



מרצה : פרופסור אבישי מנדלבאום
מתרגלת : שמרית ממך

תאריך הבחינה : 01.10.2007

מס' סטודנט _____

שם _____

הנדסת מערכות שירות – 096324

מועד א' - סמסטר אביב תשס"ז 2007

פתרון

- המבחן נמשך שלוש שעות. באחריותך להחזירו במועד.
- המבחן בחומר סגור, וללא מחשבון.
- המבחן כולל 29 עמודים, כולל עמוד זה.
- אלא אם נאמר אחרת, ניתן להשאיר תשובות מספריות ללא חישובן הסופי.
- 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.

ניקוד:	שאלה 1	_____ מתוך 10 נקודות	הרצאות/תרגולים
	שאלה 2	_____ מתוך 10 נקודות	תרגילי בית
	שאלה 3	_____ מתוך 15 נקודות	יישום
	שאלה 4	_____ מתוך 10 נקודות	תיאוריה
	שאלה 5	_____	שאלת בונוס: פתרו שאלה זו רק אחרי שסיימתם את כל השאלות הנ"ל.

סה"כ _____ מתוך 45 נקודות

בהצלחה!

שאלה 1 הרצאות/תרגולים (10 נקודות)

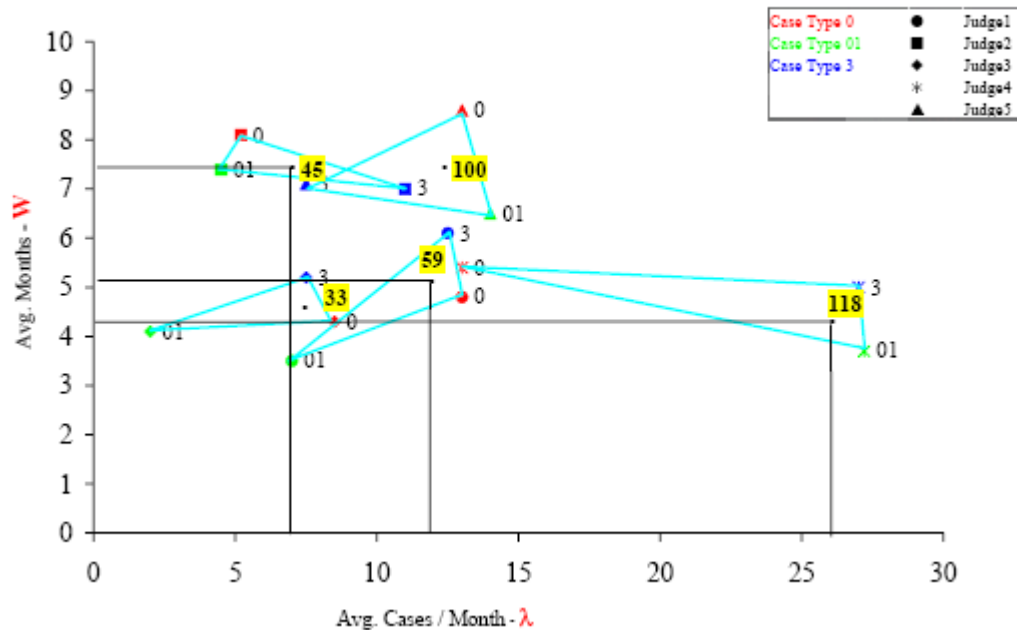
חלק 1.

בכל אחת משלוש השאלות הבאות, בחרו את התשובה הנכונה (אחת מארבע האפשרויות).
הסבירו תשובתכם בקצרה במקום המיועד, אם תדרשו לכך.

1.1

גרף 1

Judges: The Best/Worst (Operational) Performer



התרשים הנ"ל שהוצג בהרצאה מתאר את רמת הביצועים התפעולים של כמה שופטים בבית הדין לעבודה. שטחי 3 המלבנים שאנו רואים בציור מתארים:

1. מספר תיקים ממוצע ששופט מסיים במהלך חודש.

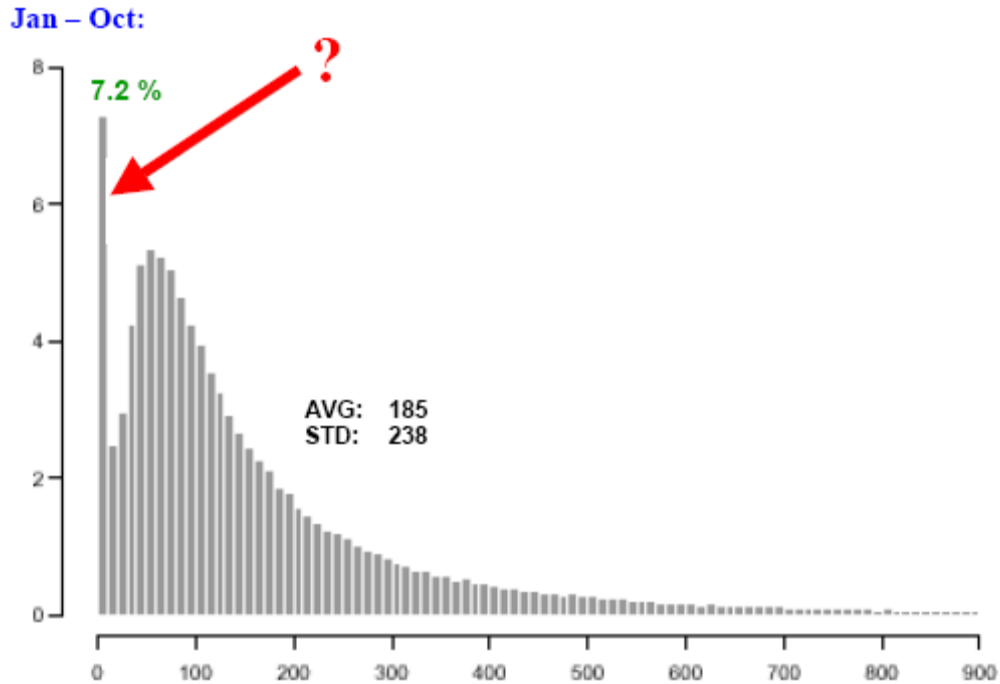
2. מספר תיקים ממוצע שנמצאים בטיפול אצל שופט בעת ובעונה אחת.

3. זמן שהיה ממוצע של תיק במערכת.

4. מספר תיקים ממוצע ששופט פותח במהלך חודש.

תשובה + הסבר: 2, נובע מנוסחת Little.

גרף 2



גרף 2 לעיל מתאר היסטוגרמה של משכי השירות במוקד שירות טלפוני בארץ. אנו רואים אחוז גדול יחסית של שיחות קצרות מאוד במוקד זה. מה הן סיבות פוטנציאליות של תופעה זו?

1. במוקד זה 7.2% מהלקוחות נוטשים את התור הטלפוני עקב חוסר סבלנות.
2. תמריץ המבוסס על מספר השיחות המטופלות גורם למוקדנים לנתק בתחילת השיחה כדי לשפר את "רמת ביצועיהם".
3. מוקדנים נמצאים תחת לחץ עבודה כבד ומנתקים מיד עם תחילת השיחה (ובטרם נוצר קשר דיבור בין המוקדן ללקוח) כדי לנוח.
4. גם 2 וגם 3 יכולים להיות סיבות פוטנציאליות לתופעה המתוארת.

