

סטודנטים יקרים

לפניכם ספר תרגילים בקורס מבוא לסטטיסטיקה והסתברות ב'.
הספר הוא חלק מקורס חדשני וראשון מסוגו בארץ בנושא זה,
המועבר ברשת האינטרנט On-line.

הקורס באתר כולל פתרונות מלאים לספר התרגילים, וכן את
התיאוריה הרלוונטית לכל נושא ונושא.

**הקורס כולו מוגש בסרטוני וידאו המלווים בהסבר קולי, כך שאתם
רואים את התהליכים בצורה מובנית, שיטתית ופשוטה, ממש כפי
שנעשה בשיעור פרטי, לדוגמה [לחצו כאן](#).**

את הקורס בנה מר ברק קנדל, מרצה מבוקש במוסדות אקדמיים
שונים ובעל ניסיון עתיר בהוראת המקצוע.

אז אם אתם עסוקים מידי בעבודה, סובלים מלקויות למידה, רוצים
להצטיין או פשוט אוהבים ללמוד בשקט בבית, אנחנו מזמינים אתכם
לחויית לימודים יוצאת דופן וחדשה לחלוטין, היכנסו עכשיו לאתר
www.gool.co.il.



אנו מאחלים לכם הצלחה מלאה בבחינות

צוות האתר GooL

תוכן

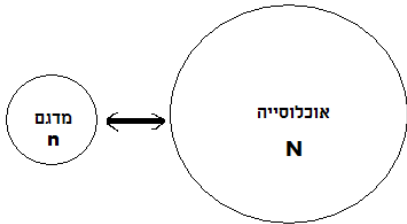
פרק 1 - הסקה סטטיסטית - הקדמה	3
פרק 2 - התפלגות הדגימה - ממוצע המדגם ומשפט הגבול המרכזי	6
פרק 3 - התפלגות הדגימה - התפלגות סכום תצפיות המדגם ומשפט הגבול המרכזי	15
פרק 4 - התפלגות הדגימה - התפלגות מספר ההצלחות במדגם - הקרוב הנורמלי להתפלגות הבינומית	19
פרק 5 - התפלגות הדגימה - התפלגות פרופורציית ההצלחות במדגם	24
פרק 6 - מושגים בסיסיים באמידה	29
פרק 7 - אמידה נקודתית	36
פרק 8 - ריוחי סמך	39
פרק 9 - בדיקת השערות - מבחנים פרמטרים	46
פרק 10 - טעויות ועוצמה בבדיקת השערות כללית	59
פרק 11 - מבחני חי בריבוע	62

פרק 1 - הסקה סטטיסטית - הקדמה

רקע:

אוכלוסייה – קבוצה שאליה מפנים שאלה מחקרית.

למשל, חברת תרופות שמעוניינת לפתח תרופה למחלת הסוכרת מתעניינת באוכלוסיית חולי הסוכרת בעולם.



מדגם – חלק מתוך האוכלוסייה.

למשל, אם נדגום באקראי 10 אנשים מתוך חולי הסוכרת אז זהו מדגם מתוך אוכלוסיית חולי הסוכרת.

במקרים רבים אין אפשרות לחקור את כל האוכלוסייה כיוון שאין גישה לכולה, היא גדולה מידי, או מוגבלים בזמן ובאמצעים טכניים ולכן מבצעים מדגם במטרה לבצע הסקה סטטיסטית מהמדגם לאוכלוסייה.

הדגימה בקורס תהייה דגימה מקרית הכוונה לדגימה שבה לכל תצפית באוכלוסייה יש את אותו סיכוי להיכלל במדגם.

סטטיסטי – גודל המחושב על המדגם.

פרמטר – גודל המתאר את האוכלוסייה.

הסימונים לפרמטר וסטטיסטי הם שונים

למשל:

סטטיסטי (מדגם)	פרמטר (אוכלוסייה)	
$\bar{X}$	μ	ממוצע
$\hat{p}$	P	פרופורציה (שכיחות יחסית)

פרמטר הוא גודל קבוע גם אם לא יודעים אותו סטטיסטי הוא משתנה ממדגם למדגם ולכן יש לו התפלגות הנקראת התפלגות הדגימה.

דוגמה (פתרון בהקלטה):

25% מאזרחי המדינה תומכים בהצעת החוק של חבר כנסת מסוים . הוחלט לדגום 200 אזרחים ומתוכם לבדוק מהו אחוז התומכים בהצעת החוק.

א. מי האוכלוסייה?

ב. מה המשתנה?

ג. מה הפרמטרים?

ד. מהו גודל המדגם?

ה. מהו הסטטיסטי שמתכננים להוציא מהמדגם?

ו. האם הפרמטר או הסטטיסטי הוא משתנה מקרי?

תרגילים :

1. מתוך כלל הסטודנטים במכללה שסיימו סטטיסטיקה א נדגמו שני סטודנטים. נתון שממוצע הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.
- מי האוכלוסייה?
 - מה המשתנה?
 - מהם הפרמטרים?
 - מהו גודל המדגם?
2. להלן התפלגות מספר מקלטי הטלויזיה למשפחה בישוב "העוגן".
נגדיר את x להיות מספר המקלטים של משפחה אקראית.
מתכננים לדגום מאוכלוסייה זו 4 משפחות ולהתבונן בממוצע מספר מקלטי הטלויזיה במדגם.

מספר המשפחות	מספר מקלטים
50	0
250	1
350	2
300	3
50	4
סך הכול $N = 1000$	

- מיהי האוכלוסייה ומהו המשתנה הנחקר?
 - מהו הסטטיסטי שיילקח מהמדגם ומה סימונו?
3. נתון כי 20% מהשכירים במדינה הם אקדמאיים. נבחרו באקראי 10 שכירים באותה אוכלוסייה ומתכננים לפרסם את מספר האקדמאיים שנדגמו.
- מהי האוכלוסייה?
 - מה המשתנה באוכלוסייה?
 - מהם הפרמטרים?
 - מהו הסטטיסטי?

פרק 2 - התפלגות הדגימה - ממוצע המדגם ומשפט הגבול המרכזי

רקע:

בפרק זה נדון בהתפלגות של ממוצע המדגם : $\bar{x} = \frac{\sum x_i}{n}$

מכיוון שממדגם למדגם אנו יכולים לקבל ממוצע מדגם שונה, אזי ממוצע המדגם הוא משתנה מקרי ויש לו התפלגות.

גדלים המתארים התפלגות כלשהי או אוכלוסייה כלשהי נקראים פרמטרים. להלן רשימה של פרמטרים החשובים לפרק זה:

ממוצע האוכלוסייה נסמן ב μ (נקרא גם תוחלת).

שונות אוכלוסייה נסמן ב- σ^2 .

סטיית תקן של אוכלוסייה: σ .

א. תכונות התפלגות

ממוצע כל ממוצעי המדגם האפשריים שווה לממוצע האוכלוסייה:

$$E(\bar{x}) = \mu_{\bar{x}} = \mu$$

שונות כל ממוצעי המדגם האפשריים שווה לשונות האוכלוסייה מחולק ב- n . תכונה זו נכונה רק במדגם מקרי:

$$V(\bar{x}) = \sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

יש יחס הפוך בין גודל המדגם לבין שונות ממוצעי המדגם.

אם נוציא שורש לשונות נקבל סטיית תקן של ממוצע המדגם שנקראת גם טעות תקן:

$$\sigma(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$

דוגמה: (פתרון בהקלטה)

השכר הממוצע במשק הינו 9000 ₪ עם סטיית תקן של 4000. דגמו באקראי 25 עובדים.

א. מי אוכלוסיית המחקר? מהו המשתנה הנחקר?

ב. מהם הפרמטרים של האוכלוסייה?

ג. מה התוחלת ומהי סטית התקן של ממוצע המדגם?

ב. דגימה מהתפלגות נורמאלית

אם נדגום מתוך אוכלוסייה שהמשתנה בה מתפלג נורמאלית עם ממוצע μ ושונות σ^2 ממוצע המדגם גם יתפלג נורמאלית:

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$Z_{\bar{x}} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

דוגמה: (פתרון בהקלטה)

משקל תינוק ביום היוולדו מתפלג נורמאלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם. מה ההסתברות שבמדגם של 4 תינוקות אקראיים בעת הולדתם המשקל הממוצע של התינוקות יהיה מתחת ל-3.5 ק"ג?

ג. משפט הגבול המרכזי

אם אוכלוסייה מתפלגת כלשהו עם ממוצע μ ושונות σ^2 אזי עבור מדגם מספיק גדול ($n \geq 30$)

$$\bar{x} \rightsquigarrow N\left(\mu, \frac{\sigma^2}{n}\right) \text{ ממוצע המדגם מתפלג בקירוב נורמאלית}$$

דוגמה: (פתרון בהקלטה)

משקל חפיסת שוקולד בקו ייצור מתפלג עם ממוצע 100 גרם וסטיית תקן של 4 גרם. דגמו מקו הייצור 36 חפיסות שוקולד אקראיות. מה ההסתברות שהמשקל הממוצע של חפיסות השוקולד שנדגמו יהיה מתחת ל 102 גרם?

תרגילים :

1. מתוך כלל הסטודנטים במכללה שסיימו סטטיסטיקה א נדגמו שני סטודנטים. נתון שממוצע הציונים של כלל הסטודנטים היה 78 עם סטיית תקן של 15.

- א. מי האוכלוסייה?
- ב. מה המשתנה?
- ג. מהם הפרמטרים?
- ד. מהו גודל המדגם?
- ה. מהו תוחלת ממוצע המדגם?
- ו. מהי טעות התקן?

2. להלן התפלגות מספר מקלטי הטלויזיה למשפחה בישוב מסוים :

מספר המשפחות	מספר מקלטים
500	0
2500	1
3500	2
3000	3
500	4
סך הכול $N = 10000$	

נגדיר את x להיות מספר המקלטים של משפחה אקראית.

- א. בנו את פונקציית ההסתברות של x .
- ב. חשבו את התוחלת, השונות וסטיית התקן של x .
- ג. אם נדגום 4 משפחות מהישוב עם החזרה מה תהיה התוחלת, מהי השונות ומהי סטיית התקן של ממוצע המדגם?

3. אם נטיל קובייה פעמיים ונתבונן בממוצע התוצאות שיתקבלו, מה תהיה התוחלת ומה תהיה סטיית התקן של ממוצע זה?

4. משקל תינוק ביום היוולדו מתפלג נורמאלית עם ממוצע 3400 גרם וסטיית תקן של 400 גרם
א. מה ההסתברות שתינוק אקראי בעת הלידה ישקול פחות מ-3800 גרם?

נתון כי ביום מסוים נולדו 4 תינוקות.

- ב. מה ההסתברות שהמשקל הממוצע שלהם יעלה על 4 ק"ג?
- ג. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה מתחת ל-2.5 ק"ג?
- ד. מה ההסתברות שהמשקל הממוצע של התינוקות יהיה רחוק מהתוחלת בלא יותר מ-50 גרם?
- ה. הסבירו ללא חישוב כיצד התשובה לסעיף הקודם הייתה משתנה אם היה מדובר על יותר מ-4 תינוקות?

5. הגובה של המתגייסים לצה"ל מתפלג נורמאלית עם תוחלת של 175 ס"מ וסטיית תקן של 10 ס"מ. ביום מסוים התגייסו 16 חיילים.

