



מרצה : פרופסור אבישי מנדלבאום
מתרגל : גנדי שייחט

תאריך הבחינה : 26.02.2006

מס' תלמיד _____

שם _____

הנדסת מערכות שירות - 096324

מועד א' - סמסטר חורף תשס"ו 2006 - פתרון

- המבחן נמשך שלוש שעות. באחריותך להחזירו במועד.
- המבחן בחומר סגור, וללא מחשבון.
- המבחן כולל 27 עמודים, כולל עמוד זה.
- 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.

ניקוד: שאלה 1	_____ מתוך 8	נקודות זמינות (רמת שירות תפעולית) בבית חולים
שאלה 2	_____ מתוך 15	נוסחת Little, מודלים אמפיריים
שאלה 3	_____ מתוך 15	פונקציית קצב הסיכון (Hazard Rates)
שאלה 4	_____ מתוך 10	מערכות תורים, 4CallCenters
שאלה 5	_____ מתוך 10	מדידות, תחומים תפעוליים
שאלה 6	_____ מתוך 10	מודלים סטוכסטיים במערכות תורים
שאלה 7	_____ מתוך 10	מודלים של נוזלים
שאלה 8	_____ מתוך 7	M/G/1 נקודות
שאלה 9	_____ מתוך 15	QED, Erlang-A, איוש אופטימלי

סה"כ _____ מתוך 100 נקודות

בהצלחה!

שאלה 1: זמינות (רמת שירות תפעולית) בבית חולים (8 נקודות)

מספר המיטות לאישפוז בבית חולים מהווה את המדד העיקרי לקיבולת (capacity) בית החולים. בהתאם, נצילות ממוצעת של מיטה בבית החולים מהווה בסיס להחלטות על הצורך בשינוי קיבולת: הרחבה או צמצום. נגדיר נצילות ממוצעת של מיטה כפרופורצית הזמן, לאורך זמן, שהמיטה מאוישת. מטרה מקובלת, היא 85% נצילות מיטה. נגדיר גם מדד לזמינות ממוצעת של מיטה במחלקה (או רמת שירות תפעולית) כאחוז החולים, לאורך זמן, שנדרש לאשפוז במחלקה, אך לא ניתן עקב תפוסה מלאה.

1.1 נניח שקובעים את מספר המיטות במחלקה גדולה מאוד (עשרות מיטות) כך שנצילותן הממוצעת תהיה 85%. מה צפויה להיות רמת השירות התפעולית במחלקה זו? תנו תשובה איכותית, דהינו רמת שירות גבוהה / נמוכה / סבירה ונמקו בקצרה.

תשובה: רמת שירות תפעולית גבוהה

לכן: שהמיטה גבוהה, כוח נצילות 85%, תהיה ממוצעת שירות.

1.2 חזרו על סעיף 1.1, אך עבור מחלקה קטנה מאוד (מספר בודד של מיטות).

תשובה: רמת שירות תפעולית נמוכה

לכן: שהמיטה נמוכה, כוח נצילות 85%, תהיה ממוצעת שירות.

1.3 נניח כי בית חולים מציב לו כמטרה לקבוע את מספר המיטות במחלקותיו כך שרמת השירות התפעולית (זמינות המיטות) תהיה אחידה בינן. (למשל, עבור כל מחלקה, לא יותר מ-5% חולים שאישפוזם במחלקה נמנע עקב תפוסה מלאה). הציעו מדיניות לקביעת מספר המיטות, תוך ציון מודל מתמטי מתאים עליו מתבססת הצעתכם.

תשובה:

$$N = \frac{\lambda}{\mu - \beta} + \beta \sqrt{\frac{\lambda}{\mu - \beta}}$$

מספר מיטות או סמן כל האנשים
 $\lambda = \text{קצת מיטות חולים}$
 $\mu = 1 = \text{מחלקה אחת שמיטה אחת}$
 $\beta = \text{כוח נצילות יותר גבוה} \quad 1.5 < \beta < 1.5$

אז צנוי שמות החולים ה"נחשמים" הוא כ- $\lambda \sqrt{N} / (\mu - \beta)$
 $\lambda = \text{כוח נצילות יותר גבוה} \quad \mu = 1$

שאלה 2: נוסחת Little, מודלים אמפיריים (15 נקודות)

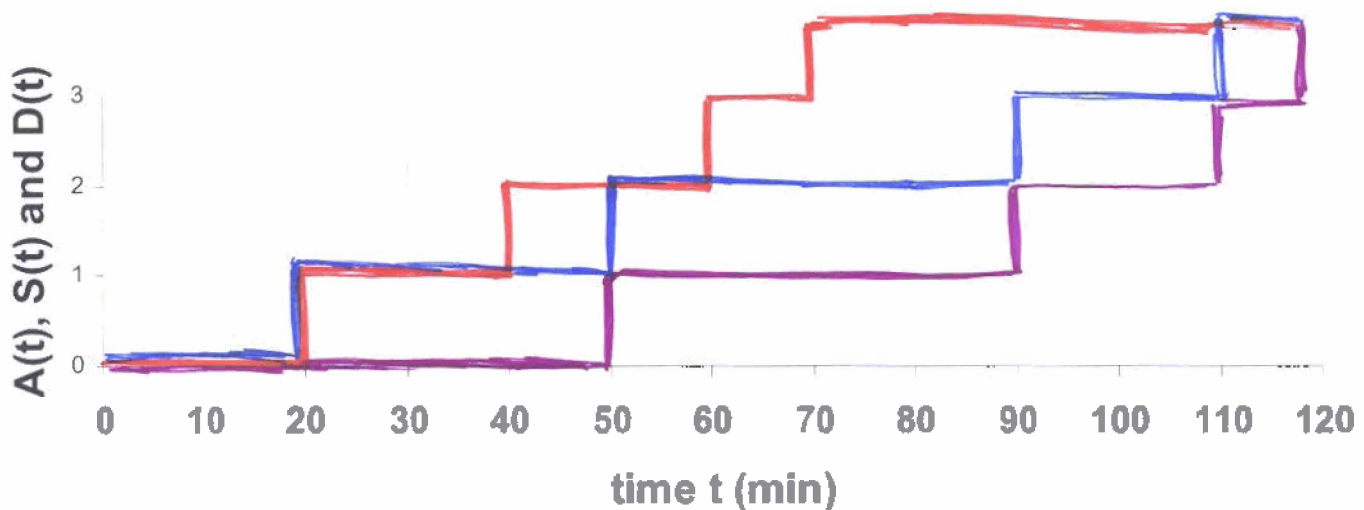
סוכנות נסיעות קטנה עם סוכן יחיד משרתת לקוחות לפי מדיניות First Come First Serve. הלקוחות המופיעים ומוצאים סוכן תפוס – מצטרפים לתור ממתינים. נניח כי לקוחות לא נוטשים. בפרק הזמן מ-8:00 עד ל-10:00 בבוקר הסוכנות עבדה כפי שמתואר בטבלה הבאה:

מספר לקוח	זמן מופע	משך שירות
לקוח 1	8:20	30 דקות
לקוח 2	8:40	40 דקות
לקוח 3	9:00	20 דקות
לקוח 4	9:10	10 דקות

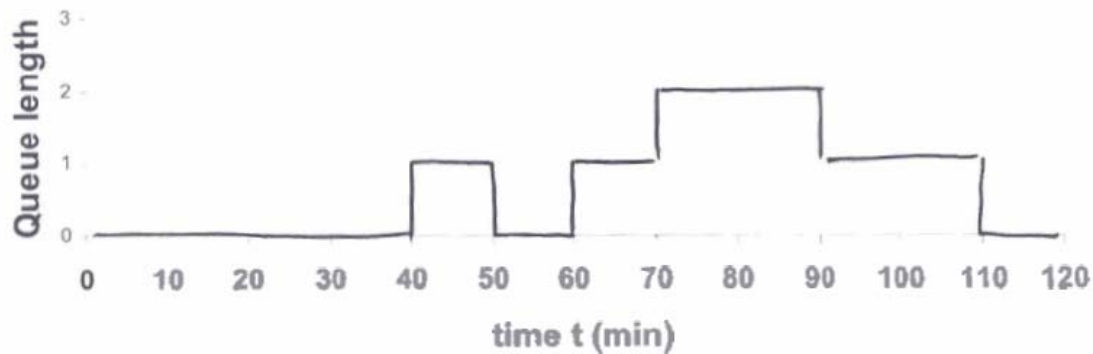
עבור זמן t כלשהו באינטרוול $[8:00, 10:00]$ נגדיר:

- $A(t)$ - סה"כ מספר המופעים מ-8:00 עד לזמן t ,
- $S(t)$ - סה"כ מספר הנכנסים לשירות מ-8:00 עד לזמן t ,
- $D(t)$ - סה"כ מספר סיומי השירות מ-8:00 עד לזמן t .
- $L_q(t)$ - מספר האנשים בתור בזמן t .