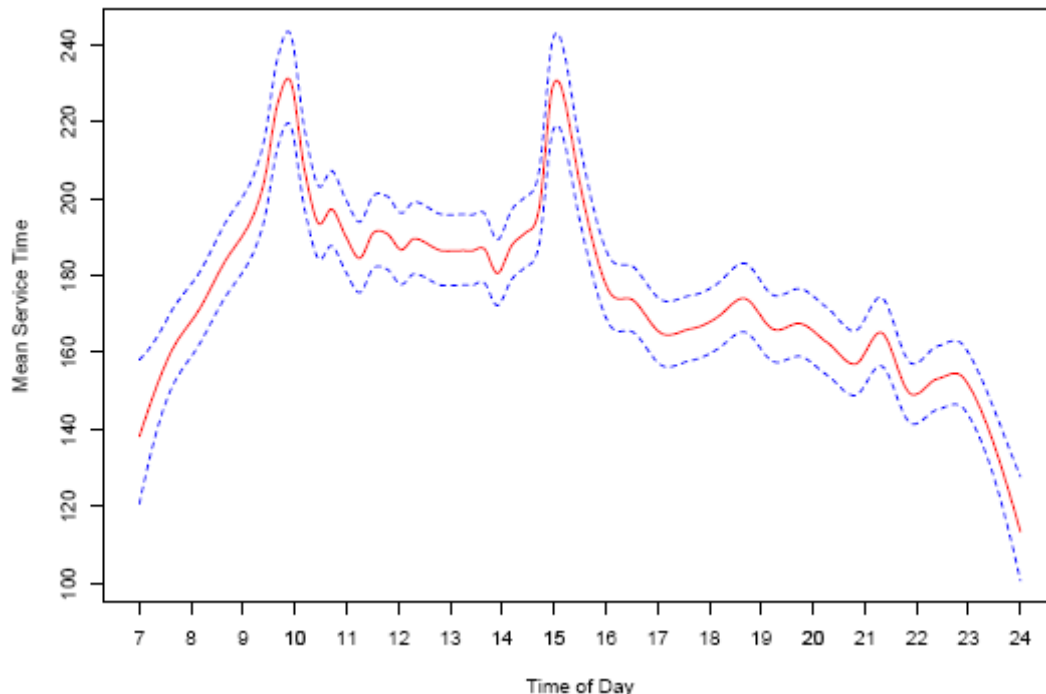
תשובה בלבד : 4.

גרף 3 להלן מתאר משך שירות ממוצע במהלך יום עבודה אופייני במוקד מסוים, גרף 4 מתאר קצב מופע במהלך כזה יום עבודה וגרף 5 משווה בין סבלנות של לקוחות רגילים (Regular) ולקוחות VIP (Priority) באותו מוקד. תנו את ההסבר המלא ביותר לסיבה שבסביבת השעות 10:00 ו-15:00 מתרחשת עליה ברורה גם בקצב המופע וגם במשך השירות הממוצע.

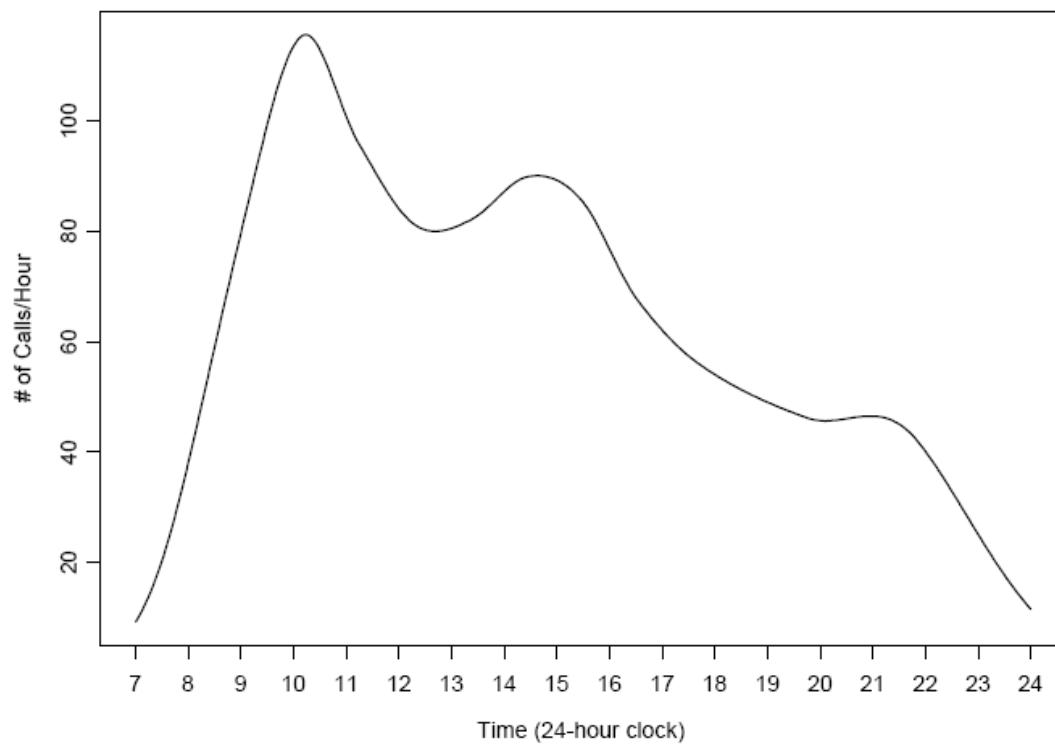
1. מכיוון שאלה שעות עמוסות, הלקוחות שמגיעים לשירות מבזבזים זמן בתלונות ולכן משך השירות מתארך.
2. מגרף 5 עולה כי לקוחות VIP הם גם לקוחות פחות סבלניים. סביר גם להניח שללקוחות VIP משך שירות יותר קצר. לכן, הלקוחות עם זמן השירות הארוך יותר הם אלה ששורדים וזמן השירות הממוצע מתארך.
3. מגרף 5 עולה כי לקוחות VIP הם גם לקוחות סבלניים יותר. סביר גם להניח שללקוחות VIP משך שירות יותר ארוך. לכן, הלקוחות עם זמן השירות הארוך יותר הם אלה ששורדים וזמן השירות הממוצע מתארך.
4. גם 1 וגם 3 יכולים להיות סיבות פוטנציאליות לתופעה המתוארת.

תשובה בלבד : 4.

