



מרצה : ד"ר גלית יום-טוב

מתרגל : אריק סנדרוביץ'

תאריך הבחינה : 17.7.2014

מס' סטודנט _____

שם _____

הנדסת מערכות שירות – 096324

מועד א' – סמסטר אביב 2014

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.

שאלה 1	_____ מתוך 20 נקודות	בחינות קודמות/הרצאות/תרגולים
שאלה 2	_____ מתוך 20 נקודות	תרגילי בית
שאלה 3	_____ מתוך 45 נקודות	יישום
שאלה 4	_____ מתוך 15 נקודות	תיאוריה

סה"כ _____ מתוך 100 נקודות

בהצלחה !

שאלה 1 - בחינות קודמות (20 נקודות).

חלק 1. מופעים למערכת שירות (10 נקודות).

1.1 (7 נקודות) בתרשימים להלן מתוארים מופעים למוקד שירות טלפוני בארץ, בימים שונים של השבוע: הגרף הראשון מתאר מספר מופעים לחצי שעה והשני את מספר מופעים זה ביחס לממוצע היומי (באחוזים).

א. באיזה מודל מתמטי אפשר להשתמש כדי לתאר את תהליך המופע למוקד? אילו הנחות צריך להניח, והאם סביר להניח הנחות אלו?

ב. הצע שיטה למידול קצב המופע $\lambda(t)$. הצע שיטת חיזוי עבור הפרמטרים של המודל שהצעת.

תשובה:

א. המודל המתמטי המקובל לתיאור תהליך המופע למוקד שירות טלפוני הינו תהליך פואסון לא-הומוגני בזמן. ההנחות הנדרשות הן: לקוחות בלתי תלויים ומגיעים בבודדים (לא בקבוצה), זמנים בין-מופעים אקספוננציאליים. כתוצאה מכך, מספר הלקוחות באינטרוול זמן מסויים (לדוגמה: שעה), מתפלג

$$N_t \sim \text{Pois}(\Lambda(t)) \text{ כאשר } \Lambda(t) = \int_{u=0}^t \lambda(u) du.$$

ב. אנו ממליצים על שימוש במודל $\lambda(t) = C\lambda_0(t)$ כאשר המ"מ C הינו מספר המופעים ביום מסוים והפונקציה $\lambda_0(t)$ היא הצורה של המופעים ששונה עבור ימי חול, ימי שבת וימי ראשון.

אפשר לחזות את C ע"פ שיטות פשוטות כגון שיטת הממוצע הנע:

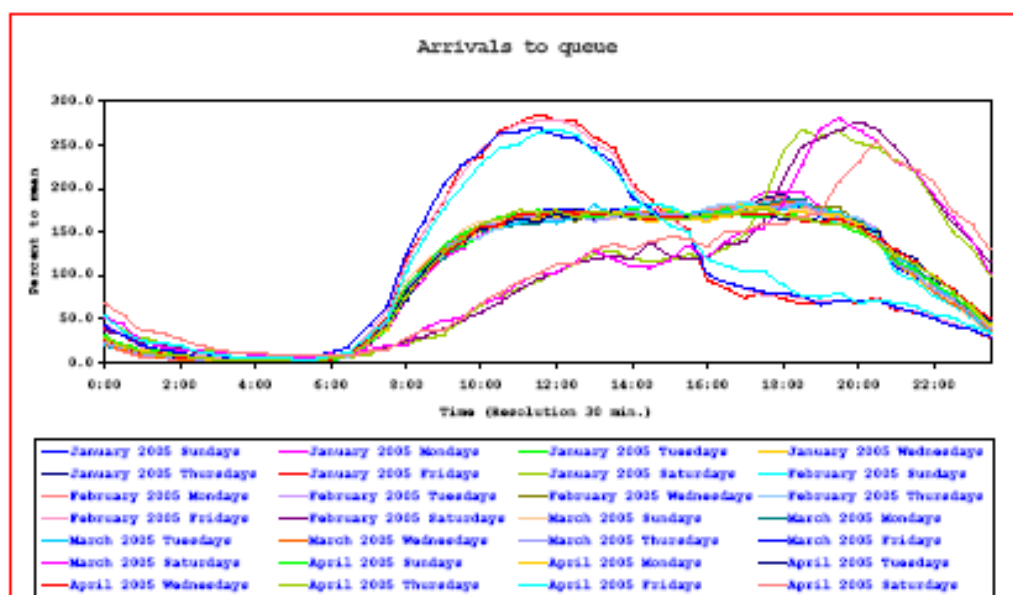
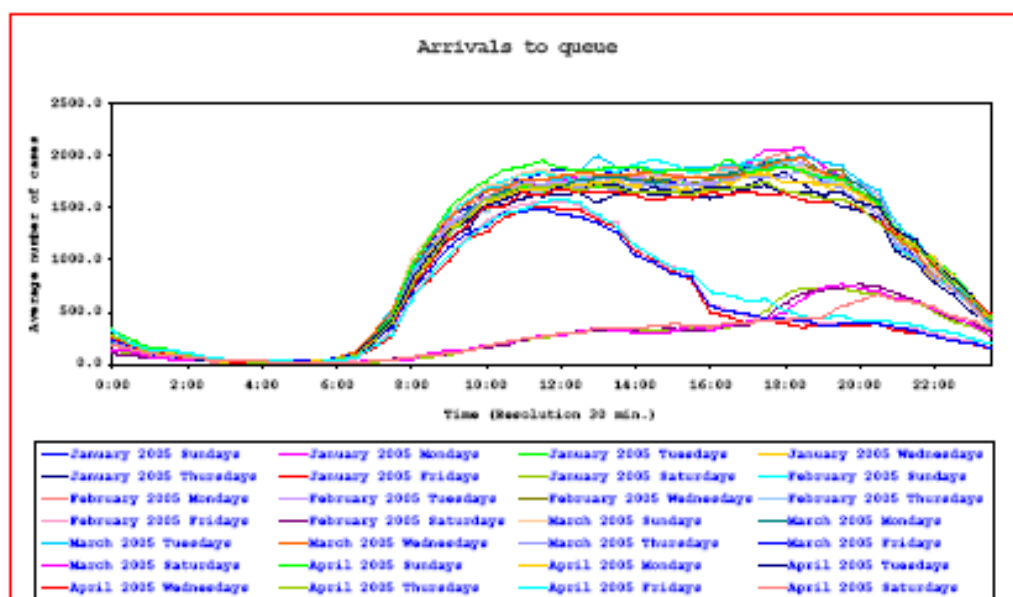
$$F_{j,k} = \frac{N_{j-7,k} + N_{j-14,k} + N_{j-21,k} + N_{j-28,k}}{4}$$

או $F_{j,k} = N_{j-7,k}$: Most Recent Observation

או ע"פ שיטות מתחום סדרות עתיות.