- א. מה ההסתברות שהגובה הממוצע שלהם יהיה לפחות 190 ס"מ?
 - ב. מה ההסתברות שהגובה הממוצע שלהם יהיה בדיוק 180 ס"מ?
 - ג. מה ההסתברות שהגובה הממוצע שלהם יסטה מתוחלת הגבהים בפחות מ-5 ס"מ?
 - ד. מהו הגובה שבהסתברות של 90% הגובה הממוצע של המדגם יהיה נמוך ממנו?
6. הזמן הממוצע שלוקח לאדם להגיע לעבודתו 30 דקות עם שונות של 16 דקות רבועות. האדם נוסע לעבודה במשך שבוע 5 פעמים. לצורך פתרון הניחו שזמן הנסיעה לעבודה מתפלג נורמאלית.
- א. מה ההסתברות שבמשך שבוע משך הנסיעה הממוצע יהיה מעל 33 דקות?
 - ב. מהו הזמן שבהסתברות של 90% ממוצע משך הנסיעה השבועי יהיה גבוה ממנו?
 - ג. מה ההסתברות שממוצע משך הנסיעה השבועי יהיה מרוחק מ-30 דקות בלפחות 2 דקות?
 - ד. כיצד התשובה לסעיף הקודם הייתה משתנה אם האדם היה נוסע לעבודה 6 פעמים בשבוע?

7. נפח היין בבקבוק מתפלג נורמאלית עם תוחלת של 750 סמ"ק וסטיית תקן של 10 סמ"ק.

- א. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה בדיוק 755 סמ"ק?
- ב. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה יותר מ-755 סמ"ק?
- ג. בארגז 4 בקבוקי יין. מה ההסתברות שהנפח הממוצע של הבקבוקים בארגז יהיה לפחות 755 סמ"ק?
- ד. בקבוקי היין שבארגז נמזגים לקערה עם קיבולת של שלושה ליטר. מה ההסתברות שהיין יגלוש מהקערה?

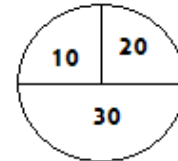
8. משתנה מתפלג נורמאלי עם תוחלת 80 וסטיית תקן 4 .

א. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה כאשר גודל המדגם הוא 9?

ב. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה שגודל המדגם הוא 16?

ג. הסבר את ההבדל בתשובות של שני הסעיפים.

9. בקזינו ישנה רולטה. על הרולטה רשומים המס' הבאים כמוראה בשרטוט :



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.

א. בנו את פונקציית ההסתברות של סכום הזכייה במשחק בודד.

ב. מה התוחלת ומה השונות של סכום הזכייה?

ג. אם האדם ישחק את המשחק 5 פעמים מה התוחלת ומה השונות של ממוצע סכום הזכייה בחמש המשחקים?

ד. אם האדם משחק את המשחק 50 פעם מה ההסתברות שבסה"כ יזכה ב-1050 ₪ ומעלה?

10. לפי הערכות הלשכה המרכזית לסטטיסטיקה השכר הממוצע במשק הוא 8000 ₪ עם סטיית תקן של 3000 ₪. מה ההסתברות שבמדגם מקרי של 100 עובדים השכר הממוצע יהיה יותר מ-8500 ₪?

11. מטילים קובייה 50 פעמים בכל פעם מתבוננים בתוצאה של הקובייה. מה ההסתברות שהממוצע של התוצאות יהיה לפחות 3.72 ב-50 ההטלות?

12. אורך צינור שמפעל מייצר הינו עם ממוצע של 70 ס"מ וסטיית תקן של 10 ס"מ .

א. נלקחו באקראי 100 מוטות, מה ההסתברות שממוצע אורך המוטות יהיה בין 68 ל 78 ס"מ?

ב. יש לחבר 2 בניינים באמצעות מוטות. המרחק בין שני הבניינים הינו 7200 ס"מ. מה ההסתברות ש 100 המוטות יספיקו למלאכה?

ג. מה צריך להיות גודל המדגם המינימאלי, כדי שבהסתברות של 5% ממוצע המדגם יהיה קטן מ-69 ס"מ. העזר במשפט הגבול המרכזי.

13. נתון משתנה מקרי בדיד בעל פונקציית ההסתברות הבאה :

2	4	6	8	X
$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	P(X)

מתוך התפלגות זו נלקח מדגם מקרי בגודל 50 . מה הסיכוי שממוצע המדגם יהיה קטן מ-5?

14. נתון ש $X \sim N(\mu, \sigma^2)$ דגמו 5 תצפיות מאותה התפלגות והתבוננו בממוצע המדגם $\bar{X}$:

לכן $P(\bar{X} > \mu)$ יהיה : (בחר בתשובה הנכונה)

- א. 0
- ב. 0.5
- ג. 1
- ד. לא ניתן לדעת.

15. נתון ש X מתפלג כלשהו עם תוחלת μ ושונות σ^2 .

החליטו לבצע מדגם בגודל 200 מתוך ההפלגות הנתונה לפי משפט הגבול המרכזי מתקיים ש :
(בחר בתשובה הנכונה)

א. $X \sim N(\mu, \frac{\sigma^2}{200})$

ב. $\mu \sim N(\mu, \frac{\sigma^2}{200})$

ג. $\bar{X} \sim N(\mu, \sigma^2)$

ד. $\bar{X} \sim N(\mu, \frac{\sigma^2}{200})$

16. נתון ש $X \sim N(\mu, \sigma^2)$. אם נדגום n תצפיות מתוך ההתפלגות ונגדיר $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ אזי :

(בחר בתשובה הנכונה)

- א. μ ו- $\bar{X}$ יהיו משתנים מקריים.
- ב. μ יהיה משתנה מקרי ו $\bar{X}$ קבוע.
- ג. $\bar{X}$ יהיה משתנה מקרי ו μ קבוע.
- ד. μ ו $\bar{X}$ יהיו קבועים.

17. משקל חפיסת שוקולד בקו ייצור מתפלג עם ממוצע 100 גרם . החפיסות נארזות בקרטון המכיל 36 חפיסות שוקולד אקראיות. ההסתברות שהמשקל הממוצע של חפיסות השוקולד בקרטון יהיה מעל 99 גרם הוא 0.9932.
 א. מהי סטיית התקן של משקל חפיסת שוקולד בודדת?
 ב. מה הסיכוי שמתוך 4 קרטונים בדיוק קרטון אחד יהיה עם משקל ממוצע לחפיסה הנמוך מ-100 גרם?

18. משתנה מקרי כלשהו מתפלג עם סטיית תקן של 20. מה הסיכוי שאם נדגום 100 תצפיות בלתי תלויות מאותה התפלגות אזי ממוצע המדגם יסטה מתוחלתו בפחות מ-2?

19. מספר המכונות הנכנסות לחניון "בציר" במשך היום מתפלג פואסונית עם קצב של מכונית אחת לדקה. שומר מסר נתונים על מספר המכוניות שנכנסות בכל שעה לגבי 40 שעות שאסף נתונים. מה ההסתברות שממוצע מספר המכוניות שנכנסו לחניון לשעה בשעות אלה יהיה לפחות 63?

20. הוכיחו שאם משתנה מתפלג כלשהו עם תוחלת μ ושונות σ^2 ומבצעים מדגם בגודל n של תצפיות בלתי תלויות מהמשתנה, אזי מתקיימות התכונות הבאות לגבי ממוצע המדגם:

$$E(\bar{x}) = \mu$$

$$V(\bar{x}) = \frac{\sigma^2}{n}$$

פתרונות:**שאלה 2**

א.

4	3	2	1	0	X
0.05	0.3	0.35	0.25	0.05	P(x)

$$\mu = 2.05 \quad \sigma^2 = 0.9475 \quad \sigma = 0.973$$

$$\mu_{\bar{x}} = 2.05 \quad \sigma_{\bar{x}}^2 = 0.2369$$

$$\sigma(\bar{X}) = 0.486$$

שאלה 3

$$\mu_x = 3.5$$

$$\sigma(\bar{X}) = 1.21$$

שאלה 4

$$0.8413 \quad \text{א.}$$

$$0.0013 \quad \text{ב.}$$

$$0 \quad \text{ג.}$$

$$0.1974 \quad \text{ד.}$$

שאלה 6

$$0.0465 \quad \text{א.}$$

$$27.71 \quad \text{ב.}$$

$$0.2628 \quad \text{ג.}$$

שאלה 7

$$0 \quad \text{א.}$$

$$0.1587 \quad \text{ב.}$$

$$0.1587 \quad \text{ג.}$$

$$0.5 \quad \text{ד.}$$

שאלה 8

א. 0.5468

ב. 0.6826

שאלה 9

א.

30	20	10	
0.5	0.25	0.25	$P(x)$

ב. התוחלת: 22.5

השונות: 68.75

ג. התוחלת: 22.5

השונות: 13.75

ד. 0.8997

שאלה 10

0.0475

שאלה 11

0.1814

שאלה 12

א. 0.9772

ב. 0.0228

ג. 271

שאלה 14

התשובה ב

שאלה 15

התשובה ד

שאלה 16

התשובה ג

שאלה 17

א. 2.429

ב. 0.25

פרק 3 - התפלגות הדגימה - התפלגות סכום תצפיות המדגם ומשפט הגבול המרכזי

רקע:

$$T = \sum_{i=1}^n X_i \text{ כעת נדון בסטטיסטי המבטא את סכום התצפיות במדגם}$$

כאשר כל התצפיות נדגמו באקראי מאותה אוכלוסייה.

כלומר, היו $X_1, \dots, X_n$ - משתנים מקריים בלתי תלויים בעלי התפלגות זהה שתוחלתה μ ושונותה σ^2 אזי:

א. התוחלת והשונות של סכום התצפיות:

$$E(T) = n\mu$$

$$V(T) = n\sigma^2$$

ב. דגימה מתוך התפלגות נורמלית:

$$T \sim N(n\mu, n\sigma^2)$$

$$Z = \frac{T - n\mu}{\sqrt{n\sigma^2}} \text{ אזי}$$

$$X \sim N(\mu, \sigma^2) \text{ אם}$$

ג. משפט הגבול המרכזי :

$$E(X) = \mu$$

$$V(X) = \sigma^2 \text{ אם } X \text{ מתפלג כלשהו וידוע}$$

אזי עבור מדגם מספיק גדול (לפחות 30)

$$T \sim N(n\mu, n\sigma^2)$$

דוגמה: (פתרון בהקלטה)

בעיר מסוימת המשכורת הממוצעת של עובד הינה 8000 ₪. עם סטיית תקן של 2000 ₪. נדגמו 100 עובדים מהעיר שמפקידים את משכורותיהם לסניף בנק.

- מה התוחלת וסטיית התקן של סך המשכורות שיופקדו לסניף הבנק על ידי העובדים הללו?
- מה ההסתברות שלסניף יופקד פחות מ-780 אלף ₪ ע"י אותם עובדים? (0.1587)

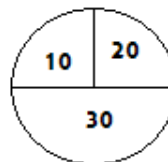
תרגילים:

1. המשקל באוכלוסייה מסוימת מתפלג נורמאלית עם תוחלת של 60 ק"ג וסטיית תקן של 10 ק"ג.
 א. מה הסיכוי שאדם אקראי מהאוכלוסייה ישקול מתחת ל-65 ק"ג?
 ב. מה הסיכוי שהמשקל הממוצע של 4 אנשים אקראיים יהיה מתחת ל-65 ק"ג?
 ג. מה הסיכוי שהמשקל הכולל של 4 אנשים אקראיים יהיה מתחת ל-240 ק"ג?
2. נפח יין בבקבוק מתפלג נורמאלית עם תוחלת של 750 מ"ל וסטיית תקן של 20 מ"ל. אדם קנה מארז של 4 בקבוקי יין.
 א. מהי התוחלת ומהי סטיית התקן של נפח היין במארז?
 ב. את היין שבמארז האדם מזג לכלי שקיבולתו 3.1 ליטר. מה ההסתברות שהיין יגלוש מהכלי?
 ג. אם לא היה נתון שנפח היין מתפלג נורמאלית. האם התשובה לסעיף א הייתה משתנה? האם התשובה לסעיף ב הייתה משתנה?
3. בספר כלשהו 500 עמודים. קצב הקריאה הממוצע הוא עמוד אחד ב 4 דקות עם סטיית תקן של 1 דקות.
 א. מה ההסתברות לסיים את הפרק הראשון (40 עמודים) תוך שעתיים וחצי?
 ב. מהו האחוזון ה-95 לזמן סיום קריאת הספר?
4. במגדל נבנו 40 יחידות דיור. כמו כן נבנו 135 מקומות חנייה לבניין. להלן פונקציית ההסתברות של מספר המכוניות ליחידת דיור:

x	1	2	3	4	5
$P(X = x)$	0.1	0.2	0.3	0.25	0.15

- נניח שמספר המכוניות ליחידת דיור בלתי תליות זו בזו ועם אותה פונקציית הסתברות לכל יחידת דיור (אין צורך בתיקון רציפות).
- א. מהי ההסתברות שיהיה מקום בחניון המגדל לכל מכוניות הבניין ?
 - ב. בהינתן ויש מקום במגדל לכל המכוניות , מה הסיכוי שבפועל מספר המכוניות נמוך מ-130?

