי שפורסם באחרונה ב"הארץ", דו"ח בתי המשפט הצבאיים בישראל לשנת 2010 לימד כי רק 0.26% מהתיקים מסתיימים בוויכוח מוחלט של הנאשם. עם זאת, בבתי המשפט טוענים כי מדובר בנתון מטעה בשל ריבוי עסקאות הטיעון.

גורם במערכת אמר ל"הארץ" כי בנייתו פריטי האישום שנשארו במחלוקת בין הצדדים, שיעור הוויכוח עומד על 29%. כך או כך, בשנת 2010, רק ב-82 תיקים נוהלו הוכחות ונכתבה הכרעת דין מפורשת.

עורך דין המרבה להופיע בבתי המשפט הצבאיים אמר ל"הארץ" כי "הנתונים מסתדרים. ידוע שמעטים מבקשים לנהל תיק על הפרות סדר. זמן ניהול המשפט, שבמהלכו יושב הנאשם במעצר עד תום ההליכים, ארוך יותר מפרק הזמן שניתן לקבל בעסקת טיעון. לכן תיק הפרות סדר מתנהל בפרות מחצי שעה. אין טעם לנהל אותו".

דובר צה"ל טירב לאשר שיחה עם עורך המחקר, רביסרן דהאן, בנימוק ששופטים אינם נוהגים לשוחח עם עיתונאים מטעמים אתיים.



תצלום: דניאל בר און / ג'ני

שכ"ח ב-45 דקות. השופטים בחנו גם את משך זמן הטיפול של תיקי מעצרים, העומד על 18 דקות. זהו נתון שיש לו מקבילה בישראל, שבו הטיפול בתיק מעצר עומד על שש דקות בלבד.

ממצאי המחקר משקפים את המציאות בבתי המשפט הצבאיים, שם הרוב המוחלט של התיקים לא מתנהלים כלל בבתי המשפט אלא מוכרעים בעסקת טיעון או בהודאה. במאמר אחר שפרסמו בעבר,

המחקר בחן כמה זמן התנהלו הדיונים בכל תיק על כל שלביו – מהקראה ועד גזר הדין

הקראה, דיונים שונים, הוכחות, טיעונים לעונש, הכרעת דין, גזר דין וכי'. הממצאים אינם כוללים זמן שיפוטי כמו הכנה לדיון וכתובת פסקי דין. הממצאים של לקח ודהאן מלמדים שהטיפול בתיקים בבתי המשפט הצבאי בעופר מתנהל במהירות שיא: 43.72 דקות

חיים לוינסון

זמן שמיעה ממוצע של תיק בבית המשפט הצבאי במחנה עופר עומד על 43 דקות – כך עולה ממחקר שנערך בבתי המשפט.

מדובר בפרק הזמן נטו שבו מתנהל תיק בבית המשפט, לא כולל תקופות המתנה בין דיון לדיון. התיקים שדיוניהם נמשכים פרק זמן טובי הם תיקים של נאשמים בפעילות חבלנית עוינת (פח"ע) בהרכב מורחב של שלושה שופטים.

המחקר נערך על ידי השופטים סגני-אלוף צבי לקח, חבר בית המשפט הצבאי לערעורים, ורביסרן אמיר דהאן, שופט בית המשפט הצבאי יהודה והאדריא על הסטיסטיקות בבתי המשפט.

השניים ערכו את המחקר ככסים למאמרם "הצעה למודל מתמטי לקביעת מספר השופטים הרצוי עבור כמות התיקים המצויה", שפורסם בגיליון האחרון של כתב העת "הפיקליט".

השניים ביקשו לערוך בחינה של פרק הזמן שבו מתנהל תיק בבתי המשפט, בהתאם למאפייניו, כדי לחשב כמה תיקים ניתן לתת לשופט. במסגרת המחקר הם בחנו 468 תיקים שונים ובאמצעות המחשב ברקו כמה זמן התנהלו הדיונים בכל תיק על כל שלביו:

מטרת המחקר תפעולית לגמרי: פיתוח מודל שישקף בבהירות מרבית ומציאותית את עומס העבודה בבתי המשפט הצבאיים. השופטים בבתי המשפט הצבאיים מטפלים במספר סוגי תיקים שונים: תיקי הפרות סדר, תיקים פליליים, תיקי תעבורה, תיקי מעצרים ועוד.

3.1 עורכי המחקר טוענים כי אי אפשר לכמת את עומס העבודה של שופט רק על סמך מספר התיקים בהם הוא מטפל. האם אתם מסכימים עם טענה זו? נמקו את תשובתכם.

תשובה:

אכן לא ניתן לקבוע את העומס על שופט רק על סמך מספר התיקים בהם הוא מטפל מאחר וכל תיק מצריך זמן טיפול אחר. חישוב העומס על השופטים צריך לקחת בחשבון גם את מספר התיקים בהם הוא מטפל וגם את משך זמן הטיפול הנדרש לכל אחד מהתיקים. טענה זו נכונה במיוחד לאור העובדה ששופטי בתי הדין הצבאיים מטפלים בתיקים מסוגים שונים וניתן להניח שכל סוג דורש משך טיפול אחר.

אפשר לחלק את הטיפול בתיק ל- 3 שלבים עוקבים:

1. עיון והכנה לפני דיונים.
 2. שמיעה – שמיעת טענות הצדדים, ראיותיהם וניהול הדיון באולם בית המשפט.
 3. כתיבה – כתיבת החלטת השופט במערכת.
- נתמקד כעת בזמן השמיעה. הטבלה הבאה מציגה את ממוצע זמן השמיעה וסטיית התקן בתיקים מסוגים שונים. מספר התיקים הכולל שנבדק במדגם זה הוא 468 תיקים.
- הערה:** בטבלה, המושג "תיקי הרכב" מתייחס לדיון ע"י הרכב של שלושה שופטים. מלבד תיקים אלו, הדיונים בכל שאר התיקים מתקיימים במעמד שופט יחיד.