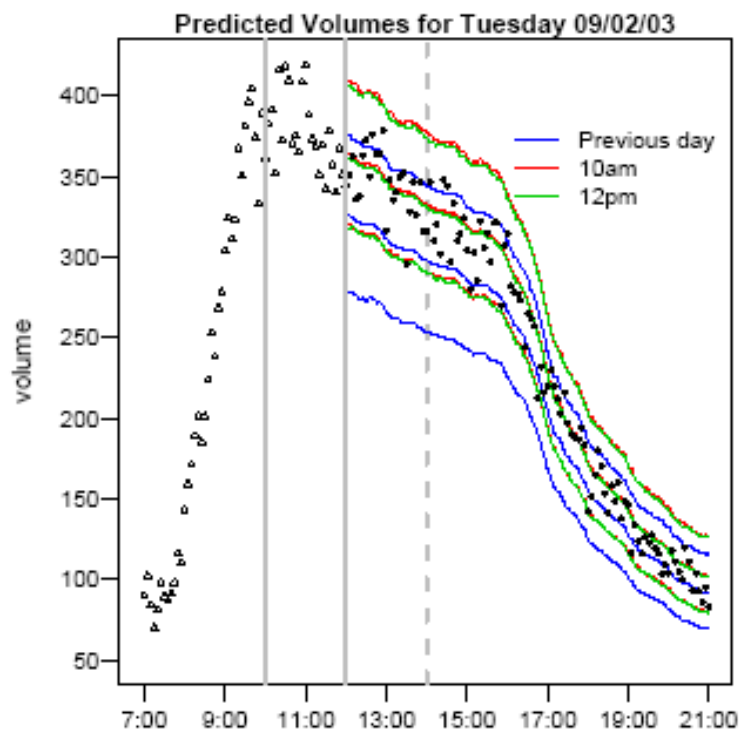
על פי "צורת" היום המתקבלת בגרף השני ניתן למצוא את $\lambda_0(t)$, בהתאם ליום המסוים בשבוע.

Arrival Patterns, Israeli Telecom, 2005



1.2 (3 נקודות) להלן מתוארים שלשה ניסיונות לחזות קצב מופע למוקד שירות טלפוני בארה"ב. התחזית היא למספר המופעים החל משעה 12:00. הנקודות מתארות את המספר האמיתי. סכמו בקצרה את המסקנות העיקריות הנובעות מדוגמה זו.

Comparison between Day-Ahead and Within-Day Predictions (Weinberg, Brown, Stroud, 2005)



תשובה:

נתונים מהבוקר (7:00-10:00) מאפשרים לשפר את איכות החיזוי בצורה משמעותית (ביחס לתחזיות ע"פ נתונים עד יום קודם). אין הבדל משמעותי בין תחזיות על סמך נתונים עד 10 בבוקר ותחזיות על סמך נתונים עד 12צהריים. ז"א, השעות 10:00-12:00 לא מוסיפות אינפורמציה.

חלק 2. משכי שירות (5 נקודות).

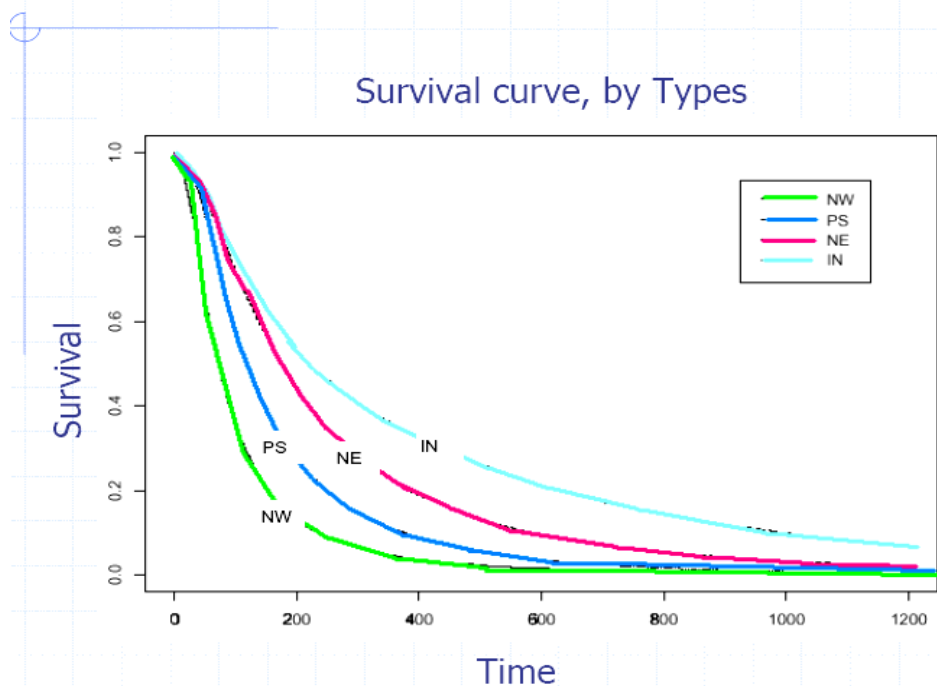
1.2 (2 נקודות) הגדירו יחס סדר סטוכסטי בין שני משתנים מקריים X ו- Y . (לצורך ההגדרה, השתמשו בסימונים F_X ו- F_Y עבור פונקציות ההתפלגות של X ו- Y בהתאמה.)

תשובה:

הגדרה: משתנה מקרי **חיובי** X גדול סטוכסטי ממשתנה מקרי **חיובי** Y אם מתקיים $\bar{F}_X(t) \geq \bar{F}_Y(t)$ לכל $t \geq 0$ כאשר $\bar{F} = 1 - F$ - היא פונקציית השרידות.

1.3 (3 נקודות) הגרף הבא מתאר את **פונקציית השרידות** (*survival function*) של משך השירות, עבור כל אחד מארבעת סוגי השירות במוקד שהכרנו בקורס.