2.1 ציירו בגרף אחד את שלוש הפונקציות A , S ו- D .



2.2 ציירו את הפונקציה L_q .



2.3 השתמשו בנוסחת Little כדי לחשב את ממוצע זמן ההשהיה בסוכנות, בין השעות [8:00, 10:00].

$$\lambda = 4 \frac{\text{מחלק}}{\text{שעה}} = 2 \frac{\text{מחלק}}{\text{שעה}}$$

תשובה:

$$E(L_q) = \frac{1}{120} \int_0^{120} L_q(t) dt = \frac{2}{3}$$

$$E(W_q) = \frac{\lambda}{\mu} = \frac{2/3}{2} = \frac{1}{3} \text{ שעה} = 20 \text{ דקות}$$

$$E(W) = 20 + E(S) = 45 \text{ דקות}$$

2.4 נניח שבמקום המידע המפורש בטבלה, ידוע כי באינטרוול [08:00, 10:00] הגיעו בדיוק 4 לקוחות, אורך התור הממוצע היה כאורך התור הממוצע מסעיף 2.2 וכולם סיימו את שירותם עד השעה 10:00. האם די בנתונים אלו כדי לחשב את זמן ההמתנה הממוצע? את זמן ההשהיה הממוצע במערכת (המתנה+שירות)?

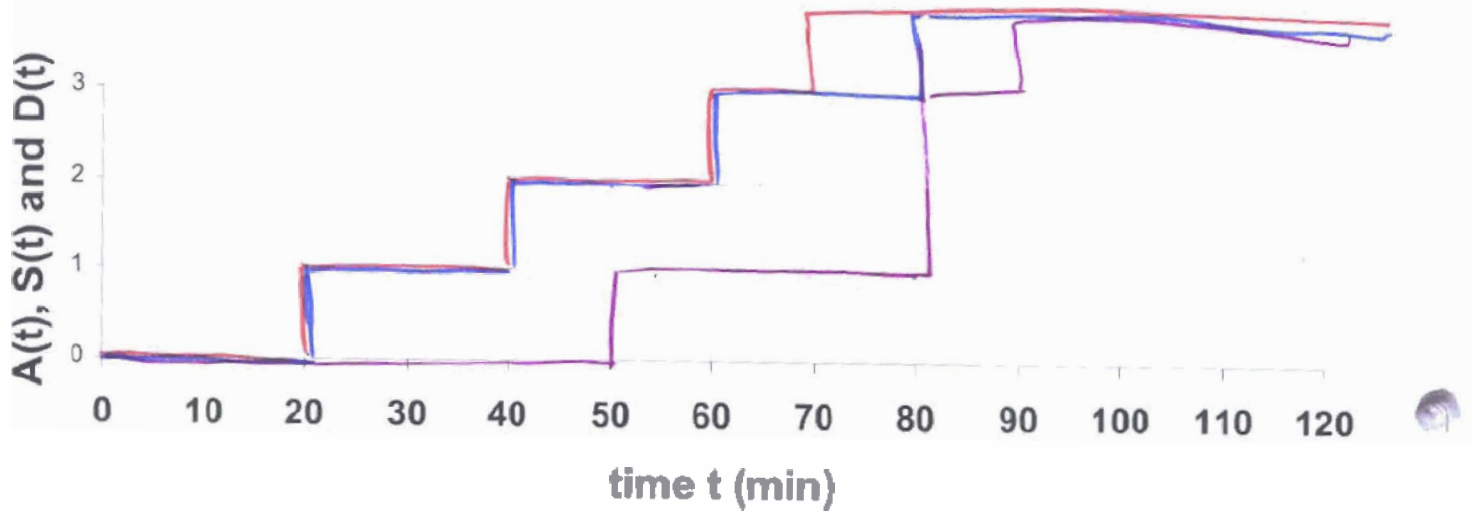
תשובה:

$$E(W_q) = 4 * 2/3$$

זמן המתנה ממוצע: $E(W_q) = 4 * 2/3$
את זמן ההשהיה הממוצע לא ניתן לחשב על סמך נתונים אלו בלבד, חסר נתון על אורך שירות ממוצע.

2.5 חזרו על סעיף 2.1 אך כאשר בסוכנות עובדים שני סוכנים . מה עתה ממוצע זמן השהיה בתחנה בשעות [8:00, 10:00] .

תשובה:

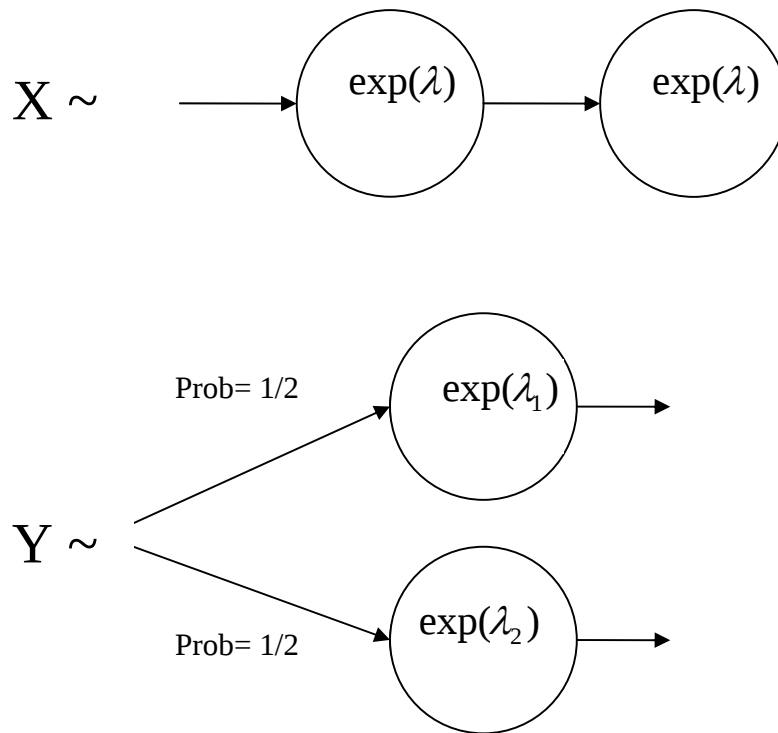


ממוצע זמן השהיה הוא:

27.5 שניות

שאלה 3: פונקציית קצב הסיכון (Hazard Rates) (15 נקודות)

נתונים שני משתנים מקריים X ו- Y המפולגים כפי שמתואר כדלהלן. דהינו, X מתפלג כמו סכום של שני מ"מ אקספוננציאלים בלתי תלויים עם פרמטר λ , ואילו Y מתפלג באופן הבא : בסיכוי $1/2$, Y מתפלג אקספוננציאלית עם פרמטר λ_1 ובסיכוי $1/2$, Y מתפלג אקספוננציאלית עם פרמטר λ_2 .



3.1 תנו שמות מדויקים ככל האפשר להתפלגויות אלו :

תשובה:

$Erlang(2, \lambda)$ מתפלג X

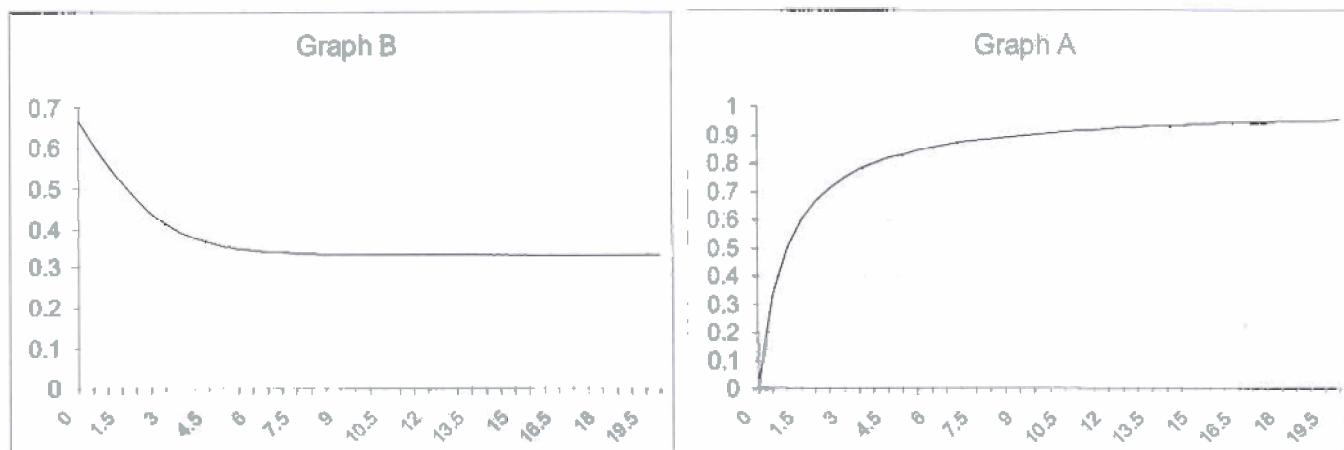
$Hyper\ exponential$ מתפלג Y

3.2 עבור משתנה מקרי Z אי-שלילי עם פונקציית צפיפות f_Z , תנו הגדרה ל- h_Z , פונקציית קצב הסיכון (ה- hazard rate) של Z . הביעו את תשובתכם בעזרת f_Z בלבד.