5. בקזינו ישנה רולטה עליה מסומנים המספרים הבאים :



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה.

א. אם האדם משחק את המשחק 50 פעמים מה ההסתברות שבסך הכול יזכה בסכום של 1050

שקלים ומעלה?

ב. האדם מגיע בכל יום לקזינו ומשחק את המשחק 50 פעם עד אשר מגיע היום בו הוא יזכה

ב- 1050 שקלים ומעלה. מה התוחלת ומהי השונות של מספר הימים שיבלה בקזינו?

6. נתון ש $X_i \sim \exp(\lambda = 1)$ כאשר $i = 1, 2, \dots, 100$,

חשבו את הסיכוי $P(\sum_i X_i \geq 115)$.

7. אורך חיי סוללה בשעות הוא בעל פונקציית הצפיפות הבאה :

$$f(x) = 2x \quad 0 < x < 1$$

ברגע שסוללה מתרוקנת מחליפים אותה במידית בסוללה אחרת. כמה סוללות יש להחזיק במלאי אם רוצים שבסיכוי של 90% לפחות המלאי יספיק עבור 35 שעות לפחות?

פתרונות:**שאלה 1**

א. 0.6915

ב. 0.8413

ג. 0.5

שאלה 2

א. תוחלת 3000 מ"ל וסטיית תקן 40 מ"ל

ב. 0.0062

שאלה 4

א. 0.883

שאלה 5

א. 0.8997

ב. תוחלת : 1.111 שונות 0.1239

שאלה 7

56

פרק 4 - התפלגות הדגימה - התפלגות מספר ההצלחות במדגם - הקרוב הנורמלי להתפלגות הבינומית

רקע:

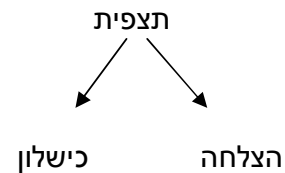
תזכורת על התפלגות בינומית

בפרק זה נדון על התפלגות מספר ההצלחות במדגם אקראי (תצפיות בלתי תלויות זו בזו).

מספר ההצלחות במדגם נסמן ב Y .

מחלקים כל תצפית במדגם להצלחה או כישלון.

כעת מה שמשתנה מתצפית לתצפית הוא משתנה דיכוטומי (משתנה שיש לו שני ערכים).



הסיכוי להצלחה יסומן עם הפרמטר p וכישלון יסומן ע"י הפרמטר $q = 1 - p$.

מבצעים מדגם אקראי בגודל n .

$$Y \sim B(n, p)$$

פונקציית ההסתברות של ההתפלגות הבינומית היא: $p(y = k) = \binom{n}{k} p^k q^{n-k}$

$$\text{תוחלת: } E(y) = np$$

$$\text{שונות: } V(y) = npq$$

קירוב נורמלי עבור התפלגות בינומית

אם לפנינו התפלגות בינומית : $Y \sim B(n, p)$ ומתקיים ש :

$$1. \quad n \cdot p \geq 5$$

$$2. \quad n \cdot (1 - p) \geq 5$$

$$y \sim N(np, npq)$$

$$Z_y = \frac{y - np}{\sqrt{npq}} \quad \text{אז :}$$

תיקון רציפות:

כאשר משתמשים בקירוב הנורמלי להתפלגות הבינומית יש לבצע תיקון רציפות .

הסיבה שעוברים כאן מהתפלגות בדידה להתפלגות נורמלית שהיא התפלגות רציפה.

על פי הכללים הבאים :

$$1. \quad p(Y = a) \cong p\left(a - \frac{1}{2} \leq Y \leq a + \frac{1}{2}\right)$$

$$2. \quad P(Y \leq a) \cong P(Y \leq a + 0.5)$$

$$3. \quad P(Y \geq a) \cong P(Y \geq a - 0.5)$$

הערות:

- התנאים למעבר מבינומי לנורמלי הם נזילים, כלומר משתנים ממרצה אחד לשני. התנאי שהצגתי כאן הוא הפופולרי ביותר:

$$1. \quad n \cdot p \geq 5$$

$$2. \quad n \cdot (1 - p) \geq 5$$

- ישנם מרצים שנותנים את התנאי המחמיר הבא:

$$1. \quad n \cdot p \geq 10$$

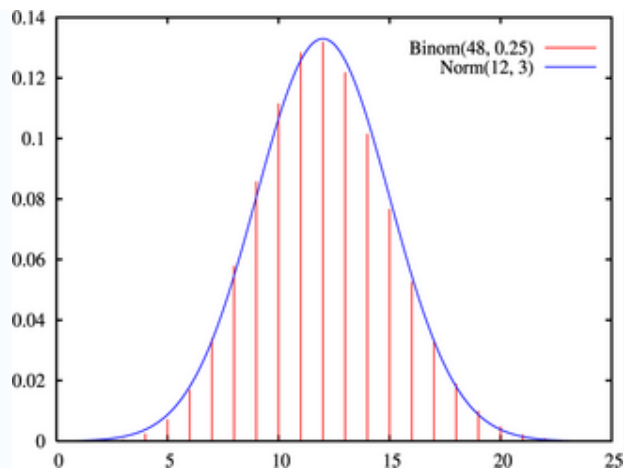
$$2. \quad n \cdot (1 - p) \geq 10$$

- וישנם מרצים שפשוט התנאי שהם נותנים הוא: $(n \geq 30)$.
- תאלצו לבדוק מהו התנאי שנתנו לכם בכיתה כדי לעבור מהתפלגות בינומית לנורמלית.
- הערה נוספת היא לגבי תיקון רציפות. ישנם מרצים שלא מחייבים לבצע תיקון רציפות שהמדגמים גדולים (בדרך כלל מעל 100 תצפיות) אני בפתרונות שאציג תמיד אבצע תיקון רציפות במעבר מבינומי לנורמלי כיוון שכך הפתרון יהיה יותר מדויק (בכל מקרה שהמדגמים גדולים העניין זניח).

דוגמה: (הפתרון בהקלטה)

נתון שבקרב אוכלוסיית הנוער 25% זקוקים למשקפיים. נדגמו באקראי 48 בני נוער.

- מה הסיכוי שבדיוק 14 מתוכם יהיו זקוקים למשקפיים?
- מה הסיכוי שלכל היותר 13 מתוכם זקוקים למשקפיים?



תרגילים:

1. נתון ש-20% מאוכלוסייה מסוימת אקדמאית. נבחרו באקראי 10 אנשים באותה אוכלוסייה.
 - א. מה ההסתברות ששלושה מהם אקדמאים?
 - ב. מה ההסתברות שלכל היותר אחד מהם אקדמאי?
 - ג. מה התוחלת ומהי סטיית התקן של מספר האקדמאים במדגם?
2. במפעל 10% מהמוצרים פגומים. נלקחו 100 מוצרים באקראי מקו הייצור.
 - א. מה ההסתברות שנדגמו לפחות 6 מוצרים פגומים?
 - ב. מה ההסתברות שמספר המוצרים הפגומים יהיה לכל היותר 11 במדגם?
3. ציוני פסיכומטרי בקרב הנרשמים למוסד מסוים מתפלגים נורמאלית עם ממוצע 500 וסטיית תקן 100. למוסד מסוים הוחלט לקבל אך ורק סטודנטים שקיבלנו מעל 600 בפסיכומטרי. 100 סטודנטים אקראיים נרשמו למוסד. מה ההסתברות שלפחות 20 יתקבלו?
4. מטילים מטבע 50 פעמים.
 - א. מה ההסתברות לקבל לכל היותר 30 עצים?
 - ב. מה ההסתברות לקבל 28 עצים לפי התפלגות הבינומית ולפי הקירוב הנורמאלי?
5. במטוס מקום ל-400 נוסעים. נרשמו לטיסה 430 אנשים (overbooking). מנתונים סטטיסטיים ידוע שהסיכוי שאדם שנרשם לטיסה אכן יגיע הוא 0.9.
 - א. מה ההסתברות שלא יהיו מקומות ישיבה לכל האנשים שהגיעו לטיסה?
 - ב. מה צריך להיות גודל המטוס כדי שבסיכוי שלפחות 95% המטוס יספיק לכמות הנרשמים?
6. מפעל לייצור ארטיקים טוען ש הסיכוי שארטיק שהוא מייצר יהיה פגום הוא 0.01. מוכר הזמין 1000 ארטיקים מהמפעל. מה ההסתברות שהמוכר יקבל לפחות 980 ארטיקים תקינים אם טענת המפעל מוצדקת?
7. מהמר מטיל קובייה הוגנת 100 פעמים. בכל הטלה, אם מתקבל תוצאה זוגית בקובייה המהמר זוכה בשקל. אחרת, המהמר משלם שקל. המהמר הטיל את הקובייה 100 פעמים מה הסיכוי שהרווח של המהמר יהיה לכל היותר 10?

פתרונות :**שאלה 1**

א. 0.201

ב. 0.3758

ג. התוחלת : 2, סטיית התקן : 1.2649

שאלה 2

א. 0.9332

ב. 0.6915

שאלה 3

0.1611

שאלה 4

א. 0.9406

שאלה 5

א. 0.015

שאלה 6

0.8544

שאלה 7

0.8643

פרק 5 - התפלגות הדגימה - התפלגות פרופורציית ההצלחות במדגם

רקע:

בפרק זה נדון על התפלגות הדגימה של פרופורציית המדגם.

Y - מספר ההצלחות במדגם (למשל, מספר המובטלים במדגם)

$$\hat{p} = \frac{y}{n} \text{ - פרופורציית ההצלחות במדגם (למשל, שיעור המובטלים במדגם)}$$

למשל,

$$n = 200$$

מספר המובטלים : $Y = 20$

$$\hat{p} = \frac{20}{200} = 0.1 \text{ פרופורציית המובטלים במדגם}$$

נסמן ב- p את שיעור ההצלחה באוכלוסייה וב- q את שיעור הכישלונות באוכלוסייה.

נבצע מדגם מקרי (הנחה שהתצפיות בלתי תלויות זו בזו) ונתבונן בהתפלגות של פרופורציית המדגם.