Service Time



נסמן ב- T_{PS} משתנה מקרי בעל התפלגות משך השירות של לקוחות מסוג PS (פעילות שוטפת) וב- T_{NE} התפלגות משך השירות של לקוחות מסוג NE (ניירות ערך).
מי גדול יותר: $E[T_{PS}]$ או $E[T_{NE}]$? נמקו במדויק, תוך ציטוט העובדות המתמטיות בהן אתם משתמשים.

תשובה: $E(T_{NE}) > E(T_{PS})$

נימוק: מהתמונה אנו רואים $\bar{F}_{NE} > \bar{F}_{PS}$ ולמשתנה מקרי חיובי Z מתקיים $E(Z) = \int_0^\infty \bar{F}_Z(t) dt$

חלק 3. סבלנות (5 נקודות).

1.5 (5 נקודות) בכיתה הוסבר שאמד' לממוצע הסבלנות של לקוח, המבוסס על נתונים קטומים (מצונזרים) ניתן על ידי

$$\text{ממוצע הסבלנות} = \frac{\text{סך הכל זמן ההמתנה (של משורתיים או נוטשיים)}}{\text{מספר הנוטשיים}}$$

הסבירו כיצד ניתן להסיק מהאמד הנ"ל את ההצגה:

	ממוצע	+	ממוצע	=	ממוצע הסבלנות
אינדקס	זמן ההמתנה	x	זמן ההמתנה		
הסבלנות	של נוטשיים		של משורתיים		

עבור הסברכם, הגדירו במדויק את אינדקס הסבלנות.

תשובה:

$$\frac{\text{מספר משורתים}}{\text{מספר הנוטשים}} = \text{הגדרת אינדקס הסבלנות:}$$

הסבר:

$$= \frac{\text{סך הכל זמן ההמתנה (של משורתים או נוטשים)}}{\text{מספר הנוטשים}}$$

$$= \frac{\text{זמן ההמתנה של נוטשים}}{\text{מספר הנוטשים}} + \frac{\text{מספר משורתים}}{\text{מספר משורתים}} \times \frac{\text{זמן ההמתנה של משורתים}}{\text{מספר הנוטשים}}$$

$$= \frac{\text{זמן ההמתנה של נוטשים}}{\text{מספר הנוטשים}} + \frac{\text{מספר משורתים}}{\text{מספר משורתים}} \times \frac{\text{זמן ההמתנה של משורתים}}{\text{מספר הנוטשים}}$$

$$\text{סבלנות ממוצעת} = \text{זמן ההמתנה של נוטשים} + \text{זמן ההמתנה של משורתים} \times \frac{\text{ממוצע מספר משורתים}}{\text{ממוצע מספר הנוטשים}}$$

שאלה 2 – תרגילי בית.

חלק 1 – "Pizza Anonymous" (10 נקודות).

בפיצרייה מסוימת ניתן לתאר את תהליך הגעת ההזמנות לפיצה באמצעות תהליך פואסון עם קצב של 12 הזמנות בשעה.

בפיצרייה מועסק מכין פיצות בודד, אשר עלות העסקתו הינה \$10 לשעה, המכין בממוצע פיצה אחת כל 3 דקות. זמן הכנת הפיצות מפולג מעריכית. לאחר הכנת הפיצה, היא מוכנסת לתנור בעל קיבולת של 10 פיצות, וזמן אפייתה הוא 12 דקות בדיוק. בפיצרייה צוות של 6 שליחים וזמן משלוח (הלוך ושוב) מפולג מעריכית עם ממוצע של 10 דקות. זמן נסיעה לכיוון אחד שווה למחצית זמן נסיעה הלוך ושוב, ונניח לפשטות כי גם זמן זה מפולג מעריכית. עלות העסקת שליח הינה \$8 לשעה. זמני הכנת הפיצות וזמני המשלוח בלתי תלויים.

הפיצרייה פועלת במדיניות לפיה היא מתחייבת לספק את הפיצה המוזמנת תוך 40 דקות מקבלת ההזמנה (תחילת הכנת הפיצה), ובמקרה של סטייה ממדיניות זו הלקוח מפוצה ע"י הוזלה של \$2 במחיר הפיצה.

2.1.1 (2 נקודות) שרטטו דיאגרמת זרימה לתהליך המתואר לעיל (Process Flow Diagram), החל מהגעת הזמנה ועד לשליחתה ללקוח. תארו כל תחנה באמצעות המודל המתמטי המתאים לה, תוך הסבר מתאים.

תשובה:
תרשים זרימה:

הזמנות פיצה
12/hour

תחנה 1:

הכנת פיצה
M/M/1

תחנה 2:

אפיית פיצה
M/D/10

תחנה 3:

משלוח פיצה
M/M/6

$$\rho = \frac{\lambda}{\mu} = \frac{12}{20} = 60\%$$
$$\rho = \frac{\lambda}{n\mu} = \frac{12}{5 \cdot 10} = 24\%$$
$$\rho = \frac{\lambda}{n\mu} = \frac{12}{6 \cdot 6} = 33.3\%$$

אנו מציגים את תהליך השירות כרשת תורים פשוטה בעלת 3 תחנות.

תחנה 1: M|M|1

הסבר: נתון כי תהליך הגעת ההזמנות הוא תהליך פואסון, זמן הכנת הפיצות מפולג מעריכית ובתחנה מכין פיצה בודד.

תחנה 2: M|D|10

הסבר: תהליך העזיבה מתחנה 1 (שהוא גם תהליך הכניסה לתחנה 2) הוא פואסוני. זמן האפייה הוא דטרמיניסטי והקיבולת הסטטית של התנור היא 10 פיצות.

תחנה 3: M|M|6

הסבר: זמן השהיה בתחנה 2 הוא דטרמיניסטי ולכן תהליך העזיבה מתחנה זו הוא פואסוני (כתהליך הכניסה לתחנה). ולכן, תהליך הכניסה לתחנה 3 הוא פואסוני. נתון כי זמן השירות (נסיעה הלוך ושוב) מפולג מעריכית וכי בתחנה עובדים 6 שליחים.

2.1.2 (2 נקודות) איזו פעולה מהווה את צוואר הבקבוק בתהליך? הסבירו בדיוק ככל שתוכלו.

תשובה :

הקיבולת של תחנה 1 היא 20 פיצות לשעה, של תחנה 2 היא 50 פיצות לשעה ושל תחנה 3 היא 36 פיצות לשעה. קצב ההגעה לכל אחת מהתחנות זהה ולכן תחנה 1, שלה הקיבולת הנמוכה ביותר, היא צוואר הבקבוק.