תשובה:

$$h_Z(t) = \frac{f_Z(t)}{\int_t^\infty f_Z(s)ds}, t \geq 0.$$

3.3 לפניכם שני גרפים המתארים את פונקציות ה-hazard rate של X ו-Y מתחילת השאלה. איזה גרף מתאים לאיזה משתנה? הסבירו בקצרה.



גרף A מתאר את ה-hazard rate של X

הסבר: ערך פונקציית הסיכון הריבה ל- $t=0$ מכיוון שהסיכוי לסיים את הפעולה בשלב ראשון הוא אקספוננציאלי (הוא אפסי). כאשר מתקדם הזמן, ההסתברות שנמצאים בפעולה השנייה עולה ל-1 ולכן פונקציית הסיכון מתקרבת לפונקציית ההסתברות הפעולה השנייה.

גרף B מתאר את ה-hazard rate של Y

הסבר: ערך פונקציית הסיכון ב- $t=0$ הוא מחצית הקצבים של שתי הפעולות הנפרדות $\frac{1}{2}g_1 + \frac{1}{2}g_2$. במידה והפעולה הראשונה היא הפעולה הראשונה, סביר להניח שהפעולה הראשונה היא הפעולה הראשונה. ולכן, פונקציית הסיכון מתקרבת לערך הפונקציית הסיכון של הפעולה הראשונה.

3.4 עבור הפרמטרים $\lambda = 1$, $\lambda_1 = 1$, $\lambda_2 = 1/3$, תארו בקצרה את ההתנהגות הגבולית (באינסוף) של הגרפים A ו-B, ז"א לאיזה ערך מתכנסים הגרפים כאשר $t \leftarrow \infty$. הסבירו בקצרה.

תשובה:

$$\lim_{t \rightarrow \infty} h_X(t) = 1$$

הסבר:

$$\lim_{t \rightarrow \infty} h_X(t) = \lim_{t \rightarrow \infty} \lambda \cdot P(X(t) \text{ in } 2 | X(t) \text{ in } 1 \text{ or } 2) = \lambda = 1$$

תשובה:

$$\lim_{t \rightarrow \infty} h_Y(t) = \frac{1}{3}$$

הסבר:

$$h_X(t) = \lambda_1 \cdot P(X(t) \text{ in } 1 | X(t) \text{ in } 1 \text{ or } 2) + \\ + \lambda_2 \cdot P(X(t) \text{ in } 2 | X(t) \text{ in } 1 \text{ or } 2) \rightarrow \lambda_2 = \frac{1}{3} \quad (\text{slowest rate})$$

3.5 (אמידת ה-hazard rate). נתונות N ריאליזציות (מדידות) של משתנה מקרי Z רציף וחיובי עם התפלגות לא ידועה. המטרה היא לאמוד את ה-hazard rate של Z על סמך מדידות אלו. נניח שכל המדידות שייכות לאינטרוול $[0, T]$. נחלק את הקטע $[0, T]$ לאינטרוולים קטנים באורך Δ כל אחד, ועבור כל אינטרוול קטן j , $(j = 1, 2, \dots)$ נגדיר:

- ♦ d_j - מספר המדידות שנפלו באינטרוול j ,
- ♦ r_{j-1} - סה"כ המדידות מתוך N , שלא שייכות ל- $j-1$ האינטרוולים הקודמים ($r_0 = 0$).

לפשטות נניח כי הצפיפות של Z נשארת קבועה בכל אינטרוול קטן.

3.5.1 תנו אמד לפונקצית ההישרדות (*survival function*) של Z באינטרוול j .

תשובה:

$$\bar{F}_Z(t) = \frac{r_{j-1} - \frac{1}{2}d_j}{n}, \quad t \in j. \text{ אינטרוול}$$

3.5.2 תנו אמד לצפיפות של Z באינטרוול j .

תשובה:

$$f_Z(t) = \frac{d_j}{n \cdot \Delta}, \quad t \in j. \text{ אינטרוול}$$

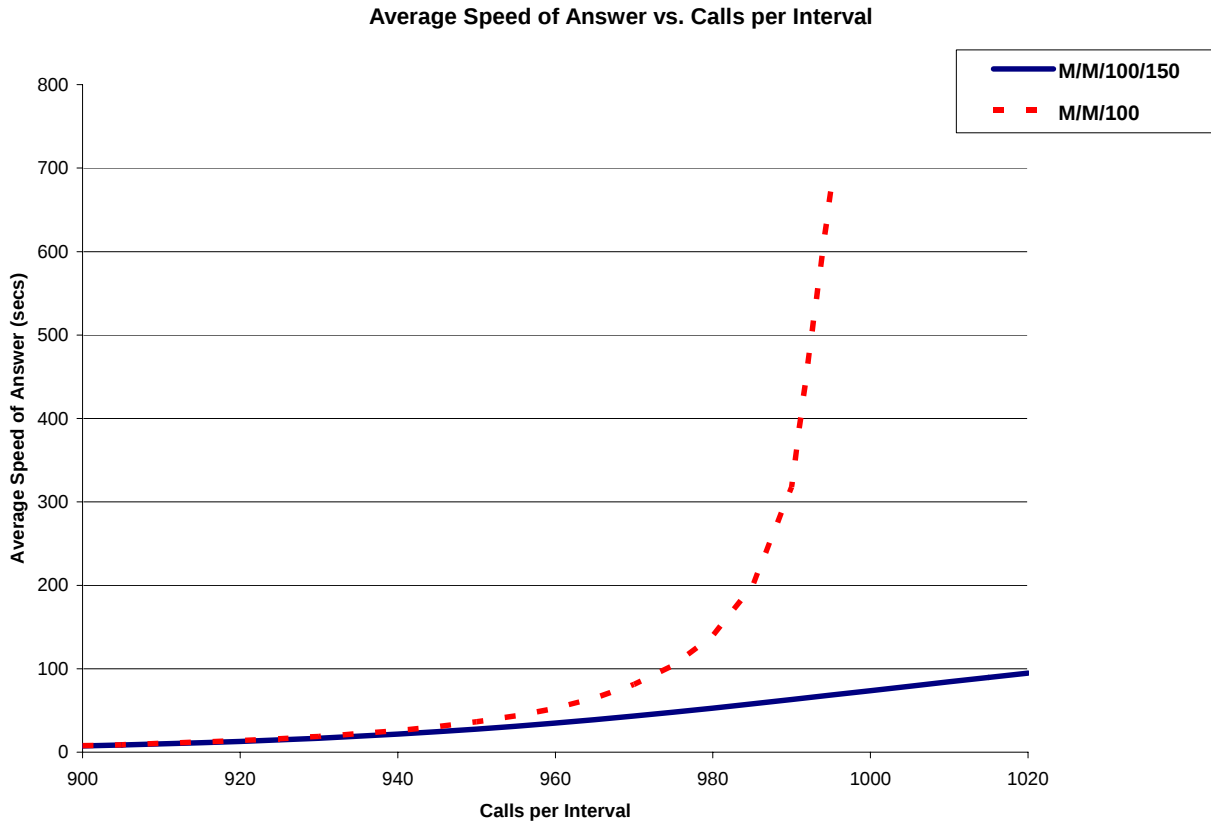
3.5.3 תנו אמד לפונקצית ה-hazard rate של Z , כפונקציה על האינטרוול $[0, T]$.

