הטכניון - מכון טכנולוגי לישראל
הפקולטה להנדסת תעשיה וניהול ע"ש ויליאם משה דוידסון

Technion - Israel Institute of Technology
The William Davidson Faculty of Industrial Engineering and Management

---

תאריך הבחינה: 20.09.12
שם המרצה: ד"ר אלה לוין

מבוא להסתברות וסטטיסטיקה - 094480
סמסטר אביב תשע"ב – בחינה סופית, מועד ב'

כולל פתרון

משך הבחינה: שלוש שעות.

חומר מותר: 2 דפי עזר בגודל A4 כתובים משני הצדדים, מחשבון כיס פשוט. בנוסף מחולקים דפי התפלגויות בדידות מיוחדות, טבלת מבחני השערות ולוחות סטטיסטיים.
לבחינה שני חלקים. **יש לענות על כל השאלות** בכל אחד מחלקי הבחינה.
החלק הראשון מכיל שאלות רב-ברירתיות (סגורות). **על חלק זה יש לענות רק בדף המיועד לכך המצורף למבחן**. אין לסמן יותר מתשובה אחת לכל שאלה. הציון יינתן על סמך הסימונים בדף התשובות בלבד!
החלק השני מכיל שאלות פתוחות, עליהן יש לענות במחברת. בחלק זה יש לפרט ולהסביר את כל השלבים שביצעתם בדרך לפתרון.
עליכם להחזיר את שאלון הבחינה עם מחברות החישובים (טיוטא).

ב ה צ ל ח ה !

חלק ראשון – יש לענות על כל השאלות בחלק זה. משקל כל שאלה: 5 נק'

שאלה 1

מפעל AAA מייצר צינורות השקיה איכותיים. קוטר הקנה של כל צינור (בס"מ) הוא משתנה מקרי Y בעל פונקצית הצפיפות הבאה:

$$f_Y(y)=\begin{cases}\dfrac{y^2}{8} & 0\le y\le 2\\ \dfrac{y\cdot(4-y)}{8} & 2\le y\le 4\\ 0 & \text{אחרת}\end{cases}$$

אין תלות בין הצינורות השונים.

יהי W משתנה מקרי המציין את מספר הצינורות אשר קוטר הקנה שלהם עולה על 3 ס"מ, מבין 33 צינורות שנדגמו באקראי.

$P(W\ge 10)$ נמצאת באינטרוול:

א. 0.01- 0.05

ב. 0.05- 0.10

ג. 0.10-0.15

ד. 0.85-0.90

פתרון:

$$P(Y>3)=\int_3^4 \frac{y}{2}-\frac{y^2}{8}\cdot dx=\frac{y^2}{4}-\frac{y^3}{24}\Big|_3^4=4-\frac{8}{3}-\frac{9}{4}+\frac{27}{24}=\frac{5}{24}=0.2083$$

מכאן נקבל כי $W\sim Bin(33,0.2083)$. לפי הקירוב הנורמלי לבינום

$$W\dot{\sim} N(33\cdot 0.2083, 33\cdot 0.2083\cdot 0.7917)=N(6.8739,5.442)$$

$$P(W\ge 10)=1-\Phi\left(\frac{10-6.8739}{\sqrt{5.442}}\right)=1-\Phi(1.34)=1-0.9909=0.0901$$

שאלה 2

באולימפיאדת לונדון 2012 משתתפות 12 נבחרות כדורגל נשים, המחולקות באופן אקראי ל-3 בתים (4 נבחרות בכל בית). בשלב הבתים, כל נבחרת משחקת רק מול הנבחרות המשובצות בבית שלה ופעם אחת מול כל נבחרת.

מספר המשחקים בשלב הבתים הוא:

א. 6

ב. 12

ג. 18

ד. 36

פתרון:

$$\underbrace{\binom{4}{1}}_{\text{Choose the 1st team}} \quad \underbrace{\binom{3}{1}}_{\text{Choose the opponent out of the rest}} \cdot \underbrace{\frac{1}{2}}_{\text{Symmetry}} \cdot 3 = 18$$

שאלה 3 (עם התיקון)

4 דגלים אדומים שונים, 2 דגלים לבנים שונים ו-2 דגלים שחורים שונים מסודרים באופן אקראי בשורה.

