



מרצה : פרופסור אבישי מנדלבאום

מתרגלת : גלית יום-טוב

תאריך הבחינה : 05.03.2009

מס' סטודנט _____

שם _____

הנדסת מערכות שירות – 096324

מועד א' – סמסטר חורף תשס"ט 2009

פתרון

- המבחן נמשך שלוש שעות. באחריותך להחזירו במועד.
- המבחן בחומר סגור.
- מותר להשתמש במחשבון.
- המבחן כולל 25 עמודים, כולל עמוד זה.
- 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	ניקוד:
תרגילי בית	מתוך 10 נקודות	שאלה 2	
יישום	מתוך 15 נקודות	שאלה 3	
תיאוריה	מתוך 10 נקודות	שאלה 4	

סה"כ _____ מתוך 45 נקודות

בהצלחה !

שאלה 1 - תרגילי כיתה/הרצאות/מבחנים קודמים

חלק 1. חוק Little.

בחברה יפנית כלשהי 36,000 עובדים, 20% מהם נשים. משך העסקה ממוצע של נשים הוא 37 חודשים, ולגברים 200 חודשים. הנח כי סך כמות המועסקים ואחוזי הגברים והנשים בתוכם קבועים לאורך זמן.

1.1 הסבירו מדוע קצב ההצטרפות של נשים חדשות לחברה הוא גדול יותר מקצב הצטרפות הגברים, למרות שפרופורציות הגברים והנשים במצב יציב הן 80% גברים ו-20% נשים. גבו את תשובתכם בחישובים.

תשובה:

צריך לחשב את λ_i (i=1 לגברים, i=2 לנשים). ידוע על פי הנתונים ש-

$$L_1 = 36,000 * 0.8 = 28,800, L_2 = 36,000 * 0.2 = 7,200$$

ולכן ע"ס חוק Little מתקיים $\lambda_1 = \frac{L_1}{W_1} = \frac{28,800}{200} = 144$, $\lambda_2 = \frac{L_2}{W_2} = \frac{7,200}{37} = 194.6$

כלומר, מכיוון שזמני שהיה של גברים הרבה יותר גדולה משל נשים, קצב ההגעה של נשים צריך להיות גבוה יותר (אם רוצים לשמור על 20% נשים בחברה).

נסמן את ממוצעי זמני ההעסקה של גברים ונשים על ידי W_1 ו- W_2 בהתאמה.

נסמן ב- W את ממוצע זמן ההעסקה הכללית בחברה.

1.2 בסעיף זה נניח כי נתונים (ע"י מדידות או מנסיון) p_1 ו- p_2 : פרופורציות מספר העובדים

שנשכרים בחודש מסוג 1 (גברים) וסוג 2 (נשים). ז"א, $p_i = \frac{\lambda_i}{\lambda_1 + \lambda_2}$, כאשר λ_i הוא קצב

הגעת עובדים חדשים לחברה מסוג i , ($i = 1, 2$).

בטאו את W בעזרת p_1, p_2, W_1 ו- W_2

תשובה:

$$W = p_1 W_1 + p_2 W_2$$

דרך חישוב: בעזרת נוסחת Little.

$$L = L_1 + L_2 \Rightarrow \lambda W = \lambda_1 W_1 + \lambda_2 W_2 \Rightarrow W = p_1 W_1 + p_2 W_2$$

ז"א W הוא ממוצע חשבונני של W_1 ו- W_2 .

1.3 בסעיף זה נניח כי נתונים q_1 ו- q_2 : פרופורציות מספר העובדים המועסקים בחברה, מסוג 1 וסוג 2.

ז"א, $q_i = \frac{L_i}{L_1 + L_2}$, כאשר L_i הוא ממוצע מספר העובדים בחברה מסוג i , $(i = 1, 2)$.

בטאו את W בעזרת q_1, q_2, W_1 ו- W_2 .

בעזרת תשובתכם תנו הסבר מילולי לגדלים $\frac{1}{W_1}, \frac{1}{W_2}, \frac{1}{W}$.

תשובה:

$$W = \frac{1}{\frac{q_1}{W_1} + \frac{q_2}{W_2}}$$

דרך חישוב:

$$W = \frac{L_1 + L_2}{\lambda_1 + \lambda_2} = \frac{L_1 + L_2}{\frac{L_1}{W_1} + \frac{L_2}{W_2}} = \frac{L_1 + L_2}{\frac{q_1(L_1 + L_2)}{W_1} + \frac{q_2(L_1 + L_2)}{W_2}} = \frac{1}{\frac{q_1}{W_1} + \frac{q_2}{W_2}}$$

ז"א W הוא ממוצע הרמוני של W_1 ו- W_2 .

הסבר מילולי:

$\frac{1}{W_1}$ - קצב עזיבה (או פיטורים) אינדיבידואלי של גברים העובדים בבנק.

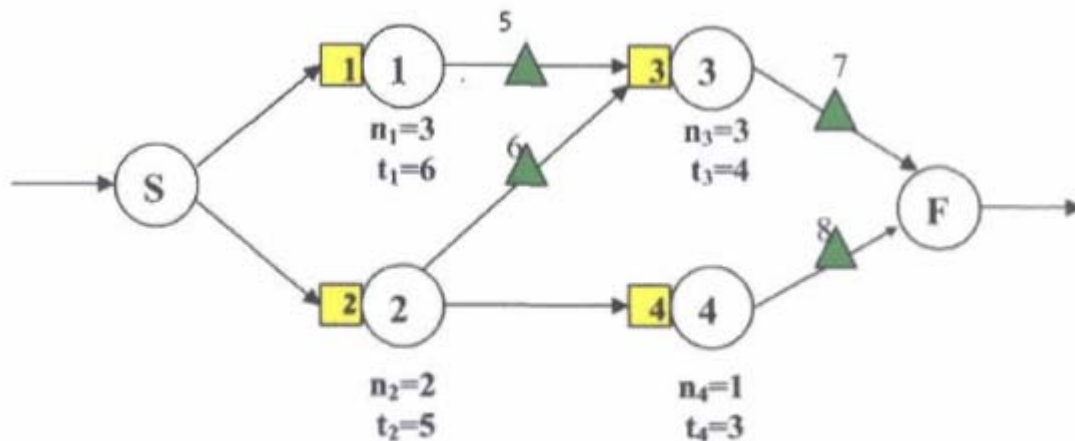
$\frac{1}{W_2}$ - קצב עזיבה (או פיטורים) אינדיבידואלי של נשים העובדות בבנק.

$\frac{1}{W}$ - קצב עזיבה (או פיטורים) אינדיבידואלי משוקלל.

ומתקיים $\frac{1}{W} = \frac{q_1}{W_1} + \frac{q_2}{W_2}$. ז"א $\frac{1}{W}$ הוא ממוצע חשבוני של $\frac{1}{W_1}$ ו- $\frac{1}{W_2}$.

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

2.1 לרשת 8 תורים הממוספרים מ-1 עד 8. מיהם התורים הנובעים מפערי סינכרון ומי הנובעים ממשאבים מוגבלים?

תשובה:

תורי סינכרון: 5,6,7,8

תורי משאבים: 1,2,3,4

2.2 מהי נצילות כל אחד משלושת השרתים המטפל בפעילות 3? הסבירו במשפט או שניים את תשובתכם.