תשובה:

$$h_Z(t) = \frac{d_j}{\Delta(r_{j-1} - \frac{1}{2}d_j)}, \quad t \in [0, T].$$

שאלה 4: מערכות תורים, 4CallCenters (10 נקודות)

לפניכם שני גרפים שנוצרו ע"י התוכנה 4CallCenters (4CC).



הנתונים בגרפים נלקחו משתי הטבלאות של 4CC המצורפות בעמוד הבא.

4.1 הסבירו את השוני והדמיון של שתי המערכות $M/M/100$ ו- $M/M/100/150$, כפי שהם באים לידי ביטוי בגרפים והטבלאות:

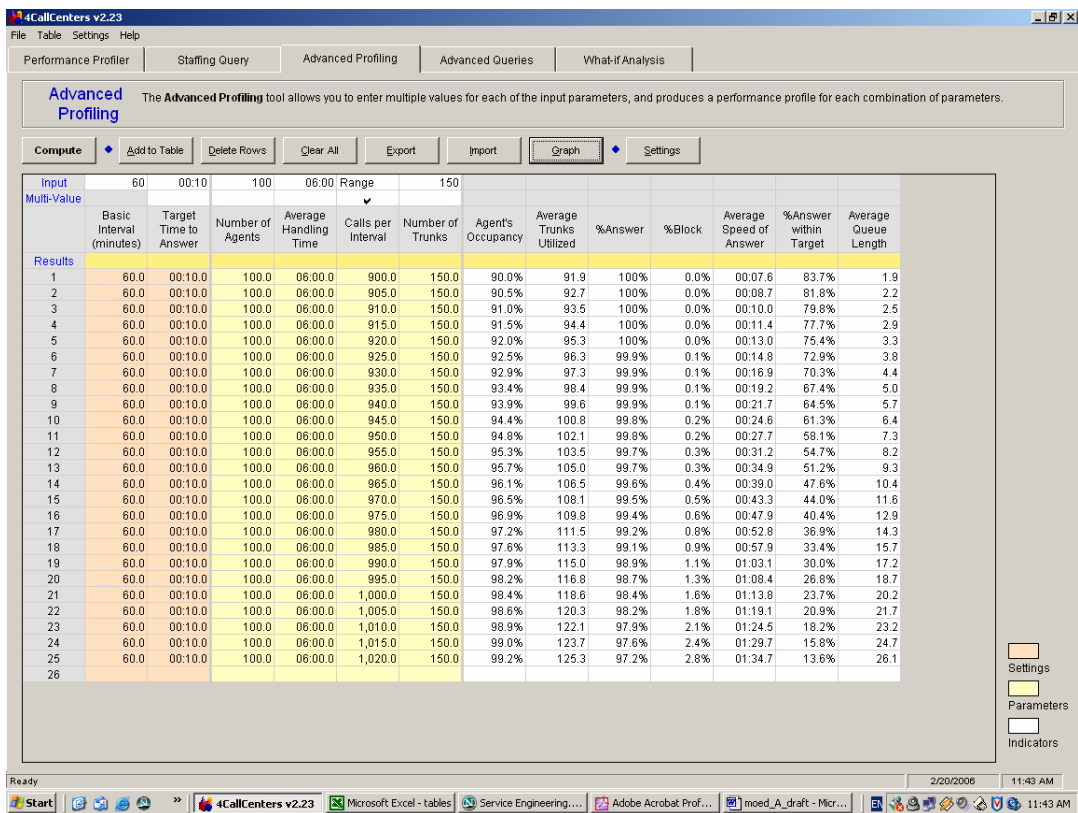
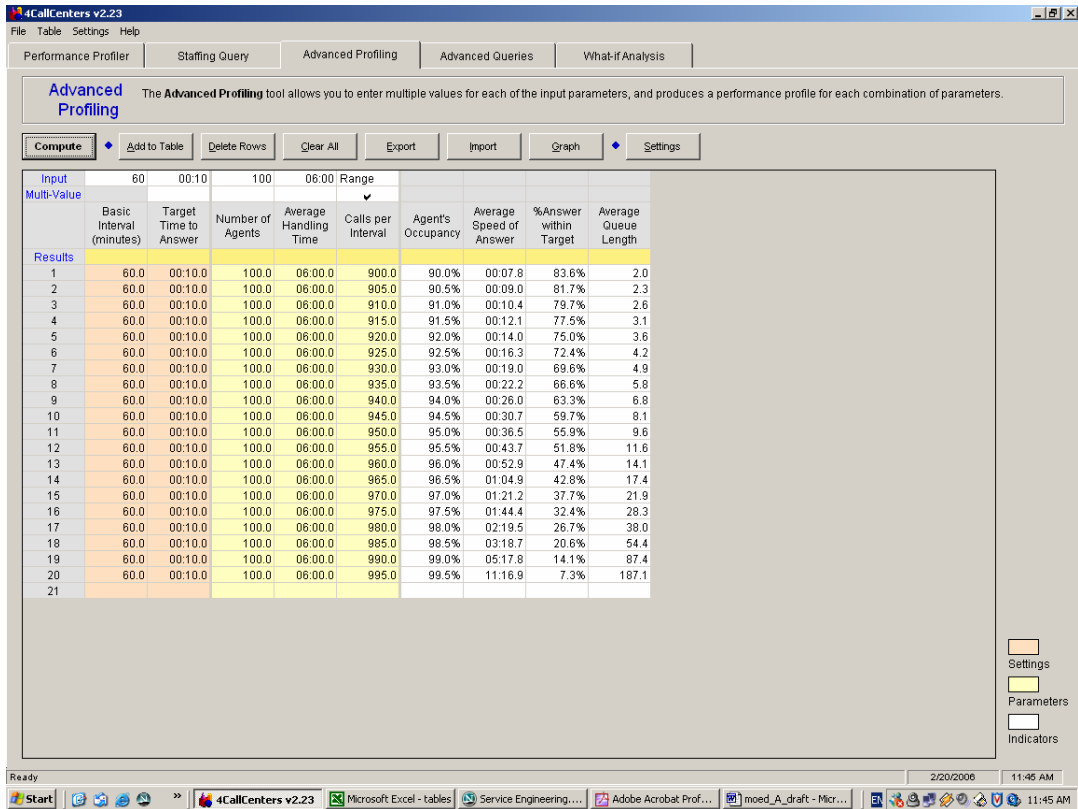
4.1.1 התנהגות דומה בעומסים נמוכים.

הסבר: כאשר העומס נמוך, ההשפעה של ה- *blocking* זניחה.

4.1.2 התנהגות שונה מאוד בעומסים גבוהים.

הסבר: כאשר העומס מגיע ל-1, מערכת $M/M/N$ מתפוצצת, ואילו מערכת

$M/M/N/K$ לא יכולה לגדול עד אינסוף, זמן ההמתנה שם תמיד יהיה חסום.



4.2 לפניכם טבלה נוספת מ-4CC.

4CallCenters v2.23

File Table Settings Help

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

Advanced Profiling The **Advanced Profiling** tool allows you to enter multiple values for each of the input parameters, and produces a performance profile for each combination of parameters.