התוחלת, השונות וסטיית התקן של פרופורציית המדגם:

$$E(\hat{P}) = p$$

$$V(\hat{P}) = \frac{pq}{n}$$

$$\sigma(\hat{p}) = \sqrt{\frac{pq}{n}}$$

משפט הגבול המרכזי עבור הפרופורציה המדגמית :

$$\hat{p} \sim N\left(p, \frac{pq}{n}\right) \text{ אם } np \geq 5 \text{ \& } nq \geq 5$$

$$Z_{\hat{p}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

הערות:

- התנאים לקרוב הנורמאלי הם נזילים, כלומר משתנים ממרצה אחד לשני. התנאי שהצגתי כאן הוא הפופולרי ביותר:

$$1. \quad n \cdot p \geq 5$$

$$2. \quad n \cdot (1 - p) \geq 5$$

- ישנם מרצים שנותנים את התנאי המחמיר הבא:

$$1. \quad n \cdot p \geq 10$$

$$2. \quad n \cdot (1 - p) \geq 10$$

- וישנם מרצים המשתמשים בתנאי: $(n \geq 30)$.

- תאלצו לבדוק מהו התנאי שנתנו לכם בכיתה כדי לעבור לנורמלית.

- כיוון שפרופורציה אינה חייבת להיות מספר שלם בהכרח לא נהוג לבצע כאן תיקון רציפות.

דוגמה: (פתרון בהקלטה)

לפי נתוני משרד החינוך בעיר ירושלים ל-60% מתלמידי התיכון זכאים לתעודת בגרות. נדגמו 200 תלמידי תיכון.

- מה ההסתברות שהשכיחות היחסית $\hat{p}$ של הזכאים לבגרות במדגם תעלה על 60%?
- מה ההסתברות שפרופורציית הזכאים לבגרות במדגם תעלה על 70%?

תרגילים:

1. במדינה מסוימת 10% מכלל האוכלוסייה הינם מובטלים. נדגמו באקראי 140 אנשים מהמדינה.
 - א. מה התוחלת ומהי השונות של פרופורציות המובטלים שנדגמו?
 - ב. מה ההסתברות שבמדגם לפחות 10% יהיו מובטלים?
 - ג. מה ההסתברות שלכל היותר 9% מהמדגם יהיו מובטלים?

2. נניח כי 30% מהאוכלוסייה תומכים בהצעת חוק מסוימת. אם נדגום מהאוכלוסייה 200 איש.
 - חשבו את ההסתברויות הבאות:
 - א. לפחות 35% יתמכו בהצעת החוק במדגם.
 - ב. לכל היותר 25% יתמכו בהצעת החוק במדגם.
 - ג. יותר מ – 27% יתמכו בהצעת החוק במדגם.

3. לפי נתוני משרד התקשורת 40% מהאוכלוסייה מחזיקים בטלפון נייד מסוג "סמארטפון". נדגמו 400 אנשים מהאוכלוסייה.
 - א. מה ההסתברות שבמדגם לכל היותר ל 40% יש סמארטפון?
 - ב. מה ההסתברות שבמדגם לרוב יש סמארטפון?
 - ג. מה ההסתברות שפרופורציית בעלי הסמארטפון במדגם תסטה מהפרופורציה באוכלוסייה בלא יותר מ-4%?
 - ד. כיצד התשובה לסעיף הקודם הייתה משתנה אם הינו מגדילים את גודל המדגם?

4. נתון כי 80% מבתי האב מחוברים לאינטרנט. נדגמו 400 בתי אב אקראיים.
 - א. מה ההסתברות שלפחות 340 מהם מחוברים לאינטרנט?
 - ב. מה ההסתברות שפרופורציית המחוברים לאינטרנט במדגם תסטה מהפרופורציה האמתית ביותר מ-4%?
 - ג. כמה בתי אב יש לדגום כדי שהסטייה בין הפרופורציה המדגמית לפרופורציה האמתית לא תעלה על 3% בהסתברות של 90%?
 - ד. מהו העשירון התחתון של התפלגות פרופורציית המדגם?

5. נתון שציוני פסיכומטרי מתפלגים נורמלית עם תוחלת 500 וסטיית תקן 100. ל"מועדון ה-700" נכללים נבחנים שמקבלים ציון מעל 700 בפסיכומטרי. מה הסיכוי שבמועד בו נבחנו 2000 נבחנים אקראיים יהיו לפחות 3% המשתייכים למועדון?

6. נתון ש $X \sim B(n, p)$ נגדיר את המשתנה הבא : $\hat{P} = \frac{X}{n}$.

א. הוכיחו ש :

$$E(\hat{P}) = p$$

$$V(\hat{P}) = \frac{p(1-p)}{n}$$

ב. מה p המביא את $V(\hat{P})$ להיות במקסימום?

פתרונות:**שאלה 1**

א. התוחלת: 0.1, השונות: 0.00064

ב. 0.5

ג. 0.3446

שאלה 2

א. 0.0618

ב. 0.0618

ג. 0.8238

שאלה 3

א. 0.5

ב. 0

ג. 0.8968

ד. גדלה

שאלה 4

א. 0.0062

ב. 0.0456

פרק 6 - מושגים בסיסיים באמידה

רקע:

כזכור מהמפגש הקודם פרמטר הוא גודל המתאר את האוכלוסייה או התפלגות מסוימת.

כמו ממוצע הגבהים בקרב מתגייסים לצה"ל- μ .

כמו פרופורציית התומכים בממשלה בקרב אזרחי המדינה - p .

בדרך כלל הפרמטרים הם גדלים שאינם ידועים באמת, ולכן מבצעים מדגמים במטרה לאמוד אותם. אין אפשרות לחשב אותם הניסיון הוא בלהעריך כמה הם שווים ככל שניתן.

• נסמן באופן כללי פרמטר באות θ ואומד ב- $\hat{\theta}$. הוא סטטיסטי המחושב על המדגם ובאמצעותו נאמוד את θ .

• שגיאת אמידה: $|\hat{\theta} - \theta|$ - ההפרש בין האומד לאמת(הפרמטר).

דוגמה: (פתרון בהקלטה)

בכנסת ה-19 קיבלה מפלגת העבודה 15 מנדטים. בערוץ 10 ברגע סגירת הקלפיות העריכו את מספר המנדטים של המפלגה להיות 17 מנדטים וזאת על סמך תוצאות מדגם של הערוץ.

מה הפרמטר בדוגמה זו?

מהי טעות האמידה של ערוץ 10?

- $E(\hat{\theta}) = \theta$: יהיה אומד חסר הטיה ל θ אם התוחלת של $\hat{\theta}$ תהיה שווה ל θ
- $\sigma(\hat{\theta}) = S.E$: כלומר, טעות התקן של אומד היא סטיית התקן שלו,

להלן פרמטרים מרכזיים והאומדים שלהם:

ממוצע האוכלוסייה: μ

האומד הנקודתי שלו יהיה: ממוצע המדגם $\bar{x} = \frac{\sum x}{n}$

$E(\bar{x}) = \mu$ לכן $\bar{x}$ הינו אומד חסר הטיות ל μ .

כמו כן טעות תקן: $\sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}} = SE$

פרופורציה באוכלוסייה: p

האומד הנקודתי שלו יהיה: פרופורציה במדגם: $\hat{p} = \frac{y}{n}$

$E(\hat{p}) = p$ לכן $\hat{p}$ הינו אומד חסר הטיות ל p .

כמו כן טעות התקן: $\sigma(\hat{p}) = \sqrt{\frac{p \cdot (1-p)}{n}}$

שונות האוכלוסייה: σ^2

האומד הנקודתי שלו יהיה: $S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$

$E(S^2) = \sigma^2$ ולכן S^2 הינו אומד חסר הטיות ל σ^2 .

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$$

הערה: אומד הוא הנוסחה הכללית לאמידת הפרמטר ואומדן הוא הערך הספציפי שהתקבל במדגם מסוים.

דוגמה: (פתרון בהקלטה)

נדגמו 10 משפחות בתל אביב ונבדק עבור כל משפחה מספר הילדים שלה. להלן התוצאות שהתקבלו :

2,1,3,2,1,4,5,2,1,3

אמדו באמצעות אומדים חסרי הטיה את הפרמטרים הבאים :

1. ממוצע מספר הילדים למשפחה בתל אביב.
2. שונות מספר הילדים למשפחה בתל אביב.
3. פרופורציית המשפחות בנות שני ילדים.

תרגילים:

1. מתוך 500 טירונים נמצאו 120 בעלי שברי הליכה. נתון שהסיכוי שטירון יהיה עם שבר הליכה הוא 0.25.
 - א. מהי האוכלוסייה המוצגת בשאלה? מהם הפרמטרים שלה?
 - ב. מהי טעות התקן של האומדן כשהמדגם בגודל 500?
 - ג. מהו האומדן לפרמטר?
 - ד. מהי טעות האמידה?

2. לפי נתוני היצרן מקרר צורך בממוצע 2400 וואט לשעה עם סטיית תקן של 500 וואט לשעה. במדגם של 25 מקררים של היצרן התקבל ממוצע של 2342 וואט לשעה.
 - א. מהי האוכלוסייה המוצגת בשאלה? מהם הפרמטרים שלה?
 - ב. מהי טעות התקן של האומדן?
 - ג. מהו האומדן לפרמטר?
 - ד. מהי טעות האמידה?

3. נדגמו עשרה מתגייסים לצה"ל. גובהם נמדד בס"מ. להלן התוצאות שהתקבלו:

$$175, 168, 184, 192, 171, 180, 177, 187, 168, 177$$
 - א. מצא אומדן חסר הטיה לגובה הממוצע של מתגייסי צה"ל.
 - ב. מצא אומדן חסר הטיה לשונות הגבהים של מתגייסי צה"ל.
 - ג. מצא אומדן חסר הטיה לפרופורציות המתגייסים בגובה של לפחות 180 ס"מ.

4. נדגמו 20 שכירים באקראי. עבור כל שכיר נמדד השכר באלפי שקלים. להלן התוצאות שהתקבלו:

$$\sum_{i=1}^{20} X_i^2 = 1502.2 \quad \sum_{i=1}^{20} X_i = 162$$
 - א. אמדו את השכר הממוצע של השכירים במשק.
 - ב. אמדו את סטיית התקן של שכר השכירים במשק.

5. במטרה לאמוד את ממוצע האוכלוסייה. דגמו תצפיות בלתי תלויות מהאוכלוסייה וחישבו את

הממוצע שלהם. מהי טעות התקן?

א. סטיית התקן של האוכלוסייה.

ב. סטיית התקן של ממוצע האוכלוסייה.

ג. סטיית התקן של המדגם.

ד. סטיית התקן של ממוצע המדגם.

6. משקל הממוצע של אוכלוסייה מסוימת הוא 75 ק"ג עם שונות של 25. אם יבחרו כל המדגמים

האפשריים בגודל 10 מאוכלוסייה זו סטיית התקן של ממוצעי המדגמים תהיה:

א. 3

ב. 2.5

ג. 1.581

ד. אין מספיק נתונים לדעת.

7. במדגם מקרי, מתי סכום ריבועי הסטיות מהממוצע, $\sum_{i=1}^n (x_i - \bar{x})^2$, מחולק ב- $n-1$?

א. כאשר n קטן.

ב. כאשר תצפיות המדגם אינן בלתי תלויות.

ג. כאשר האוכלוסייה אינה מתפלגת נורמאלית.

ד. כאשר מעוניינים באומדן חסר הטיה לשונות האוכלוסייה ממנה הוצא המדגם.

ה. כאשר מעוניינים לחשב את שונות התפלגות הדגימה של ממוצע המדגם.

8. $X_1, X_2, \dots, X_{16}$ מדגם מקרי מתוך אוכלוסייה בעלת ממוצע μ לא ידוע ושונות

$\sigma^2 = 64$. טעות התקן של האומדן ל- μ היא:

א. 16

ב. 8

ג. 4

ד. 2

9. מהו אומד חסר הטיה?

- א. אומד שערכו שווה לממוצע התפלגות הדגימה שלו.
- ב. אומד שערכו שווה לערך הפרמטר באוכלוסייה.
- ג. אומד שממוצע התפלגות הדגימה שלו שווה לערך הפרמטר באוכלוסייה.
- ד. אומד שהסיכוי שערכו יהיה גבוה מערך הפרמטר באוכלוסייה שווה לסיכוי שיהיה נמוך ממנו.