נגדיר משתנה מקרי X הסופר את מספר הדגלים הלבנים **לפני** הדגל האדום הראשון.

$P(X=0)$ שווה ל:

א. 0.0666

ב. 0.1666

ג. 0.2666

ד. 0.6666

פתרון:

כל אפשרויות הסידור: $8!$

נפריד ל-3 מקרים אפשריים עבור $X=0$:

הדגל הראשון הוא אדום: $\binom{4}{1}7!$ (בחירת הדגל האדום הראשון וסידור השאר)

הדגל הראשון הוא שחור והשני הוא אדום: $\binom{2}{1}\binom{4}{1}6!$ (בחירת השחור הראשון, בחירת האדום למקום השני וסידור השאר.)

שני הדגלים הראשונים הם שחורים והדגל השלישי הוא אדום: $\binom{4}{1}5!2!$ (סידור 2 השחורים, בחירת האדום למקום השלישי וסידור שאר הדגלים)

מכאן:

$$P(X=0)=\frac{\binom{4}{1}7!+\binom{2}{1}\binom{4}{1}6!+2!\binom{4}{1}5!}{8!}=\frac{26880}{40320}=\frac{2}{3}$$

שאלה 4

עבור אוכלוסיה שתוחלתה μ ושונותה $\sigma^2=1$, רוצים לאמוד את μ^2 . לשם כך נלקח מדגם מקרי בגודל 20 והוצעו שלושת האומדים הבאים:

1. $\overline{X}^2+\frac{S^2}{20}$
2. $\overline{X}^2-\frac{S^2}{20}$
3. $\overline{X}^2-\frac{1}{20}$

כאשר S^2 היא השונות המדגמית.

מהי הטענה הנכונה?

א. האומד הראשון חסר הטיה, 2 ו-3 מוטים.

ב. האומד הראשון מוטה, 2 ו-3 חסרי הטיה.

ג. האומד השלישי מוטה, 1 ו-2 חסרי הטיה.

ד. האומד השלישי חסר הטיה, 1 ו-2 מוטים.

פתרון:

נבדוק אם האומדים ח"ה:

1. $E\left(\bar{X}^2+\frac{S^2}{20}\right)=E\left(\bar{X}^2\right)+E\left(\frac{S^2}{20}\right)=Var\left(\bar{X}\right)+E\left(\bar{X}\right)^2+\frac{1}{20}E\left(S^2\right)=\frac{\sigma^2}{20}+\mu^2+\frac{\sigma^2}{20}\underset{\sigma=1}{=}\mu^2+\frac{1}{10}$

2. $E\left(\bar{X}^2-\frac{S^2}{20}\right)=E\left(\bar{X}^2\right)-E\left(\frac{S^2}{20}\right)=\frac{\sigma^2}{20}+\mu^2-\frac{\sigma^2}{20}=\mu^2$

3. $E\left(\bar{X}^2-\frac{1}{20}\right)=E\left(\bar{X}^2\right)-\frac{1}{20}=\frac{\sigma^2}{20}+\mu^2-\frac{1}{20}\underset{\sigma=1}{=}\mu^2$

התשובה היא ב'!

שאלה 5

יהי $X_1, X_2, ..., X_n$ מדגם מקרי מתוך אוכלוסייה כלשהי.

יהי $\bar{X} = \frac{X_1 + X_2 + ... + X_n}{n}$

$\bar{X}$ הוא:

א. פרמטר של האוכלוסייה

ב. הממוצע של האוכלוסייה

ג. תוחלת

ד. סטטיסטי

פתרון:

סטטיסטי הוא פונקציה של התצפיות במדגם וזה בדיוק מה ש- $\bar{X}$ מקיים.

שאלה 6

יהיו A ו-B מאורעות זרים במרחב מדגם Ω.

אזי הם:

א. תמיד תלויים.

ב. אף פעם לא תלויים.

ג. בלתי מתואמים.

ד. אין מספיק נתונים בשאלה.

פתרון:

אם A ו-B מאורעות זרים **ולא ריקים** אזי הם תלויים, אחרת הם ב"ת. לא נאמר אם המאורעות ריקים או לא ולכן אין מספיק מידע כדי לקבוע משהו לגבי התלות ביניהם.