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

Input	60	00:10	130	06:00	Range	130							
Multi-Value	Basic Interval (minutes)	Target Time to Answer	Number of Agents	Average Handling Time	Calls per Interval	Number of Trunks	Agent's Occupancy	Average Trunks Utilized	%Answer	%Block	Average Speed of Answer	%Answer within Target	Average Queue Length
Results													
1	60.0	00:10.0	130.0	06:00.0	900.0	130.0	69.2%	90.0	100%	0.0%	00:00.0	100%	0.0
2	60.0	00:10.0	130.0	06:00.0	905.0	130.0	69.6%	90.5	100%	0.0%	00:00.0	100%	0.0
3	60.0	00:10.0	130.0	06:00.0	910.0	130.0	70.0%	91.0	100%	0.0%	00:00.0	100%	0.0
4	60.0	00:10.0	130.0	06:00.0	915.0	130.0	70.4%	91.5	100%	0.0%	00:00.0	100%	0.0
5	60.0	00:10.0	130.0	06:00.0	920.0	130.0	70.8%	92.0	100%	0.0%	00:00.0	100%	0.0
6	60.0	00:10.0	130.0	06:00.0	925.0	130.0	71.2%	92.5	100%	0.0%	00:00.0	100%	0.0
7	60.0	00:10.0	130.0	06:00.0	930.0	130.0	71.5%	93.0	100%	0.0%	00:00.0	100%	0.0
8	60.0	00:10.0	130.0	06:00.0	935.0	130.0	71.9%	93.5	100%	0.0%	00:00.0	100%	0.0
9	60.0	00:10.0	130.0	06:00.0	940.0	130.0	72.3%	94.0	100%	0.0%	00:00.0	100%	0.0
10	60.0	00:10.0	130.0	06:00.0	945.0	130.0	72.7%	94.5	100%	0.0%	00:00.0	100%	0.0
11	60.0	00:10.0	130.0	06:00.0	950.0	130.0	73.1%	95.0	100%	0.0%	00:00.0	100%	0.0
12	60.0	00:10.0	130.0	06:00.0	955.0	130.0	73.5%	95.5	100%	0.0%	00:00.0	100%	0.0
13	60.0	00:10.0	130.0	06:00.0	960.0	130.0	73.8%	96.0	100%	0.0%	00:00.0	100%	0.0
14	60.0	00:10.0	130.0	06:00.0	965.0	130.0	74.2%	96.5	100%	0.0%	00:00.0	100%	0.0
15	60.0	00:10.0	130.0	06:00.0	970.0	130.0	74.6%	97.0	100%	0.0%	00:00.0	100%	0.0
16	60.0	00:10.0	130.0	06:00.0	975.0	130.0	75.0%	97.5	100%	0.0%	00:00.0	100%	0.0
17	60.0	00:10.0	130.0	06:00.0	980.0	130.0	75.4%	98.0	100%	0.0%	00:00.0	100%	0.0
18	60.0	00:10.0	130.0	06:00.0	985.0	130.0	75.7%	98.5	100%	0.0%	00:00.0	100%	0.0
19	60.0	00:10.0	130.0	06:00.0	990.0	130.0	76.1%	99.0	100%	0.0%	00:00.0	100%	0.0
20	60.0	00:10.0	130.0	06:00.0	995.0	130.0	76.5%	99.5	100%	0.0%	00:00.0	100%	0.0
21													

Settings
Parameters
Indicators

Ready 2/20/2006 11:53 AM

Start | 4CallCenters v2.23 | Microsoft Excel - tables | Service Engineering... | Adobe Acrobat Prof... | moed_A_draft - Micr... | 11:53 AM

לאיזון מערכת מהמערכות הבאות מתאימה הטבלה. הסבירו את תשובתכם.

א. $(Erlang - C) \ M / M / N$

ב. $(Erlang - A) \ M / M / N + M$

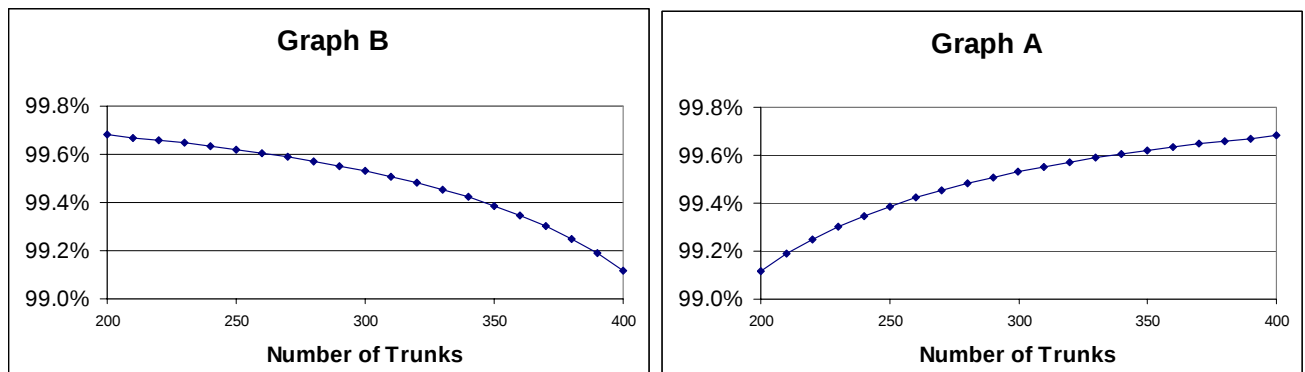
ג. $\Rightarrow \underline{\underline{(Erlang - B) \ M / M / N / N}}$

ד. $N < K , \ M / M / N / K$

תשובה:

הסבר:

4.3 לפניכם שני גרפים כך שאחד מהם מתאר נצילות השרתים במודל $M/M/N/K$ כתלות במספר הקווים (trunks). יתר הפרמטרים אינם משתנים.



ציינו, איזה גרף הוא נכון. נמקו בקצרה

הגרף שמתאר את נצילות השרתים הוא _____ גרף A _____.

הסבר:

כאשר K גדל – אזי $P(block)$ קטנה ו-זה גורר $\rho = \frac{\lambda}{N\mu}(1 - P(block))$ גדלה !!!

שאלה 5: מדידות, תחומים תפעוליים (10 נקודות)

בעמוד הבא מצורף דו"ח ACD של מוקד טלפוני אותו ניתחנו בכיתה. (בעמודה האחרונה של כל חצי שעה מופיע ה- offered load).

5.1 נתמקד בשלוש חצאי השעות 13:30-14:00, 14:30-15:00 ו- 17:00-17:30. התאימו אליהן את המדיניות התפעולית המתאימה מבין: QED, ED ו- QD ונמקו בקצרה אך במדויק ככל שתוכלו.

$$\beta = \frac{N-R}{\sqrt{R}}$$

$$\approx \frac{163-180}{14} \approx -1.25$$

$$\frac{206-204}{\cdot} \approx 0$$

$$\frac{135-142}{10.5} \approx -2$$

אינטרוול	מדיניות	נימוק קצר
13:30-14:00	ED	נטיית 10% ASA סביב קוץ לא מתן איכות $\beta \approx -1.25$
14:30-15:00	QED	נטיית 2-3% ASA סביב קוץ פחות מתן איכות $\beta \approx 0$
17:00-17:30	QD	נטיית 2+ נטיית 2+ מתן איכות $\beta \approx 2+$

5.2 עבור כל אחד מהאינטרוולים בסעיף 5.1, נניח כי מספר המופעים היה גדל פי 4 וגם מספר השרתים היה גדל פי 4. תנו הערכתכם לאחוז הנוטשים תחת תרחיש חדש זה. נמקו בקצרה.

$$\delta(1-\rho) = N\mu$$

$$\beta = \frac{N-R}{\sqrt{R}}$$

אינטרוול	% הנוטשים הנוכחי	% הנוטשים החדש	נימוק קצר
13:30-14:00	9.4%	9.4% לוי מילר	$P(ACT) \approx 1 - \frac{1}{\beta}$ $\rho = \frac{R}{N}$ לוי מילר
14:30-15:00	2.7%	1.35% לוי מילר	סינוי $\sqrt{N} = 2$ (לשינוי גודל N) $P(ACT) \approx 1/\sqrt{N}$
17:00-17:30	0%	0	משיגים קוץ יחיד סביב קוץ יחיד

(א/ה גם קצ' 400)
חזרת לאט

Charlotte – Center

Time	Recvd	Answ	Abn %	ASA	AHT	Occ %	On Prod%	On Prod FTE	R
0	20,577	19,860	3.5%	30	307	95.1%	85.4%	222.7	
8:00	332	308	7.2%	27	302	87.1%	79.5%	59.3	55.70
8:30	653	615	5.8%	58	293	96.1%	81.1%	104.1	106.29
9:00	866	796	8.1%	63	308	97.1%	84.7%	140.4	148.18
9:30	1,152	1,138	1.2%	218	303	90.8%	81.6%	211.1	193.92
10:00	1,330	1,286	3.3%	22	307	98.4%	84.3%	223.1	226.84
10:30	1,364	1,338	1.9%	33	296	99.0%	84.1%	222.5	224.30
11:00	1,380	1,280	7.2%	34	306	98.2%	84.0%	222.0	234.60
11:30	1,272	1,247	2.0%	44	298	94.6%	82.8%	218.0	210.59
12:00	1,179	1,177	0.2%	1	306	91.6%	88.6%	218.3	200.43
12:30	1,174	1,160	1.2%	10	302	95.5%	93.6%	203.8	196.97
13:00	1,018	999	1.9%	9	314	95.4%	91.2%	182.9	177.58
13:30	1,061	961	9.4%	67	306	100.0%	88.9%	163.4	180.37
14:00	1,173	1,082	7.8%	78	313	99.5%	85.7%	188.9	203.97
14:30	1,212	1,179	2.7%	23	304	96.6%	86.0%	206.1	204.69
15:00	1,137	1,122	1.3%	15	320	96.9%	83.5%	205.8	202.13
15:30	1,169	1,137	2.7%	17	311	97.1%	84.6%	202.2	201.98
16:00	1,107	1,059	4.3%	46	315	99.2%	79.4%	187.1	193.73
16:30	914	892	2.4%	22	307	95.2%	81.8%	160.0	155.89
17:00	615	615	0.0%	2	328	83.0%	93.6%	135.0	112.07
17:30	420	420	0.0%	0	328	73.8%	95.4%	103.5	76.53
18:00	49	49	0.0%	14	180	84.2%	89.1%	5.8	4.90

- 5.3 נניח עתה כי סך מספר המופעים היומי היה גדל פי 4, וגם מספר השרתים בכל חצי שעה היה גדל פי 4. מדוע לא ניתן לתת הערכה מדויקת לאחוז הנוטשים היומי תחת תרחיש חדש זה? נמקו בקצרה.

