מבוא לסטטיסטיקה בניהול

תוכן העניינים

1. סטטיסטיקה תיאורית-הקדמה 1
2. סטטיסטיקה תיאורית- הצגה של נתונים 4
3. סטטיסטיקה תיאורית-גבולות מדומים ואמיתיים 15
4. סטטיסטיקה תיאורית- סכימה 17
5. סטטיסטיקה תיאורית -מדדי מרכז ..... (ללא ספר)
6. סטטיסטיקה תיאורית - מדדי פיזור - הטווח, השונות וסטיית התקן 21
7. סטטיסטיקה תיאורית- מדדי מיקום יחסי-ציון תקן 24
8. סטטיסטיקה תיאורית-אחוזונים בטבלה בדידה 26
9. סטטיסטיקה תיאורית - טרנספורמציה לינארית 28
10. סטטיסטיקה תיאורית-מקדם ההשתנות 31
11. סטטיסטיקה תיאורית- תרשים קופסא 33
12. סטטיסטיקה תיאורית- ניתוח פלטים 35
13. סטטיסטיקה תיאורית - שאלות מסכמות 37
14. יסודות ההסתברות 41
15. פעולות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכילים 45
16. הסתברות מותנית-במרחב מדגם אחיד 53
17. הסתברות מותנית - מרחב לא אחיד 56
18. דיאגרמת עצים - נוסחת בייס ונוסחת ההסתברות השלמה 59
19. תלות ואי תלות בין מאורעות 64
20. שאלות מסכמות בהסתברות 66
21. המשתנה המקרי הבדיד - פונקציית ההסתברות 69
22. המשתנה המקרי הבדיד - תוחלת - שונות וסטיית תקן 72
23. המשתנה המקרי הבדיד- טרנספורמציה לינארית 75

תוכן העניינים

24. תוחלת ושונות של סכום משתנים מקריים 78
25. התפלגויות בדידות מיוחדות -התפלגות בינומית 81
26. התפלגויות רציפות מיוחדות - התפלגות נורמלית 85
27. הסקה סטטיסטית - הקדמה 93
28. התפלגות הדגימה ומשפט הגבול המרכזי 96
29. מושגי יסוד באמידה 112
30. רווח סמך לתוחלת (ממוצע) 117
31. מבוא לבדיקת השערות על פרמטרים 127
32. בדיקת השערות על תוחלת (ממוצע) 133
33. רווח סמך להפרש תוחלות (ממוצעים) במדגמים בלתי תלויים 160
34. בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים 164
35. רווח סמך לתוחלת (ממוצע) ההפרשים במדגמים מזווגים 172
36. בדיקת השערות לתוחלת ההפרש במדגמים מזווגים 174
37. הקשר בין רווח סמך לבדיקת השערות להפרש תוחלות 178
38. רווח סמך לשונות וסטיית תקן 181
39. רווח סמך ליחס שונויות 186
40. בדיקת השערות ..... (ללא ספר)
41. רווח סמך לפרופורציה 190
42. רווח סמך להפרש פרופורציות 193
43. בדיקת השערות על פרופורציה 195
44. בדיקת השערות על הפרש פרופורציות 206
45. שאלות מסכמות על רווחי סמך 210
46. בדיקת השערות על שונויות 213
47. שאלות מסכמות בבדיקת השערות 223
48. מבחני חי בריבוע 240

מבוא לסטטיסטיקה בניהול

פרק 1 - סטטיסטיקה תיאורית-הקדמה

תוכן העניינים

1. כללי 1

סטטיסטיקה תיאורית – הקדמה:

רקע:

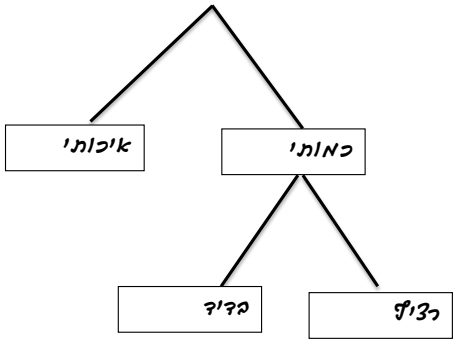
בסטטיסטיקה תיאורית אנו חוקרים קבוצה מסוימת, שיכולה להיות קבוצת ילדים בגן, קבוצת מניות בתיק, כלל התושבים בעיר מסוימת וכולי.
בין ישות לישות בקבוצה ישנם גורמים היכולים לקבל מספר ערכים. גורמים אלה נקראים משתנים. למשל, בין מניה למניה בתיק משתנה התשואה היומית של המניה, הוותק של המניה, תחום המניה וכדומה.
בסטטיסטיקה תיאורית אנחנו נתבונן בקבוצה מסוימת ובתוך הקבוצה הזו נאסוף נתונים לגבי משתנה מסוים ונלמד להציג את הנתונים ולנתח אותם מכל מיני אספקטים.