שתי השאלות הבאות מתייחסות לנתונים הבאים:

יהי X ציון בבחינת הבגרות בספרות. X מתפלג נורמלית כך ש- 72.57% מהתלמידים קיבלו ציון גבוה מ-70 ו- 91.92% קיבלו ציון נמוך מ-88.

שאלה 7

בכיתה כלשהי יש 100 תלמידים. תוחלת מספר התלמידים שקיבלו ציון בין 66 ל-84 היא (בערך) :

א. 59.7

ב. 68.1

ג. 71.6

ד. 75.4

פתרון:

$X \sim N(\mu,\sigma^2)$ ציון בבחינת הבגרות בספרות. ראשית יש למצוא את הפרמטרים של ההתפלגות.

נבנה מערכת של שתי משוואות בשני נעלמים :

$$\begin{cases} P(X>70)=0.7257 \\ P(X<88)=0.9192 \end{cases} \Rightarrow \begin{cases} P\left(Z>\dfrac{70-\mu}{\sigma}\right)=0.7257 \\ P\left(Z<\dfrac{88-\mu}{\sigma}\right)=0.9192 \end{cases} \Rightarrow \begin{cases} \dfrac{70-\mu}{\sigma}=-0.6 \\ \dfrac{88-\mu}{\sigma}=1.4 \end{cases}$$

$$\Rightarrow \begin{cases} \mu=75.4 \\ \sigma=9 \end{cases}$$

מכאן : $X \sim N(75.4,9^2)$

נסמן ב-Y את מספר התלמידים מתוך 100 שציונם נע בין 66 ל-84.

$Y \sim Bin(100,p)$.

$$p=P(66<X<84)=P(X<84)-P(X<66)=P\left(Z<\frac{84-75.4}{9}\right)-P\left(Z<\frac{66-75.4}{9}\right)$$

$$=\Phi(0.95)-(1-\Phi(1.04))=0.8303-(1-0.8508)=0.681$$

$$E(Y)=100\cdot 0.681=68.1 \quad \Leftarrow \quad Y\sim Bin(100,0.681)$$

שאלה 8

מהי הטענה ש**אינה** נכונה?

א. השברון ה-0.1492 של ההתפלגות הנורמלית סטנדרטית הוא -1.044

ב. עבור $x=66$, הערך של פונקצית ההתפלגות המצטברת של המשתנה המקרי הנורמלי הסטנדרטי שווה לערך פונקצית ההתפלגות המצטברת של המשתנה המקרי הנורמלי ושווה ל-0.1492.

ג. השברון ה-0.8508 של ההתפלגות הנורמלית הוא 84

ד. $P(X\le 66)=0.1492$

שאלה 9

בניתוח קשר בין שני משתנים X ו-Y נבדקו 35 תצפיות ונמצא כי $\hat{y} = -0.85x + C$.

כאשר C מספר קבוע. כמו כן נמצא כי $R^2 = 0.9$

מהי הטענה הנכונה?

א. מקדם המתאם הוא 0.9486-

ב. מקדם המתאם הוא 0.9486

ג. אין אפשרות לדעת מהו מקדם המתאם, כי C הוא מספר לא ידוע.

ד. אין אפשרות לדעת מהו מקדם המתאם, כי השיפוע של הקו לא ידוע.

פתרון:

$R = \sqrt{R^2} \Rightarrow R = \pm\sqrt{0.9} = \pm 0.9486$. השיפוע שלילי ולכן $R = -0.9486$

שאלה 10

עבור המשתנים X ו-Y נמצאו הערכים המסכמים הבאים:

$\sum x_i y_i = 16731 \quad S_{yy} = 465.34 \quad S_{xx} = 3756.96 \quad n = 26 \quad \sum x_i = 1613 \quad \sum y_i = 281.9$

בדקו את ההשערה:

$H_0 : \rho = -0.5$

$H_1 : \rho < -0.5$

מהי מסקנתכם, ברמת מובהקות של 5%?

א. לא נדחה את H_0

ב. נדחה את H_1

ג. נדחה את H_0

ד. אין מספיק נתונים כדי לפתור את השאלה.