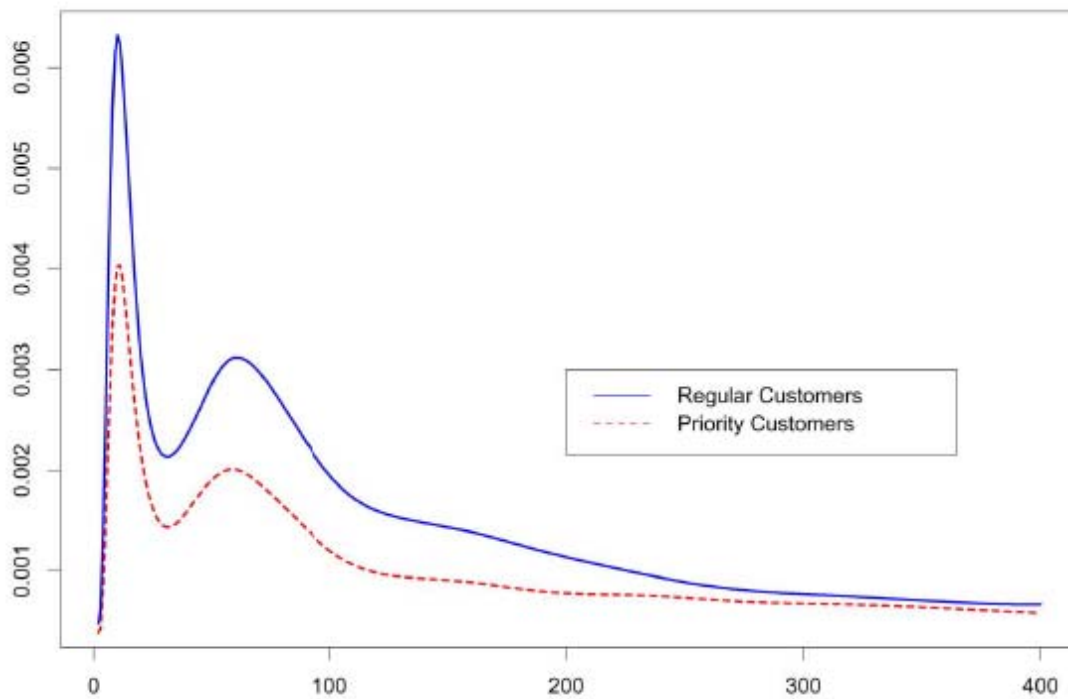
גרף 3



גרף 4



גרף 5



חלק 2. מערכת עם פעילויות עזר

נתבונן במערכת תורים עם פעילויות עזר (*ancillary activities*), אותה למדנו בתרגול. המערכת ניתנת לתיאור ע"י "קרוב משפחה" של מודל $M/M/N$, כאשר N - מספר השרתים - יכול להשתנות כתלות במספר הלקוחות בתור. בתחילת יום העבודה נמצא במערכת שרת יחיד. נגדיר את משתנה הבקרה J ע"י

$$J = \frac{\text{מספר הממתינים}}{\text{מספר השרתים}}$$

(כמובן, J משתנה עם מצב המערכת). נוקטים במדיניות הבאה: למערכת מוסיפים שרת נוסף כל פעם כאשר הערך J עולה על מספר מסוים K שנקבע מראש ($K=0,1,2,\dots$, $K < J$). בנוסף, שרת מוצא מהמערכת (מוחזר לפעילות עזר השונה משירות) אם מייד אחרי סיום השירות עם לקוח כלשהו הוא לא מוצא אף לקוח ממתיין (התור ריק). כלומר, השרתים המתאספים ממשיכים לשרת כל עוד יש ממתינים בתור, וחוזרים לפעילויות העזר בזה אחרי זה אם התור ריק בסיום שירותם. נניח כי במערכת נמצא תמיד לפחות שרת אחד.

נניח שמופע הלקוחות למערכת הוא פואסוני עם קצב λ , ואילו משך השירות – אקספוננציאלי עם פרמטר μ . נסמן

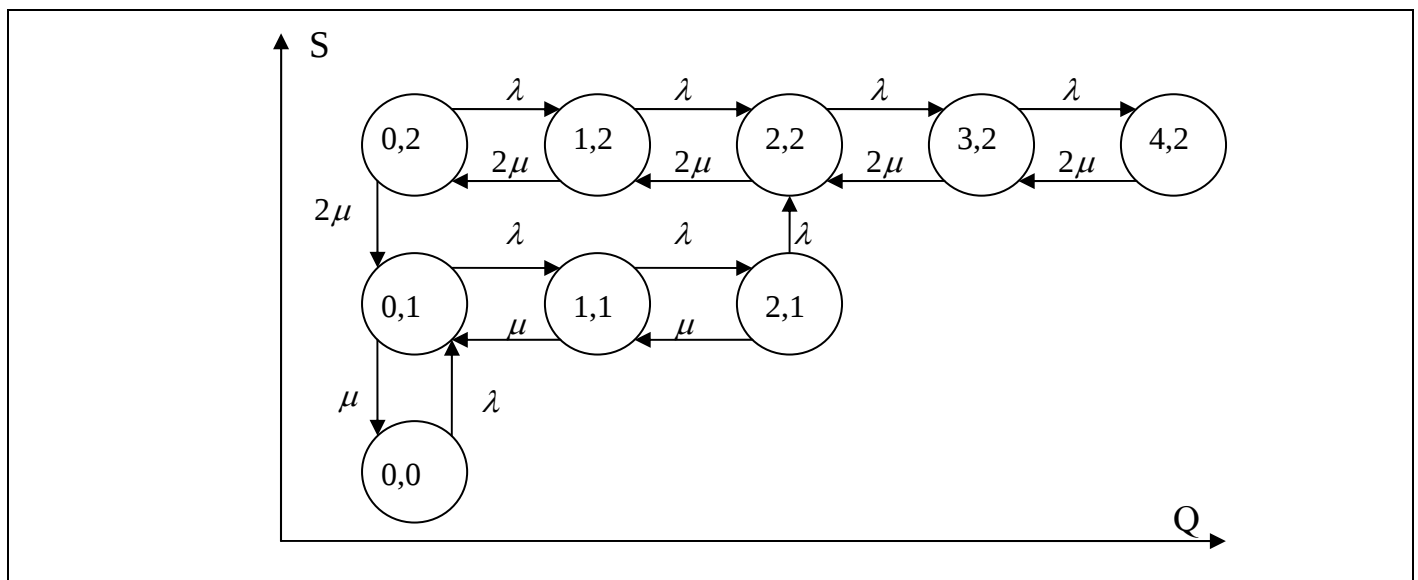
$$R = \frac{\lambda}{\mu} \text{ כרגיל ב-}$$

נגדיר תק"מ דו-ממדי (S, Q) , כאשר S מתאר את מספר הלקוחות הנמצאים **בשירות** ואילו Q מתאר את מספר הלקוחות

בתור. נסמן ב- $\pi_{s,q}$ את ערך ההתפלגות הסטציונרית של המערכת במצב (s, q) .

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

תשובה:



1.2.2 מהי ההסתברות שיהיה צורך בהוספת שרת כאשר מגיע לקוח למערכת? הביעו תשובתכם, כאן ובהמשך, בעזרת ה-

$$\pi_{s,q}$$

תשובה:

$$P(\text{interruption}) = \sum_{i=1}^{\infty} \pi_{i,k \times i}$$

$\pi_{i,k \times i}$ - לפי PASTA, זוהי ההסתברות שלקוח מגיע ומוצא את המערכת כאשר נמצאים בה i לקוחות בתור ו- $k \times i$ לקוחות בשירות. הסכום אינו כולל את המצב בו $i = 0$ מכיוון שהפרעה אינה מתרחשת כאשר מגיע לקוח ומוצא את המערכת במצב $(0, 0)$ (במערכת נמצא תמיד לפחות שרת אחד)

נתבונן בגרף המתאר את התלות של N - ממוצע השרתים במערכת - בערך של הסף K עבור ערכים שונים של R .