טבלה 3

חותר מהות	גודל המדגם (מספר תיקים)	ממוצע זמן שמיעה בדקות	סטיית תקן
כללי	468	43.72	68.34
תיקי הרכב		189.20 (זמן נטו – יש להכפיל תוצאה זו במס' השופטים היושבים בהרכב)	175.32
תיקי דן יחיד פח"ע ¹⁹		60.12	47.94
תיקי דן יחיד פלילי ²⁰		46.096	43.7
תיקי הפרות סדר ²¹		29.34	24.55
תיקי תעבורה		18.8	6.7
תיקי שב"ח ²²		45.31	13.37
תיקי מעצרים ²³		18.17	13.19

3.2 בשנת 2010 נפתחו בבית משפט מסוים 1000 **תיקי הפרות סדר** ו-500 **תיקי הרכב**. בשנת 2011 נפתחו באותו בית משפט 1500 **תיקי הפרות סדר** ורק 400 **תיקי הרכב**. מה יהיה הגידול/קטון בעומס העבודה הנובע מניהול הדיונים בבית המשפט? פרטו את חישוביכם.

תשובה:

בשנת 2010 עומס העבודה הנגרם משני סוגי התיקים הנ"ל הינו:

$$R = \lambda_1 E[S_1] + \lambda_2 E[S_2] = 1000 \cdot 29.34 + 500 \cdot 189.2 \cdot 3 = 313,140 \text{ min} = 5219 \text{ hours}$$

בשנת 2011 עומס העבודה הנגרם משני סוגי התיקים הנ"ל הינו:

$$R = \lambda_1 E[S_1] + \lambda_2 E[S_2] = 1500 \cdot 29.34 + 400 \cdot 189.2 \cdot 3 = 271,050 \text{ min} = 4,517.5 \text{ hours}$$

לכן הקטון בעומס העבודה הינו:

$$5219 - 4517.5 = 701.5 \text{ hours}$$

3.3 האם ההנחה כי זמן השמיעה בכל אחד מסוגי התיקים מתפלג מעריכית סבירה או לא סבירה? ליתר דיוק, באילו סוגי תיקים סבירה הנחה זו ובאילו לא? נמקו.

תשובה:

נחשב את מקדם ההשתנות בכל אחד מסוגי התיקים:

$$CV = \frac{47.94}{60.12} = 0.797 \text{ - תיקי פח"ע, } CV = \frac{175.32}{189.2} = 0.927 \text{ - תיקי הרכב}$$

$$CV = \frac{24.55}{29.34} = 0.837 \text{ - תיקי הפרות סדר, } CV = \frac{43.7}{46.096} = 0.948 \text{ - תיקים פליליים}$$

$$CV = \frac{13.37}{45.31} = 0.295 \text{ - תיקי שב"ח, } CV = \frac{6.7}{18.8} = 0.356 \text{ - תיקי תעבורה}$$

$$CV = \frac{13.19}{18.17} = 0.726 \text{ - תיקי מעצרים}$$

להתפלגות מעריכית מקדם השתנות השווה ל-1, לפיכך ההנחה תיתכן כסבירה רק עבור סוגי התיקים בהם מקדם ההשתנות קרוב ל-1 – תיקי הרכב ותיקים פליליים.

3.4 במהלך חודש ינואר 2012 נפתחו בבית משפט מסוים 30 תיקי פח"ע ונסגרו 25. בסוף החודש היו בסה"כ 75 תיקי פח"ע פתוחים באותו בית משפט. מהו אחוז הזמן, מתוך כלל הזמן שתיק פח"ע נמצא בבית משפט שבו הוא נמצא בשלב השמיעה (כלומר בשלב שבו מתקיים דיון לגביו בבית משפט)?

הערה: משך השהות, בשעות, של תיק בבית משפט יחושב רק על פי שעות העבודה תקניות בחודש – 20 ימי עבודה, 8 שעות עבודה ביום, סה"כ 160 שעות עבודה בחודש.

תשובה:

$$\lambda = \frac{\lambda_{opened} + \lambda_{closed}}{2} = \frac{30 + 25}{2} = 27.5$$

קרוב לקצב מעבר תיקי פח"ע בבית המשפט: 27.5

קרוב לכמות תיקי הפח"ע בבית המשפט:

$$L_{end} = 75, L_{start} = L_{end} + \lambda_{closed} - \lambda_{opened} = 70$$

$$L = \frac{L_{end} + L_{start}}{2} = \frac{70 + 75}{2} = 72.5$$

$$E[W] = \frac{L}{\lambda} = \frac{72.5}{27.5} = 2.63 \text{ months}$$

מחוק Little, משך השהייה של תיק פח"ע בבית המשפט: 2.63 months

כלומר, אחוז הזמן שתיק פח"ע נמצא בדיון מתוך סך השהות שלו הוא כשעה מתוך למעלה מחודשיים וחצי, ליתר דיוק:

$$\frac{60.12}{2.63 \cdot 20 \cdot 8 \cdot 60} = 0.0024 = 0.24\%$$

חלק זה של העמוד נשאר ריק

בכתבה נטען כי מרבית התיקים שמגיעים לבתי הדין הצבאיים מסתיימים בהסדר או בהודאה, כלל ללא התערבות של בית המשפט הצבאי. לגבי תיקי הפרות סדר, למשל, נאמר כי: "ידוע שמעטים מבקשים לנהל תיק על הפרות סדר. זמן ניהול המשפט, שבמהלכו יושב הנאשם במעצר עד תום ההליכים, ארוך יותר מפרק הזמן שניתן לקבל בעסקת טיעון".

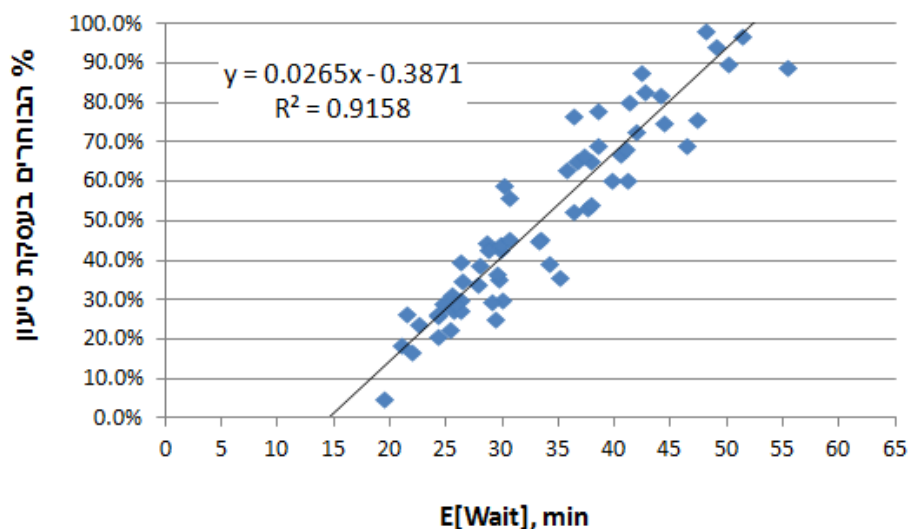
3.5 נניח כי נאשמים בתיקי הפרות סדר נכנסים לבית המשפט ע"פ תהליך פואסון עם קצב λ . במהלך ההמתנה לדיון בתיקים, כלומר, בטרם החל הדיון בתיק בבית המשפט, הנאשם יכול לבחור לגשת לעסקת טיעון. במקרה זה העסקה תיסגר בין עורך דינו לתובע, ללא דיון בבית המשפט. ניתן להניח כי ברגע שהנאשם בוחר לגשת לעסקת טיעון הוא יוצא מיד מהמערכת ואינו ממתיך עוד לטיפול בתיקו. בנוסף, נניח כי משך זמן הדיון בתיקים אלו מתפלג מעריכית עם תוחלת $E[S]$. איזה מודל מתמטי מתאים לתיאור מערכת הטיפול בתיקים אלו? מהן הנחות המודל?