תשובה:

נצילות השרתים: $4/(3.5*3)=0.3809$

הסבר: מכיוון שהרשת נמצאת במצב יציב, קצב המופע לכל תחנה הוא $1/3.5$ פעילויות (פרויקטים) ליום. כל פרויקט דורש 4 ימי עבודה של פעילות 3 והעבודה מתחלקת בין 3 שרתים.

2.3 מהו המספר המינימלי של שרתים הנדרשים לטפל בפעילות 3 כדי שנצילות כל אחד לא תעלה על 33%?

תשובה:

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

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

תשובה:

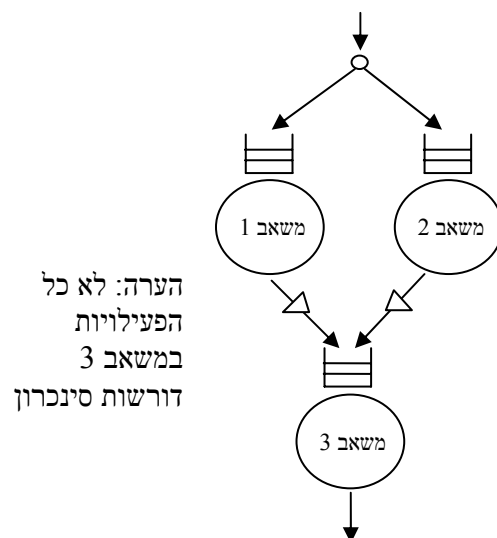
ההסבר כולל שני חלקים. מכיוון שהרשת סטוכסטית ועל סמך מבנה הקדימויות, אנחנו כבר יודעים ש- $ET_{DS} \leq 10$ (ראו הוכחה בהמשך). בנוסף, הדינאמיות של המערכת (ריבוי הפרויקטים) גורמת לתורי משאבים שמאריכים עוד יותר את משך ביצוע הפרויקט.

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

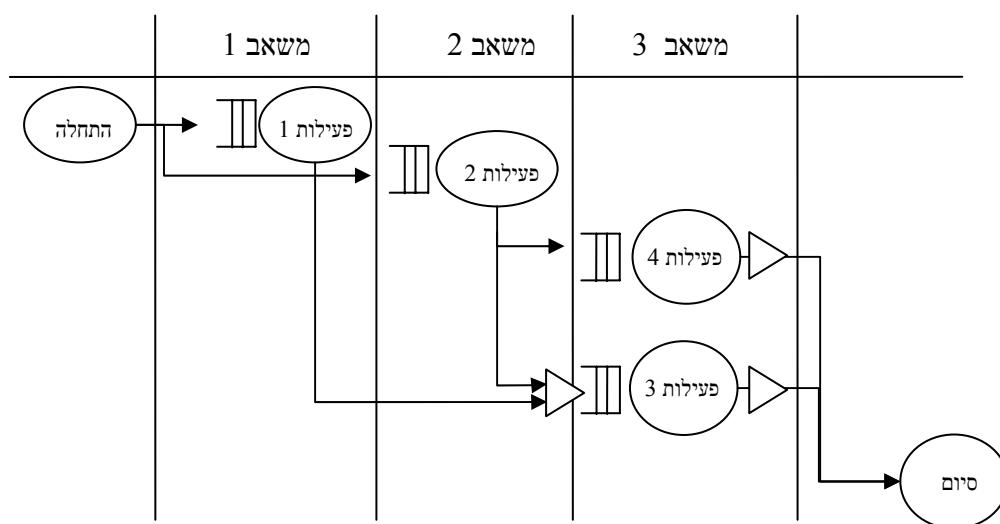
ברשת.

2.5 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:



תרגיל 2 – תרגילי בית

למרכז שרות US Bank פונים לקוחות מכמה סוגים, ביניהם לקוחות המוגדרים Business (1) ו- Platinum (2). לקוחות Platinum נחשבים לקוחות VIP המקבלים עדיפות בנגישות לשרות. השרות עצמו, הניתן לשני סוגי הלקוחות, זהה, ואותם מוקדנים יכולים לתת שרות לשני סוגי הלקוחות. למרות זאת הוגדרו שתי קבוצות נציגי שירות נפרדים: "מוקדני Business" ו"מוקדני Platinum", שכל אחת מהן נותנת שרות לקבוצת הלקוחות שלה. הפרמטרים של 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_1 (N_2) - מספר המוקדנים בקבוצת שירות 1 (2).

מנהל השיווק החליט להשתמש בתוכנת SEESat על מנת לבחון את רמת השירות הניתנת לשני סוגי לקוחות אלו. הוא טוען שבחודש יולי 2001 רמת השירות לא היתה טובה וזאת על סמך הגרפים הבאים:

Figure 1: Average waiting time for Business and Platinum calls (July-01)

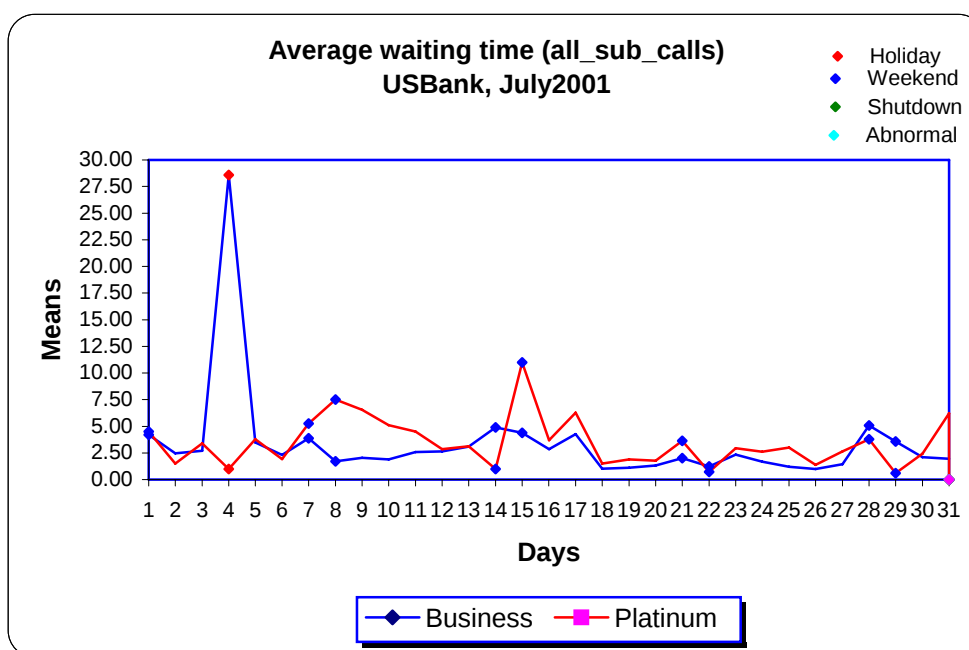
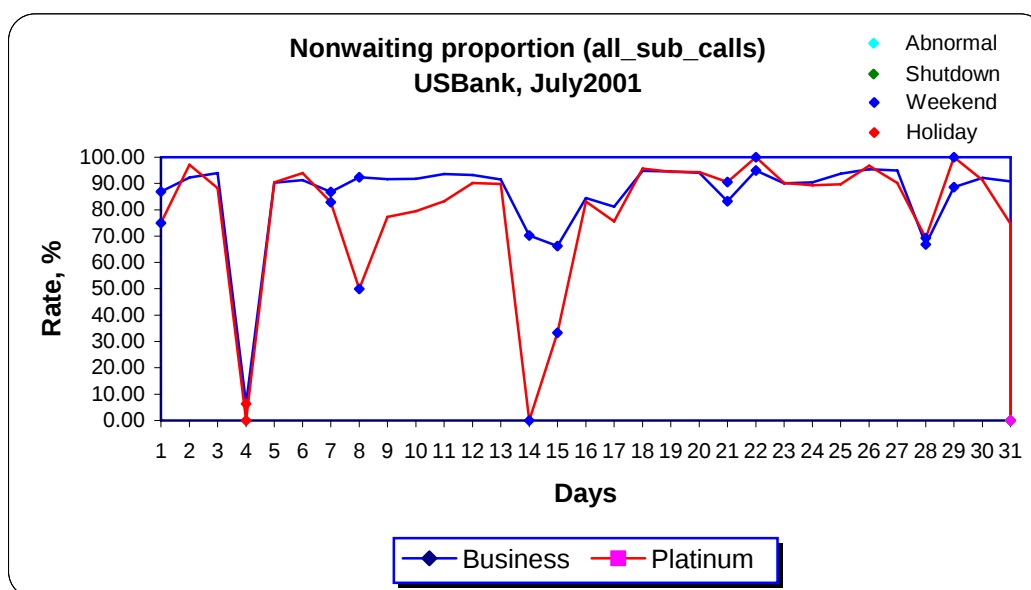