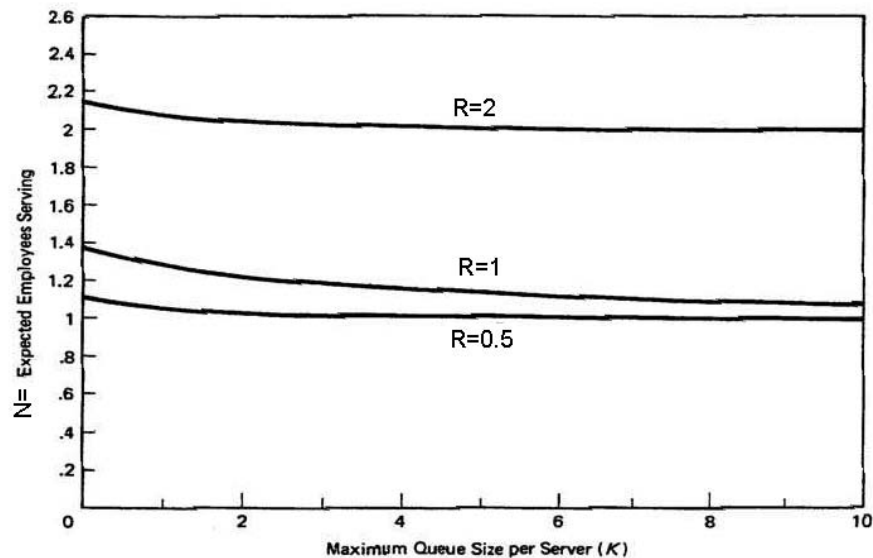


Figure 5.12 The expected number of employees serving declines as the maximum queue size per server increases. For $R > 1$, the limit is R . For $R \leq 1$, the limit is one.

1.2.3 כיצד ניתן להסביר את הערכים של N כאשר $K = 0$?

רמז: במצב סטציונרי, מספר הלקוחות X במערכת $M/M/\infty$ עם פרמטרים λ ו- μ מפולג פואסונית עם פרמטר $R = \frac{\lambda}{\mu}$.

$$P(X = m) = \frac{e^{-R} R^m}{m!}, \quad m = 0, 1, 2, \dots$$

תשובה:

$$\underbrace{\frac{\lambda}{\mu}}_{\text{Average number in service (by Little)}} + \underbrace{\pi_{0,0}}_{\text{idleness of the "must be in" server}} = R + \pi_{0,0}$$

ממוצע מספר השרתים במערכת הוא

כאשר $K = 0$, מוסיפים שרת עבור כל לקוח שמגיע למערכת ולכן זוהי מערכת $M/M/\infty$, ולכן $N = R + \pi_{0,0} = R + e^{-R}$

שאלה 2 תרגילי בית (10 נקודות)

ניתן לתאר את תהליך הגעת ההזמנות לפיצה בפיצרייה מסוימת באמצעות תהליך פואסון עם קצב של 12 הזמנות בשעה. בפיצרייה מועסק מכין פיצות בודד, אשר עלות העסקתו הינה 36\$ 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 36\$ לשעה. זמני הכנת הפיצות וזמני המשלוח ב"ת. 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 8\$ במחיר הפיצה.

2.1 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:



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

תחנה 3: M|M|6

הסבר: זמן השליחה 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.

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.

2.3 חשבו את ההסתברות כי לקוח יפוצה עקב איחור בהגעת פיצה.

לצורך החישוב, היעזרו בחומר המופיע בעמודים 10-12, הכולל: פלט של תוכנת ה-4CallCenter, 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 $M|M|1$)

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

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

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

$Z =$ זמן משלוח (זמן שהיה בתחנה 3)

$$Z \sim \exp(\text{mean} = 5 \text{ minutes}) = \exp(12)$$

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

$$P(X + 12 + Y + Z > 40) = P(X + Y + Z > 28) =$$

$$= P(\underbrace{X+Z}_{\text{sum of two exponential random variables}} > 28) * P(Y=0) + P(\underbrace{X+Y+Z}_{\text{sum of three exponential random variables, first scenario}} > 28 | Y > 0) * P(Y > 0)$$

$$= 0.064 * 0.982 + 0.093 + 0.018 = 0.065$$

Input	60	00:00	Range	10:00	12	
Multi-Value			✓			
	Basic Interval (minutes)	Target Time to Answer	Number of Agents	Average Handling Time	Calls per Interval	%Answer within Target
Results						
1	60.0	00:00.0	3.0	10:00.0	12.0	55.6%
2	60.0	00:00.0	4.0	10:00.0	12.0	82.6%
3	60.0	00:00.0	5.0	10:00.0	12.0	94.0%
4	60.0	00:00.0	6.0	10:00.0	12.0	98.2%
5	60.0	00:00.0	7.0	10:00.0	12.0	99.5%
6	60.0	00:00.0	8.0	10:00.0	12.0	99.9%
7	60.0	00:00.0	9.0	10:00.0	12.0	100%
8	60.0	00:00.0	10.0	10:00.0	12.0	100%

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.

First scenario: 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:

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

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

$$F(x) = 1 + \left[\frac{x \lambda_1 \lambda_2}{\lambda_1 - \lambda_2} - \frac{\lambda_2 (\lambda_2 - 2\lambda_1)}{(\lambda_1 - \lambda_2)^2} \right] e^{-\lambda_1 x} - \frac{\lambda_1^2}{(\lambda_1 - \lambda_2)^2} e^{-\lambda_2 x}.$$

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:

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

Compare with M/M/1!

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

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

2.4 תארו 3 קריטריונים תפעוליים שונים לפיהם ניתן לקבוע את מספר השליחים המינימלי הנדרש.

תשובה :

קריטריון 1 : הזמן ל"התקררות" הפיצה (זמן המתנה לשליח).

קריטריון 2 : עלות התפעול (קנסות, עלות שליחים).

קריטריון 3 : משך המתנת לקוח לפיצה.

שאלה 3 יישום (15 נקודות)

להלן נתון דו"ח ACD המתאר פעילות יומית של מוקד טלפוני בארץ. הדו"ח מתאר פעילות ממוצעת על פני 20 חצאי שעות. לצורך השאלות שלהלן, ניתן לתאר את המוקד כמערכת Erlang-A. לנחיותכם, הדו"ח מצורף גם בעמוד האחרון של הבחינה, כדי שתוכלו לתלוש עמוד זה ולהשתמש בו למענה על השאלות הבאות.