תשובה:

המודל המתאים הוא מודל $M/M/N+M$ – Erlang A. הנחות המודל והפרמטרים:

- הגעות לפי תהליך פואסון עם קצב λ .
- משכי שירות אקספוננציאליים עם פרמטר $\mu = \frac{1}{E[S]}$.
- N שרתים (שופטים)
- יש נטישות מהמערכת (בחירה בעסקת טיעון בטרם התיק נדון בבית משפט נחשבת כנטישה). משך הזמן עד לנטישה מתפלג אקספוננציאלי עם פרמטר θ .
- אין תלות בין מרכיבי המערכת, דהיינו משכי השירות, קצבי ההגעה וקצבי הנטישה.

3.6 על סמך הגרף הבא, מהי תוחלת הזמן מרגע כניסת נאשם בתיק הפרות סדר לבית המשפט ועד שהנאשם בוחר לגשת לעסקת טיעון? נמקו את תשובתכם בעמוד הבא.



תשובה:

ניתן לראות כי מדובר במודל Two Phases of Patience, כאשר אנו למעשה מפרשים את הבחירה בעסקת טיעון כנטישה מהמערכת. עד 15 דקות מרגע פתיחת התיק אין כלל נטישות מהמערכת. מ-15

$$P(Ab) = \theta E[W_q] \text{ : הקשר}$$

כלומר קצב הנטישה הינו למעשה שיפוע הגרף ולכן, תוחלת הזמן עד לבחירה בעסקת טיעון, או

$$\text{Average Patience} = 15 + \frac{1}{\theta} = 15 + \frac{1}{0.0265} = 52.73 \text{ min}$$

3.7 על סמך טבלה 3 ניתן לראות כי משך הדיון הממוצע בתיקי הפרות סדר הוא 29.34 דקות, או כחצי שעה. לעיתים, בשל הפגנות גדולות חלה עליה משמעותית במספר תיקי הפרות הסדר. נניח כי בעת תקופות אלו, נאשמים בתיקי הפרות סדר נכנסים לבית המשפט ע"פ תהליך פואסון בקצב של שני נאשמים בשעה. בנוסף, נניח כי לטיפול בתיקי הפרות סדר מוקצה שופט אחד לאורך היום וכן נניח כי בתקופות אלו התפלגות הזמן מרגע כניסת נאשם בתיק הפרות סדר למערכת ועד שהנאשם בוחר לגשת לעסקת טיעון מתפלגת מעריכית עם ממוצע של כשעה.

על סמך הפלט הבא של 4CC, מה יהיה אחוז הנאשמים הבוחרים בעסקת טיעון? מה יהיה אחוז

הנאשמים אשר נאלצים לחכות בטרם יחל הדיון בעניינם?

File Table Settings Help

Performance Profiler | Staffing Query | Advanced Profiling | Advanced Queries | What-if Analysis

Performance Profiler Performance Profiler allows you to determine and optimize the Performance Level of your Call Center. Enter your call center's

Your Call Center's Parameters

- Number of Agents Answering Calls: 1
- Average Time to Handle One Call (mm:ss): 30:00
- Calls per 60 minute Interval: 2
- Average Callers' Patience (mm:ss): 59:00

Settings

- Features: Abandons
- Basic Interval: 60 minutes
- Target Time: 00:00 (mm:ss)

Change Settings

Compute | Add to Table | Delete Rows | Clear All | Export | Import | Graph

	Basic Interval (minutes)	Target Time to Answer	Number of Agents	Average Handling Time	Calls per Interval	Average Patience	Agent's Occupancy	%Answer	%Abandon	%Answer within Target	Average Queue Length
Results	60.0	00:00.0	1.0	30:00.0	2.0	59:00.0	68.6%	68.6%	31.4%	31.4%	.6
1											

תשובה:

אחוז הבוחרים בעסקת טיעון הינו למעשה אחוז הנוטשים. ע"פ הפלט אחוז הנוטשים הינו 31.4%.

אחוז הנאשמים הנאלצים לחכות לפני הדיון בעניינם הינו:

$$100 - \%Answer \text{ within target} = 100 - 31.4 = 68.6\%$$

3.8 על סמך גרף הפונקציות של Garnett והגרף המתאר את $P(Ab)\sqrt{n}$ כפונקציה של β , המופיעים בסוף השאלה, וכן על סמך הנתונים בסעיף 3.7, מהו אחוז הנאשמים הבוחרים בעסקת טיעון? מה יהיה אחוז הנאשמים אשר נאלצים להמתין בתור בטרם יחל הדיון בעניינם? האם יש הבדל בין תשובתכם בסעיף זה לתשובתכם בסעיף 3.7? אם כן, ממה הוא נובע?

תשובה:

$$R = \lambda E[S] = \frac{2 \cdot 30}{60} = 1$$
 ראשית נחשב את העומס המוצע:

$$\beta = \frac{1-R}{\sqrt{R}} = \frac{1-1}{\sqrt{1}} = 0$$
 נחשב את פרמטר דרגת השירות:

$$\text{לפי הסעיף הקודם } \theta = \frac{1}{60} = 0.0166 \text{ וכן } \mu = \frac{1}{E[S]} = \frac{1}{30} = 0.033$$
 , ולכן , ע"פ הסעיף הקודם,

$$\frac{\theta}{\mu} = \frac{0.0166}{0.033} = 0.5$$

על פי הגרף המתאר את $P(Ab)\sqrt{n}$ כפונקציה של β עם $GMR(0.5)$ ו- $\beta = 0$, נקבל כי $P(Ab)\sqrt{n} \approx 35\%$, ומכיון ש- $n=1$, $P(Ab) \approx 35\%$. על סמך פונקציית Garnett עם $GMR(0.5)$ ו- $\beta = 0$, $P\{Wait > 0\} \approx 60\%$.