Figure 2: % immediately answered for Business and Platinum calls (July-01)



2.1 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-Business קיבלו רמת שירות טובה יותר מלקוחות ה-Platinum, כיוון שברוב הימים (ובייחוד בשבוע שבין ה-7 ל-16 ביולי) לקוחות ה-VIP מחכים יותר וגם ההסתברות שיקבלו שירות מיידי נמוכה יותר.

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

תשובה:

- ניתן להעביר קבוצת מוקדנים מ-Business ל-Platinum.
- ניתן לאחד את שתי קבוצות המוקדנים ולהגדיר מדיניות ניתוב שתיתן עדיפות ללקוחות ה-VIP.
- ניתן לאפשר למוקדני ה-Business לתת שירות גם ללקוחות ה-Platinum, אם למשל יש יותר מ-X לקוחות Platinum בתור.

2.3 תארו, במדויק ככל שתוכלו, כיצד ניתן לייצר את גרף 2 בעזרת תוכנת SEEStat. (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 שרות Business ו- Platinum.
3. נעבור לבחירת התאריכים ונסמן את יולי 2001, נבחר באופציית ה- Days for one month.

Variables
Options
Select Categories

Select by

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 (all_sub_calls)	Counts	service	
Handled proportion (all_sub_calls)	Counts	service	
Unhandlded (all_sub_calls)	Rate	service	
Unhandlded proportion (all_sub_calls)	Counts	service	
Short abandons (all_sub_calls)	Rate	service	
Short abandons proportion (all_sub_calls)	Counts	service	
Abandons (all_sub_calls)	Rate	service	
Abandons proportion (all_sub_calls)	Counts	service	
Other unhandlded (all_sub_calls)	Rate	service	
Other unhandlded proportion (all_sub_calls)	Counts	service	
Disconnected in service (all_sub_calls)	Rate	service	
Disconnected in service proportion (all_sub_calls)	Counts	service	
Nonwaiting proportion (all_sub_calls)	Rate	service	
Average waiting time given waiting (all_sub_calls)	Means	service	
Average waiting time (all_sub_calls)	Means	service	
Average customer service time (all_sub_calls)	Means	service	

DB Tables
Dictionaries
Calendar
Summaries
Variables
Procedures
Dates ->

New Model
OK

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

Time series type

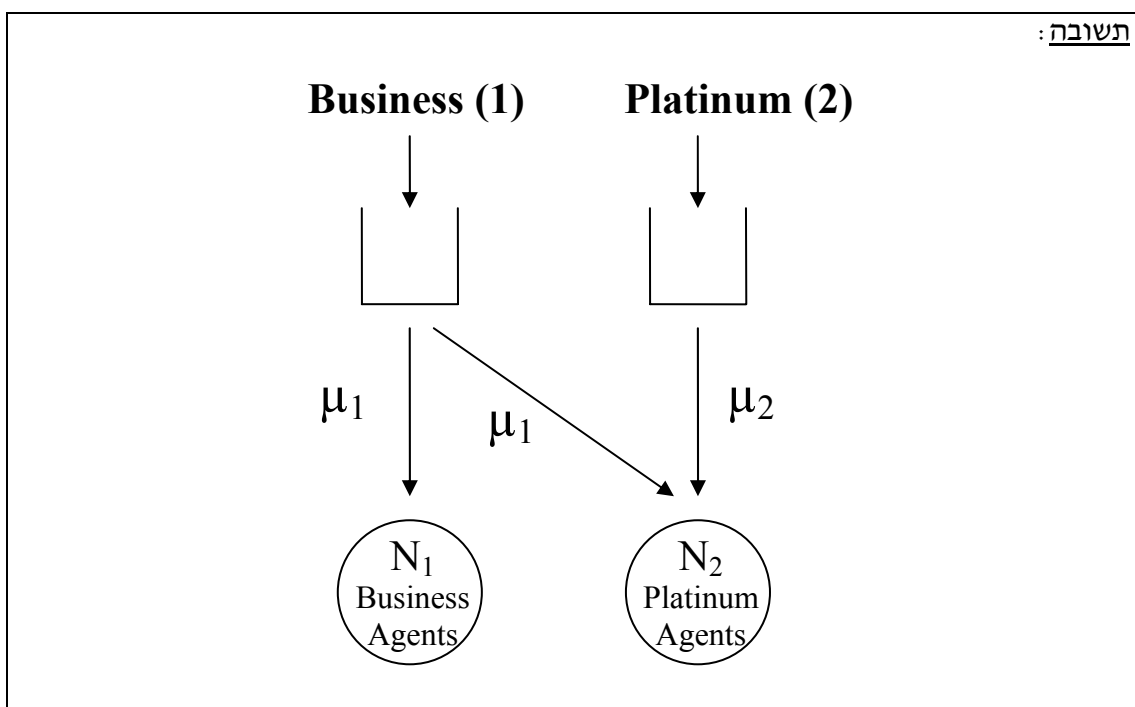
☒ Days for one month

☐ Aggregated days by months

כדי לשפר את רמת השירות ללקוחות Platinum, מציע מנהל השיווק להוסיף מוקדנים לקבוצת מוקדני ה-Platinum. כדי להבטיח את יעילות מוקדני ה-Platinum, הוא מציע בנוסף גם לאפשר להם לתת שירות ללקוחות ה-Business. לשם כך הוא הגדיר את מדיניות השירות הבאה:

- א. קבוצת מוקדני ה-Business נותנת שירות רק ללקוחות Business.
- ב. קבוצת מוקדני ה-Platinum נותנת שירות לכולם, תוך מתן עדיפות ללקוחות Platinum.
- ג. אם לקוח Business מגיע ויש מוקדנים פנויים בשתי הקבוצות, הוא יועבר למוקדני ה-Business.
- ד. 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.
- ה. המדיניות הינה *work conserving*, ז"א לא ייתכן מצב שבו יש שרתים פנויים ובעת ובעונה אחת יש גם לקוחות הממתינים בתור.