תשובה:

טיפ: אם מס' המופעים היומי אינו מספר רב, חלש או אחוז הנוטשים היומי, שכן לא ידוע הכמה/פי כמה גם מס' המופעים באינטרוו = 4 לא ניתן לחשב את אחוז הנוטשים באינטרוו ולפיכך גם לא את אחוז הנוטשים היומי.

- 5.4 בשורה הראשונה של דו"ח ה-ACD מסוכמים מדדי המוקד היומיים. למשל,

- 20,577 מופעים ליום, או $\frac{20,577}{21}$, ז"א $\lambda = 990$ מופעים לחצי שעה, בערך.
- 307 שניות משך שירות ממוצע, או כ- $E(S) = 5$ דקות, בערך.
- 222.7 מוקדנים בממוצע, או כ- $n = 220$ מוקדנים, בערך.
- 95.1% נצילות ממוצעת, או כ-95% נצילות.

בשורה ראשונה זו נפלה טעות! מהי טעות זו? הסבירו בקצרה.

רמז: כדי לאתר את הטעות, שימו לב לחישוב הפשוט הבא:

$$\frac{\lambda \cdot E(S)}{n} = \frac{\frac{990}{30} \cdot 5}{220} \approx 0.75$$

הטעות היא: 222.7 נקודת מבט לא נכונה, אלא 75.7 נקודת מבט.

הסבר: זו ימין זמנה FTE, 222.7 נקודת מבט נכונה ונאמרה. שם המה נכונה, אלא שם נכונה אמרתי וזה לא נכון.

(לדוגמה נניח, $E(S)$ ימין שם 307 שניות, אלא שם אמרתי וזה לא נכון).
למחרת, ימין שם הנקודת הממוצעת 95%.

שאלה 6: מודלים סטוכסטיים במערכות תורים (10 נקודות)

נתבונן במערכת תורים עם פעילויות עזר (ancillary activities), אותה למדנו בתרגול. המערכת ניתנת לתיאור ע"י מודל $M/M/n$, כאשר n - מספר השרתים - יכול להשתנות כתלות במספר הלקוחות בתור.

תזכורת לדינמיקת המערכת: בתחילת יום העבודה נמצא במערכת שרת יחיד. נגדיר את ערך הסף J ע"י

$$J = \frac{\text{מספר ממתנים בתור}}{\text{מספר השרתים העובדים}}$$

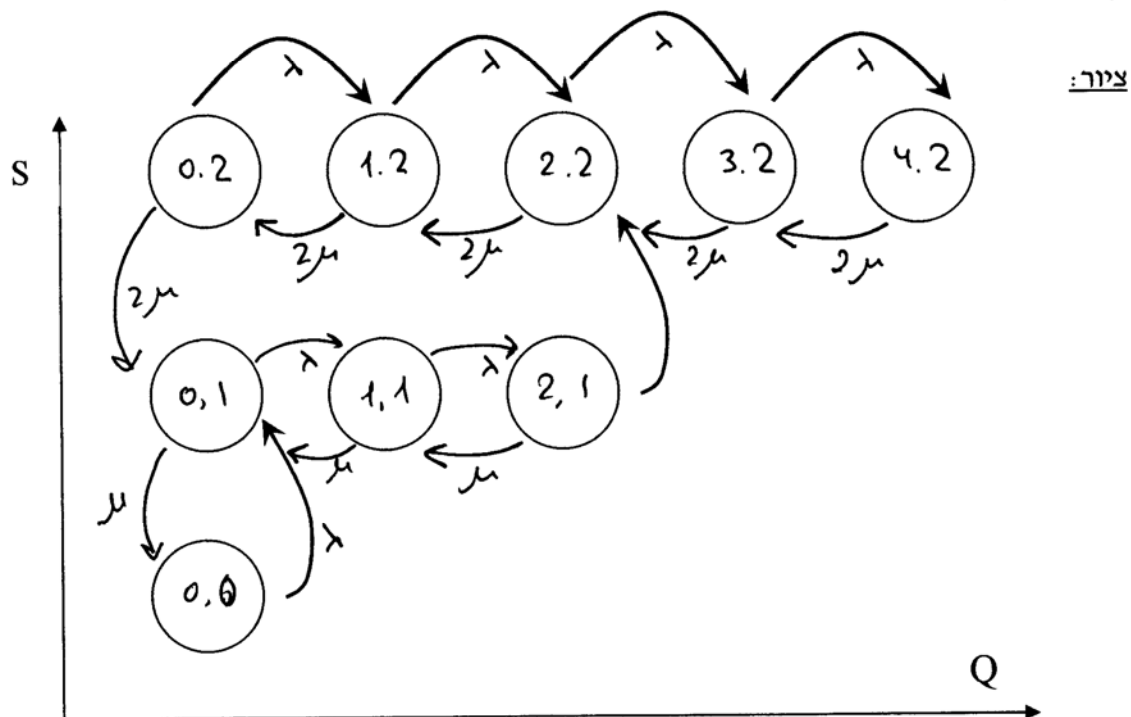
(כמובן, J משתנה עם מצב המערכת). במערכת נוקטים במדיניות הבאה: מוסיפים שרת אחד כל פעם כאשר הערך J עולה על מספר מסוים K שנקבע מראש ($K < J$). בנוסף, שרת מוחזר לפעילות עזר (שונה משירות) אם מייד אחרי סיום השירות עם לקוח כלשהו הוא לא מוצא אף לקוח ממתין. כלומר, השרתים המתאספים ממשיכים לשרת כל עוד יש ממתנים בתור, וחוזרים לפעילויות העזר בזה אחר זה אם התור ריק בסיום שירותם. שימו לב כי במערכת נמצא תמיד לפחות שרת אחד.

נניח שמופע הלקוחות למערכת הוא פואסוני עם קצב λ , ואילו זמן השירות - אקספוננציאלי עם

$$\text{פרמטר } \mu. \text{ נסמן כרגיל } R = \frac{\lambda}{\mu}.$$

נגדיר תק"מ דו-ממדי (Q, S) , כאשר S מתאר את מספר אנשים הנמצאים בשירות ואילו Q מתאר את מספר אנשים בתור.

6.1 נניח $K = 2$. ציירו את דיאגרמת קצבי המעברים של התהליך (Q, S) . המשיכו את הדיאגרמה עד ל- $S = 2$.



6.2 נתבונן בגרף המתאר את התלות של L - ממוצע מספר הלקוחות במערכת - בערך של הסף K עבור ערכים שונים של R .

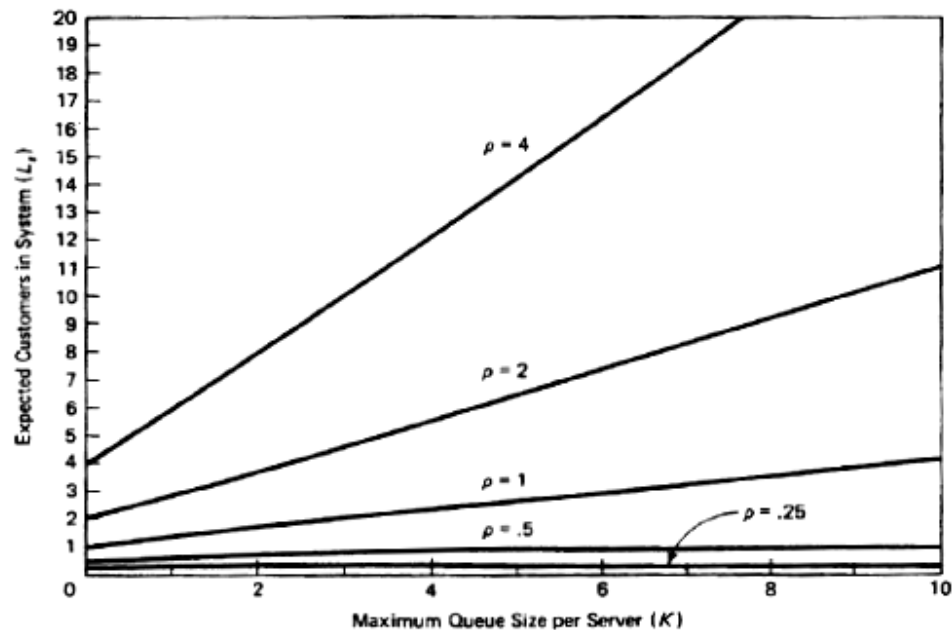


Figure 5.10 Increasing the maximum queue size per server increases L_s at a nearly linear rate when ρ is greater than 1. For $\rho < 1$, L_s approaches $\rho/(1 - \rho)$ as K becomes large.