2.1.3 (6 נקודות) חשבו את ההסתברות כי לקוח יפוצה עקב איחור בהגעת פיצה.

לצורך החישוב, היעזרו בחומר המופיע בעמודים 13-14, הכולל: נספח שהופיע בתרגיל הבית והעמודים הרלוונטיים שנלקחו מההרצאות. הסבירו כל שלב בחישוביכם. **אין צורך בחישובים מספריים - מספיק שימוש בנוסחאות הרלוונטיות והצבת הפרמטרים בהן.**

תשובה :

זמן שהיה בתחנה 2 הוא 12 דקות.

נסמן :

$X =$ זמן שהיה בתחנה 1 (תור $M|M|1$)

$$X \sim \exp\left(\text{mean} = \frac{1}{\mu(1-\rho_1)} = \frac{1}{8} \text{ hour}\right) = \exp(8)$$

$Y =$ זמן המתנה בתור לתחנה 3 (תור $M|M|6$)

$$Y = \begin{cases} 0 & \text{w.p. } 1 - E_{2,N} \\ \exp\left(\text{mean} = \frac{1}{n\mu} \cdot \frac{1}{1-\rho_3} \text{ hour}\right) = \exp(24) & \text{w.p. } E_{2,N} \end{cases}$$

$Z =$ זמן משלוח (זמן שהיה בתחנה 3)

$$Z \sim \exp\left(\text{mean} = 5 \text{ minutes} = \frac{1}{12} \text{ hour}\right) = \exp(12)$$

(הנצילות בתחנה 2 נמוכה ולכן נניח כי אין תור לתחנה זו, כלומר פיצה מוכנסת לתנור מיד בתום הכנתה בתחנה 1).

X, Y, Z ב"ת.

$$\begin{aligned} P(X + 12 + Y + Z > 40) &= P(X + Y + Z > 28) = \\ &= P(\underbrace{X + Z}_{\text{sum of two exponential random variables}} > 28) * P(Y = 0) + P(\underbrace{X + Y + Z}_{\substack{\text{sum of three exponential random variables,} \\ \text{first scenario}}} > 28 | Y > 0) * P(Y > 0) \\ &= 0.064 * 0.982 + 0.093 * 0.018 = 0.065 \end{aligned}$$

Sum of two exponential random variables.

Assume that $X \sim \exp(\lambda_1)$, $Y \sim \exp(\lambda_2)$, X and Y are independent and $\lambda_1 \neq \lambda_2$. Then straightforward calculations provide us with the following cumulative distribution function of $X + Y$:

$$F(x) = 1 - \frac{\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 x} - \frac{\lambda_1}{\lambda_1 - \lambda_2} e^{-\lambda_2 x}.$$

Sum of three exponential random variables.

Assume that $X \sim \exp(\lambda_1)$, $Y \sim \exp(\lambda_2)$, $Z \sim \exp(\lambda_3)$, X , Y and Z are independent and $\lambda_1 \neq \lambda_2 \neq \lambda_3$. Then the cumulative distribution function of $X + Y + Z$ is given by:

$$F(x) = 1 - \frac{\lambda_2 \lambda_3}{(\lambda_3 - \lambda_1)(\lambda_2 - \lambda_1)} e^{-\lambda_1 x} - \frac{\lambda_1 \lambda_3}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)} e^{-\lambda_2 x} - \frac{\lambda_1 \lambda_2}{(\lambda_1 - \lambda_3)(\lambda_2 - \lambda_3)} e^{-\lambda_3 x}.$$

M/M/1: Sojourn Time

Sojourn time is **exponentially** distributed:

$$W \stackrel{d}{=} \exp \left(\text{mean} = \frac{1}{\mu(1 - \rho)} = \frac{1}{\mu} \left[1 + \frac{\rho}{1 - \rho} \right] \right).$$

Proof: By PASTA and the memoryless-property of Exponentials,

$$W \stackrel{d}{=} \sum_{i=1}^N X_i, \quad X_i \sim \exp(\mu) \text{ i.i.d.},$$

$N \triangleq L + 1 \stackrel{d}{=} \text{Geom}(1 - \rho)$ (from 1);

N and $\{X_i\}$ are all independent.

Conclude by recalling: Geometric sum of iid Exponentials is Exponentially distributed. (The parameter is calculated via Wald.)

Aside: The latter property is essentially Poisson-Splitting. A self-contained proof can be give via *Moment generating functions*:

$$\begin{aligned} \phi_W(t) &\triangleq \mathbb{E}[\exp\{tW\}] = \mathbb{E} \left[\exp \left\{ t \cdot \sum_{i=1}^N X_i \right\} \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\exp \left\{ t \cdot \sum_{i=1}^N X_i \right\} \middle| N \right] \right] \\ &\quad (X_i \text{ independent with } \mathbb{E}[\exp\{tX_i\}] = \frac{\mu}{\mu - t}) \\ &= \mathbb{E} \left[\left(\frac{\mu}{\mu - t} \right)^N \right] = \sum_{k=1}^{\infty} (1 - \rho) \rho^{k-1} \left(\frac{\mu}{\mu - t} \right)^k \\ &= \frac{\mu(1 - \rho)}{\mu - t} \cdot \sum_{k=0}^{\infty} \left(\frac{\mu \rho}{\mu - t} \right)^k = \frac{\mu(1 - \rho)}{\mu(1 - \rho) - t} \\ &= \phi_{\exp(\mu(1 - \rho))}(t). \end{aligned}$$

M/M/n Queue: Properties

Erlang-C Formula (1917) for the Delay probability:

$$P\{W_q > 0\} \triangleq E_{2,n} = \sum_{i \geq n} \pi_i = \frac{R^n}{n!} \frac{1}{1 - \rho} \cdot \pi_0.$$

Erlang-C computation: via recursion, see Erlang-B below.

Erlang-C approximations: important later in course.

Number-in-queue:

$$P\{L_q = i\} = E_{2,n} \cdot (1 - \rho)\rho^i, \quad i > 0,$$

or

$$L_q = \begin{cases} 0 & \text{wp } 1 - E_{2,n} \\ \text{Geom}_{\geq 0}(1 - \rho) & \text{wp } E_{2,n} \end{cases}$$

Waiting time distribution:

$$\frac{W_q}{1/\mu} = \begin{cases} 0 & \text{wp } 1 - E_{2,n} \\ \text{Exp}\left(\text{mean} = \frac{1}{n} \cdot \frac{1}{1-\rho}\right) & \text{wp } E_{2,n} \end{cases}$$

Compare with M/M/1!