Time	Arrivals	Served	% Abn	E(Wq)	E(Wq Wq>0)	% Wait=0	Service Time (sec)	No. of Agents	Number in Queue	Offered Load	% Wait>30
10:00	387	386	0.26	2.06	7.8	73.58	194.9	42.60	0.51	41.90	0.00
10:30	456	435	4.61	28.15	31.9	11.84	185.6	44.21	6.86	47.02	35.18
11:00	448	441	1.56	7.94	18.1	56.09	173.3	40.93	2.07	43.13	8.53
11:30	500	463	7.40	33.44	40.9	18.27	192.2	45.97	8.78	53.38	54.66
12:00	520	474	8.85	45.92	46.6	1.40	178.8	47.28	12.03	51.66	67.67
12:30	491	471	4.07	27.34	32.7	16.41	177.9	44.66	7.34	48.54	49.01
13:00	498	487	2.21	17.90	29.5	39.41	194.6	50.04	4.90	53.85	24.13
13:30	478	472	1.26		20.3	78.27	170.7	49.81			8.10
14:00	475	471	0.84	3.13	13.9	77.56	178.4	45.47	0.92	47.07	1.97
14:30	462	455	1.52	5.92	12.7	53.45	195.6	48.97	1.57	50.20	1.82
15:00	434	427	1.61	7.00	15.7	55.34	185.4	43.48	1.72	44.70	4.09
15:30	403	389	3.47	14.71	30.7	52.13	207.8	43.03	3.33	46.52	23.18
16:00	400	400	0.00	0.28	5.8	95.26	166.8	39.95	0.16	37.07	0.00
16:30	432	417	3.47	21.03	37.5	43.92	186.5	41.34	5.04	44.77	29.57
17:00	489	473	3.27	14.68	30.0	51.09	199.5	60.23	4.09	54.20	21.30
17:30	478	465	2.72	20.09	34.6	41.94	194.2	49.72	5.29	51.56	32.97
18:00	439	437	0.46	2.73	27.0	89.91	168.9	43.52	0.82	41.18	2.19
18:30	573	539	5.93	33.22	44.3	25.10	209.0	57.14	10.21	66.54	52.08
19:00	418	416	0.48	0.18	5.7	96.87	192.1	56.93	0.14	44.60	0.00
19:30	478	472	1.26	11.28	29.6	61.88	171.7	48.17	3.10	45.59	19.91

3.1 עבור האינטרוול 13:30-14:00, חשבו את העומס המוצע (Offered Load) ואת תוחלת זמן ההמתנה, ואימדו את ממוצע אורך התור. מדוע ערך האמד עבור ממוצע אורך התור יכול להיות רחוק מהערך האמיתי?

תשובה:

$$R = \frac{\lambda * E(S)}{1800} = \frac{478 * 170.7}{1800} = 45.32$$

$$E(W_q) = E(W_q | W_q > 0) P(W_q > 0) = 20.3 * (1 - 0.7827) = 4.41 \text{ sec}$$

$$E(L_q) = \lambda E(W_q) = \frac{478 * 4.41}{1800} = 1.17$$

ייתכן והאמד לאורך התור הממוצע רחוק מהערך האמיתי מכיוון שהמערכת אינה במצב יציב.

3.2 הסבירו כיצד תחשבו (בעזרת דו"ח ה-ACD) את הממוצע היומי של מספר השרתים במערכת ואת ההסתברות היומית להמתנה (דהיינו, אחוזי הלקוחות שנאלץ להמתין בתור לפני קבלת שירות, לאורך היום). הגדירו את הסימונים בהם אתם משתמשים.

תשובה:

נסמן:

λ_i - מספר ההגעות באינטרוול ה- i ($i=1, 2, 3, \dots, 20$).

N_i - ממוצע מספר השרתים במערכת באינטרוול ה- i .

$P_i(W_q > 0)$ את ההסתברות להמתנה באינטרוול ה- i .

$$P(N_{\text{daily}}) = \frac{\sum_{i=1}^{20} N_i}{20}$$

$$P(W_q > 0) = \frac{\sum_{i=1}^{20} \lambda_i * P_i(W_q > 0)}{\sum_{i=1}^{20} \lambda_i}$$

3.3 נתמקד ב-3 חצאי השעות המודגשים בדו"ח (12:00-12:30, 17:00-17:30, 19:00-19:30). עבור כל אחד מהאינטרוולים התאימו את המדיניות התפעולית המתאימה (ED, QD או QED). נמקו במדויק ככל שתוכלו, תוך התייחסות למדד תפעולי נוסף שאינו מופיע בדו"ח ה-ACD.

תשובה:

12:00-12:30 -

מדיניות תפעולית: ED

הסבר:

אחוז הממתינים $\approx 100\%$

אחוז נוטשים גבוה $\approx 9\%$

זמן המתנה ארוך יחסית (למעלה מ-1/4 משך שירות).

$$\rho = \frac{R}{N} = \frac{51.66}{47.28} = 100\%$$

$$\beta = \frac{47.28 - 51.66}{\sqrt{51.66}} \approx -0.6$$

17:00-17:30 -

מדיניות תפעולית: QED

הסבר:

אחוז הממתינים $\approx 49\%$

אחוז נוטשים לא גבוה $\approx 3\%$

$$\rho = \frac{R}{N} = \frac{54.2}{60.23} \approx 90\%$$

$$\beta = \frac{60.23 - 54.2}{\sqrt{54.2}} \approx 0.8$$

19:00-19:30 -

מדיניות תפעולית: QD

הסבר:

נטישות והמתנות זניחות

נצילות השרתים $\approx 80\%$, $\rho = \frac{R}{N}$, יחסית נמוכה עבור "עשרות" שרתים.

$$\beta = \frac{56.93 - 44.6}{\sqrt{44.6}} \approx 1.84$$

3.4 עבור כל אחד משלושת האינטרוולים המסומנים, נניח כי מספר המופעים גדל פי 4 וגם מספר המוקדנים גדל פי 4. תנו הערכה לאחוז הנטישות תחת תרחיש אלטרנטיבי זה ונמקו בקצרה.

תשובה:

- 12:00-12:30

אחוז נוטשים חדש: 8.85% (ללא שינוי)

הסבר:

$$\lambda(1 - P\{Ab\}) \approx N\mu \Rightarrow P\{Ab\} \approx 1 - \frac{1}{\rho}$$

ו- $\rho = \frac{\lambda}{N\mu}$ לא משתנה.

- 17:00-17:30

אחוז נוטשים חדש: 1.63% (קטן פי 2)

הסבר:

$$P\{Ab\} \text{ פרופורציונאלי ל- } \frac{1}{\sqrt{N}}$$

- 19:00-19:30

אחוז נוטשים חדש: 0%

הסבר:

המערכת גדלה ולכן לא סביר שתהיה גרועה יותר

תשובה:

הנחנו מערכת Erlang-A ולכן

$$\frac{1}{\theta} = \frac{E(W_q)}{P\{Ab\}} = \frac{4.41}{0.0126} = 350 \text{ sec}$$

3.6 בשאלה זו נתייחס לגרף המוצג בעמוד הבא.

3.6.1 כפי שהוצג בהרצאות, כל אחת מהעקומות בגרף מתאימה לערך אחר של θ/μ . כתבו בתיבות המצוינות בגרף את הערך θ/μ המתאים עבור כל אחת משלוש העקומות המסומנות. הסבירו את קביעתכם.