פתרונות:**שאלה 3**

א. 177.9

ב. 64.1

ג. 0.4

שאלה 4

א. 8.1

ב. 3.16

שאלה 5

התשובה היא ד.

שאלה 6

התשובה היא ג.

שאלה 7

התשובה היא ד.

שאלה 8

התשובה היא ד.

שאלה 9

התשובה היא ג.

פרק 7 - אמידה נקודתית

אומדים חסרי הטיה

1. הציון במבחן מסוים של תלמידי כיתה ח' הנו משתנה מקרי בעל תוחלת μ וסטיית תקן 10. כדי לאמוד את התוחלת - μ , נלקח מדגם של 5 ציונים $X_1, \dots, X_5$. שלושה חוקרים הציעו אומדים לתוחלת על סמך מדגם זה:

$$T_1 = \frac{X_1 + \dots + X_5}{5} \quad \text{חוקר א' הציע:}$$

$$T_2 = \frac{2X_1 - X_3 + X_4}{2} \quad \text{חוקר ב' הציע:}$$

$$T_3 = \frac{2X_1 + X_3}{2} \quad \text{חוקר ג' הציע:}$$

- (א) איזה מן האומדים הוא חסר הטיה?
 (ב) במדגם התקבלו הציונים הבאים: 65, 78, 58, 82, 100. חשבו את האומדנים המתקבלים עבור האומדים חסרי הטיה.
 (ג) איזה מבין שני האומדים חסרי ההטיה עדיף? נמקו.

2. כדי לאמוד את המשקל הממוצע של הנשים בארה"ב, נבחר מדגם של $2n$ נשים. נסמן את שונות הגובה ב- σ^2 . הוצעו שני אומדים לממוצע המשקל על סמך מדגם זה:

$$T_1 = \frac{1}{n} \sum_{i=1}^n X_i \quad T_2 = \frac{1}{2n} \sum_{i=1}^{2n} X_i$$

- (א) בדקו לגבי כל אומד אם הוא בלתי מוטה.
 (ב) איזה אומד עדיף? נמקו.
3. $X \approx B(n, p)$ כלומר X הינו משתנה מקרי המתפלג בינומית עם פרמטר P (סיכוי להצלחה בניסיון בודד) במדגם בגודל n .

- (א) פתחו אומד חסר הטיה ל- P .
 (ב) מהו אומד חסר הטיה לסיכוי לכישלון בניסיון בודד.
 (ג) מהו אומד חסר הטיה ל- $E(X)$.

4. בתיק מניות שתי מניות . מספר המניות שיעלו ביום מסוים הוא משתנה מקרי התלוי בפרמטר פרמטר לא ידוע θ , $0 \leq \theta \leq 2$.

פונקציית ההסתברות של X – מספר המניות שיעלו ביום מסוים :

$$P(X=0) = 1 - \frac{\theta}{2} \quad P(X=1) = \frac{\theta}{3} \quad P(X=2) = \frac{\theta}{6}$$

- א. מצאו אומד בלתי מוטה ל- θ שמתבסס על מספר המניות שיעלו ביום מסוים.
 ב. מצאו אומד בלתי מוטה ל- θ שמתבסס על מספר המניות שעלו ביום במשך שלושה ימים X_1, X_2, X_3 (לכל אחד מהם אותה התפלגות כנ"ל והם בלתי תלויים).

5. בקרב המטפלות בת"א מספר התינוקות שבטיפולן הוא משתנה מיקרי בעל התפלגות התלויה בפרמטר θ באופן הבא :

הסיכוי שמטפלת תטפל בתינוק אחת בלבד הוא 3θ ,
 הסיכוי שמטפלת תטפל ב-2 תינוקות הוא $1 - 4\theta$,
 הסיכוי שמטפלת תטפל ב-3 תינוקות הוא θ .
 במדגם מיקרי של 4 מטפלות מת"א, נמצא כי שתיים מהם מטפלות בתינוק אחד בלבד, אחת מהן בשתיים ואחת השלושה תינוקות.

- א. מצא אומד חסר הטיה לפרמטר θ על סמך תצפית בודדת.
 ב. מצאו אומד חסר הטיה לפרמטר θ על סמך 4 תצפיות .
 ג. מהו האומדן לפרמטר θ על סמך תוצאות המדגם .
 ד. מצאו אומד חסר הטיה לסיכוי שלמטפלת בת"א תטפל בתינוק בודד אחד.
 ה. מצאו אומדים חסרי הטיה לתוחלת ולשונות של מספר התינוקות בטיפול אצל מטפלת מת"א.

תשובות סופיות - אמידה נקודתית

שאלה 1

א. $T_2 T_1$

ב. $T_2 = 110 T_1 = 76.6$

ג. T_1

שאלה 4

א. $\frac{3x}{2}$

ב. $\frac{3\bar{x}}{2}$

שאלה 3

א. $\frac{x}{n}$

ב. $1 - \frac{x}{n}$

ג. x

שאלה 5

א. $1 - \frac{x}{2}$

ג. 0.125

ה. לשונות 0.917

פרק 8 - רווחי סמך

א. אמידה נקודתית (אומדים חסרי הטיה)

1. נדגמו עשרה מתגייסים לצה"ל. גובהם נמדד בס"מ. להלן התוצאות שהתקבלו :
 $168, 184, 192, 171, 180, 177, 187, 168, 177$ ו- 175 .
 ד. מצא אומדן חסר הטיה לגובה הממוצע של מתגייסי צה"ל.
 ה. מצא אומדן חסר הטיה לשונות הגבהים של מתגייסי צה"ל.
 ו. מצא אומדן חסר הטיה לפרופורציות המתגייסים בגובה של לפחות 180 ס"מ.

2. נדגמו 20 שכירים באקראי. עבור כל שכיר נמדד השכר באלפי שקלים. להלן

$$\sum_{i=1}^{20} X_i^2 = 1502.2 \quad \sum_{i=1}^{20} X_i = 162$$

התוצאות שהתקבלו :

- ג. אמדו את השכר הממוצע של השכירים במשק.
- ד. אמדו את סטיית התקן של שכר השכירים במשק.

ב. רווח סמך לתוחלת

ב.1. שונות האוכלוסיה ידועה

1. מעוניינים לאמוד את ממוצע אורך החיים של מכשיר. מנתוני היצרן ידוע שאורך החיים מתפלג נורמאלית עם סטיית תקן של 20 שעות. נדגמו 25 מכשירים ונמצא כי ממוצע אורך החיים שלהם היה 230 שעות.
 א. בנו רווח סמך ברמת סמך של 90% לאורך החיים הממוצע של מכשיר.
 ב. בנו רווח סמך ברמת סמך של 95% לאורך החיים הממוצע של מכשיר.
 ג. הסבר כיצד ומדוע השתנה רווח הסמך.
 ד. מהי טעות התקן של האומד?

2. בנו רווח סמך לממוצע הציונים של מבחן אינטליגנציה. ידוע שסטיית התקן היא 15 והמדגם מתבסס על 100 תצפיות. רווח הסמך שהתקבל הוא (99,105). שחזרו את :
 א. ממוצע המדגם.
 ב. שגיאת האמידה המקסימאלית.
 ג. רמת הסמך.

3. מעוניינים לאמוד את המשקל הממוצע של רכיב מסוים. ידוע שהמשקל מתפלג נורמאלי עם סטיית תקן 12.
- א. מהו גודל המדגם המינימאלי שיש לקחת אם מעוניינים ששגיאת האמידה המקסימאלית תהיה 2 גרם ברמת סמך של 95%.
- ב. ביצעו מדגם שאת גודלו מצאתם בסעיף א והתקבל ממוצע של 23 גרם. בנו רווח סמך ברמת סמך של 95% למשקל הממוצע של הרכיב.
4. זמן החלמה מאנגינה מתפלג עם סטיית תקן של יומיים. חברת תרופות מעוניינת לחקור אנטיביוטיקה חדשה שהיא פיתחה. במחקר השתתפו 60 אנשים שחלו באנגינה וקיבלו את האנטיביוטיקה החדשה. בממוצע הם החלימו לאחר 4 ימים.
- א. בנו רווח סמך לתוחלת זמן ההחלמה תחת האנטיביוטיקה החדשה ברמת סמך של 90%.
- ב. מה היה קורה לאורך רווח הסמך אם היה תקציב להגדלת גודל המדגם פי 4? הסבירו.
- ג. מה היה קורה לאורך רווח הסמך אם היינו בונים את רווח הסמך ברמת סמך גדולה יותר? הסבירו.

5. משתנה מקרי מתפלג נורמאלי עם סטיית תקן ידועה 12. מה צריך להיות גודל המדגם כדי לבנות רווח סמך ברמת סמך של 98% שאורכו לא יעלה על 2?

2.ב. שונות האוכלוסיה לא ידועה

1. זמן התגובה מתפלג נורמאלי. במטרה לאמוד את תוחלת זמן התגובה נדגמו 4 תצפיות. להלן התוצאות בשניות: 3, 4, 5, 6, 7, 4, 5, 2. בנו רווח סמך ברמת סמך של 95% לממוצע זמן התגובה באוכלוסיה.
2. ציוני מבחן אינטליגנציה מתפלגים נורמאלי. נדגמו 25 מבחנים והתקבל ממוצע ציונים 102 ו סטיית תקן מדגמית 13.
- א. בנו רווח סמך לממוצע הציונים באוכלוסיה ברמת ביטחון של 95%.
- ב. חזרו על סעיף א' אם סטיית התקן הנתונה היא סטיית התקן האמיתית של כלל הנבחנים.
- ג. הסבירו את ההבדלים בין שני הסעיפים הנ"ל.

3. נשקלו 60 תינוקות אשר נולדו בשבוע ה-40 של ההיריון. המשקל נמדד בקילוגרמים. להלן התוצאות שהתקבלו: $\sum_{i=1}^{60} X_i = 195$, $\sum_{i=1}^{60} X_i^2 = 643.19$. בנו רווח סמך ברמת סמך של 95% לתוחלת משקל תינוק ביום היוולדו.

ג. רווח סמך לפרופורציה

1. במטרה לאמוד את אחוז המובטלים במשק נדגמו 200 אזרחים. מתוכם התקבל ש 24 היו מובטלים.
 - א. בנו רווח סמך לאחוז המובטלים באוכלוסיה ברמת סמך של 95%.
 - ב. מהו האומד לטעות התקן?
 - ג. מה יקרה לאורך רווח הסמך אם הוא יבנה ברמת סמך של 99%?
2. נבנה רווח סמך לאחוז בעלי רישיון נהיגה באוכלוסיה (69%, 79%) הוא התבסס על מדגם שכלל 400 נשאלים.
 - א. לכמה נשאלים לא היה רישיון נהיגה במדגם?
 - ב. באיזו רמת סמך נבנה רווח הסמך?
3. מעוניינים לאמוד את הרייטינג של תוכנית מסוימת. האמידה צריכה להתבצע ברמת סמך של 90% ועם שגיאת אמידה שלא תעלה על 4%.
 - א. מהו גודל המדגם המינימאלי שיש לקחת?
 - ב. חזור לסעיף א' אם ידוע שאחוז הרייטינג של התוכנית לא אמור לעלות על 20%.
4. במדגם של 300 נשים בגילאי 35-40 נמצא ש-140 היו נשואות, 80 היו גרושות, 60 רווקות והיתר אלמנות.
 - א. מצאו רווח סמך ברמה של 90% לאחוז הגרושות באוכלוסיה.
 - ב. מצאו רווח סמך ברמה של 99% לאחוז הלא נשואות באוכלוסיה.
5. הממשלה אומדת מדי חודש את אחוז התמיכה בה. מהו גודל המדגם אשר יש לקחת אם דורשים שהאומדן לא יסטה מהאחוז האמיתי באוכלוסיה ביותר מ-3%, וזאת בביטחון של 95%?