בסעיף הקודם קיבלנו $P(Ab) \approx 30\%$ וכי $P\{Wait > 0\} \approx 68\%$. הסיבה להבדל היא שכאשר אנו משתמשים בכלל השורש ובפונקציית Garnett אנו למעשה משתמשים בקירובים אסימפטוטיים. במקרה שלנו מספר השרתים נמוך מאוד ולכן הקירובים האסימפטוטיים אינם נותנים מענה מספיק טוב במקרה זה.

3.9 היעזרו בדפי ההרצאה המופיעים בסוף השאלה על מנת לחשב את נצילות השופטים? האם תוצאה זו קרובה לנצילות השופטים המופיעה בפלט ה- 4CC בסעיף 3.7?

תשובה:

על סמך נוסחאות Erlang A ב- QED :

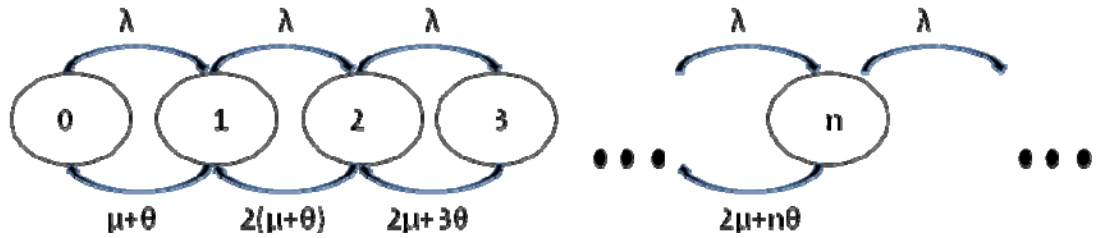
$$P\{Ab\} \approx \frac{\gamma}{\sqrt{n}} \Rightarrow \gamma \approx 0.35$$

$$Occupancy \approx 1 - \frac{\beta + \gamma}{\sqrt{n}} \approx 1 - \frac{0 + 0.35}{\sqrt{1}} = 0.65 = 65\%$$

אנו רואים כי כאן יש הבדל קטן (פער של כ- 5%) בין התוצאה שהתקבלה ע"י שימוש בקירובים לתוצאה שקיבלנו מ- 4CC.

3.10 כעת נניח כי הנאשם יכול לבחור לגשת לעסקת טיעון גם אם כבר החל הטיפול בתיקו. כיצד ניתן לתאר בעזרת תק"מ את מספר תיקי הפרות הסדר הנמצאים במערכת? בתשובתכם הגדירו במדויק את הפרמטרים של התק"מ. הניחו שמדובר במערכת עם 2 שופטים, לגבי הפרמטרים האחרים: λ , μ , θ , אין צורך לרשום את הערכים המספריים שלהם, רק להגדירם.

תשובה:



פרמטרים:

λ - קצב כניסת תיקי הפרות סדר

μ - קצב טיפול בתיקי הפרות סדר

θ - קצב נטישה מהמערכת עקב בחירה בעסקת טיעון.

חלק זה של העמוד נשאר ריק

3.11 נמצא שפונקצית הסיכון (Hazard Rate) של זמן השהייה במעצר בתיקי הפרות סדר הינה

מהצורה: $h_x(t) = \frac{2}{1+t}$, $0 \leq t < \infty$. עבור התפלגות זו ניתן להראות שממוצע משך השהות במעצר

הוא יום אחד (אין צורך להוכיח טענה זו).

בהינתן שעברו a ימים במעצר, אמד נאיבי לשארית הזמן במעצר הוא $\max\{1-a, 0\}$. מה דעתכם

על אמד זה? בתשובתכם השתמשו בדוגמה הבאה:

בהינתן שעבר יום אחד במעצר (התוחלת הנ"ל), מהי תוחלת משך השהייה הנוטר במעצר?

תשובה:

ע"פ האמד לעיל, אם עבר יום אחד במעצר ($a=1$) תוחלת משך השהייה הנוטר תהיה

$\max\{1-1, 0\} = 0$. ע"פ אמד זה כל עוד עבר יותר מיום אחד במעצר (כלומר משך השהייה שעבר

עולה על תוחלת השהייה) תוחלת משך השהייה הנוטר תהיה 0, ללא תלות בכמה ימים כבר עברו במעצר. אמד למשך השהייה הנוטר במעצר צריך לחשב את ממוצע (או התפלגות) משך השהייה הנוטר בהינתן שמשך השהייה עד כה היה a ימים. אם X הוא משך השהייה הכולל במעצר, האמד הטבעי

שנציע לשארית הזמן במעצר הוא $E(X-a | X \geq a)$.

במקרה הספציפי שלפנינו: פונקצית הסיכון של זמן השהייה במעצר הינה פונקציה יורדת, כלומר, ככל שעובר הזמן, הסיכוי שהשהייה במעצר תסתיים, בהינתן שנמשכה עד כה, הולך וקטן. לפיכך, תוחלת משך השהייה במעצר, בהינתן שעברו כבר a ימים הולכת ועולה ככל ש- a עולה. כלומר לאחר יום במעצר תוחלת משך השהייה הנוטר במעצר תהיה גבוהה מיום אחד. תופעה זו למעשה הפוכה לאינטואיציה שאומרת כי ככל שעובר הזמן תוחלת משך השהייה הנוטר תלך ותקטן.

$$h_x(t) = \frac{2}{1+t} = \frac{f_x(t)}{1-F_x(t)}$$

$$\Rightarrow S_x(t) = 1 - F_x(t) = e^{-\int_0^t h(x) dx} = e^{-\int_0^t \frac{2}{1+x} dx} = e^{-2 \ln(1+t)} = \frac{1}{(1+t)^2}$$

$$E(X) = \int_0^\infty S(x) dx = \int_0^\infty \frac{1}{(1+x)^2} dx = -\frac{1}{(1+x)} \Big|_0^\infty = 1$$