תשובה:

$$\frac{\theta}{\mu} = \frac{E(\text{Service})}{E(\text{Patience})}$$

נניח N (מספר השרתים) קבוע וממוצע זמן שירות קבוע.

נשנה את סבלנות הלקוחות - אם הסבלנות גדלה, אזי פחות לקוחות ינטשו, כלומר עקומה נמוכה יותר $\leq$ העקומה

התחתונה מתאימה ל- $\theta/\mu = 0.1$, העליונה ל- $\theta/\mu = 100$ והשלישית ל- $\theta/\mu = 1$.

3.6.2 להבדיל מסעיף 3.5, אימדו עתה את סבלנות הלקוחות באינטרוול 13:30-14:00 תוך שימוש בגרף שבעמוד הבא. הסבירו בפירוט את דרך החישוב.

תשובה:

$$\sqrt{N} * P\{Ab\} = \sqrt{50} * 0.0126 = 0.09$$

$$R = 45$$

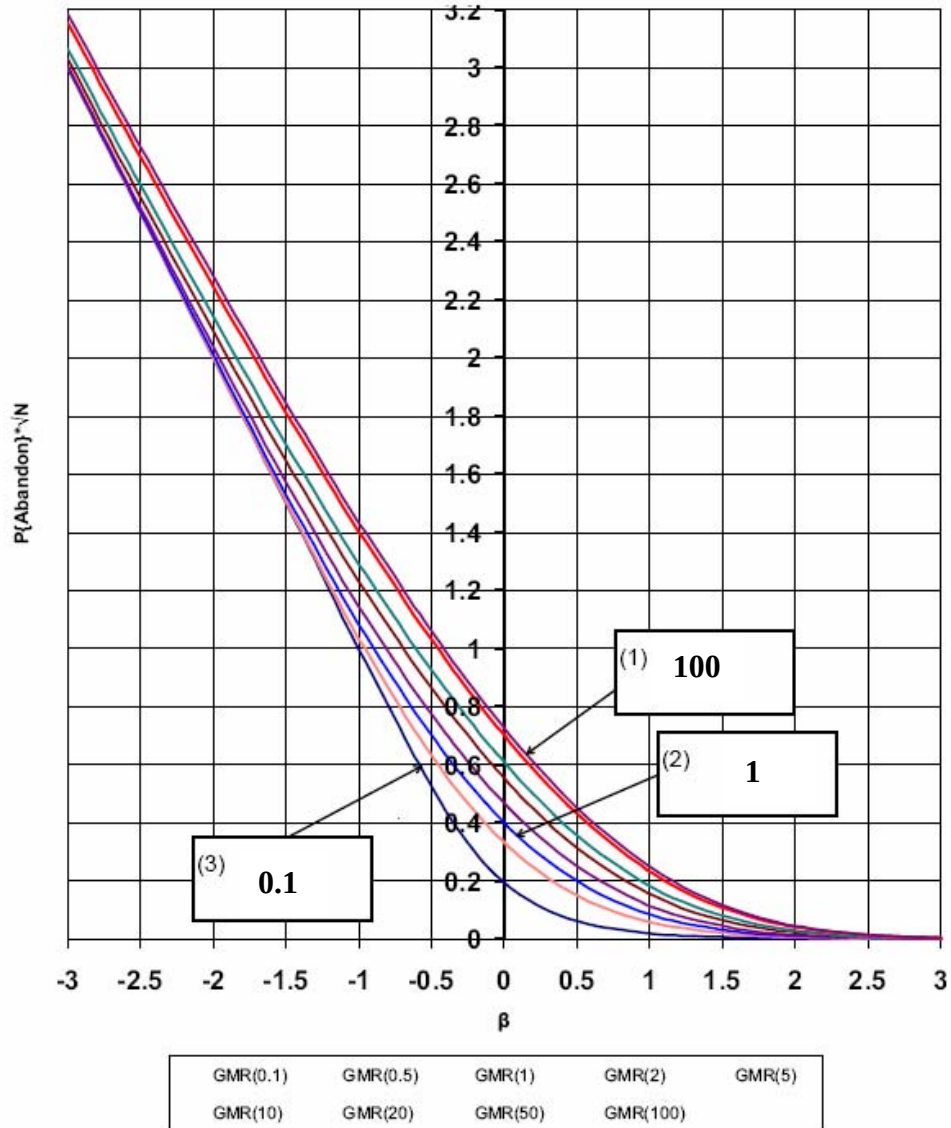
$$\beta = \frac{50 - 45}{\sqrt{45}} = 0.75$$

לפי הגרף, הנקודה (0.75, 0.09) נמצאת על הגרף של $\theta/\mu = 0.1$ ולכן

$$\frac{1}{\theta} = 10 * \frac{1}{\mu} = 1700 \text{ sec}$$

Erlang-A: % Abandonment

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



Note the behavior like $-\beta/\sqrt{n}$, 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.

חלק 2. איוש אופטימלי.

בחלק זה נניח כי את מוקד השירות הטלפוני ניתן לתאר כמערכת $M|M|N+M$ (Erlang-A) עם קצבי המופע המופיעים בדו"ח ה-ACD מתחילת השאלה, ממוצע זמן שירות של 3 דקות וממוצע סבלנות של $4\frac{1}{2}$ דקות.

3.7 בהרצאה תואר כלל ה-20-80 (לפחות 80% מהלקוחות יקבלו שירות תוך 20 שניות), ככלל נפוץ אך ללא הצדקה מדעית.

באמצעות תוכנת *4CallCenter*, נמצא כי עבור אינטרוול הזמן 12:30-13:00 כלל זה יושג ע"י איוש מינימלי של 52 שרתים.

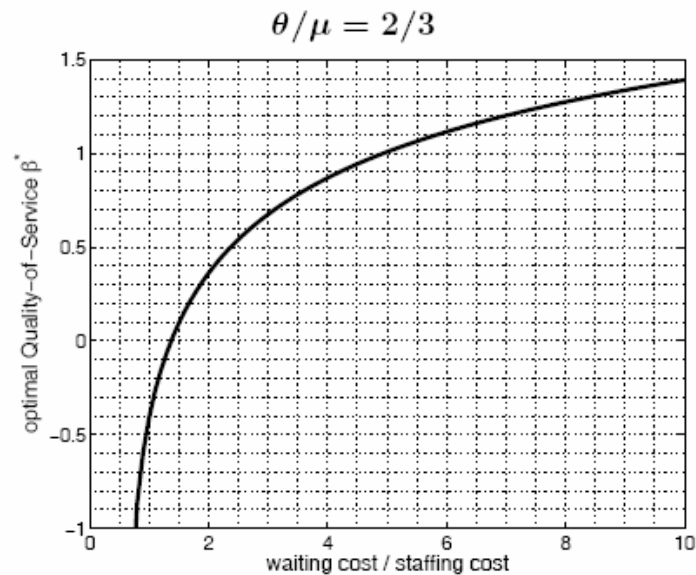
3.7.1 תארו, במדויק ככל שתוכלו, כיצד ניתן בעזרת הנתונים והתוכנה לקבל את רמת האיוש הנ"ל (לשם תזכורת הפונקציות של התוכנה, צרפנו להלן את העמוד הראשי שלה).