Departure process: Poisson(λ) in steady-state.

Proof via reversibility, as with M/M/1.

חלק 2. המאמר של David Maister (The psychology of waiting lines) 10 נקודות.

2.2. (10 נקודות) מחבר המאמר דן בהבדל בין זמן ההמתנה בפועל להמתנה הנתפסת. על פי חוק השירות הראשון במאמר – שביעות הרצון של לקוחות תלויה בפער בין רמת השירות הנתפסת לבין הציפייה לרמת שירות עמה הגיע הלקוח. תן ארבע דרכים להשפיע על שביעות הרצון של הלקוחות וציין על איזה מהפרמטרים הנ"ל היא משפיעה (תפיסה/ציפייה). יש לציין לפחות דרך אחת להשפיע על כל אחד משני הפרמטרים.

תשובה:

1. הודעה על זמן המתנה משוער – משפיע על הציפייה.
2. שירות תוך כדי המתנה (Tracking, IVR) – משפיע על תפיסה.
3. תעסוקה תוך כדי המתנה (עיתונים בחדר המתנה של רופא) – משפיע על תפיסת הזמן שעובר.
4. הסבר מדוע מחכים לדוגמא: הודעה בחדר מיון שהרופאים עסוקים בחדר הטראומה – משפיע על ציפייה.

שאלה 3 – יישום (45 נקודות)

להלן קטע קצר מתוך כתבה שדנה במסקנות ועדת גרמן. בסוף השאלה מופיעות נוסחאות שיכולות לעזור בפתרון השאלה.

גרמן: מיליארד שקל לקיצור תורים

לראשונה בישראל: מערכת הבריאות תקבל תוספת של למעלה ממיליארד שקלים במטרה לקצר תורים בבתי"ח ולחזק את קופות החולים. לא יהיה שר"פ בבתי"ח הציבוריים

הציבור ישלם פחות, יחכה פחות, ויקבל יותר אחרי שנה של עבודה מאומצת, שרת הבריאות, יעל גרמן ומנכ"ל משרד הבריאות, פרופ' ארנון אפק, הציגו היום (ד') את עיקרי מסקנות ועדת גרמן לחיזוק מערכת הבריאות הציבורית. כשהחשובה שבהן היא תוספת של מיליארד שקלים למערכת הבריאות לטובת קיצור תורים במערכת הציבורית. תוספת זו מורכבת מ-300 מיליון שקלים שיועברו בשנה זו למערכת בתי החולים לקיצור משכי ההמתנה לניתוחים ושיפור השירות בחדרי המיון: יורחבו חדרי ניתוח, יוכשרו רופאים ואחיות ועוד. ו-700 מיליון שקלים מיועדים לקופות החולים במהלך 2015-2016 לטובת קיצור תורים.

בשאלה זו נתמקד בתורים לחדרי הניתוח בבתי החולים הציבוריים. בבית חולים "בריאות הציבור" ישנן כיום, ערב פרסום מסקנות גרמן, 30 חדרי ניתוח, בהם מבוצעים ניתוחים מתוכננים מסוגים שונים. חדרי הניתוח פועלים בין השעות 8:00 בבוקר ל-17:00 אחר הצהריים. קצב המופע לקביעת תור לניתוח הינו 70 מטופלים ליום. אחוז הביטולים בגין סיבות רפואיות או בגין מעבר לתורים ברפואה פרטית הינו 20% בממוצע.

חלק 1. הבנת מצב קיים (18 נקודות)

3.1. (2 נקודות) כיום, בבית חולים "בריאות הציבור", זמן ההמתנה הממוצע לניתוח מתוכנן הינו שנה. כמה לקוחות בממוצע נמצאים ברשימת ההמתנה לניתוחים מתוכננים? (ניתן להניח כי חדרי הניתוח עובדים 5 ימים בשבוע, 50 שבועות בשנה – כלומר 250 ימי עבודה).

תשובה:

על פי חוק ליטל:

$$EL_q = \lambda_{eff} \times EW_q = (70 \times 250) \times 0.8 \times 1 = 14000$$

ישנם 14000 פציינטים המחכים בתור לניתוח מתוכנן.

3.2. (3 נקודות) הצע מודל תורים אשר יתאר את מערך חדרי הניתוח של בית החולים במצב יציב. פרט את ההנחות שהנחת ותן פרשנות מילולית לפרמטרים של המודל. (ניתן להניח הנחות על התפלגויות אבני הבנין של המודל). אין צורך לנקוב בערך הפרמטרים.

תשובה:

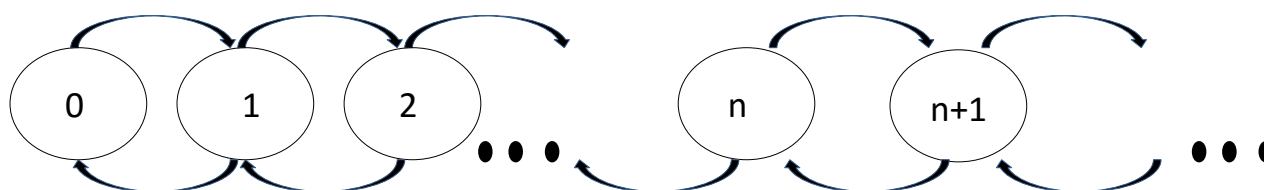
מודל מתאים לתיאור המערכת הנ"ל הוא מודל התורים Erlang-A, כאשר מספר חדרי הניתוח אנלוגי למספר השרתים.

הנחות המודל: תהליך המופע הוא תהליך פואסון הומוגני בזמן, משך השירות מתפלג מעריכית וזמני השירות ב"ת, אי תלות בין משך השירות לזמנים הבין-מופעיים, סבלנות הלקוחות מתפלגת מעריכית (אי תלות בין הסבלנות לשאר הפרמטרים).

פרשנות:

משך ניתוח = משך שירות, הגעות הן למעשה רישום לניתוח מתוכנן, נטישה = מעבר לרפואה פרטית או שינוי במצב הבריאותי של המטופל.

3.3. (2 נקודות) השלם את דיאגרמת קצבי המעבר המתאימה למודל שבחרת בסעיף 3.2.:



תשובה:

$$q(x, x+1) = \lambda$$

$$q(x+1, x) = \mu(N \wedge x) + \theta[x - N]^+$$

כל מעבר אחר לא אפשרי.