2.4 תארו את הארכיטקטורה של המערכת המוצעת בעזרת סכמה המתארת את זרימת הלקוחות במערכת, תחת ההנחה שאין נטישות. (השלימו את התרשים)



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

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

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

2.5 כתבו משוואה דיפרנציאלית עבור $X_1(t)$ ו- $X_2(t)$, $0 \leq t$.

תשובה:

$$\begin{aligned} \frac{dX_1(t)}{dt} &= \lambda_1 - \mu_1 (X_1(t) \wedge N_1) - \mu_1 \left((X_1(t) - N_1)^+ \wedge (N_2 - X_2(t))^+ \right) = \\ &= \lambda_1 - \mu_1 (X_1(t) \wedge N_1) - \mu_1 \left(X_1(t) \wedge \left(N_1 + (N_2 - X_2(t))^+ \right) \right) \\ \frac{dX_2(t)}{dt} &= \lambda_2 - \mu_2 (X_2(t) \wedge N_2) \end{aligned}$$

2.6 תארו בקיצור כיצד ניתן, בעזרת גליון אלקטרוני בלבד, לפתור בקירוב (מדויק ככל הנדרש) את המשוואה הדיפרנציאלית בסעיף 2.5.

תשובה:

נתחיל עם $X(0) = (X_1(0), X_2(0))$ נתון. לאחר מכן, עבור $n = 1, 2, \dots$, $t_n = t_{n-1} + \Delta$ נחשב,

$$X_1(t_n) = X_1(t_{n-1}) + \lambda_1 \Delta - \mu_1 (X_1(t) \wedge N_1) \Delta - \mu_1 \left((X_1(t) - N_1)^+ \wedge (X_2(t) - N_2)^+ \right) \Delta$$

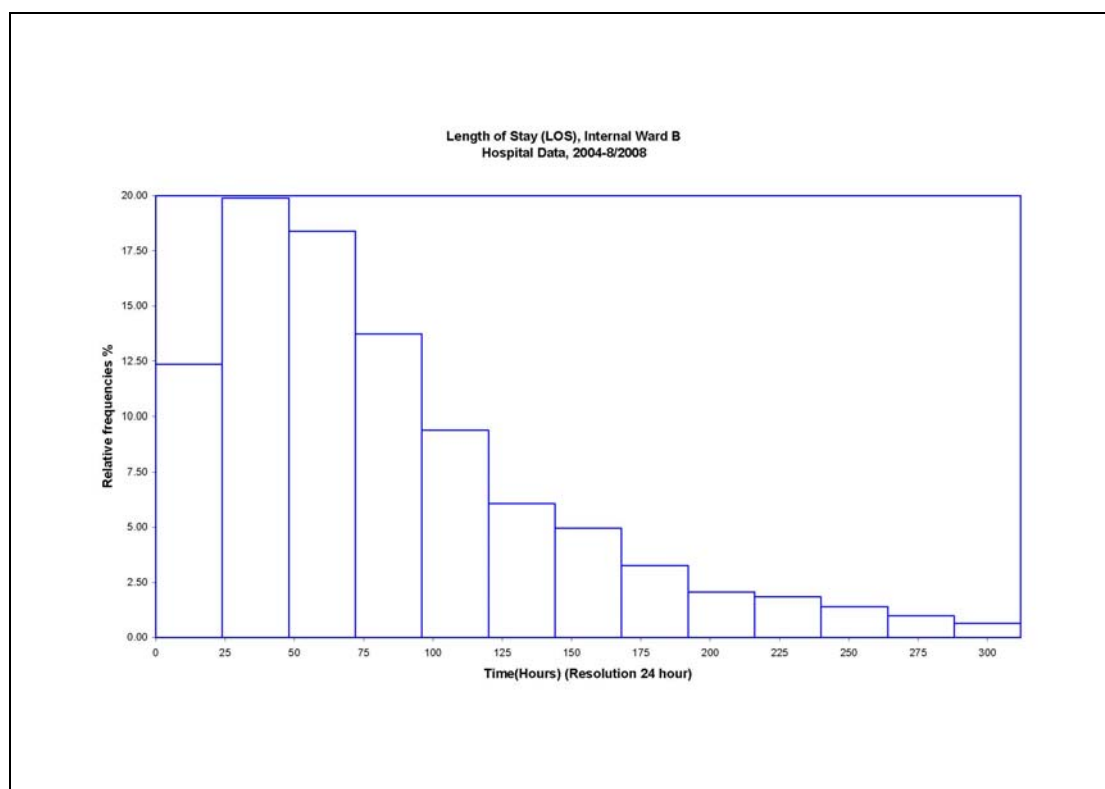
$$X_2(t_n) = X_2(t_{n-1}) + \lambda_2 \Delta - \mu_2 (X_2(t) \wedge N_2) \Delta$$

ונקבל את $X(0), X(\Delta), X(2\Delta), \dots$ ככל ש- Δ קטן יותר, כך הקירוב מדויק יותר.

שאלה 3 – יישום: בתי חולים

בית חולים כולל מחלקות אשפוז שונות. בכל מחלקה מספר קבוע של מיטות - N (בין 10 ל-50). כאשר מוחלט במיון לאשפוז חולה במחלקה כלשהי, מתבצע ברור האם אכן יש מקום במחלקה. במידה וכל המיטות תפוסות (המחלקה מלאה), החולה מועבר לאשפוז במחלקה אחרת, בה מטפלים בו. בדרך כלל רמת הטיפול במחלקה החלופית פחות טובה לחולה, כיוון שהצוות הסייעודי בה פחות מיומן בטיפול בבעיותיו הספציפיות. מצב זה אינו רצוי ולכן הוחלט לעקוב אחרי אחוז החולים שאין להם מקום במחלקה אליה יועדו לראשונה.

3.1 הגרף הבא מראה את התפלגות משך השהיה של חולה במחלקה כלשהי בימים. לאיזו משפחה של התפלגויות נראה שהגרף מתאים? כיצד הייתם בודקים השערה זו?



תשובה:

נראה כי זמן האשפוז נלקח מהתפלגות לוג-נורמלית. כדי לבדוק זאת ניתן לצייר גרף Q-Q plot של מדגם זמני השהיה לעומת שברונים של התפלגות לוג-נורמלית עם הפרמטרים המתאימים. או לחלופין, ואולי עדיף, להתאים את היסטוגרמת לוג-משך-הזמן לפעמון הנורמאלי.

לצרכים מסוימים, ניתן להניח שחולה שוהה במחלקה מסוימת, שבה 30 מיטות, זמן המתפלג מעריכית עם ממוצע 6 ימים, ושתהליך סך הגעות חולים למחלקה זו הוא פואסוני, עם קצב הגעה קבוע של 6 חולים ביום.

3.2. איזה מודל נראה לכם מתאים לתיאור מספר החולים הנמצאים במחלקה הנידונה? מדוע? הגדירו את ערכי הפרמטרים של המודל.

תשובה:

כאשר מסתכלים על מחלקה אחת, עפ"י התיאור הנ"ל, המודל המתאים הוא Erlang-B $(M/M/N/N)$, שבו מספר ה"שרתים" הוא מספר המיטות. הפרמטרים המגדירים מודל זה הם: $\lambda = 6 \text{ patients / day}$; $\mu = 1/6$ ו- $N=30$.