	Basic Interval (minutes)	Target Time to Answer	Number of Agents	Average Handling Time	Calls per Interval	Agent's Occupancy	Average Speed of Answer	%Answer within Target	Average Queue Length
Results									
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									

תשובה:

באופציית ה-Advanced Queries או ב-Staffing Query ניתן להכניס את הפרמטרים של הבעיה, לקבוע Target Time = 00:20, לקבוע כמטרה : 80% -ב- %Answer within Target ולעשות Query על מספר השרתים. בתוצאת ההרצה יש לבחור את ערך ה-Upper (לפחות 80%).

3.7.2 תוך שימוש בגרף הנתון, ותחת הנחת כלל ה-20-80, העריכו את שוויה של דקת המתנה של לקוח יחסית לדקת עבודה של מוקדן באינטרוול הנידון.

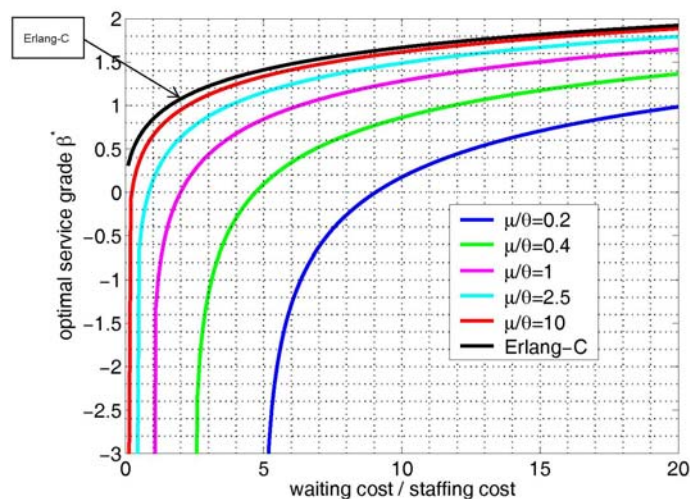


תשובה :

$$\beta \approx \frac{52 - 49}{\sqrt{49}} = 0.42 \Rightarrow \frac{\text{waiting_cost}}{\text{staffing_cost}} \approx 2$$

3.8 הוסיפו בגרף הנתון לעיל את העקומה המתאימה למודל שבו סבלנות הלקוחות אינסופית (ציירו אותה כמיטב יכולתכם) והסבירו בקצרה.

תשובה :



הערה: שימו לב כי בגרף זה העקומות מתייחסות ליחס- μ/θ , בעוד שהגרף המוצג לעיל מתייחס ל- θ/μ

במודל Erlang-C, כדי שהמערכת תהיה במצב יציב נדרוש כי $\beta > 0$ (עבור כל ערך של $r = \frac{\text{waiting_cost}}{\text{staffing_cost}}$ בנוסף, כאשר עלות ההמתנה גדלה (או עלות המוקדנים קטנה) נוסף מוקדנים למערכת ולכן הפונקציה עולה.

3.9 כעת נניח כי המוקד מאויש לפי כלל ה- 20-80, כאשר איוש אופטימלי מחושב לכל **שעה** (לפי איחוד נתוני שני חצאי השעה המופיעים בדו"ח). האיוש האופטימלי מחושב עבור כל שעה, באופן שתואר בסעיף 3.7.

נסמן ב- $N_1, \dots, N_{10}$ את רמות האיוש האופטימליות עבור כל שעה. בעיית מנהלת המוקד היא שלא ניתן להזמין מוקדנים רק לשעה מסוימת: כל מוקדן מוזמן **למשמרת** שאורכה 8 שעות, כולל **שעת הפסקה** לאחר 4 שעות עבודה. נסמן עלות מוקדן במשמרת ה- j - ית ב- C_j . על המנהלת לכסות את 10 שעות פעילות המוקד. הינכם מתבקשים לעזור למנהלת לקבוע את רמת האיוש האופטימלית למשמרת תחת האילוצים הנ"ל, וכפי שנלמד בכיתה.

3.9.1 הגדירו מטריצה $A = [a_{ij}]$ המתארת את המשמרות האפשריות (ז"א הגדירו את אברי המטריצה a_{ij}).

תשובה:

$$a_{ij} = \begin{cases} 1, & \text{if an agent working according to schedule } j \text{ is available during interval } i; \\ 0, & \text{otherwise.} \end{cases}$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

3.9.2 נסחו בעיית תכנות בשלמים, שפתרונה נותן איוש אופטימלי למשמרת עבור כל המשמרות האפשריות, וזאת במטרה להביא למינימום את עלויות האיוש של המוקד.

תשובה:

משתנה ההחלטה: X_j - מספר המוקדנים במשמרת ה- j .

$$\min \sum_{j=1}^3 C_j X_j$$

s.t.

$$\sum_{j=1}^3 a_{ij} X_j \geq N_i \quad i = 1, 2, \dots, 10$$

$$X_j \geq 0 \quad \text{integer}$$

שאלה 4 תאוריה (10 נקודות)

4.1 סטודנט שלא למד הנדסת מערכות שירות נדרש לתת עצה למנהל מוקד טלפוני שבו, בממוצע, קצב המופע הוא 100 שיחות לדקה ומשך השירות הוא 4 דקות. הסטודנט התבקש להמליץ על מספר המוקדנים הנדרשים להתמודד עם העומס הנ"ל. הסטודנט עשה חישוב נאיבי והמליץ על $400 = 4 \times 100$ מוקדנים.

תארו במדויק ככל שתוכלו, כולל הערכה מספרית של לפחות מדד ביצועים אחד, מה יקרה במוקד הנ"ל כאשר:

4.1.1 הלקוחות בעלי סבלנות רבה מאוד -

תשובה:

המוקד יפעל כמערכת ללא נטישות (Erlang-C), אחוז הממתינים גבוה, ממוצע זמן ההמתנה קרוב לממוצע זמן שירות ונצילות המוקדנים גבוהה (קרובה ל- 100%).

4.1.2 הלקוחות מאוד קצרי רוח (סבלנות נמוכה מאוד) -

תשובה:

המוקד יפעל בקירוב כמערכת עם מספר קווים כמספר המוקדנים (Erlang-B) וזאת כי לקוח המגיע למערכת ומוצא את כל המוקדנים עסוקים ינטוש מיד, נצילות המוקדנים גבוהה (קרוב ל- 100%), אחוז נטישות נמוך.

4.1.3 הלקוחות בעלי סבלנות סבירה (הסבלנות הממוצעת שווה בערך למשך השירות הממוצע) -

תשובה:

המוקד יפעל כמערכת $M | M | \infty$.

נשים לב כי $\theta = \mu$ ו- $\beta = 0$ ולכן אחוז הממתינים יהיה 50%.

כמוכן, מהגרף המוצג בעמוד 19 ניתן לחשב את אחוז הנוטשים - $P\{Ab\} = \frac{0.4}{\sqrt{400}} = 0.2$, $\theta/\mu = 1$.