3.4. (4 נקודות) באיזה תחום תפעולי פועל מערך חדרי הניתוח של בית החולים? הסבר על סמך אילו נתונים הגעת למסקנה זו. תחת תחום זה, מהו פרמטר רמת השירות הנוכחית?

תשובה:

על פי הנתונים נראה כי חדרי הניתוח בבית החולים עובדים בתחום התפעולי "מכוון יעילות" (Efficiency Driven). הנתונים התומכים במסקנה זו הם זמני המתנה ארוכים (בסדרי גודל יותר מאשר זמן שרות – זמן השירות הוא שעות, אך זמני ההמתנה הם חודשים), והסתברות נטישה גבוהה (כ- 20%).

בתחום תפעולי זה פרמטר רמת השירות (γ), הוא בדיוק ההסתברות לנטישה ועל כן ערכו 0.2.

3.5. (3 נקודות) על סמך הנתונים שהופיעו בשאלה, תן אמד לסבלנות המטופלים להמתנה לניתוח מתוכנן.

תשובה:

במודל Erlang-A:

$$EW_q \theta = P_{ab}$$

בקירוב (עבור ה-ED Regime):

$$EW_q \theta = P_{ab}$$

$$\frac{1}{\theta} = \frac{EW_q}{\gamma} = \frac{250}{0.2} = 1250 \text{ days}$$

3.6. (4 נקודות) תן קירוב לעומס המוצע (Offered-Load) במערכת הנוכחית. מהו בקרוב הממוצע של משך ניתוח יחיד?

תשובה:

העומס המוצע הוא :

$$R \approx \frac{N}{(1-\gamma)} = \frac{30}{1-0.2} = 37.5 \text{ Erlang}$$

ממוצע משך ניתוח הוא :

$$m \approx \frac{R}{\lambda} = \frac{37.5}{\frac{70}{9}} \approx 4.82 \text{ hr}$$

חלק 2 – ניתוח מצב מוצע. (12 נקודות)

ועדת גרמן שמה לה כמטרה להביא לקיצור תורים לניתוחים. אחת השיטות היא הוספת חדרי ניתוח וצוותים רפואיים שיתפעלו אותם.

3.7. (4 נקודות) מהי כמות שעות הפעילות ליום שיש להקצות על מנת להוריד את ההמתנה לחצי שנה? הסבר.

תשובה:

כדי להוריד את זמן ההמתנה מ- 250 יום ל- 125 יום, יש להוריד את אחוז הנוטשים מ- 20% ל-

$$P_{ab} = EW_q \theta = 125 \times \frac{1}{1250} = 0.1$$

כיוון שאנו עדיין בתחום התפעולי "מכוון יעילות", הפרמטר המתאים בנוסחת האיוש הוא $\gamma = 0.1$. כלומר, יש להקצות :

$$N = R(1-\gamma) = 37.5 \times 0.9 = 33.75$$

בערך 34 חדרים, שהם כ- 306 שעות פעילות.

3.8. (1 נקודות) מהו ממוצע אורך התור במערכת החדשה?

תשובה:

על פי חוק ליטל:

$$EL_q = \lambda_{eff} \times EW_q = (70 \times 250) \times 0.9 \times 0.5 = 7875$$

ישנם 14000 פציינטים הממכים בתור לניתוח מתוכנן.

3.9. (5 נקודות) בהינתן שכמות חדרי הניתוח תשתנה היום, המערכת לא תתייצב על רמת השירות החדשה באופן מיידי. כיצד ניתן לחשב את הזמן שיעבור עד להתיצבות המערכת על רמת השירות הרצויה? כתבו את הנוסחאות המתאימות על מנת לבצע חישוב זה, אין צורך לבצע את חישוב הזמן בפועל.

תשובה:

ניתן להגיע לזמן עד להתיצבות המערכת בעזרת מודל נוזלים.

מודל הנוזלים של מודל Erlang-A היא:

$$\dot{Q}(t) = \lambda t - (N \wedge Q(t))\mu - \theta(Q(t) - N)^+$$

יש להציב את תנאי ההתחלה הבאים:

מספר האנשים במערכת הוא מספר האנשים בתור (מסעיף 3.1) ועוד מספר האנשים בשרות.

$$EL_s = \lambda_{eff} \times EW_s = (70/9) \times 0.8 \times 4.82 = 30$$

(בתחום תפעול זה)

$$Q(0) = 14000 + 30 = 14030$$

הזמן עד התיצבות הוא ה- t המינימלי שמקיים $Q(t) = 7875 + 34 = 7909$ (שזה מספר הממתינים במערכת החדשה (מסעיף 3.8 וכמות האנשים שנמצאים בשירות).

3.10 (2 נקודות) נניח שמשרד הבריאות ביצע חישוב לזמן הרצוי, והתקבל שדרושים 4 שנים להתיצבות המערכת. הועדה מעוניינת שהזמן יקוצר לשנתיים. כיצד היית ממליץ לועדה להגיע למטרה זו?

תשובה:

ניתן להפעיל בשנתיים הקרובות את חדרי הניתוח באופן מתוגבר (בשעות הערב).

חלק 3. חיזוי זמני המתנה. (15 נקודות)

כחלק מהחלטות ועדת גרמן, כל מטופל אשר ממתין לניתוח מתוכנן יוכל לדעת בזמן אמת את זמן ההמתנה המשוער שלו. בשאלה זו תצטרכו לאמוד את זמן ההמתנה הוירטואלי (Virtual waiting time) של לקוח, ולהשתמש בזמן זה כתחזית לזמן ההמתנה.

3.11 (4 נקודות) מהו זמן ההמתנה הוירטואלי ומה ההבדל בינו לבין זמן ההמתנה ממוצע? האם הם זהים במודל שהגדרת בסעיף 3.2?

תשובה:

זמן המתנה וירטואלי הוא זמן המתנה שלקוח שמגיע עכשיו למערכת יצטרך להמתין (בהנחה שלא ינטוש את המערכת). זמן המתנה ממוצע הוא זמן המתנה של כלל הלקוחות במערכת. מכיוון שבמודל Erlang-A יש נטיות, שני הגדלים אינם זהים.