3.3. שוקלים לשפץ את המחלקה הנידונה. כיצד תקבעו את מספר המיטות הנדרש על סמך הנתונים הנ"ל, כך שההסתברות שלחולה לא יהיה מקום במחלקה תהיה בסדר גודל של עד 7%? מהו פחות או יותר תחום מספר המיטות הנדרש?

תשובה:

יש להשתמש במשטר QED, לכן נשתמש בכלל השורש:

$$N = R + \beta \sqrt{R}$$

$$R = \frac{\lambda}{\mu} = \frac{6}{1/6} = 36$$

$$\Rightarrow N = 36 + \beta \sqrt{36}$$

בפועל תחום ה- β הרלוונטי שצריך לבדוק הוא $[-1,1]$, ולכן $N \in [30, 42]$. (חישובים מדויקים יותר בעזרת 4CC, שאינם נדרשים בבחינה, מראים שמספיק לבחון בטא חיובי בתחום $[0,1]$)
(הערה: שימו לב כי לא סביר להשתמש בשיוויון: $N = \frac{\lambda(1 - P(\text{block}))}{\mu}$, כיוון שהוא מוביל למדיניות ED שבה אחוז הנחסמים גבוה יותר)

--

3.4. בבדיקה מעמיקה יותר התגלה שקיימים הבדלים מהותיים בין רמת השרות התפעולית הניתנת במחלקות השונות. בטבלה הבאה נתונים כמה מנתוני המחלקות, כפי שנאספו ממערכות המידע של בית החולים.

3.4.1. התאימו לכל מחלקה את משטר השרות בה היא מתפקדת (ED, QED, QD) והסבירו תשובתכם. חשבו את העומס המוצע (offered-load) של כל מחלקה. קבעו מהו פרמטר איכות השרות המתאים לכל מחלקה. (למשל, כאשר $N \approx R + \beta\sqrt{R}$, β הוא פרמטר איכות השרות).

מחלקה	מספר מיטות	ממוצע כמות חולים שמופנים למחלקה ביום	זמן אשפוז ממוצע בימים	אחוז התפוסה (ניצולת המיטות)	אחוז חולים שאין להם מקום
טיפול נמרץ	13	1.54	9.47	85%	24%
פנימיות (5 מחלקות)	180	34.32	5.24	94%	5.9%
יולדות (2 מחלקות)	60	15.52	2.8	67.5%	0.2%

תשובה:

מחלקת טיפול נמרץ מתפקדת במשטר ED, אחוז התפוסה גבוה ואחוז החסימה גבוה. המחלקות הפנימיות מתפקדות במשטר QED, כיוון שאחוז התפוסה סביב 90% ואחוז החסימה נמוך מ-10%, אך לא 0.

מחלקת יולדות מתפקדת במשטר QD, כיוון שאחוז התפוסה נמוך ואין כמעט חסימות. פרמטר השרות של המחלקות השונות הוא:

מחלקה	$\text{offered-load} = \frac{\lambda}{\mu}$	נוסחה לחישוב פרמטר השרות	פרמטר איכות השרות
טיפול נמרץ	14.6	$\gamma = \frac{N - R}{R}, \gamma < 0$	$\gamma = \frac{13 - 14.6}{14.6} = -0.11$
פנימיות	179.8	$\beta = \frac{N - R}{\sqrt{R}}$	$\beta = \frac{180 - 179.8}{\sqrt{179.8}} \approx 0$
יולדות	43.5	$\delta = \frac{N - R}{R}, \delta > 0$	$\delta = \frac{60 - 43.5}{43.5} = 0.38$

הערה: שימו לב שב-Erlang-B, התחומים התפעוליים (רמות האיוש) נקבעים ע"י נוסחאות זהות למודל Erlang-A.

3.4.2. מדוע למחלקת טיפול נמרץ אחוז גבוה של חולים שאין עבורם מקום למרות שאחוז תפוסתה בינוני (באופן יחסי, למשל ביחס למחלקות הפנימיות).

תשובה:

למחלקת טיפול נמרץ מספר מיטות קטן וכתוצאה מכך לא ניתן ליהנות מהיתרון לגודל. לכן למרות תפוסה יחסית נמוכה, אחוז החולים הנדחים גבוה.

3.4.3. חוו דעתכם – האם הגיוני שנבחר משטר כמתואר למחלקת טיפול נמרץ? אם לא, לאיזה משטר הייתם מצפים ומדוע? מה לדעתכם היו הנסיבות שהביאו לבחירתו של משטר מסוג זה?

תשובה:

משטר זה למחלקת טיפול נמרץ הוא לא הגיוני. היינו מצפים למשטר QD, כך שלחולים שזקוקים למיטת טיפול נמרץ כמעט תמיד תהיה מיטה פנויה. יש לפחות שתי סיבות אפשריות: הראשונה והסבירה ביותר היא שהסיבה למדיניות זו היא כלכלית - מיטות טיפול נמרץ הן יקרות כיוון שדורשות צוות אחיות מיומן מאוד ומחייבות מספר קטן של חולים פר מיטה, ולכן מאילוצי תקציב לכח אדם, המחלקה קטנה מדי. סיבה שניה אפשרית היא, שמשאירים במחלקה חולים שלא חייבים להיות שם (אך עדיף להם קלינית להיות שם), וברגע שמישהו במצב יותר קריטי זקוק למקום הם מפונים; במקרה זה המודל Erlang-B לא מתאים.

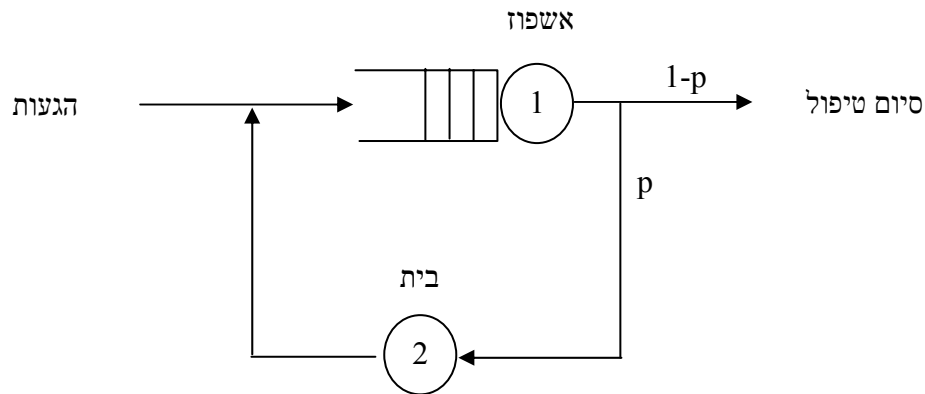
3.5. לאורך השנים, אוכלוסיית האזור שמשרת בית החולים גדלה. בית החולים מתעתד לבצע תכנון לטווח ארוך ולשם כך נדרש לאמוד את מספר המיטות הנחוצות בעוד 3 שנים במחלקות הפנימיות. לצורך כך, נניח כי בעוד 3 שנים קצב ההגעה למחלקות אלו יעלה פי 1.5. בית החולים רוצה לבחון מה יקרה לרמות השרות ואחוז התפוסה אם מספר המיטות יוגדל גם הוא פי 1.5.

ביחס למצב הקיים, מה יקרה למדד אחוז התפוסה ומדד איכות השרות בעוד 3 שנים?