4.2 נתון מודל למערכת שירות (Erlang-A או Erlang-C) עם הפרמטרים:

λ = קצב המופע ;

μ = קצב השירות (של שרת יחיד) ; או $E(S) = \frac{1}{\mu}$ ממוצע משך שירות ;

θ = קצב הנטישה (של לקוח יחיד) , במקרה של Erlang-A ;

N = מספר השרתים ;

בכיתה הוגדרו 3 תחומים תפעוליים עבור המערכות הנ"ל, המתוארים בשתי הטבלאות הבאות (בעמוד הנוכחי ובעמוד הבא).

השלימו את החלקים החסרים בטבלאות, תוך הגדרת הסימונים בהם אתם משתמשים (במידת הצורך):

$M | M | N$:

אינדקס הצפף $E[Wait]/E[S]$	רמת איש N	$P\{Wait > 0\}$	תחום תפעולי
$\frac{1}{\gamma}$	$0 < \gamma, R + \gamma$	≈ 1	ED
≈ 0	$0 < \delta, R + \delta R$	≈ 0	QD
$\frac{\beta}{\sqrt{N}}$	$0 < \beta, R + \beta\sqrt{R}$	$\alpha, 0 < \alpha < 1$	QED

סימונים :

- $R = \frac{\lambda}{\mu} = \lambda \cdot E(S)$ העומס המוגדר כרגיל ע"י $Offered-Load = R$
-
-

$$: \underline{M \mid M \mid N + M}$$

תחום תפעולי	$P\{Wait > 0\}$	רמת איוש N	אינדקס הצפף $E[Wait]/E[S]$	$P\{Ab\}$
ED	≈ 1	$0 < \gamma < 1, R - \gamma R$	$\frac{\gamma}{\theta E(S)} = \frac{\gamma \mu}{\theta}$	γ
QD	≈ 0	$0 < \delta, R + \delta R$	≈ 0	≈ 0
QED	$\alpha, 0 < \alpha < 1$	$-\infty < \beta < \infty, R + \beta \sqrt{R}$	$\frac{c \cdot \mu}{\theta \sqrt{N}}$	$\frac{c}{\sqrt{N}}$

סימונים:

- $c = c\left(\beta, \theta/\mu\right)$ – קבוע הניתן לחישוב בעזרת הפרמטרים של המודל.
-

4.3 נתון מודל Erlang-C בתחום מכון יעילות (ED), ז"א $N \approx R + \gamma$, $0 < \gamma$.

הוכיחו כי אינדקס הצפף נתון על ידי $\frac{1}{\gamma}$, ז"א:

$$\frac{E(W_q)}{E(S)} \approx \frac{1}{\gamma}$$

רמז:

$$W_q | W_q > 0 \stackrel{d}{=} \exp \left(\text{mean} = E(S) \cdot \frac{1}{N} \cdot \frac{1}{1-\rho} \right)$$

$$\text{כאשר } \rho = \frac{\lambda}{N\mu}$$

תשובה:

תחת מדיניות ED, $P(W_q > 0) \approx 1$ ולכן $E(W_q | W_q > 0) \approx E(W_q)$.

$$\frac{E(W_q | W_q > 0)}{E(S)} \approx \frac{1}{N} \cdot \frac{1}{1-\rho} = \frac{1}{\gamma}$$

4.4 נתון מודל Erlang-A בתחום מכון יעילות (ED), ז"א $N \approx R - \gamma R$, $0 < \gamma < 1$.

הוכיחו שאחוז הנוטשים נתון על ידי γ , ז"א:

$$P(\text{Abandon}) \approx \gamma$$

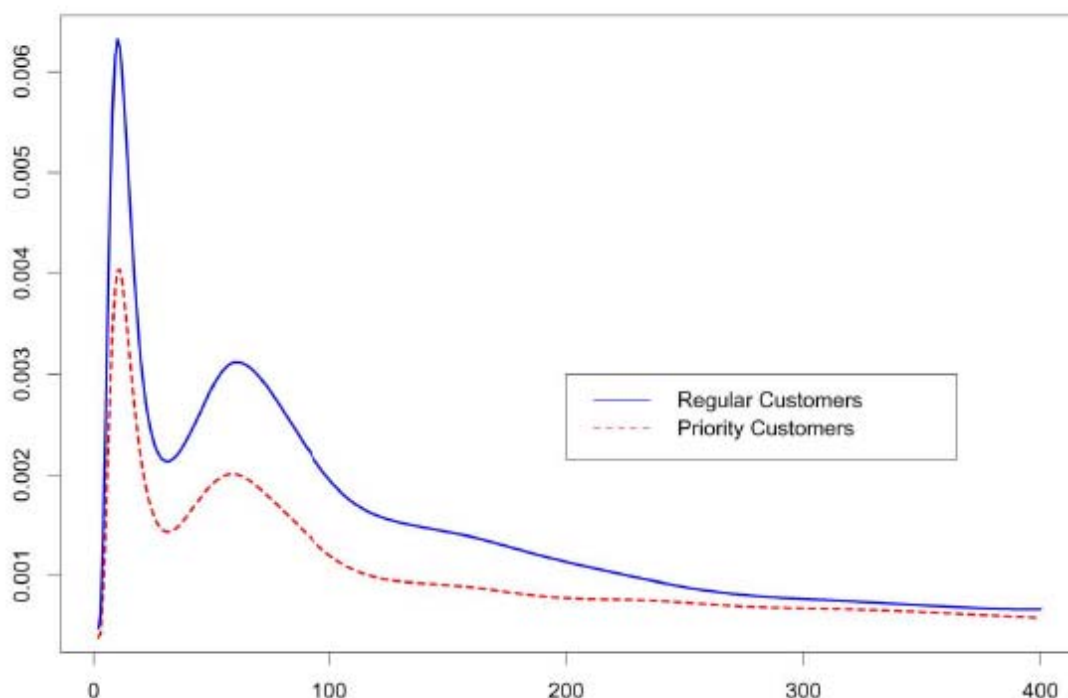
תשובה:

$$\lambda(1 - P\{Ab\}) \approx N\mu \Rightarrow P\{Ab\} = 1 - \frac{1}{\rho}$$

$$N = R - \gamma R \Rightarrow \gamma = \frac{R - N}{R} = 1 - \frac{1}{\rho} = P\{Ab\}$$

שאלה 5 שאלת בונוס

המאמר המצורף בעמוד הבא פורסם במגזין הטכניון, אביב 2007. כדרכם של עיתונאים, עריכה עיתונאית של הטקסט המקורי הובילה לאי דיוקים. למשל, הפיסקה המסומנת במאמר היא קיצור דרסטי של תאור הגרף הבא:



מה מדויק יותר ומה פחות בפיסקה המסומנת?

תשובה:

הנאמר בפיסקה המסומנת מתייחס למוקד טלפוני מסוים, עבורו בוצע ניתוח סבלנות הלקוחות, בעוד שמידע זה מוצג במאמר כמסקנות כלליות עבור כלל מוקדי הטלפונים. ה"שיא" ב- 60 שניות בפונקציית הסיכון הוא תוצאה של הודעה המושמעת ללקוח כעבור זמן זה במוקד שנבחן, ולכן לא נכון להציג נטישה כעבור 60 שניות כתופעה כללית. עם זאת, הצגת לקוחות VIP כסבלניים יותר מלקוחות רגילים מדויקת יותר.