מכאן והלאה נחשב את תוחלת זמן ההמתנה הוירטואלי של מטופל, **בהינתן** מספר האנשים בתור ברגע הגעתו. לצורך שאלה זו, נניח כי n שרתים (חדרי ניתוח) נותנים שירות בקצב אינדיווידואלי μ , ומשכי

השירות מתפלגים אקספוננציאלית עם תוחלת $\frac{1}{\mu}$.

3.12. (2 נקודות) מהי תוחלת זמן ההמתנה הוירטואלי של מטופל אשר בזמן הגעתו אין לקוחות בתור, אך כל חדרי הניתוח תפוסים? (תחת הנחת FCFS).

תשובה:

$$\frac{1}{n\mu}$$

3.13. (3 נקודות) מהי תוחלת זמן ההמתנה הוירטואלי של מטופל בהינתן שבזמן הגעתו יש לפניו 2 מטופלים בתור? (תחת הנחת FCFS).

תשובה:

$$\frac{1}{n\mu} + \frac{1}{n\mu + \alpha} + \frac{1}{n\mu + 2\alpha}$$

שימו לב שהפתרון מניח נטישות אקספוננציאליות עם בקצב α , אבל במבחן אפשר היה להניח שאין נטישות, כלומר $\alpha = 0$.

3.14. (4 נקודות) בהינתן שתוחלת זמן ההמתנה הוירטואלי תהווה אמד למשך ההמתנה של לקוח שנמצא במקום ה-k בתור, תן נוסחה כללית לאמד זה. הסבר.

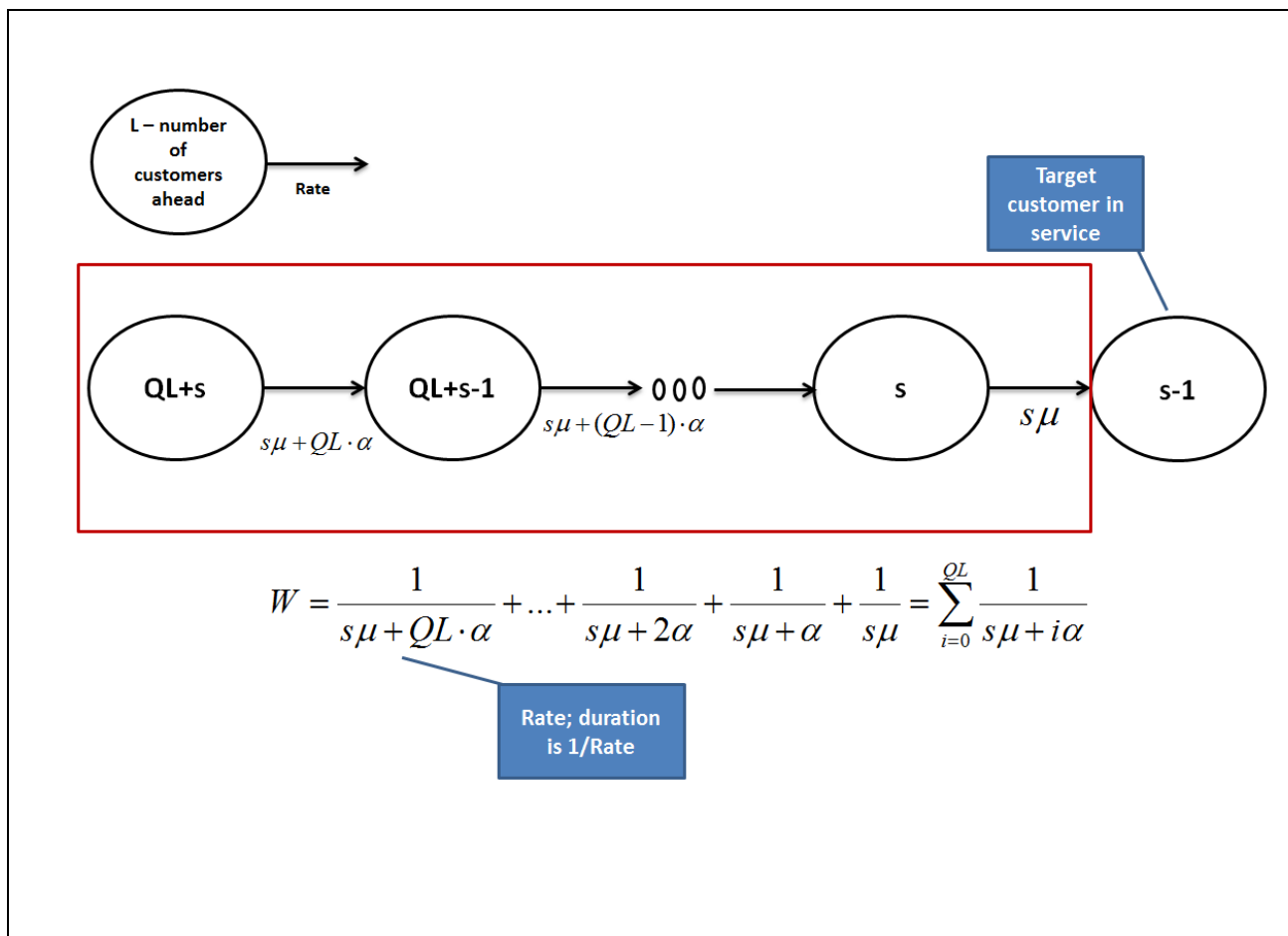
תשובה:

$$\sum_{i=1}^k \frac{1}{n\mu + i\alpha}$$

הסבר לחישובים בסעיפים 3.11-3.13 :

נתבונן בתהליך היציאה מהמערכת של המטופלים, בהינתן שאורך התור הוא k :

(ראה עמוד הבא)



3.15 (2 נקודות) הצע שיטה אלטרנטיבית לאמידת זמן ההמתנה של מטופל (אין צורך בביטוי מפורש).
הסבר האם ובאילו תנאים הוא צפוי לעבוד טוב יותר מאמד זמן המתנה וירטואלי?

תשובה:

זמן ההמתנה של האחרון שנכנס לשירות.

אמד זה מניח כי המערכת לא משתנה הרבה בזמן המתנה של לקוח מסוים, וכן שגישות של כל לקוח לא מושפעים מלקוחות אחרים (הגיוני בתור בלתי נראה, טלפוני). אמד זה צפוי לעבוד טוב כאשר המערכת עמוסה (ED), עם הרבה שרתים.