לפי נוסחת הזנב: $= 1$

$$F_{X|X \geq a} = 1 - P(X \geq x | X \geq a) = \begin{cases} 1 - \frac{P(X \geq x)}{P(X \geq a)} = 1 - \frac{(1+a)^2}{(1+x)^2}, & x \geq a \\ 0, & 0 \leq x < a \end{cases}$$

$$E(X | X \geq a) = \int_0^\infty 1 - F_{X|X \geq a}(x) dx = \int_0^a 1 dx + \int_a^\infty \frac{(1+a)^2}{(1+x)^2} dx = a + \left[-\frac{(1+a)^2}{(1+x)} \right]_a^\infty = a + 1 + a = 2a + 1$$

$$E(X-a | X \geq a) = a + 1 \Leftarrow$$

The Erlang-A Queue in the QED-Regime

Theorem (with Garnett & Reiman, 2002)

The following **points of view** are equivalent:

- 0. **QED:** $P\{W_q > 0\} \approx \alpha$, for some $0 < \alpha < 1$;
- 1. **Manager:** $n \approx R + \beta\sqrt{R}$, for some $-\infty < \beta < \infty$;
- 2. **Servers:** Occupancy $\approx 1 - \frac{\beta + \gamma}{\sqrt{n}}$;
- 3. **Customers:** $P\{Ab\} \approx \frac{\gamma}{\sqrt{n}}$, for some $0 < \gamma < \infty$;

in which case

$$\alpha = \alpha\left(\beta, \frac{\mu}{\theta}\right) = \left[1 + \sqrt{\frac{\theta}{\mu}} \cdot \frac{h(\hat{\beta})}{h(-\beta)}\right]^{-1},$$

which we call the **Garnett Delay-Function(s)**;

here $\hat{\beta} \triangleq \beta\sqrt{\frac{\mu}{\theta}}$, and

$$\gamma = \alpha \cdot \sqrt{\frac{\theta}{\mu}} \cdot [h(\hat{\beta}) - \hat{\beta}].$$

The Erlang-B Queue in the QED-Regime

Recall the Erlang-B Formula:

$$E_{1,n} \triangleq \text{P}\{\text{Blocked}\} = \frac{R^n}{n!} \bigg/ \sum_{j=0}^n \frac{R^j}{j!}$$

Theorem (QED Erlang-B; Jagerman, 1974)

As $n \rightarrow \infty$, the following 3 statements are equivalent:

1. **Customers:** $E_{1,n} \approx \frac{\gamma}{\sqrt{n}}$, for some $\gamma > 0$;
2. **Servers:** $\rho \approx 1 - \frac{\beta}{\sqrt{n}}$, for some $-\infty < \beta < \infty$;
3. **Manager:** $n \approx R + \beta\sqrt{R}$ (square-root “staffing”);

in which case

$$\gamma = h(-\beta) = \frac{\phi(-\beta)}{\bar{\Phi}(-\beta)} = \frac{\phi(\beta)}{\Phi(\beta)},$$

where $\phi, \Phi, \bar{\Phi}$ and h are the density, cdf, survival function and hazard rate of $N(0, 1)$ (standard-normal), respectively.

Note: **Servers’ Occupancy** $\approx 1 - \frac{\beta + \gamma}{\sqrt{n}}$, accounting for blocking.

The Erlang-C Queue in the QED-Regime

Theorem (QED Erlang-C; Halfin & Whitt, 1981)

As $n \rightarrow \infty$, the following 4 statements are equivalent:

0. **QED:** $E_{2,n} \approx \alpha$, for some $0 < \alpha < 1$;
1. **Manager:** $n \approx R + \beta\sqrt{R}$, for some $0 < \beta < \infty$;
2. **Servers:** $\rho \approx 1 - \frac{\beta}{\sqrt{n}}$;
3. **Customers:** $E[W_q | W_q > 0] \approx \frac{1}{\sqrt{n}} \cdot \frac{1}{\mu\beta}$;

in which case

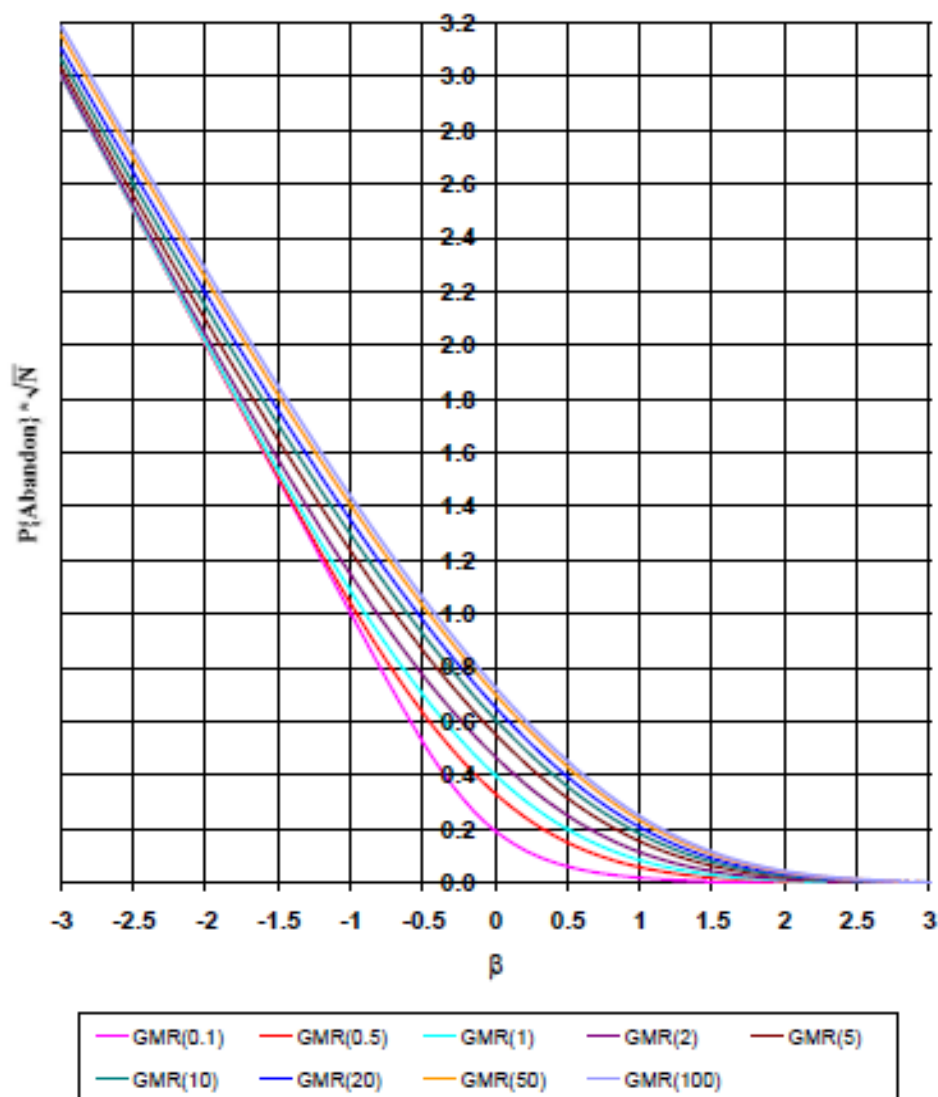
$$\alpha = \alpha(\beta) = \left[1 + \frac{\beta}{h(-\beta)} \right]^{-1},$$

which we call the **Halfin-Whitt Delay-Function**.