6.2.1 כיצד ניתן להסביר את הערכים של L כאשר $K = 0$?

רמז: מספר האנשים במערכת $M/M/\infty$ עם פרמטרים λ ו- μ מפולג פואסוני עם פרמטר $R = \frac{\lambda}{\mu}$.

תשובה:

$$E(\text{Poiss}(R)) = R$$

6.2.2 עבור $1 \leq R$ כיצד ניתן להסביר את ההתנהגות הגבולית (דהינו-עליה) של L כאשר K גדל?

תשובה:

כאשר K גדל לאינסוף – המערכת הופכת להיות $M/M/1$ והיא מתפוצצת אם $\lambda \geq \mu$.
אזי L גדל!

6.2.3 עבור $R > 1$, כיצד ניתן להסביר את ההתנהגות הגבולית של L כאשר K גדל?

רמז: מספר האנשים במערכת $M/M/1$ עם פרמטרים λ ו- μ מפולג גיאומטרי (כולל ערך 0)

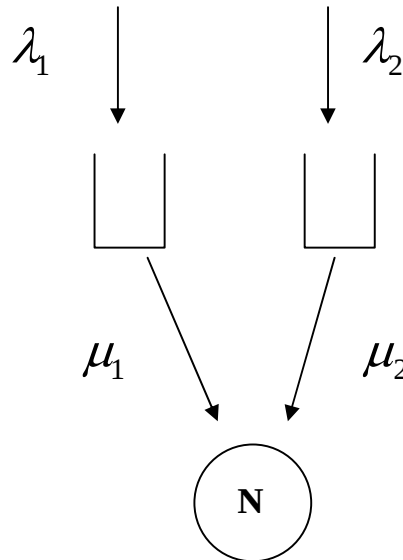
עם פרמטר $1-R$.

תשובה:

$$E(Geom_{with\ zero}(R)) = \frac{R}{1-R}$$

שאלה 7 : מודלים של נוזלים (10 נקודות)

נתונה מערכת תורים בארכיטקטורת V, המתוארת בציור הבא :



הפרמטרים של מודל המערכת נתונים ע"י :

- ♦ λ_1 (λ_2) - קצבי מופע לקוחות מסוג 1 (2),
- ♦ μ_1 (μ_2) - קצבי שירות לקוחות מסוג 1 (2),
- ♦ N - מספר השרתים בתחנה

המערכת פועלת לפי מדיניות שהיא *preemptive*, ז"א ניתן בכל רגע להפסיק שירות של לקוח ולהחזירו להמתנה בתור מהסוג שלו, אבל אסור "לזרוק" לקוחות מהמערכת. בנוסף, המדיניות הינה *work conserving*, דהיינו, השרתים נותנים שירות כאשר יש לקוחות בתור, (ז"א לא ייתכן מצב שבו יש שרתים פנויים ובעת ובעונה אחת יש גם לקוחות הממתינים בתור).

לצורך בנית **מודל נוזלים** למערכת הנ"ל, נגדיר את המשתנים הבאים :

$$X_1(t) = \text{סה"כ מספר הלקוחות מסוג 1 במערכת (בשירות או בתור), בזמן } 0 \leq t.$$

$$X_2(t) = \text{סה"כ מספר הלקוחות מסוג 2 במערכת (בשירות או בתור), בזמן } 0 \leq t.$$

7.1 כתבו משוואה דיפרנציאלית עבור $X_1(t)$ ו- $X_2(t)$, $0 \leq t$, תחת ההנחה שהתחנה נותנת עדיפות ללקוחות מסוג 1.

תשובה:

$$\frac{dX_1(t)}{dt} = \lambda_1 - \mu_1(X_1 \wedge N)$$

$$\frac{dX_2(t)}{dt} = \lambda_2 - \mu_2(X_2 \wedge (N - X_1)^+)$$

7.2 כתבו ביטוי למספר הכולל של הממתינים בתור בזמן t . (השתמשו ב- N , $X_1(t)$ ו- $X_2(t)$)

תשובה:

$$(X_1(t) + X_2(t) - N)^+$$

7.3 כאלטרנטיבה לעדיפות הנ"ל ללקוחות מסוג 1, נגדיר מדיניות חדשה ששומרת על תורים שווים בשני הסוגים. ז"א, בכל רגע $0 \leq t$, מספר הממתינים בתור מסוג 1 שווה למספר הממתינים בתור של סוג 2.

7.3.1 כתבו משוואה דיפרנציאלית עבור $X_1(t)$ ו- $X_2(t)$ תחת מדיניות אלטרנטיבית זו.

$$\frac{dX_1(t)}{dt} = \lambda_1 - \mu_1 \left(X_1 - \frac{1}{2} (X_1 + X_2 - N)^+ \right)$$

$$\frac{dX_2(t)}{dt} = \lambda_2 - \mu_2 \left(X_2 - \frac{1}{2} (X_1 + X_2 - N)^+ \right)$$

7.3.2 לא תמיד ניתן לממש את המדיניות של תורים שווים, אך בתנאים מסויימים על X_1 ו- X_2 ניתן לממשה.

ציינו תנאים על X_1 ו- X_2 , המבטיחים שניתן לממשה.
תנו, בנוסף, דוגמא נגדית להראות שלא תמיד ניתנת המדיניות למימוש.

רמז: כדי לענות על השאלה נדרשים רק X_1, X_2 ו- N .

תשובה:

$$X_1 \wedge X_2 \geq \frac{1}{2}(X_1 + X_2 - N)^+$$

שאלה 8: M/G/1 (7 נקודות)

נתונה מערכת שירות הניתנת למידול בעזרת רשת תורים $M / G / 1$ העובדת לפי שיטה First Come First Serve

8.1 נניח כי כל הלקוחות זהים סטטיסטית ומגיעים למערכת במופע פואסוני עם קצב λ . נגדיר כרגיל

$$\rho = \lambda \cdot E(S) \quad \text{אילו ערכים של } \rho \text{ מבטיחים סטביליות של המערכת.}$$

תשובה:

$$\rho < 1$$

8.2 כעת נניח כי למערכת מגיעים K סוגים של לקוחות, כאשר סוג k , $(k = 1, 2, \dots, K)$, מגיע במופע

פואסוני עם קצב λ_k עם ממוצע זמן שירות m_k .

$$\text{נגדיר } \rho_k = m_k \lambda_k$$

$$\text{ו- } \rho = \rho_1 + \dots + \rho_K$$

על סמך סעיף 8.1 ציינו עתה את התנאי לסטביליות של המערכת? הסבירו תשובתכם.

תשובה:

$$\rho < 1$$

הסבר:

עדיין ניתן להסתכל על המערכת חדשה כמו על $M / G / 1$, אך כאשר בהסתברות $\frac{\lambda_k}{\sum_i \lambda_i}$ ממוצע משך

שירות הוא m_k .