תשובה:

אחוז ה"אי-תעסוקה" והסתברות החסימה במשטר QED הן פרופורציוניים ל- $1/\sqrt{n}$. מתוכננות למשטר QED, אזי ההסתברות לחסימה תרד פי $\sqrt{1.5} \approx 1.22$ ו- % התפוסה יעלה. (הערה: ניתן לראות שאחוז התפוסה ירד גם ע"ס הנוסחה: $\rho = \frac{\lambda(1 - P(Block))}{N\mu}$ ולכן, כיוון שהמחלקות הפנימיות $P(block) \approx \frac{\gamma}{\sqrt{n}} = \frac{\phi(\beta)}{\Phi(\beta)\sqrt{n}}$, $Occupancy \approx 1 - \frac{\beta + \gamma}{\sqrt{n}}$)

3.6. מחלקה אונקולוגית שונה ממחלקות אשפוז רגילות בכך שרוב החולים מתאשפזים לא לתקופה אחת רצופה אלא מגיעים לסדרות של אשפוז. כלומר, בכל פעם החולה מגיע לאשפוז של כמה ימים ואם מצבו מאפשר זאת, חוזר לביתו עד לסדרת הטיפולים הבאה (בדרך כלל אחרי מספר שבועות לא גדול). כדי לתאר את מספר החולים בטיפול / בבית, נשתמש ברשת התורים הבאה:



תחנה 1 מתארת את האשפוז בבית החולים לקבלת טיפול. בתחנה זו מספר ה"שרתים" הוא מספר המיטות במחלקה (N), והתור לפני התחנה הוא ההמתנה לאשפוז במיון. תחנה 2 מתארת את שהיית החולה בבית בין הטיפולים. לאחר כל טיפול מחליטים הרופאים אם יש צורך בטיפול נוסף או לא. % החולים שעבורם נדרש טיפול נוסף הוא $100 \cdot p$.

3.6.1 נגדיר משתנה מקרי K הסופר את מספר הטיפולים שכל חולה עובר. כיצד מתפלג K? מהי תוחלתו? (ניתן להיעזר בטבלת ההתפלגויות המצורפת בסוף המבחן)

תשובה:

$$K \text{ מתפלג גיאומטרית (עם הפרמטר } (1-p) \text{) ותוחלתו } \frac{1}{1-p}.$$

3.6.2 כמה "שרתים" סביר שיהיו במודל של תחנה 2? הסבירו תשובתכם.

תשובה:

כיוון שלכל חולה בית משלו, מספר המשאבים בשלב זה אינו מוגבל, כלומר אינסוף שרתים.

3.6.3 הסבירו מדוע ה- offered-load הוא $R = \frac{\lambda}{(1-p)\mu}$?

תשובה: כיוון שכל חולה מבקר בממוצע $E[K]$ פעמים, קצב ההגעה בפועל הוא $\frac{\lambda}{1-p}$ ולכן R נתון כנ"ל. (שימו לב ש- R הוא גם מספר הלקוחות בתחנה 1 כאשר "שמים" בה אינסוף שרתים) הסבר אלטרנטיבי: כל חולה שוהה במערכת כולה זמן שהיה המתפלג מעריכית עם פרמטר $(1-p)\mu$, וזאת כיוון שסכום גיאומטרי של מ"מ מערכיים שווי התפלגות מתפלג מעריכית.

3.6.4 במודל הנ"ל, כל חולה המגיע למחלקה מחכה עד שתתפנה עבורו מיטה. הציעו מדד בו ניתן להשתמש על מנת לבדוק רמת שרות תפעולית?

תשובה: אחוז החולים הממתינים למיטה. עבור אמידתו ניתן למדוד את ההסתברות שמספר החולים בתחנה 1 יהיה גדול מ- $N-1$. או אחוז החולים הממתינים יותר מאשר זמן נתון (למשל שעה).

3.7. ידוע כי מספר התקנים של רופאים ואחיות בבתי חולים פרופורציוני למספר המיטות במחלקה. יחס הפרופורציה שונה בין המחלקות השונות בהתאם לחומרה הקלינית וזמני הטיפול הנדרשים. לדוגמה: לכל 6 מיטות במחלקה פנימית יש תקן של אחות אחת, אך במחלקת טיפול נמרץ לכל 2 מיטות אחות אחת.

3.7.1 על איזה מדדי שרות תפעוליים משפיע מספר האחיות במחלקה? האם אלו המדדים החשובים ביותר לצורך תמיכה בקבלת החלטות?

תשובה: כמות האחיות משפיע על זמינות השרות הסיעודי. לדוגמה על המדד: תוחלת זמן המתנה לאחות. אבל עם כל הכבוד למדדים תפעוליים, מדדי איכות קליניים חשובים יותר. ואז כמובן מעניין לבדוק את השילוב ביניהם: למשל איכות טיפול קליני יורדת ככל שהמתנה לאחות עולה.

3.7.2. מתוך המודלים שנלמדו בכיתה, באיזה מודל הייתם משתמשים על מנת לקבוע את מספר האחיות הדרוש? מדוע?

תשובה:

Erlang-C הוא המודל המתאים ביותר, כיוון שבמציאות בה אנו דנים נטישות וחסיונות נראים כיחסית לא משמעותיים.

3.7.3. נניח כי קצב בקשות חולים לטיפול הניתן ע"י אחות הוא 20 בקשות לשעה (מכל החולים במחלקה), ומשך טיפול אחות מתפלג מעריכית עם ממוצע של 15 דקות. כמה אחיות צריכות להיות במחלקה אם רוצים שההסתברות שחולה ימתין לאחות לא תעלה על 30%? ענו על השאלה בעזרת הגרף המצורף. הסבירו את תשובתכם.

תשובה:

הפרמטרים של המודל הם: $\lambda = 20 \text{ requests / hour}$, $\mu = 4 \text{ requests / hour}$. על כן

$$R = \frac{\lambda}{\mu} = \frac{20}{4} = 5$$

על סמך הגרף, אם רוצים שההסתברות להמתנה לא תעלה על 30%, פרמטר

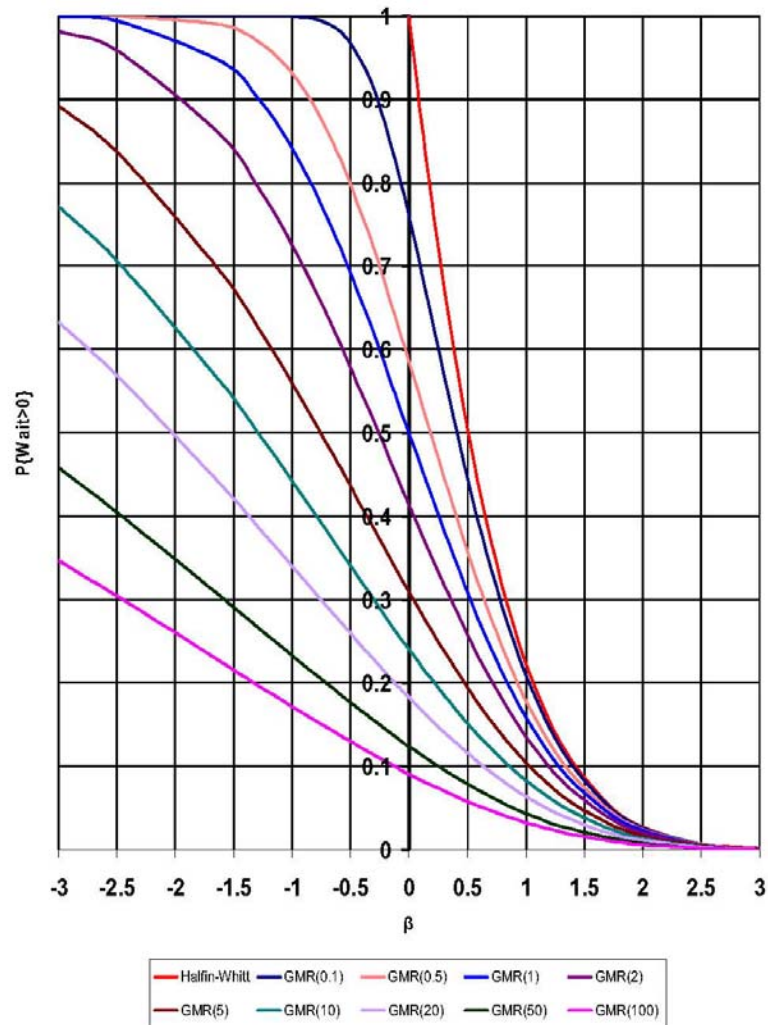
איכות השירות צריך להיות $\beta = 0.8$. ולכן ע"ס כלל השורש:

$$S = \left\lceil R + \beta \sqrt{R} \right\rceil = \left\lceil 5 + 0.8 \sqrt{5} \right\rceil = 7$$

כלומר צריך לפחות 7 אחיות.

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 $\theta/\mu = x$. 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 ושימושיה: עומס מוצע (Offered-Load)

חלק 1. ייצוגים.

במערכת שירות $M_t | G | \infty$, המופע הוא פואסוני לא הומוגני בזמן עם קצב מופע $\{\lambda(t), t \geq 0\}$, ומשכי השירות S הם בלתי-תלויים, שווי התפלגות, ובלתי תלויים במופע. נסמן ב- $L(t)$ את מספר האנשים במערכת בזמן t ויהי $R(t)$ ממוצע מספרם: $R(t) = EL(t)$. ניתן להראות ש- $R(t)$ נתון על ידי כל אחד משלושת הביטויים הבאים:

$$R(t) = \int_{-\infty}^t \lambda(u) P(S > t-u) du = E \left(\int_{t-S}^t \lambda(u) du \right) = E(S) \cdot E(\lambda(t-S_e))$$

S_e מסמן את "זמן שארית השירות" כפי שהוגדר בכיתה. (בייצוגים הנ"ל, $\lambda(t) = 0$ עבור t שליליים.)

4.1.1 נניח שמשך השירות S הוא **דטרמיניסטי** $D : S \equiv D$. הוכיחו את השוויון בין שלושת ההצגות של $R(t)$ למקרה זה.

תשובה:

S_e מתפלג אחיד על $[0, D]$. $S_e \stackrel{d}{=} U(0, D)$.

$$\text{Representation 3} = D \cdot E(\lambda(t-S_e)) = D \cdot \int_0^D \frac{1}{D} \cdot \lambda(t-x) dx = \int_0^D \lambda(t-x) dx \stackrel{u=t-x}{=} \int_{t-D}^t \lambda(u) du = \text{Rep. 2}$$

$$\text{Rep. 1} = \int_{-\infty}^t \lambda(u) 1_{\{u > t-D\}} du = \int_{t-D}^t \lambda(u) du = \text{Rep. 2}$$

4.1.2 הסבירו כיצד ניתן להרחיב את המקרה הדטרמיניסטי ב- 4.1.1 הנ"ל למקרה של S המקבל

מספר סופי של ערכים, דהיינו כיצד מוכיחים שאכן :

$$\int_{-\infty}^t \lambda(u) P(S > t-u) du = E \left(\int_{t-S}^t \lambda(u) du \right) = E(S) \cdot E(\lambda(t-S_e))$$

הערה : מאחר וניתן לקרב כל משתנה מקרי, בדיוק רב כרצונכם, על ידי משתנה מקרי המקבל מספר סופי של ערכים (מ"מ דיסקרטי), סעיף זה מוביל להוכחת שלושת הזהויות הנ"ל למקרה של S כללי.

תשובה :

נניח p_i w.p. $S = D_i$, מ"מ דיסקרטי.

על סמך עמידה מוטה, $\frac{p_i D_i}{E(S)}$ w.p. $S_e = U_i$, כאשר $U_i \stackrel{d}{=} U(0, D_i)$,
לכן,

$$\begin{aligned} \text{Representation 2} &= \sum_i p_i \int_{t-D_i}^t \lambda(u) du \stackrel{\lambda_i = \lambda \cdot p_i}{=} \sum_i \int_{t-D_i}^t \lambda_i(u) du \stackrel{4.1.1}{=} \sum_i D_i E(\lambda_i(t-U_i)) = \\ &= E(S) \sum_i \frac{p_i D_i}{E(S)} E(\lambda(t-U_i)) = E(S) E(\lambda(t-S_e)) = \text{Representation 3} \end{aligned}$$

$$\begin{aligned} \text{Representation 2} &= \sum_i p_i \int_{t-D_i}^t \lambda(u) du = \sum_i p_i \int_{-\infty}^t \lambda(u) 1_{\{u > t-D_i\}} du = \\ &= \int_{-\infty}^t \lambda(u) \sum_i p_i P(D_i > t-u) du = \text{Representation 1} \end{aligned}$$

חלק 2 . שימושים.

בכתה הוסבר ש- $R(t)$, כפי שהוגדר לעיל, משמש כהגדרת העומס המוצע (Offered Load) למערכת שירות בזמן t , כאשר המופע אליה ומשכי השירות בה הם כמתואר בחלק 1. (כזכור, $R(t)$ נמדד ב-Erlang).

4.2.1 הסבירו בקצרה מדוע אכן ההגדרה מתאימה. עשו זאת בהסתמך על הייצוג הרביעי הבא של $R(t)$:

נסמן $A(t)$ – מספר המופעים הכולל עד זמן t (כולל).

$E[A(t)] = \Lambda(t)$ – תוחלת מספר המופעים המצטבר עד זמן t .

אזי, מאחת ההצגות של $R(t)$ נובע שמתקיים: $R(t) = E[\Lambda(t) - \Lambda(t - S)]$.

תשובה:

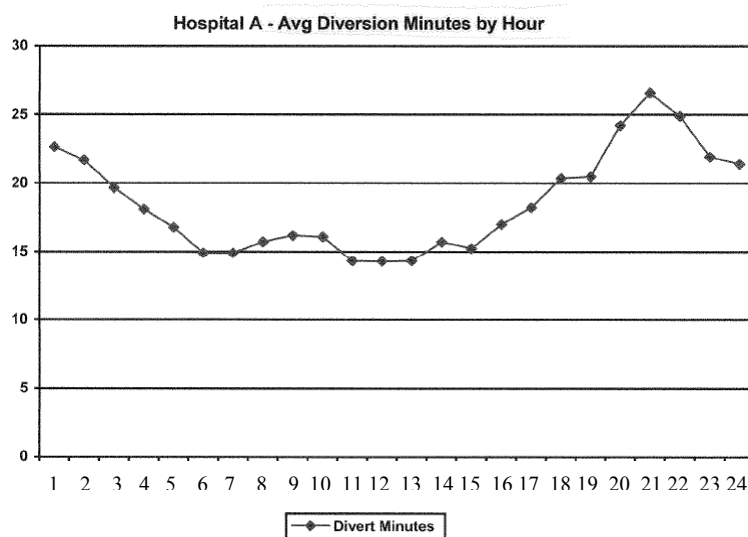
$R(t)$ = כמות העבודה הנמצאת במערכת בזמן t ; עבודה נמדדת ביחידות זמן שירות, או לחילופין

מספר יחידות זמן-שרת הנדרשות לטיפול. לכן $R(t) = E(L(t))$.