ד. רווח סמך לשונות וסטיית תקן

1. זמן התגובה מתפלג נורמאלי. במטרה לאמוד את תוחלת זמן התגובה נדגמו 4 תצפיות. להלן התוצאות בשניות: 3.5, 4.6, 5.2, 4.7. בנו רווח סמך, ברמת סמך של 95% לשונות זמן התגובה באוכלוסיה.
2. נדגמו 20 ימים אקראיים מחודשי יולי-אוגוסט ונמדדה בהם הטמפ' במעלות צלזיוס בת"א. במדגם התקבל טמפ' ממוצעת 30.8 וסטיית תקן מדגמית 1.1. בהנחה והטמפ' מתפלגת נורמאלי:
 - א. בנו רווח סמך לתוחלת הטמפ' בחודשים אלה בצהרי ת"א ברמת סמך של 95%.
 - ב. בנו רווח סמך לסטיית התקן של הטמפ' בחודשים אלה בצהרי ת"א ברמת סמך של 95%.

ה. רווח סמך להפרשים

ה.1. הפרשי פרופורציות

1. במטרה להשוות בין שתי תרופות נדגמו 200 איש שלקחו תרופה x. מתוכם 180 טענו שהתרופה עזרה להם. כמו כן נלקחו 300 איש שלקחו את תרופה y. מתוכם 150 טענו שהתרופה עזרה להם. בנו רווח סמך להפרש אחוזי ההצלחה של התרופות ברמת סמך של 95%. מה ניתן לומר על סמך רווח הסמך על ההבדלים בין התרופות?
2. מתוך 150 נשים שנדגמו באקראי 30% תמכו בהצעת חוק מסוימת. מתוך 200 גברים שנדגמו באקראי 25% תמכו בהצעת החוק.
 - א. בנו רווח סמך לפער בין אחוזי התמיכה של הנשים לעומת הגברים ברמת סמך של 96%.
 - ב. בנו רווח סמך ברמת סמך של 95% לאחוז התמיכה בהצעת החוק.

ה. הפרשי תוחלות (שונות ידועות)

1. מעוניינים לבדוק האם קיים הבדל בין ממוצע ציוני הפסיכומטרי של חיילים לממוצע ציוני הפסיכומטרי של תלמידי תיכון. ידוע שציוני הפסיכומטרי מתפלגים נורמאלית עם סטיית תקן 100. במדגם של 16 נבחנים חיילים התקבל ממוצע 543. במדגם של 20 תלמידי תיכון התקבל ממוצע 508. בנו רווח סמך לפער תוחלות הציונים בין חיילים לתלמידי תיכון ברמת סמך של 90%. מה ניתן להסיק מרווח סמך זה?
2. בנק מתלבט האם לפתוח סניף באזור A או באזור B. לצורך פתרון נניח שסטית התקן של המשכורת באזור A היא 1200 ובאזור B 1500. הבנק דגם 50 אנשים מאזור A, המשכורת הממוצעת שהתקבלה במדגם היא 6,800 ₪. כמו כן נדגמו 40 אנשים מאזור B, המשכורת הממוצעת שהתקבלה במדגם היא 6,600 ₪.
 - א. בנו רווח סמך ברמת סמך של 95% להפרש הממוצעים של המשכורות בשני האזורים. האם על סמך רווח הסמך ניתן להמליץ לבנק היכן לפתוח את הסניף. אם כן, היכן?
 - ב. בנו רווח סמך לתוחלת המשכורת באזור A ברמת סמך של 95%.

תשובות סופיות - רווחי סמך

פרק א' - אמידה נקודתית

שאלה 2

א. 8.1

ב. 3.16

שאלה 1

ג. 177.9

ד. 64.1

ה. 0.4

פרק ב' - רווח סמך לתוחלת

שונות האוכלוסיה ידועה

<u>שאלה 2</u>	<u>שאלה 1</u>
א. 102	א. $223.42 < \mu < 236.58$
ב. 3	ב. $222.16 < \mu < 237.84$
ג. 0.9544	ד. 4
<u>שאלה 4</u>	<u>שאלה 3</u>
א. $3.58 < \mu < 4.42$	א. 139
ב. יקטן פי 2	ב. $21 < \mu < 25$
ג. גדל	
	<u>שאלה 5</u>
	780

שונות האוכלוסיה לא ידועה

<u>שאלה 2</u>	<u>שאלה 1</u>
א. $96.63 < \mu < 107.37$	$4.39 < \mu < 5.51$
ב. $96.90 < \mu < 107.10$	
	<u>שאלה 3</u>
	$3.149 < \mu < 3.351$

פרק ג' - רווח סמך לפרופורציה

<u>שאלה 1</u>	<u>שאלה 2</u>
א. $7.5\% < p < 16.5\%$	א. 104
ב. 2.29%	ב. 0.9774
ג. יגדל	
<u>שאלה 3</u>	<u>שאלה 4</u>
א. 423	א. $22.5\% < p < 30.9\%$
ב. 271	ב. $45.91\% < p < 60.72\%$
<u>שאלה 5</u>	
1068	

פרק ד' - רווח סמך לשונות וסטיית תקן

<u>שאלה 1</u>	<u>שאלה 2</u>
$0.039 < \sigma^2 < 1.708$	א. $30.285 < \mu < 31.315$
	ב. $0.837 < \sigma < 1.607$

פרק ה' - רווח סמך להפרשים**הפרשי פרופורציות**

<u>שאלה 1</u>	<u>שאלה 2</u>
(33%, 47%)	א. (4.93%, 14.93%)
	ב. (22.5%, 31.8%)

הפרשי תוחלות

<u>שאלה 1</u>	<u>שאלה 2</u>
(-20, 90)	א. (-372, 772)
	ב. (6467, 7133)

פרק 9 - בדיקת השערות - מבחנים פרמטרים

א. בדיקת השערות על תוחלת

שונות האוכלוסיה ידועה

1. ממוצע הציונים בבחינת הבגרות באנגלית הנו 72 עם סטיית תקן 15 נקודות. מורה טוען שפיתח שיטת לימוד חדשה שתעלה את ממוצע הציונים. משרד החינוך החליט לתת למורה 36 תלמידים אקראיים. ממוצע הציונים של אותם תלמידים לאחר שלמדו בשיטתו היה 75.5. בהנחה שגם בשיטתו סטיית התקן תהייה 15 מה מסקנתכם ברמת מובהקות של 5%?

2. לפי הצהרת היצרן של חברת משקאות מסוימת נפח הנוזל בבקבוק מתפלג נורמלית עם תוחלת 500 סמ"ק וסטיית תקן 20 סמ"ק. אגודת הצרכנים מתלוננת על הפחתת נפח המשקה בבקבוק מהכמות המוצהרת. במדגם שעשתה אגודת הצרכנים התקבל נפח ממוצע של 492 סמ"ק במדגם בגודל 25.

א. מה מסקנתכם ברמת מובהקות של 2.5%?

ב. האם ניתן לדעת מה תהיה המסקנה עבור רמת מובהקות הגבוהה מ-5%?

3. מהנדס האיכות מעוניין לבדוק אם מכונה מכוילת (מאופסת). המכונה כוונה לחתוך מוטות באורך 50 ס"מ. לפי נתוני היצרן סטיית התקן בחיתוך המוטות היא 0.5 ס"מ. במדגם של 50 מוטות התקבל ממוצע אורך המוט 50.93 ס"מ. מה מסקנתכם ברמת מובהקות של 5%?

שונות האוכלוסיה לא ידועה

1. משך זמן ההחלמה בלקיחת אנטיביוטיקה מסוימת הוא 120 שעות בממוצע עם סטיית תקן לא ידועה. מעוניינים לבדוק האם אנטיביוטיקה אחרת מקטינה את משך זמן ההחלמה. במדגם של 5 חולים שלקחו את האנטיביוטיקה האחרת התקבלו זמני ההחלמה הבאים: 90, 95, 100, 80, 125 שעות. מה מסקנתכם ברמת מובהקות של 5%. מהי ההנחה הדרושה לצורך הפתרון.

2. ציוני מבחן IQ מתפלגים נורמלית בארצות הברית עם ממוצע 100. במדגם על 20 נבחנים ישראלים התקבלו התוצאות הבאות.

$$\bar{x} = 102$$

$$\hat{S} = 16$$

האם בישראל ממוצע הציונים הוא כמו בארצות הברית? בדקו ברמת מובהקות של 2% הניחו הנחות אם יש צורך.

ב. בדיקת השערות על פרופורצייה

1. במשך שנים אחוז המועמדים שהתקבל לפקולטה מסוימת היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. ברמת מובהקות של 5% האם השנה הקשו על תנאי הקבלה?
2. במדגם של 300 אזרחים 57% מתנגדים להצעת חוק מסוימת. לאור נתונים אלה האם רוב האזרחים מתנגדים להצעת החוק? בדקו ברמת מובהקות של 10%.
3. הטילו מטבע 50 פעמים וקיבלו 28 פעמים עץ. האם המטבע הוגן ברמת מובהקות של 5%?

ג. בדיקת השערות להפרש פרופורציות

1. במדגם של 200 גברים. 8% מהם היו מובטלים. המדגם של 180 נשים 10% מהן היו מובטלות האם קיים הבדל מובהק בין פרופורציית המובטלים לפרופורציית המובטלות. בדוק ברמת מובהקות של 5%.
2. אחוז בעלי רישיון נהיגה בקרב האוכלוסייה הבוגרת הינו 60%. במדגם של 300 בוגרים מתל אביב 204 היו בעלי רישיון נהיגה. במדגם של 220 בוגרים מירושלים 100 היו בעלי רישיון נהיגה.
 - א. ברמת מובהקות של 5% האם תקבלו את הטענה שאחוז בעלי הרישיון בתל אביב גבוה מהאחוז הארצי?
 - ב. ברמת מובהקות של 10% האם תקבלו את הטענה שאחוז בעלי הרישיון נהיגה בתל אביב גבוה מאחוז בעלי רישיון הנהיגה בירושלים?

ד. בדיקת השערות להפרש תוחלות במדגמים בלתי תלויים

שוניות האוכלוסייה ידועות

1. ציוני פסיכומטרי מתפלגים נורמלית עם סטיית תקן 100. מכון ללימוד פסיכומטרי טוען שהוא יכול לשפר את ממוצע הציונים ביותר מ-30 נקודות. במדגם של 20 נבחנים שניגשו למבחן ללא הכנה במכון התקבל ממוצע 508. במדגם של 25 נבחנים שעברו הכנה במכון התקבל ממוצע ציונים 561. מה מסקנתכם ברמת מובהקות של 5%.

2. במדגם אקראי של 20 ימים נבדקה התפוקה של מפעל ביום. התפוקה הממוצעת הייתה של 340 מוצרים ליום. במדגם אקראי של 20 ימים אחרים נבדקה התפוקה של המפעל בלילה והתפוקה הממוצעת הייתה 295. לצורך פתרון נניח שסטיית התקן של התפוקה ביום היא 40 מוצרים ובלילה 30 מוצרים. האם התפוקה הממוצעת היומית גבוהה מהתפוקה הממוצעת הלילית בדקו ברמת מובהקות של 10%. מה ההנחה הדרושה לצורך פתרון?