Note: $\beta\sqrt{R}$ = **Safety-Staffing**, in analogy to Safety-Stock.

Erlang-A: % Abandonment

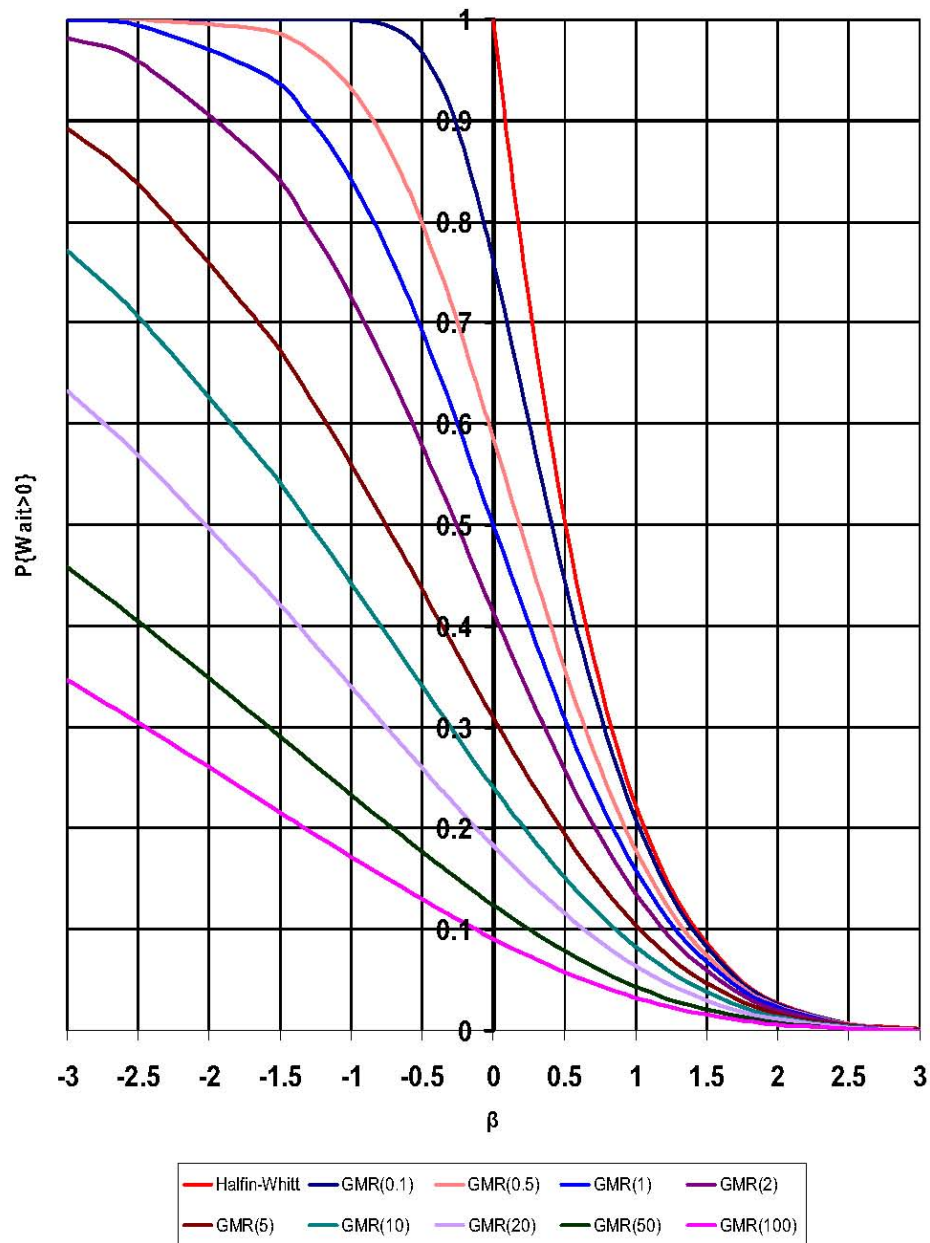
$\%Ab \times \sqrt{n}$ vs. β , for varying (im)patience (θ/μ):



Note the behavior: slope $-\beta$, for (relatively) large negative β and over all (im)patience levels. For an explanation, think **ED**: $n = R + \beta\sqrt{R} = R - \gamma R$; hence $\gamma \approx -\beta/\sqrt{R} \approx -\beta/\sqrt{n}$, and γ is $P\{Ab\}$ in the ED-Regime.

Erlang-A: The Garnett Delay-Functions

$P\{W_q > 0\}$ vs. the QOS parameter β , for varying patience θ/μ .



GMR(x) describes the asymptotic probability of delay as a function of β when $\frac{\theta}{\mu} = x$. Here, θ and μ are the abandonment and service rate, respectively.

Note: **Erlang-C** = limit of **Erlang-A**, as patience $\uparrow$ indefinitely.

שאלה 4 – תיאוריה (10 נקודות)

נניח מודל Erlang A עם פרמטרים N, θ, μ, λ , המתאר מערכת שירות. נניח בנוסף, בשאלה זו, ש- $\mu = \theta$ (אלא אם נאמר אחרת).

נסמן ב- $L(t)$ את מספר האנשים במערכת בזמן t . אזי $L = \{L(t), t \geq 0\}$ הוא תהליך לידה ומוות.

4.1 ניתן להראות שלתהליך L התפלגות סטציונארית פואסונית (אין צורך להוכיח).

נסמן את הפרמטר של ההתפלגות הפואסונית ב- R .

מה משמעותו של R כמדד תפעולי של מערכת השירות? מהו ערכו של R (כפונקציה של הפרמטרים בבעיה)? חשבו ונמקו.