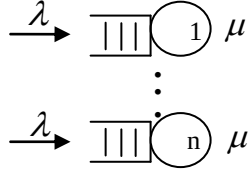
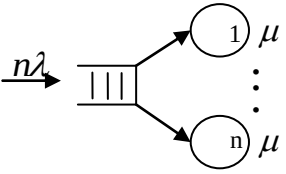
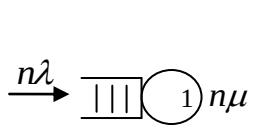
שאלה 4 – תיאוריה ושימושים (15 נקודות)

בשאלה זו נערוך השוואה בין מדדי ביצועים של שלוש המערכות הבאות:

מערכת 1: $M|M|1$ תורי n הפועלים במקביל, כאשר קצב המופע לכל אחד מהתורים הוא λ וקצב השירות μ .

מערכת 2: $M|M|n$ מערכת עם קצב מופע $n\lambda$ וקצב שירות μ (איחוד תורי מערכת 1).

מערכת 3: $M|M|1$ מערכת עם קצב מופע $n\lambda$ וקצב שירות $n\mu$ (שרת מהיר במקום n שרתים איטיים במערכת 2).

	$n \times M M 1$ λ, μ	$M M n$ $n\lambda, \mu$	$M M 1$ $n\lambda, n\mu$
			
Utilization	$\frac{\lambda}{\mu} = \rho$	$\frac{(n\lambda)}{(n)\mu} = \rho$	$\frac{(n\lambda)}{(n\mu)} = \rho$
$P\{W_q > 0\}$	ρ	$E_{2,n}$	ρ
$E[W_q W_q > 0]$	$\frac{1}{\mu} \cdot \frac{1}{1-\rho}$	$\frac{1}{n\mu} \cdot \frac{1}{1-\rho}$	$\frac{1}{n\mu} \cdot \frac{1}{1-\rho}$
$E[W_q]$	$\frac{1}{\mu} \cdot \frac{\rho}{1-\rho}$	$\frac{1}{\mu} \cdot \frac{E_{2,n}}{n(1-\rho)}$	$\frac{1}{\mu} \cdot \frac{\rho}{n(1-\rho)}$
$E[S]$	$\frac{1}{\mu}$	$\frac{1}{\mu}$	$\frac{1}{n\mu}$
$E[W_{sys}]$	$\frac{1}{\mu} \cdot \frac{1}{1-\rho}$	$\frac{1}{\mu} \cdot \left[\frac{E_{2,n}}{n(1-\rho)} + 1 \right]$	$\frac{1}{n\mu} \cdot \frac{1}{1-\rho}$

כאשר

- Utilization - נצילות שרת,
- $P\{W_q > 0\}$ - ההסתברות להמתנה,
- $E[W_q | W_q > 0]$ - תוחלת זמן המתנה בהינתן המתנה,

- $E[W_q]$ - תוחלת זמן המתנה,
- $E[S]$ - תוחלת משך שירות,
- $E[W_{sys}]$ - תוחלת זמן שהיה במערכת.

$$\text{כמו כן, ידוע כי } n \leq \frac{1-E_{2,n}}{1-\rho} \leq 1.$$

4.1 (5 נקודות) קבעו מה היחס המתאים בין ההסתברות להמתין, $P\{W_q > 0\}$, במערכת 1 ובמערכת 2. הוכיחו תשובתכם מתמטית והסבירו את התוצאה באופן אינטואיטיבי.

תשובה:

בחרו ביחס המתאים: $\geq$ **א.** $=$ **ב.** $\leq$ **ג.** $\geq$ **ד.** לא ניתן לקבוע בוודאות

לנוחיותכם, מוצג החלק מהטבלה עליו אתם נשאלים,

	$n \times M M 1$ λ, μ	$M M n$ $n\lambda, \mu$
$P\{W_q > 0\}$	ρ	$E_{2,n}$

הוכחה:

מהעובדה שנרשמה לעיל:

$$1 - E_{2,n} \geq 1 - \rho \Rightarrow \rho \geq E_{2,n}$$

הסבר אינטואיטיבי:

במערכת 2, כאשר מגיע מופע למערכת כאשר ישנו שרת פנוי, המופע ייכנס לשירות באופן מיידי, ללא המתנה. לעומת זאת, במערכת 1 יתכן מצב בו מופע המגיע למערכת, נכנס לתור בו השרת עסוק ולכן נאלץ להמתין, בעוד שישנו שרת פנוי בתור אחר.

4.2. (5 נקודות) קבעו מה היחס המתאים בין תוחלת זמן ההמתנה, $E[W_q]$, במערכת 2 ובמערכת 3. הוכיחו תשובתכם מתמטית.

להלן מוצג החלק מהטבלה עליו אתם נשאלים,

תשובה:

בחרו ביחס המתאים: א. $\geq$ ב. $=$ ג. $\leq$ ד. לא ניתן לקבוע בוודאות

לנוחיותכם, מוצג החלק מהטבלה עליו אתם נשאלים,

	$M M n$ $n\lambda, \mu$	$M M 1$ $n\lambda, n\mu$
$E[W_q]$	$\frac{1}{\mu} \cdot \frac{E_{2,n}}{n(1-\rho)}$	$\frac{1}{\mu} \cdot \frac{\rho}{n(1-\rho)}$

הוכחה:

משאלה 4.2.1, $E_{2,n} \leq \rho$

$$\frac{1}{\mu} \cdot \frac{E_{2,n}}{n(1-\rho)} \leq \frac{1}{\mu} \cdot \frac{\rho}{n(1-\rho)}$$

4.3 (5 נקודות) השוו בין היחס שמצאתם בשאלה 4.2 לבין היחס הנתון בין תוחלות זמן השהיה, $E[W_{sys}]$, במערכת 2 ובמערכת 3. נסחו את התופעה שגיליתם והסבירו מהן הסיבות לה.

תשובה:

כאשר עוברים ממערכת בה יש שרת מהיר למערכת בה מספר שרתים איטיים, ממוצע זמן ההמתנה מתקצר וממוצע משך השהיה מתארך. אמנם זמן ההמתנה במערכת 3 גבוה יותר מזמן ההמתנה במערכת 2, אך ביחד עם זמן השירות הקצר יותר (במערכת 3), נקבל כי זמן השהיה הכולל במערכת 3 קצר יותר. אם נסתכל על דיאגרמות קצבי המעבר של שתי המערכות, נראה כי החל ממצב n תהליך העזיבה זהה בשתי המערכות, ועבור מצבים $1, 2, \dots, n-1$ קצבי העזיבה במערכת 3 גדולים יותר.