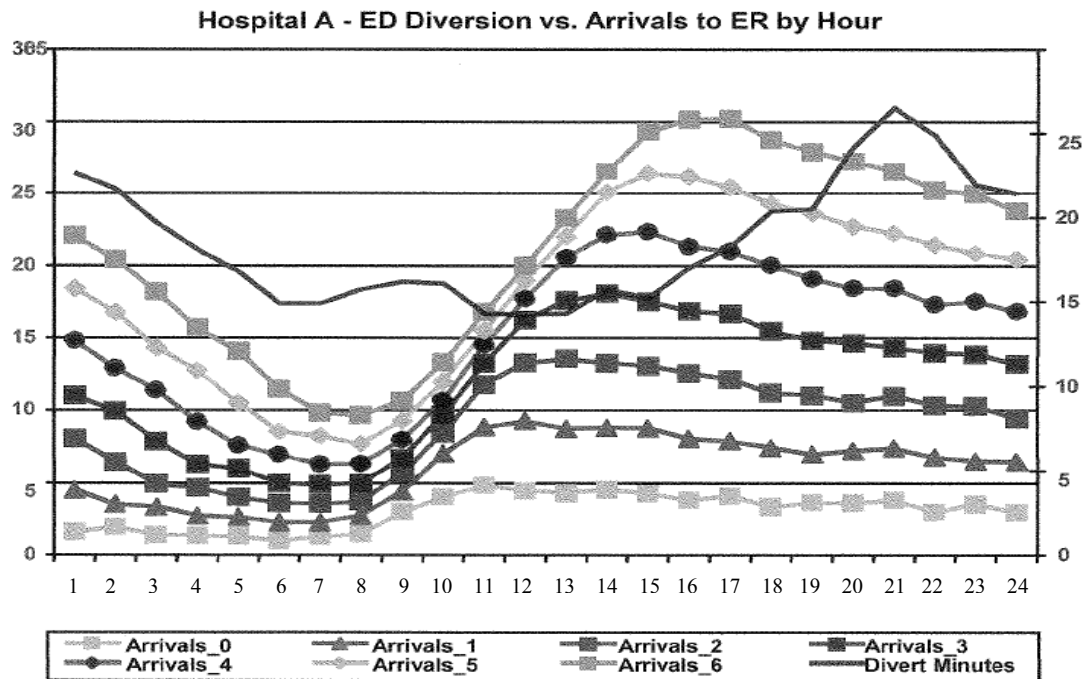
הביטוי $R(t) = E[\Lambda(t) - \Lambda(t - S)]$ ממחיש יפה את העובדה שכמות העבודה בזמן t כוללת

עבודה שהגיעה לפני כן, בהתאם לתהליך ההגעות הממוצע $\Lambda(t)$.

4.2.2 הגרף הבא מתאר את תופעת ה-Ambulance Diversion (הסטת אמבולנסים) שתוארה בכיתה (ונפוצה בארה"ב). למשל ערך הגרף בשעה 21:00 הוא אחוז הפעמים בתקופה הנמדדת שהשעה 20:00-21:00 הוכרזה כשעה שבה המיון חסום לכניסת חולים חדשים דרך אמבולנסים (במקרה שלנו 27%).



הגרף הבא מנסה לקשר בין שעת שיא הסטת האמבולנסים לשיא העומס במיון. הסבירו את הגרף. כיצד ניתן היה לשפר את ההתאמה בין הזמנים בהם מתרחשים השיאים? שימו לב: תוחלת זמן השהיה במיון הנדון היא 6 שעות.



תשובה:

הגרף Arrivals_0 מתאר את כמות החולים הממוצעת לשעה שהגיעו למיון לאורך היום, עם שיא הגעות ב- 11 בבוקר. הגרף Arrivals_1 בשעה t הוא כמות ההגעות בשעה $t + 1$, כלומר הוא סוכם את ההגעות בשעתיים האחרונות. וכי... כלומר הגרף Arrivals_6 סוכם את ההגעות ב- 6 השעות האחרונות ע"י הנוסחה: $Arrivals_6(t) = \Lambda(t) + \Lambda(t-1) + \dots + \Lambda(t-6)$. לכן גרף 6 מחשב בפועל את העומס המוצע ($R(t)$) לפי הקירוב הבא:

$$R(t) = E[\Lambda(t) - \Lambda(t-S)] \approx \Lambda(t) - \Lambda(t-E[S]) = \Lambda(t) + \Lambda(t-1) + \dots + \Lambda(t-6)$$

מראים ששיא העומס מתקבל בשעה 15-17, כלומר הוא מוזז בזמן לעומת שיא ההגעות שקורה ב- 10-11.

אך עדיין יש פער זמן לא מוסבר בין שיא העומס הנ"ל לשיא ההסטות (שמתרחש בשעה 21). הפרש זה יכול לנבוע, למשל, מאחד ההסברים הבאים (או משילובם): א. האמד ל- $R(t)$ לא מספיק מדויק. במקרה זה אנו ממליצים על השיפור הבא: כדאי להשתמש בנוסחת העומס המוצע המדויקת, או באחד הקירובים היותר מדויקים שנלמדו בכיתה.

ב. הסטת האמבולנסים, כפי שהוסבר בכיתה, בדרך כלל מופעלת כאשר מחלקות האשפוז חסומות, וזה קורה בשעה יותר מאוחרת, לעומת שעת שיא העומס במיון.

ג. יתכן שפער הזמנים מצביע על כך שהפעלת נוהל הסטת האמבולנסים מופעל מאוחר מדי, כלומר רק לאחר שרואים עומס בלתי סביר במיון, מתחילים להפעיל אותו ועד שהוא מגיע לאפקטיביות מירבית עובר זמן ניכר.

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	$X \sim NB(m, p)$	$p(k) = \binom{k-1}{m-1} p^m q^{k-m}$	$0 \leq p = 1 - q \leq 1; m \geq 1$	$k = m, m+1, \dots$	$\left[\frac{p \cdot e^t}{1 - q \cdot e^t} \right]^m$	$\frac{m}{p}$	$\frac{mq}{p^2}$	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	$X \sim HG(N, M, n)$	$p(k) = \frac{\binom{N}{k} \binom{M}{n-k}}{\binom{N+M}{n}}$	$N, M > 0;$ $1 \leq n \leq N + M$	$\max(0, n - M) \leq k$ $k \leq \min(n, N)$	$\sum_k e^{tk} p(k)$	$\frac{n \cdot N}{N+M}$	$n \cdot \frac{N}{N+M} \cdot \frac{N}{N+M} \cdot \frac{M}{N+M} \cdot \frac{M}{N+M}$ $\cdot \frac{N+M-n}{N+M-1}$	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	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$	$\alpha = \beta = 1 \rightarrow \text{Uniform}$
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	