תשובה:

כאשר $\mu = \theta$, Erlang A מתנהג כמו $M / M / \infty$ ולכן, כאמור בשאלה, ההתפלגות הסטציונארית היא פואסונית. $\text{Offered Load} = R$, העומס המוצע למערכת.

$$R = \lambda \cdot E[S] = \frac{\lambda}{\mu}$$

4.2 מהו מספר השרתים המינימלי שנדרש כדי להבטיח שלפחות 50% מהלקוחות יקבלו שירות מיידי (ללא המתנה בתור).

תשובה:

נסמן $L(\infty) = \text{Poisson}(R)^d$, ההתפלגות הסטציונארית הנ"ל. רוצים N מינימאלי עבורו $P(L(\infty) \geq N) \leq 0.5$.

$$P(L(\infty) \geq N) = P\left(\frac{L(\infty) - R}{\sqrt{R}} \geq \frac{N - R}{\sqrt{R}}\right) \approx P\left(N(0,1) \geq \frac{N - R}{\sqrt{R}}\right) \approx 0.5 \Leftrightarrow \frac{N - R}{\sqrt{R}} \approx 0$$

זאת אומרת, $N = \lceil R \rceil = \left\lceil \frac{\lambda}{\mu} \right\rceil$ הוא המינימום הנדרש (מספר השרתים שווה לעומס המוצע).

4.3 כיצד תשתנה תשובתכם לסעיף 4.2 אם ללקוחות המערכת יש סבלנות אינסופית (אין נטישות)? הסבירו תשובתכם תוך שימוש בגרף הפונקציות של Garnett המופיעות בעמוד 29.

תשובה:

סבלנות אינסופית (שפירושה $\theta = 0$ או $1/\theta = \infty$) מביאה למודל Erlang C. הפונקציה של Garnett המתאימה היא הפונקציה של Halfin-Whitt (העליונה בגרף). לשרות מיידי של 50% מהפונים נדרשת רמת איוש של $N = R + 0.5\sqrt{R}$ ($\beta = 1/2$). ז"א, נדרשים $0.5\sqrt{R}$ יותר שרתים כדי להבטיח 50% שירות מיידי (מכיוון שלקוחות אינם נוטשים בעצמם).

טבלה 1

call_id	q_start	q_exit	ser_start	ser_exit	outcome	wait_time	service_time	Agent
510013309	06:23:00	06:23:07	06:23:08	06:25:01	Agent	7	113	23015
510013355	06:23:19	06:24:30	06:24:31	06:24:51	Agent	71	20	23007
510013559	06:24:52	06:24:52	06:24:52	06:26:57	Agent	0	125	23007
510013814	06:26:32	06:26:33	06:26:33	06:27:27	Agent	1	54	23015
510013830	06:26:36	06:26:38	00:00:00	00:00:00	Abandon	2	0	0
510013871	06:26:47	06:26:57	06:26:58	06:28:45	Agent	10	107	23007
510014006	06:27:53	06:27:53	06:27:53	06:29:46	Agent	0	113	23015
510014071	06:28:15	06:28:45	06:28:46	06:29:04	Agent	30	18	23007
510014116	06:28:35	06:29:04	06:29:05	06:30:31	Agent	29	86	23007
510014132	06:28:43	06:29:46	06:29:47	06:31:27	Agent	63	100	23015
510014196	06:29:05	06:30:31	06:30:32	06:34:32	Agent	86	240	23007
510014425	06:30:38	06:31:27	06:31:27	06:34:31	Agent	49	184	23015
510014578	06:31:32	06:34:31	06:34:32	06:35:54	Agent	179	82	23015
510014674	06:32:06	06:34:32	06:34:33	06:37:04	Agent	146	151	23007
510014702	06:32:18	06:33:00	00:00:00	00:00:00	Abandon	42	0	0
510014712	06:32:26	06:35:54	06:35:55	06:37:37	Agent	208	102	23015
510014750	06:32:37	06:36:08	06:36:09	06:47:58	Agent	211	709	23041
510014927	06:33:46	06:37:04	06:37:05	06:43:49	Agent	198	404	23007
510015175	06:35:19	06:36:04	00:00:00	00:00:00	Abandon	45	0	0
510015368	06:36:28	06:37:37	06:37:38	06:39:53	Agent	69	135	23015
510015393	06:36:35	06:37:54	00:00:00	00:00:00	Abandon	79	0	0
510015405	06:36:39	06:39:53	06:39:54	06:51:04	Agent	194	670	23015
510015441	06:36:55	06:43:49	06:43:50	06:45:58	Agent	414	128	23007
510015539	06:37:41	06:45:58	06:45:59	06:47:44	Agent	497	105	23007
510015752	06:38:54	06:39:14	00:00:00	00:00:00	Abandon	20	0	0
510015802	06:39:13	06:40:31	00:00:00	00:00:00	Abandon	78	0	0
510016057	06:40:42	06:40:59	00:00:00	00:00:00	Abandon	17	0	0
510016208	06:41:36	06:47:44	06:47:45	06:51:29	Agent	368	224	23007
510016446	06:42:56	06:47:58	06:47:59	06:48:31	Agent	302	32	23041
510016552	06:43:35	06:45:36	00:00:00	00:00:00	Abandon	121	0	0
510016619	06:44:02	06:48:31	06:48:32	06:50:56	Agent	269	144	23041
510016625	06:44:05	06:46:13	00:00:00	00:00:00	Abandon	128	0	0
510016700	06:44:32	06:50:56	06:50:57	06:56:20	Agent	384	323	23041
510016708	06:44:36	06:51:04	06:51:05	06:53:01	Agent	388	116	23015
510016925	06:45:53	06:51:29	06:51:30	06:53:14	Agent	336	104	23007
510017302	06:48:04	06:53:01	06:53:02	06:54:12	Agent	297	70	23015
510017389	06:48:31	06:49:30	00:00:00	00:00:00	Abandon	59	0	0
510017423	06:48:42	06:53:14	06:53:14	06:54:56	Agent	272	102	23007
510017606	06:49:55	06:54:12	06:54:12	06:55:49	Agent	257	97	23015
510017636	06:50:03	06:51:57	00:00:00	00:00:00	Abandon	114	0	0
510017717	06:50:38	06:54:56	06:54:57	06:56:50	Agent	258	113	23007