פתרון:

$H_0 : \rho = -0.5$

$H_1 : \rho < -0.5$

סטטיסטי המבחן הוא $T = \frac{R(n-2)}{\sqrt{1-R^2}} \sim t_{(n-2)}$

נדחה את H_0 אם $T < -t_{(24),0.95} = -1.711$

נמצא את R:

$$R = \frac{s_{xy}}{\sqrt{s_{xx} \cdot s_{yy}}} = \frac{\sum_{i=1}^{26} x_i y_i - 26\bar{x} \cdot \bar{y}}{\sqrt{3756.96 \cdot 465.34}} = \frac{\sum_{i=1}^{26} x_i y_i - 26\bar{x} \cdot \bar{y}}{\sqrt{3756.96 \cdot 465.34}} = \frac{16731 - 26 \cdot \frac{281.9}{26} \cdot \frac{1613}{26}}{1322.219}$$

$$= \frac{-757.642}{1322.219} = -0.573$$

ערך סטטיסטי המבחן הוא: $T = \frac{R\sqrt{(n-2)}}{\sqrt{1-R^2}} = \frac{-0.573 \cdot \sqrt{24}}{\sqrt{1-0.573^2}} = \frac{-2.807}{0.8195} = -3.425$

$T = -3.425 < -1.711 = t_{(24),0.95}$ ולכן נדחה את השערת האפס בר״מ 0.05

חלק שני – שאלות פתוחות. יש לענות על כל השאלות בחלק זה.

שאלה מספר 11 (עם התיקון) (20 נק')

הזמן הדרוש לתיקון מכונה מקולקלת מסוג A הוא משתנה מקרי אקספוננציאלי עם תוחלת 70 דקות.

הזמן הדרוש לתיקון מכונה מקולקלת מסוג B הוא משתנה מקרי אחיד (רציף) בקטע [100 , 25] דקות.

ביום מסוים נמצאים במפעל "תיק-תיקון" 6 מכונות מקולקלות מסוג A ו- 4 מכונות מקולקלות מסוג B (לא הגיעו מכונות מקולקלות נוספות במהלך היום). במפעל עובד טכנאי אחד שתפקידו לתקן את המכונות המקולקלות. הטכנאי דוגם באקראי (וללא החזרה) 2 מכונות מקולקלות בתחילת היום.

אם הטכנאי מספיק לתקן לפחות אחת מהמכונות המקולקלות שדגם בפרק זמן של שעה לכל היותר **לכל מכונה**, זוכה הטכנאי בארוחת ערב זוגית.

א. אם ידוע שהטכנאי זכה בארוחת הערב, מה ההסתברות שבתחילת היום הוא דגם בדיוק מכונה מקולקלת אחת מסוג A?

ב. נניח שמדי תחילת יום נמצאים במפעל 6 מכונות מקולקלות מסוג A ו- 4 מכונות מקולקלות מסוג B. חשבו את ההסתברות שמספר הימים שיעברו (כולל היום) עד הזכייה השלישית בארוחת ערב זוגית יהיה 4 ימים לכל היותר.

פתרון:

א. נגדיר את המשתנים הבאים:

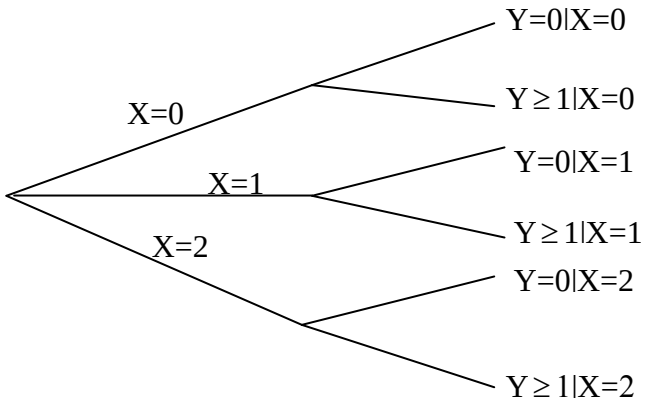
$R_A \sim \exp(1/70)$ - משך הזמן בדקות לתיקון מכונה מסוג A

$R_B \sim U(25,100)$ - משך הזמן בדקות לתיקון מכונה מסוג B

$X \sim HG(10,6,2)$ - מספר מכונות מסוג A שדגם

Y - מספר המכונות שמשך תיקונן היה פחות משעה. $Y=\{0,1,2\}$

נראה באמצעות דיאגרמת עץ:



מחפשים את ההסתברות:

$$P(X=1|Y\ge 1)\underset{Bayes}{=}\frac{P(Y\ge 1|X=1)P(X=1)}{P(Y\ge 1)}$$

נמצא את ההסתברות שיזכה בארוחת ערב בעזרת נוסחת ההסתברות השלמה:

$$P(Y\ge 1)=P(Y\ge 1|X=0)P(X=0)+P(Y\ge 1|X=1)P(X=1)+P(Y\ge 1|X=2)P(X=2)$$
$$=(1-P(Y=0|X=0))P(X=0)+(1-P(Y=0|X=1))P(X=1)+(1-P(Y=0|X=2))P(X=2)$$
$$=\left(1-P(R_B>60)^2\right)\frac{\binom{4}{2}}{\binom{10}{2}}+\left(1-P(R_B>60)P(R_A>60)\right)\frac{\binom{4}{1}\binom{6}{1}}{\binom{10}{2}}+\left(1-P(R_A>60)^2\right)\frac{\binom{6}{2}}{\binom{10}{2}}$$
$$=\left(1-\left(\frac{100-60}{100-25}\right)^2\right)\cdot\frac{2}{15}+\left(1-\left(\frac{40}{75}\right)\cdot\left(e^{-\frac{60}{70}}\right)\right)\cdot\frac{8}{15}+\left(1-\left(e^{-\frac{60}{70}}\right)^2\right)\cdot\frac{1}{3}$$
$$=0.7155\cdot\frac{2}{15}+0.773\cdot\frac{8}{15}+0.819\cdot\frac{1}{3}=0.0954+0.412+0.273=0.7804$$

ההסתברות המבוקשת היא:

$$P(X=1|Y\ge 1)=\frac{P(Y\ge 1|X=1)P(X=1)}{P(Y\ge 1)}=\frac{0.412}{0.7804}=0.5279$$

ב.

$W\sim NB(3,P(Y\ge 1)=NB(3,0.7804)$ - מספר הימים עד לזכייה שלישית בארוחת ערב.

ההסתברות המבוקשת היא:

$$P(W\le 4)=P(W=3)+P(W=4)=0.7804^3+\binom{4-1}{3-1}0.7804^3\cdot(1-0.2196)$$
$$=0.475+0.3131=0.7881$$

<u>שאלה 12</u> (30 נק')

<u>I. (20 נק')</u>

שני סטטיסטיקאים דגמו מאוכלוסיית חיפה מדגמים מקריים באופן בלתי תלוי אחד בשני, כדי לאמוד את פרופורציית תושבי חיפה המצדדים בנושא מסוים.

הסטטיסטיקאי הראשון דגם 100 תושבים וקיבל רווח סמך לפרופורציה ברמת בטחון של 95%:

$$0.304\le p\le 0.496$$

הסטטיסטיקאי השני דגם 150 תושבים וקיבל שהפרופורציה המדגמית שווה ל-48%.

א. על סמך המידע של שני הסטטיסטיקאים על 250 התושבים שנדגמו בנו רווח סמך **אחד** לפרופורציית התושבים באוכלוסיית חיפה המצדדים בנושא ברמת בטחון של 95%.

ב. ציינו מהו הפרמטר הנבדק ומהו האומד הנקודתי. כיצד האומד מתפלג (בקירוב)? מהי התוחלת והשונות שלו? האם הוא חסר הטיה או מוטה? הוכיחו.

ג. מהו גודל המדגם שהסטטיסטיקאים צריכים לדגום כדי לאמוד את הפרמטר, כך שבהסתברות של 95% שגיאת האמידה לא תעלה על 5%?

ד. ידוע מניסיון עבר כי אחוז תושבי חיפה המצדדים בנושא הנ"ל הוא 52%. רוצים לבדוק אם חלה ירידה בהעדפת תושבי חיפה:

1. נסחו את מערכת ההשערות המתאימה.
2. בדקו את ההשערות ברמת מובהקות של 0.05, לפי סטטיסטי המבחן ולפי P_value

<u>פתרון:</u>

א. רווח סמך מקורב לפרופורציה ברמת סמך של $1-\alpha$:

$$\hat{p}-Z_{1-\frac{\alpha}{2}}\cdot\sqrt{\frac{\hat{p}\hat{q}}{n}}\le p\le\hat{p}+Z_{1-\frac{\alpha}{2}}\cdot\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