שוניות האוכלוסייה לא ידועות (בהנחה שהן שוות)

1. במדגם של 10 ישראלים שנבחנו במבחן ה-IQ נתקבלו התוצאות הבאות:

$$\begin{aligned}n_1 &= 10 \\ \sum x_i &= 1020 \\ \sum x_i^2 &= 105120\end{aligned}$$

במדגם של 14 אמריקאים שנבחנו במבחן ה-IQ נתקבלו התוצאות הבאות:

$$\begin{aligned}n_1 &= 14 \\ \sum x_i &= 1386 \\ \sum x_i^2 &= 138644\end{aligned}$$

בדקו האם קיים הבדל בממוצע הציונים בבחינת ה-IQ בין ישראל לארה"ב. רשמו את כל ההנחות הדרושות. ברמת מובהקות של 5%.

2. להלן תוצאות מדגם הבדק אורך חיים של נורות מסוג W60 ומסוג W100. אורך החיים נמדד בשעות.

1-100W	2-60W	הקבוצה
956	1007	$\bar{x}$
72	80	$\hat{S}$
15	13	n

בדקו ברמת מובהקות של 5% האם נורות מסוג W60 דולקות בממוצע יותר מאשר נורות מסוג W100. רשמו את כל ההנחות הדרושות לפיתרון.

ה. בדיקת השערות על מדגמים מזווגים

1. במטרה לבדוק האם קיים הבדל בין חברת X לחברת Y מבחינת המחירים לשיחות בינ"ל. נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקת שיחה. להלן התוצאות:

המדינה	X	Y
ארה"ב	1.5	1.4
קנדה	2.1	2
הולנד	2.2	1.9
פולין	3	3.1
מצרים	3.5	3.2
סין	3.2	3.2
יפן	4.2	4.2

בהנחה והמחירים מתפלגים נורמלית בכל חברה, בדקו ברמת מובהקות של 5% האם קיים הבדל בין החברות מבחינת המחירים בממוצע.

2. מכון המכין לפסיכומטרי טוען שהוא מעלה את ממוצע הציונים ביותר מ-30 נקודות. 8 נבחנים נבדקו לפני ואחרי שהם למדו במכון. להלן התוצאות שהתקבלו:

לפני	506	470	420	640	670	390	500	590
אחרי	570	540	430	610	680	510	520	580

מה מסקנתכם ברמת מובהקות 5%? הניחו שציוני פסיכומטרי מתפלג נורמלית.

ו. בדיקת השערות על שונות

1. זמן ההחלמה ממחלה מסוימת כאשר משתמשים בטיפול מסוים מתפלג נורמלית עם סטיית תקן של 80 שעות. תרופה חדשה נוסתה על 5 חולים. זמני ההחלמה שלהם בשעות היו 38,72,90,110,50. ברמת מובהקות של 5% בדקו האם סטיית התקן של זמן החלמה של התרופה החדשה נמוכה מהתרופה המקורית.

2. הגובה של אוכלוסייה מסוימת נחשב כמתפלג נורמלית על ממוצע של 174 ס"מ וסטיית תקן 12. במדגם של 20 אנשים מהאוכלוסייה התקבל ממוצע 171 וסטיית תקן מדגמית 23.

- א. בדקו ברמת מובהקות של 5% האם חל שינוי בשונות הגבהים באוכלוסייה.
- ב. בדקו ברמת מובהקות של 5% האם חל שינוי בתוחלת הגבהים באוכלוסייה.

ז. בדיקת השערות על שתי שונויות

1. מעוניינים להשוות בין נשים וגברים מבחינת השונות בזמנים שלהם לבצע משימה מסוימת. במדגם של 10 גברים התקבלו התוצאות הבאות לגבי זמני ביצוע המשימה:

$$\sum (y_i - \bar{y})^2 = 204$$

במדגם של 13 נשים התקבלו התוצאות הבאות:

$$\sum (x_i - \bar{x})^2 = 200$$

בדקו ברמת מובהקות של 2% האם קיים הבדל בין השונויות? מה יש להניח?

2. ציוני בחינת הבגרות במתמטיקה מתפלגים נורמלית עם שונות 150. במדגם של 16 נבחנים מתל אביב התקבלה שונות חסרת הטיה- 190. במדגם של 25 ירושלמים התקבלה שונות חסרת הטיה- 118.

- א. בדקו ברמת מובהקות של 5% האם שונות הציונים במתמטיקה בקרב נבחני תל אביב גבוהה מהשונות בכלל הארץ.
- ב. בדקו ברמת מובהקות של 5% האם שונות ציונים במתמטיקה בקרב תלמידי תל אביב גבוהה מאשר בקרב תלמידי ירושלים.

ח. הגישה המעשית - מובהקות התוצאה - P-VALUE - $\alpha_{\min}$

1. השכר הממוצע במשק בשנת 2006 היה 7800 ₪ עם סטיית תקן 2000 ₪. במדגם כיום של מאה עובדים התקבל שכר ממוצע 7500 ₪. מטרת המחקר היא לבדוק האם כיום חלה שחיקה בשכר כיום. עבור אילו רמות מובהקות שיבחר החוקר יוחלט לקבל את טענת המחקר?

$$H_0 : \mu = 70$$

$$H_1 : \mu > 70$$

2. לפניך השערות של מחקר :

המשתנה הנחקר מתפלג נורמלית עם סטיית תקן 20. במדגם מאותה אוכלוסייה התקבלו התוצאות הבאות :

$$n = 100$$

$$\bar{x} = 74$$

מהי האלפא המנימלית לדחיית השערת האפס?

3. במשך שנים אחוז המועמדים שהתקבל לפקולטה מסוימת היה 25%. השנה מתוך מדגם של 120 מועמדים התקבלו 22. רוצים לבדוק האם השנה הקשו על תנאי הקבלה.
א. מהי מובהקות התוצאה?

ב. מה תהיה המסקנה ברמת מובהקות של 1% וברמת מובהקות של 5%?

4. במטרה לבדוק האם מטבע הוא הוגן הוא הוטל 100 פעמים. התקבל 58 פעמים "עץ". האם יתכן שחוקר אחד יחליט שהמטבע הוגן ואחר יחליט שהמטבע לא הוגן?

5. הגובה של מתגייסים לצה"ל מתפלג נורמלית. במדגם של 25 מתגייסים מדדו את הגבהים שלהם בס"מ והתקבלו התוצאות הבאות :

$$\bar{x} = 176.2$$

$$\sum (x_i - \bar{x})^2 = 2832$$

מטרת המחקר היא לבדוק האם תוחלת הגבהים של המתגייסים גבוהה מ-174 ס"מ באופן מובהק. מהי בקרוב מובהקות התוצאה ועל פיה מה תהיה המסקנה ברמת מובהקות של 6% ?

ט. טעויות ועוצמה

1. לפי נתוני משרד הפנים בשנת 1980 למשפחה ממוצעת היה 2.3 ילדים למשפחה עם סטיית תקן 0.4. מעוניינים לבדוק אם כיום ממוצע מספר הילדים למשפחה קטן יותר. לצורך כך הוחלט לדגום 121 משפחות. במדגם התקבל ממוצע 2.17 ילדים למשפחה.
א. רשמו כלל הכרעה במונחי ממוצע מדגם קריטי ברמת מובהקות של 5%.
ב. בהמשך לסעיף א מה תהיה המסקנה ומהי הטעות האפשרית במסקנה?
ג. אם באמת ממוצע מספר הילדים במשפחה פחת לכדי 2.1 מהי העוצמה של הכלל מסעיף א?

2. להלן נתונים על תהליך של בדיקת השערות על תוחלת:

$$H_0 : \mu = 200$$

$$H_1 : \mu \neq 200$$

$$\sigma = 30$$

$$n = 225$$

- א. רשום כלל הכרעה במונחי ממוצע מדגם קריטי וברמת מובהקות של 10%.
ב. בהמשך לסעיף א מהי העוצמה אם התוחלת שווה ל-195?
ג. הסבר ללא חישוב איך העוצמה תשתנה אם רמת המובהקות תהיה 5%?
3. אחוז הסובלים מתופעות הלוואי מתרופה מסוימת הוא 15%. חברת תרופות טוענת שפיתחה תרופה שאמורה לצמצם את אחוז הסובלים מתופעות לוואי. לצורך בדיקת הטענה הוחלט לבצע מחקר שיכלול 120 חולים שיקבלו את התרופה הנבדקת.
א. נניח שהתרופה נבדקת אכן מורידה את פרופורציות הסובלים מתופעות הלוואי ל-10%. מהי עצמת המבחן עבור רמת מובהקות של 5%?
ב. אם המדגם יתבסס על יותר תצפיות כיצד הדבר ישנה את התשובה לשל סעיף א'?
4. בעיר מסוימת היו 20% אקדמאים. בעקבות פתיחת מכללה בעיר לפני כמה שנים מעוניינים לבדוק האם אחוז האקדמאים גדל. מעוניינים שהמחקר יכלול 200 אנשים והוא יהיה ברמת מובהקות של 5%. חשבו את הסיכוי לבצע טעות מסוג שני בהנחה והיום יש 28% אקדמאים.

5. מפעל לייצור צינורות מייצר צינור שקוטרו מתפלג נורמלית עם תוחלת של 50 מ"מ וסטית תקן של 6 מ"מ. במחלקת ביקורת האיכות דוגמים בכל יום 81 צינורות ומודדים את קוטרם, בכדי לבדוק, בעזרת מבחן סטטיסטי, האם מכונת הייצור מכוילת כנדרש או שקוטר הצינורות קטן מהדרוש.
- א. רשום את ההשערות ואת ככל ההכרעה ברמת מובהקות של 5% .
- ב. אם ביום כלשהו מכונת הייצור התקלקלה והיא מייצרת את הצינורות בקוטר שתוחלתו 48 מ"מ בלבד (סטית התקן לא השתנתה), מה ההסתברות שהתקלה לא תתגלה בביקורת האיכות? כיצד נקראת הסתברות זו?
- ג. הסבר ללא חישוב כיצד התשובה לסעיף ב תשתנה אם רמת המובהקות תגדל.
- ד. הסבר ללא חישוב כיצד התשובה לסעיף ב תשתנה אם התוחלת האמיתית היא 47 ולא 48 מ"מ.

י. הקשר בין רווח סמך לבדיקת השערות

1. חוקר מעוניין לבדוק השפעת דיאטה חדשה על רמת הסוכר בדם. ידוע כי מספר מיליגרם הסוכר בסמ"ק דם הוא משתנה מקרי שמתפלג נורמלית עם סטיית תקן 10 מ"ג. נלקח מדגם של 100 נבדקים שניזונו מדיאטה זו. נמצא כי ממוצע מספר המיליגרם סוכר היא 112.6 מ"ג לסמ"ק.

א. בנה רווח סמך ברמת סמך 95% לתוחלת רמת הסוכר בדם אצל הניזונים מדיאטה זו.

ב. ידוע שתוחלת רמת הסוכר בדם באוכלוסיה היא 113 מ"ג לסמ"ק. האם לדעתך ניתן להסיק על סמך תוצאת סעיף א שהדיאטה משפיעה על רמת הסוכר בדם? הסבר.

2. חוקר רצה לבדוק את ההשערות הבאות :

$$H_0 : \mu = 90$$

$$H_1 : \mu \neq 90$$

החוקר בנה רווח סמך לתוחלת ברמת סמך של 95% וקיבל את רווח הסמך הבא: (87,97). אם החוקר מעוניין לבצע בדיקת השערות ברמת מובהקות של 1% האם ניתן להגיע למסקנה ע"ס רווח הסמך? נמקו.