טבלה 2

time	Arr_sys	Dep_sys	Arr_q	Dep_q	Arr_ser	Dep_ser
06:15	6	6	6	6	5	5
06:16	6	6	6	6	5	5
06:17	6	6	6	6	5	5
06:18	6	6	6	6	5	5
06:19	7	6	7	7	6	5
06:20	7	6	7	7	6	5
06:21	7	6	7	7	6	5
06:22	8	6	8	8	7	5
06:23	9	6	9	9	8	5
06:24	11	7	11	10	9	6
06:25	12	9	12	12	11	8
06:26	12	10	12	12	11	9
06:27	15	12	15	15	13	10
06:28	16	13	16	16	14	11
06:29	19	14	19	17	15	12
06:30	20	16	20	19	17	14
06:31	21	17	21	20	18	15
06:32	22	18	22	21	19	16
06:33	26	18	26	21	19	16
06:34	27	19	27	22	19	16
06:35	27	21	27	24	21	18
06:36	28	22	28	25	22	19
06:37	32	24	32	27	23	20
06:38	33	27	33	30	25	22
06:39	34	27	34	30	25	22
06:40	35	29	35	32	26	23
06:41	36	31	36	34	26	23
06:42	37	31	37	34	26	23
06:43	38	31	38	34	26	23
06:44	39	32	39	35	27	24
06:45	43	32	43	35	27	24
06:46	44	34	44	37	28	25
06:47	44	35	44	38	28	25
06:48	44	37	44	40	30	27
06:49	47	38	47	41	31	28
06:50	48	39	48	42	31	28
06:51	51	40	51	43	32	29
06:52	53	43	53	46	34	31
06:53	54	43	54	46	34	31
06:54	54	45	54	48	36	33
06:55	54	47	54	50	38	35
06:56	56	49	56	52	39	36
06:57	59	52	59	55	41	38
06:58	59	53	59	56	42	39
06:59	60	54	60	57	43	40
07:00	60	55	60	59	44	40
07:01	62	57	62	62	46	41
07:02	64	57	64	64	48	41
07:03	70	62	70	68	51	45
07:04	70	63	70	70	53	46
07:05	71	65	71	71	54	48
07:06	74	67	74	74	57	50
07:07	79	69	79	78	61	52
07:08	80	72	80	80	63	55

Table 1: Discrete and Continuous Distributions

Distribution	Notation	Probability Function	Parameters	Support	Moment Gen. Function	Expectation	Variance	Comments
Binomial	$N \sim \text{Bin}(n, p)$	$p(k) = \binom{n}{k} p^k q^{n-k}$	$0 \leq p = 1 - q \leq 1; n \geq 1$	$k = 0, 1, \dots, n$	$[p \cdot e^t + q]^n$	np	npq	
Multi-nomial	$X \sim \text{Mult}(n, p)$	$p(x) = \frac{n!}{x_1! \dots x_k!} \prod_{i=1}^k p_i^{x_i}$ $x = (x_1, \dots, x_k)$	$0 \leq p_i; \sum_{i=1}^k p_i = 1; n \geq 1$ $p = (p_1, \dots, p_k)$	$0 \leq x_i; \sum_{i=1}^k x_i = n$		$E(X_i) = np_i$ $i = 1, \dots, k$	$\text{cov}(X_i, X_j) = -np_i p_j$ $(i \neq j)$	$X_i \sim \text{Bin}(n, p_i)$
Negative Binomial	$X \sim NB(m, p)$	$p(k) = \binom{k-1}{m-1} p^m q^{k-m}$	$0 \leq p = 1 - q \leq 1; m \geq 1$	$k = m, m+1, \dots$	$\left[\frac{p \cdot e^t}{1 - q \cdot e^t} \right]^m$	$\frac{mp}{p}$	$\frac{mq}{p^2}$	Number of trials till m -th success
Geometric	$X \sim G(p)$	$p(k) = q^{k-1} \cdot p$	$0 \leq p \leq 1$	$k = 1, 2, \dots$	$\frac{p \cdot e^t}{1 - q \cdot e^t}$	$\frac{1}{p}$	$\frac{q}{p^2}$	$X \sim NB(1, p)$
Poisson	$X \sim \mathcal{P}(\lambda)$	$p(k) = \frac{e^{-\lambda} \lambda^k}{k!}$	$\lambda \geq 0$	$k = 0, 1, \dots$	$\exp\{-\lambda(1 - e^t)\}$	λ	λ	
Hyper-geometric	$X \sim HG(N, M, n)$	$p(k) = \frac{\binom{N}{k} \binom{M}{n-k}}{\binom{N+M}{n}}$	$N, M > 0;$ $1 \leq n \leq N + M$	$\max(0, n - M) \leq k$ $k \leq \min(n, N)$	$\sum_k e^{tk} p(k)$	$\frac{n \cdot N}{N+M}$	$n \cdot \frac{N}{N+M} \cdot \frac{M}{N+M} \cdot \frac{N+M-n}{N+M-1}$	N – special elements M – ordinary ones n – sample size k – special in sample
		Density Function						
Beta	$X \sim \text{Beta}(\alpha, \beta)$	$f(t) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} t^{\alpha-1} (1-t)^{\beta-1}$	$\alpha, \beta > 0$	$0 \leq t \leq 1$	Complicated	$\frac{\alpha \beta}{\alpha + \beta}$	$\frac{\alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$	$\alpha = \beta = 1 \rightarrow \text{Uniform}$
Gamma	$X \sim \text{Gamma}(\alpha, \lambda)$	$f(t) = \frac{\lambda^\alpha t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha)}$	$\alpha, \lambda > 0$	$t \geq 0$	$(1 - \frac{t}{\lambda})^{-\alpha}$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$	$\Gamma(x) = \int_0^\infty u^{x-1} e^{-u} du$ $= (x-1)! \Gamma(x-1)$
Erlang	$X \sim \text{Erlang}(n, \lambda)$	$f(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}$	$n = 1, 2, \dots; \lambda > 0$	$t \geq 0$	$(1 - \frac{t}{\lambda})^{-n}$	$\frac{n}{\lambda}$	$\frac{n}{\lambda^2}$	$X \sim \text{Gamma}(n, \lambda)$
Exponential	$X \sim \text{Exp}(\lambda)$	$f(t) = \lambda e^{-\lambda t}$	$\lambda > 0$	$t \geq 0$	$(1 - \frac{t}{\lambda})^{-1}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$X \sim \text{Gamma}(1, \lambda)$
Normal	$X \sim N(\mu, \sigma^2)$	$f(t) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(t-\mu)^2}{2\sigma^2} \right\}$	$-\infty < \mu < \infty; \sigma > 0$	$-\infty < t < \infty$	$\exp \left\{ t\mu + \frac{1}{2} t^2 \sigma^2 \right\}$	μ	σ^2	