לפי המדגם הראשון:

$$0.304\le p\le 0.496$$
$$\Rightarrow \hat{p}_1=\frac{0.304+0.496}{2}=0.4=\frac{S_1}{100}\Rightarrow S_1=40$$

לפי המדגם השני: $\hat{p}_2=0.48=\frac{S_2}{150}\Rightarrow S_2=72$

משני המדגמים נקבל:

$$\hat{p}=\frac{40+72}{100+150}=0.448$$

נציב ברווח הסמך ונקבל:

$$1-\alpha=0.95\Rightarrow\alpha=0.05$$
$$Z_{1-\frac{\alpha}{2}}=Z_{1-\frac{0.05}{2}}=Z_{0.975}=1.96$$
$$\hat{p}-1.96\cdot\sqrt{\frac{\hat{p}*(1-\hat{p})}{n}}\le p\le\hat{p}+1.96\cdot\sqrt{\frac{\hat{p}*(1-\hat{p})}{n}}$$
$$0.448-1.96\cdot\sqrt{\frac{0.448*(1-0.448)}{250}}\le p\le 0.448+1.96\cdot\sqrt{\frac{0.448*(1-0.448)}{250}}$$
$$\Rightarrow 0.3864\le p\le 0.5096$$

ב.

הפרמטר הנבדק הוא הפרופורציה באוכלוסייה: p

אומד נקודתי לפרופורציה: $\hat{p}$ ← $\hat{p}=\frac{X}{n}$ כאשר X הוא מספר המצדדים בנושא במדגם בגודל n ו-

$X \sim Bin(n,p)$

לפי מג"מ, כאשר המדגם גדול מספיק אזי: $\hat{p} \dot{\sim} N\left(p,\frac{p\cdot(1-p)}{n}\right)$

תוחלת: $E(\hat{p})=E\left(\frac{X}{n}\right)=\frac{1}{n}E(X)\underset{X\sim Bin(n,p)}{=}\frac{1}{n}np=p$

שונות: $Var(\hat{p})=Var\left(\frac{X}{n}\right)=\frac{1}{n^2}Var(X)\underset{X\sim Bin(n,p)}{=}\frac{1}{n^2}np(1-p)=\frac{p(1-p)}{n}$

האומד חסר הטייה: $Bias_{\hat{p}}(p)=E(\hat{p})-p=0 \leftarrow E(\hat{p})=p$

ג.

$$n \geq \left(\frac{z_{1-\alpha/2}}{2d}\right)^2=\frac{1.96^2}{4\cdot 0.05^2}=384.16$$
$$\Rightarrow n_{\min}=385$$

ד.1. מערכת ההשערות:

$$H_0: p=0.52$$
$$H_1: p<0.52$$

2. בדיקת מערכת ההשערות לפי סטטיסטי המבחן:

ראשית יש לבדוק כי אכן מותר להשתמש בקירוב בינומי לנורמלי:

$$\begin{cases} n\cdot p_0=250\cdot 0.52=130>>5 \\ n\cdot q_0=250\cdot 0.48=120>>5 \end{cases} \quad or \quad n\cdot p_0\cdot q_0=250\cdot 0.52\cdot 0.48=62.4>10$$

נדחה את השערת האפס אם מתקיים:

$$Z<-z_{1-\alpha}=-z_{0.95}=-1.645$$

סטטיסטי המבחן:

$$Z=\frac{0.448-0.52}{\sqrt{\frac{0.52\cdot 0.48}{250}}}=\frac{-0.072}{0.0316}=-2.278$$

$Z=-2.278<-1.645=-z_{0.95}$ ← נדחה את השערת האפס ברמת מובהקות 0.05.

חישוב P-value:

$$P-value = P(\hat{p} < 0.448) \approx P(Z < -2.278) = \Phi(-2.278) = 1-\Phi(2.278) \approx 1-0.9886 = 0.0114$$

$\alpha = 0.05 > 0.0114 = P-Value$ ← נדחה את השערת האפס ברמת מובהקות 0.05.

מסקנה: חלה ירידה באחוז תושבי חיפה המצדדים בנושא מסוים, ברמת מובהקות 0.05.