3. במטרה לבדוק האם קיים הבדל בין חברת X לחברת Y מבחינת ממוצע המחירים לשיחות בינ"ל. נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקת שיחה. בהנחה והמחירים מתפלגים נורמלית בנו רווח סמך לממוצע ההפרשים שיבדוק את השערות המחקר ברמת מובהקות של 5%

להלן התוצאות :

המדינה	X	Y
ארה"ב	1.5	1.4
קנדה	2.1	2
הולנד	2.2	1.9
פולין	3	3.1
מצרים	3.5	3.2
סין	3.2	3.2
יפן	4.2	4.2

תשובות סופיות - בדיקת השערות

פרק א' - בדיקת השערות לתוחלת

שונות האוכלוסיה ידועה

<u>שאלה 1</u>	<u>שאלה 2</u>
נקבל H_0	נדחה H_0
<u>שאלה 3</u>	
נדחה H_0	

שונות האוכלוסיה לא ידועה

<u>שאלה 1</u>	<u>שאלה 2</u>
נדחה H_0	נקבל H_0

פרק ב' - בדיקת השערות על פרופורציה

<u>שאלה 1</u>	<u>שאלה 2</u>
נדחה H_0	נדחה H_0
<u>שאלה 3</u>	
נקבל H_0	

פרק ג' - בדיקת השערות להפרש פרופורציות

<u>שאלה 1</u>	<u>שאלה 2</u>
נקבל H_0	א. נדחה H_0 ב. נדחה H_0

פרק ד' - בדיקת השערות להפרש תוחלותשונויות האוכלוסייה ידועות

<u>שאלה 1</u>	<u>שאלה 2</u>
נקבל H_0	נדחה H_0

שונויות האוכלוסייה לא ידועות

<u>שאלה 1</u>	<u>שאלה 2</u>
נקבל H_0	נדחה H_0

פרק ה' - בדיקת השערות במדגמים מזווגים

<u>שאלה 1</u>	<u>שאלה 2</u>
נקבל H_0	נקבל H_0

פרק ו' - בדיקת השערות על שונות

<u>שאלה 1</u>	<u>שאלה 2</u>
נדחה H_0	א. נדחה H_0

פרק ז' - בדיקת השערות לשתי שונות

<u>שאלה 1</u>	<u>שאלה 2</u>
נקבל H_0	א. נקבל H_0
	ב. נקבל H_0

פרק ח' - מובהקות התוצאה

<u>שאלה 1</u>	<u>שאלה 2</u>
מעל 0.0668	0.0228
<u>שאלה 3</u>	<u>שאלה 5</u>
א. 0.0455	נקבל H_0

פרק ט' - טעויות ועוצמה

<u>שאלה 1</u>	<u>שאלה 2</u>
ב. נדחה H_0	ב. 0.8051
ג. 1	
<u>שאלה 3</u>	<u>שאלה 4</u>
א. 0.4404	0.1446

פרק י' - הקשר בין בדיקת השערות לרווח סמך

<u>שאלה 1</u>	<u>שאלה 3</u>
(110.64, 114.56)	(-0.241, 0.041)

פרק 10 - טעויות ועוצמה בבדיקת השערות כללית

1. אדם חשוד בביצוע פשע. מהן הטעויות האפשריות בהכרעת הדין?

2. ילד קנה שקית סוכריות אטומה שבה ציפה ל-10 סוכריות תות ו-5 לימון. ישנה שקית אחרת אותה הוא לא רצה בה 6 סוכריות תות ו-9 לימון. הוא החליט להוציא באקראי סוכרייה אם היא תהיה לימון הוא יחזיר את השקית לחנות. מה הסיכויים לכל סוג של טעות בהכרעתו?

3. יהי X מספר שלם הנבחר באקראי מבין המספרים השלמים. הסיכוי ש- X יקבל ערך

$$p(X = k) = \frac{1}{n} \quad \text{עבור } k = 1, 2, \dots, n$$

נתונות ההשערות הבאות לגבי התפלגות של X :

$$H_0 : n = 4$$

$$H_1 : n = 6$$

כמו כן נתון כלל ההכרעה הבא: נדחה את השערת האפס אם $X > 3$.

חשבו את הסיכוי לטעות מסוג ראשון וטעות מסוג שני ואת העוצמה?

4. איכות של מוצר מסווגת ל-4 רמות איכות: מצוין, טוב, בינוני וירוד. להלן התפלגות טיב

המוצר בשני מפעלים:

מפעל	מצוין	טוב	בינוני	ירוד
"היוצר"	0.6	0.2	0.2	0
"שמשון"	0.1	0.2	0.3	0.4

בוחרים ממשלוח מוצר באקראי, אך לא יודעים מאיזה מפעל המשלוח הגיע. על סמך בדיקת האיכות מנסים להכריע האם מדובר במפעל "היוצר" (השערת האפס) או במפעל "שמשון" (השערה אלטרנטיבית).

א. להלן כלל החלטה: אם מדובר במוצר שטיבו "טוב" נכריע שהמוצר בא ממפעל

"שמשון", מהן ההסתברויות לסוגי הטעויות השונים?

ב. להלן כלל החלטה: אם מדובר במוצר שטיבו "בינוני" וגרוע מכך נכריע שהמוצר בא

ממפעל "שמשון", מה מהן ההסתברויות לסוגי הטעויות השונים?

ג. איזה כלל החלטה עדיף? נמק!

5. במטרה לבדוק האם מטבע תקין הטילו אותו 8 פעמים. הוחלט שאם מספר העצים יהיה בין 1 ל 7 כולל יוחלט שהמטבע תקין, אחרת נחליט שהמטבע מזויף.
- א. רשמו את השערות המחקר.
- ב. מה ההסתברות לטעות מסוג ראשון?
- ג. מהי עצמת המבחן אם במציאות אכן המטבע אינו תקין כי הסיכוי לעץ בו הוא 20%.

6. להלן השערות:

$$H_0: X \sim t(5) \quad (\text{התפלגות } t \text{ עם 5 דרגות חופש})$$

$$H_1: X \sim Z \quad (\text{התפלגות נורמאלית סטנדרטית})$$

כלל החלטה: נדחה את השערת האפס אם X גדול מ-2.015.

א. מהי רמת המובהקות של כלל ההחלטה?

ב. מהי העוצמה של כלל ההחלטה?

7. במפעל מסוים נפלטים לאוויר חומרים רעילים. במצב שיגרה העוצמה הממוצעת של החומר הרעיל אמורה להיות 6,000 יחידות עם סטיית תקן 900. במצב חירום העוצמה הממוצעת היא 7,000 עם סטיית תקן 900. במפעל מערכת התראה נתמכת על ידי 9 חיישנים. אם ממוצע העוצמה של החומר הרעיל לפי תשעת החיישנים עולה על 6600 יחידות מופעלת מערכת ההתראה. נתון שעוצמת הזיהום מתפלגת נורמאלית.
- א. מה הסיכוי להתראת שווא? (באיזה סוג טעות מדובר)?
- ב. מה הסיכוי שבמצב חירום מערכת ההתראה לא תפעל? (באיזה סוג טעות מדובר)?
- ג. מה ההסתברות שאם המצב הוא מצב חירום מערכת ההתראה תפעל? (איך קוראים להסתברות זו)?
- ד. בסעיפים הבאים נשנה בכל סעיף נתון מסוים. כל סעיף עומד בפני עצמו, כיצד השינוי ישנה את הסיכוי לטעות מסוג ראשון ושני?
1. המפעל יקנה עוד 4 חיישנים.
2. מצב חרום מוגדר כעת בתוחלת של 7500 יחידות.
3. מערכת ההתראה תופעל אם ממוצע של תשעת החיישנים יהיה מעל 6700.

תשובות סופיות - טעויות ועוצמה בבדיקת השערות כללית

<u>שאלה 1</u>	<u>שאלה 2</u>
	$\beta = \frac{2}{5} \quad \alpha = \frac{1}{3}$
<u>שאלה 3</u>	<u>שאלה 4</u>
$\beta = 0.5 \quad \alpha = 0.25$	א. $\beta = 0.8 \quad \alpha = 0.2$ ב. $\beta = 0.3 \quad \alpha = 0.2$ ג. כלל ב'.
<u>שאלה 5</u>	<u>שאלה 6</u>
ב. 0.00781 ג. 0.1678	א. 0.05 ב. 0.022
<u>שאלה 7</u>	
א. 0.0228 ב. 0.0918 ג. 0.9082	

פרק 11 - מבחני חי בריבוע

א. מבחן טיב התאמה

1. במטרה לבדוק האם קובייה הוגנת, מטילים אותה 120 פעמים. התקבל 17 פעמים 1, 23 פעמים 2, 20 פעמים 3, 25 פעמים 4, 18 פעמים 5 ו- 17 פעמים 6. מה מסקנתכם ברמת מובהקות של 5%?

2. מפעל מייצר סוכריות בצבעים כחול, אדום, ירוק וכתום. מעוניינים לבדוק שפרופורציית הסוכריות הכחולות גדולה פי 2 מכל צבע אחר. לצורך כך נדגמו באקראי 200 סוכריות והתקבל: 70 כחולות, 50 אדומות, 40 ירוקות והיתר כתומות. מה מסקנתכם ברמת מובהקות של 5%?

3. משרד החינוך טוען שבקרב השכירים במשק היחס בין השכירים בעלי השכלה נמוכה, תיכונית ואקדמאית הוא 1:2:1 בהתאמה. במדגם של 200 שכירים התקבלו 56 אנשים בעלי השכלה נמוכה, 105 בעלי השכלה תיכונית והיתר בעלי השכלה גבוהה. ע"ס תוצאות המדגם האם התפלגות ההשכלה היא כמו שמשרד החינוך מפרסם? בדוק ברמת מובהקות של 5%.

4. בפנס יש 4 סוללות. בבדיקה שנערכה ב-400 פנסים נמצאו סוללות פגומות לפי השכיחויות הבאות:

מספר הסוללות הפגומות	0	1	2	3 ומעלה
שכיחות	276	104	12	8

מעוניינים לבדוק על סמך תוצאות מדגם אלה האם הסיכוי לסוללה פגומה הוא 20%. בדוק ברמת מובהקות של 5%.

ב. מבחן אי תלות

1. במפעל עובד בשלוש משמרות. להלן מספר המוצרים הפגומים והתקינים בכל אחת מן המשמרות לפי מדגם שנעשה :

	יום	ערב	לילה
פגומים	50	60	70
תקינים	600	700	800

האם קיים קשר בין טיב המוצר למשמרת שלו? הסיקו עבור רמת מובהקות $\alpha = 0.05$.

2. בקרב מדגם של 200 נשים טענו שהן תצבענה למועמד R לראשות העיר. בקרב מדגם של 200 גברים 80 טענו שהם יצביעו למועמד R האם קיים הבדל בין דפוס ההצבעה של הנשים ושל הגברים? בדוק ברמת מובהקות של 5%.

3. בחנות בגדים A בדקו את התפלגות הצבעים של הבגדים הנמכרים ביום מסוים:

צבע	שחור	לבן	אדום	כחול
מספר הפריטים	15	20	15	50

כמו כן בדקו את התפלגות הצבעים בחנות שכנה B :

צבע	שחור	לבן	אדום	כחול
מספר הכדורים	60	20	10	20

- א. בדוק ברמת מובהקות של 5% האם התפלגות הצבעים בחנות A היא ביחס של 1:1:1:3 לטובת הכחול.
- ב. בדוק ברמת מובהקות של 2.5% האם קיים הבדל בין החנויות מבחינת התפלגות הצבעים של הפריטים הנמכרים.

תשובות סופיות - מבחני חי בריבוע

פרק א' - מבחן טיב התאמה

<u>שאלה 1</u>	<u>שאלה 2</u>
נקבל H_0	נקבל H_0
<u>שאלה 3</u>	<u>שאלה 4</u>
נקבל H_0	נדחה H_0

פרק ב' - מבחן לאי תלות

<u>שאלה 1</u>	<u>שאלה 2</u>
נקבל H_0	נדחה H_0
<u>שאלה 3</u>	
ג. נקבל H_0	
ד. נדחה H_0	