אזי

$$S(S) = \sum_{k=1}^K \left(\frac{\lambda_k}{\sum_i \lambda_i} \right) \cdot m_k = \left(\sum_{k=1}^K \rho_k \right) \cdot \frac{1}{\sum_i \lambda_i}$$

זה גורר

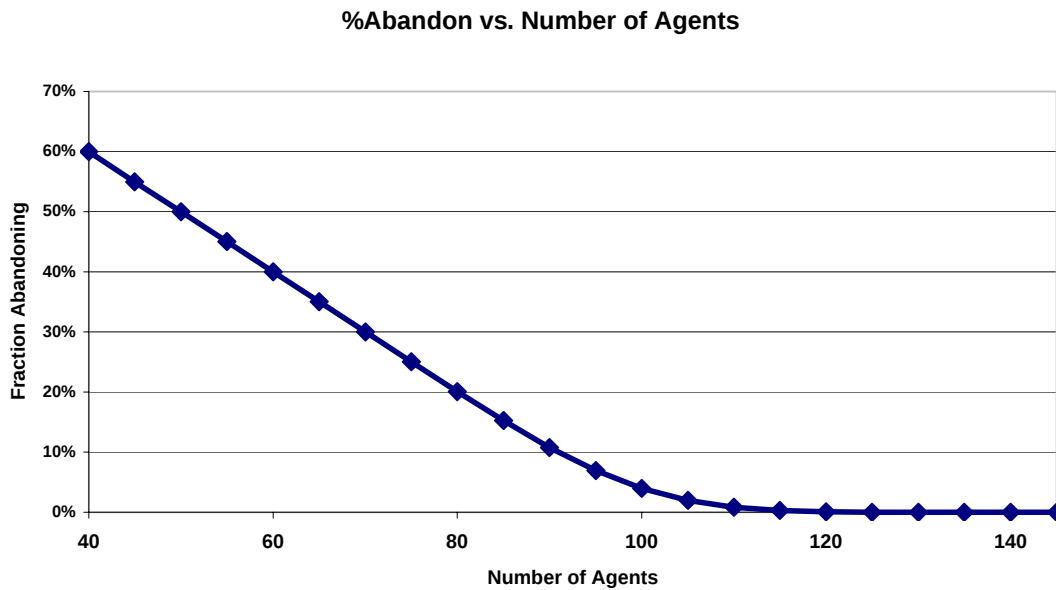
$$\text{new utilization } \bar{\rho} = \left(\sum_i \lambda_i \right) \cdot ES = \sum_k \rho_k = \rho$$

$$\bar{\rho} < 1 \Leftrightarrow \rho < 1$$

שאלה 9 : Erlang-A , QED , איוש אופטימלי (15 נקודות)

נתונה מערכת תורים $M/M/N+M$ (Erlang-A). המערכת מאופיינת על ידי ממוצע משך שירות של 3 דקות, ממוצע סבלנות שגם הוא 3 דקות, וקצב מופע של 2,000 לקוחות לשעה. כל לקוח שמקבל שירות מביא למערכת רווח ממוצע של 5 ₪ ודקת עבודה של שרת יחיד עולה 1 ₪.

9.1 נניח כי מספר השרתים $N=60$. ע"י שימוש בגרף הבא, שנוצר בעזרת התוכנה *4CallCenters* ומתאים למערכת הנ"ל, חשבו את הרווח הממוצע של המערכת לשעה.



תשובה:

רווח של המערכת שווה ל $2,000 \cdot 5 \cdot (1 - P\{Ab\}) - 60 \cdot N$.
הסתברות נטישה היא 40%.
 $2,000 \cdot 5 \cdot (1 - 0.4) - 60 \cdot 60 = 6,000 - 3,600 = 2,400.$

9.2 בהמשך לסעיף 9.1, חשבו את זמן ההמתנה הממוצע של לקוחות במערכת.

תשובה:

קצב נטישה $\theta = 1/3$. בתור Erlang-A $P\{Ab\} = \theta \cdot E[W]$, לכן
 $E[W] = 0.4 / (1/3) = 1.2 \text{ min}$

9.3 נניח שסבלנות הלקוחות עלתה מ-3 ל-5 דקות בממוצע וכל יתר הפרמטרים נשארים ללא שינוי. האם רווח המערכת יעלה/ירד/לא ישתנה? נמקו בקצרה.

תשובה:

הרווח של המערכת יעלה משום שהסתברות נטישה ירד.

9.4 נניח שזמן השירות הממוצע ירד באופן דרסטי לדקה אחת (מה שגורם להקלה בעומס, ולכן לפחות נטישות), אולם גם הסבלנות ירדה לדקה אחת בממוצע (מה שגורם ליותר נטישות). האם הרווח של המערכת יעלה/ירד/לא ישתנה? נמקו בקצרה.

תשובה:

הרווח של המערכת יעלה. ρ Offered load per server יהי שווה ל

$$\rho = \frac{\lambda}{N\mu} = \frac{2,000/60}{60 \cdot 1} = 55.6\%$$

במערכת גדול עם הנצילות נמוכה כל כך, אחוז הנטישות יהי זניח.

בכתה למדנו כיצד לאייש אופטימלית מודל $(Erlang - C) M/M/N$, בהינתן עלויות המתנה ומשכורות שרתים. באנלוגיה, נניח שקצבי ההגעה, שירות ונטישה במערכת $Erlang-A$ נתונים ע"י μ, λ , ו- θ בהתאמה. אזי גם כאן ניתן לחשב את רמת האיוש האופטימלית N^* (המביאה את הרווח למקסימום) בעזרת כלל השורש הבא:

$$N^* \approx R + \beta^*(r)\sqrt{R}$$

$$\frac{\text{קצב הגעה}}{\text{קצב שירות}} = \frac{\lambda}{\mu} = R \quad \text{כאשר כרגיל}$$

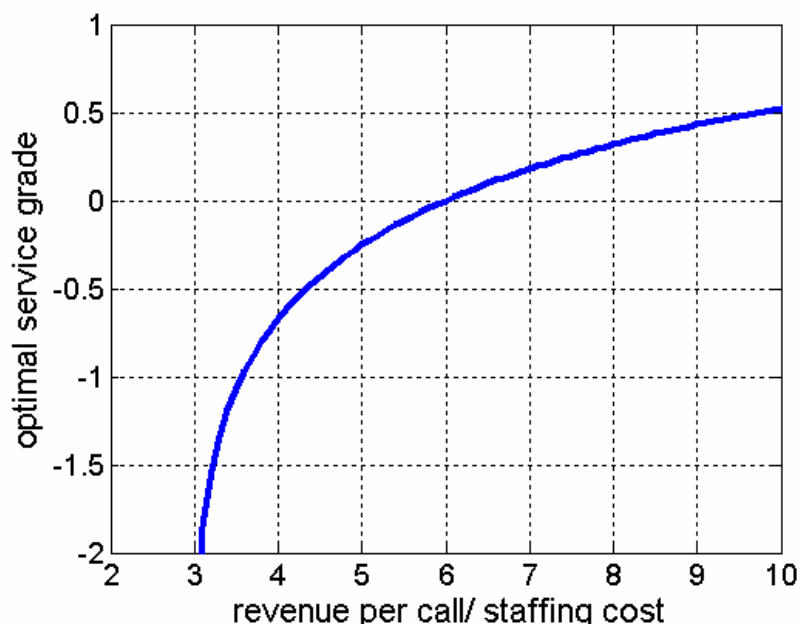
ופרמטר רמת השירות β^* (service grade) הוא פונקציה של המנה r הנתונה ע"י

$$\frac{\text{revenue per call}}{\text{staffing cost per server per minute}} = \frac{\text{הרווח משירות בודד}}{\text{משכורת לדקה}} = r$$

9.5 חשבו את רמת האיוש האופטימלית N^* תחת ההנחות בתחילת השאלה. השתמשו בגרף הבא המתאר את הקשר בין β^* ו- r תחת הנחות אלו.

תשובה:

$$r = 5, \text{ לכו מהגרף } \beta^* \approx -0.25. \text{ Offered load } R = 2,000/20 = 100. \\ N^* = 100 - 0.25\sqrt{100} \approx 98$$



9.6 נניח שהרווח הממוצע לשירות בודד ירד ל- 2 ₪. מה תהיה עתה רמת האיוש האופטימלית?

רמז: אין צורך בחישובים כדי לענות על שאלה זו.

תשובה:

$$N^* = 0. \text{ עלות שירות של לקוח (3 ₪) יותר גדולה מרווח לשירות אחד.}$$

9.7 נניח עתה שלקוחות לא מביאים שום רווח, אולם כל נטישה מביאה לחברה הפסד של 5 ₪. כל

ההנחות האחרות מתחילת השאלה נשארות תקפות.

חשבו את רמת האיוש N^* המביאה את עלות תפעול המערכת למינימום.

תשובה:

אנו צריכים למזער $60 \cdot N + 2,000 \cdot 5 \cdot P\{Ab\}$ וקודם היינו צריכים להביא למקסימום את

$2,000 \cdot 5 \cdot (1 - P\{Ab\}) - 60 \cdot N$. שתי בעיות הן שקולות ולכן $N^* = 98$ כמו בסעיף 5.