<u>II. (10 נקודות)</u>

במסגרת הבקרה על האיכות במפעל כלשהו בודקים כל יום מדגם מקרי של 25 מוצרים. לכל מוצר רושמים את משקלו בגרמים. נסמן משקל זה ב-X. נניח כי $X \sim N(\mu,\sigma^2)$

א. ביום מסוים התקבל $\bar{X} = 24$, $s = 0.9$. בנו רווח סמך של ברמת סמך של 95% לתוחלת המשקל של המוצר בקו הייצור? מהו אורך רווח הסמך שבניתם?

ב. ביום אחר התקבל $\bar{X} = 23.9$, $s = 1.01$. מה יהיה אורך רווח הסמך לתוחלת ברמת סמך של 95% ביום זה? הסבירו מדוע אורך רווח הסמך מסעיף ב' שונה מאורך רווח הסמך מסעיף א'.

ג. האם בהשוואת התוצאות של הממוצעים בשני הימים ניתן להסיק שחל שינוי בתוחלת? בצעו את המבחן הסטטיסטי המתאים ברמת מובהקות של 0.05 באמצעות סטטיסטי המבחן ו- P-value.

<u>פתרון:</u>

א. רווח סמך:

$$\mu \in \left[\bar{x} \pm t_{0.975}^{24} \cdot \frac{s}{\sqrt{n}}\right] \Rightarrow \mu \in \left[24 \pm 2.064 \cdot \frac{0.9}{5}\right] \Rightarrow [23.628, 24.37]$$

$$2d_1 = 0.743$$

ב.

$$\mu \in \left[\bar{x} \pm t_{0.975}^{24} \frac{s}{\sqrt{n}}\right] \Rightarrow \mu \in \left[23.9 \pm 2.064 \cdot \frac{1.01}{5}\right] \Rightarrow [23.48, 24.317]$$

$$2d_2 = 0.833$$

לא נקבל את אותו אורך רווח סמך (ובאופן כללי את אותו רווח סמך) כי בכל יום המדגם שונה – מורכב מפריטים שונים. מתוך המדגם השונה יש לחשב סטיית תקן מדגמית וממוצע מדגמי ואלו שונים בכל מדגם. על סמך האומדים הללו, השונים בכל יום, נחשב רווח סמך. על כן רווח הסמך ואורכו שונים.

ג. מערכת ההשערות:

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 - \mu_2 \neq 0$$

מדובר בבדיקת השערת על הפרש תוחלות, כאשר המדגמים ב"ת, השונויות לא ידועות ומניחים כי הן שוות.

נחשב את s_p^2 : $s_p^2 = \frac{24*0.9^2 + 24*1.01^2}{25+25-2} = 0.91505$ או $s_p^2 = \frac{0.9^2 + 1.01^2}{2} = 0.91505$ (משום שגדלי שני המדגמים שווים)

נבדוק את מערכת ההשערות דרך **סטטיסטי המבחן** :

$$T = \frac{(24-23.9)-0}{\sqrt{0.91505*(\frac{1}{25}+\frac{1}{25})}} = 0.3696$$

נשים לב שאין בטבלה 48 דרגות חופש ולכן נעבוד לפי 40 דרגות חופש (הכי קרוב שניתן)

עבור $\alpha = 0.05$ נדחה את השערת האפס אם מתקיים : $|T| > t_{(48),0.975} \approx t_{(40),0.975} = 2.021$

$\leftarrow -2.021 < 0.3696 < 2.021$

ולכן לא נדחה את השערת האפס עבור רמת מובהקות 0.05.

P-Value :

$$\frac{1}{2}P-Value = P\left(T^{(48)} > 0.3696\right) = 1-P\left(T^{(48)} < 0.3696\right) \approx 1-P\left(T^{(40)} < 0.3696\right)$$

$$\rightarrow 1-0.65 < \frac{1}{2}P-Value < 1-0.6$$

$$\rightarrow 0.7 < P-Value < 0.8$$

$P-Value$ מקבל ערכים גבוהים יחסית ולכן בכל ר״מ סבירה לא נדחה את השערת האפס, ובפרט עבור $\alpha = 0.05$.

מסקנה : אין הבדל בין תוחלות משקלי המוצרים בשני הימים שנבדקו, בכל רמת מובהקות סבירה ובפרט עבור ר״מ 0.05