דוגמה:

בתיק מניות 10 מניות. מנהל התיק פרסם את התשואה של כל מניה בשנת 2011.

1) מי הקבוצה הנחקרת?
2) מה גודל הקבוצה?
3) מה המשתנה הנחקר?

סוגי משתנים:



משתנה איכותי

משתנה שלערכיו אין משמעות של יותר או פחות, אין עניין כמותי לערכים המתקבלים.
כמו: מקום מגורים של אדם (רעננה, תל אביב, אשדוד..), מין האדם (זכר, נקבה) ומצב משפחתי (רווק, נשוי, גרוש, אלמן).

© כל הזכויות שמורות לברק קנדל

משתנה כמותי

משתנה שערכיו הם מספרים, להם יש משמעות כמותית כמו: גובה אדם בס"מ, ציון בבחינה וכדומה.
את המשתנה הכמותי נסווג לשני סוגים:

1. משתנה בדיד – משתנה שערכיו מתקבלים מתוך סידרה של ערכים אפשריים. כמו: מספר ילדים למשפחה (1,2,3..) וציון בבחינה (מ-0 ועד 100 בקפיצות של 1).

2. משתנה רציף – משתנה שערכיו מתקבלים מתוך אינסוף ערכים בתחום מסוים. הערכים מתקבלים ברצף וללא קפיצות של ערכים. כמו: גובה בס"מ – אם למשל, הגובה הנמוך ביותר הוא 150 ועד 190 ס"מ בקבוצה הגבהים הם ברצף. גם בין 160 ל-161 ס"מ יש רצף אינסופי של ערכים אפשריים לגובה (160.33 ס"מ הוא גם גובה אפשרי), משקל בק"ג, מהירות בקמ"ש וכולי.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) סווגו את המשתנים הבאים לפי: איכותי / כמותי בדיד / כמותי רציף:

א. מספר הדירות בבניין.

ב. גיל אדם בשנים.

ג. אחוז האבטלה בעיר.

ד. מקצוע לימוד מועדף.

2) להלן התפלגות מספר האיחורים לעבודה בחודש של העובדים בחברת "סטארר":

מספר האיחורים	מספר העובדים
0	17
1	23
2	85
3	50
4	25

בחברה 200 עובדים.

א. מהו המשתנה הנחקר כאן?

ב. האם מדובר במשתנה איכותי או כמותי ?
אם הוא כמותי האם הוא בדיד או רציף?

3) להלן רשימה של משתנים כמותיים, ציינו ליד כל אחד אם הוא רציף או בדיד:

א. שכר עובד ב-₪.

ב. ציון בחינת בגרות.

ג. תוצאה בהטלת קובייה.

ד. מהירות ריצה בתחרות.

ה. שיעור התמיכה בממשלה.

תשובות סופיות:

1) א. כמותי בדיד. ב. כמותי רציף. ג. כמותי רציף.
ד. איכותי.

2) א. מספר איחורים. ב. כמותי בדיד.

3) א. רציף. ב. בדיד. ג. בדיד.
ד. רציף. ה. רציף.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 2 - סטטיסטיקה תיאורית- הצגה של נתונים

תוכן העניינים

1. כללי 4

סטטיסטיקה תיאורית – הצגה של נתונים:

רקע:

דרכים להצגת נתונים שנאספו:

רשימה של תצפיות:

התצפית היא הערך שנצפה עבור ישות מסוימת בקבוצה. רושמים את התצפיות שהתקבלו כרשומה, יעיל שיש מספר מועט של תצפיות. ההצגה הזו רלבנטית לכל סוגי המשתנים. למשל, להלן מספר החדרים בבניין בן 5 דירות: 4, 5, 3, 4, 3.

טבלת שכיחויות בדידה:

שם המשתנה- X	שכיחות – $f(x)$	שכיחות יחסית באחוזים
X_1	f_1	$\frac{f_1}{N}\cdot 100$
X_2	f_2	$\frac{f_2}{N}\cdot 100$
X_3	f_3	$\frac{f_3}{N}\cdot 100$
⋮	⋮	⋮
X_k	f_k	$\frac{f_x}{N}\cdot 100$
סה"כ	$N=\sum_{i=1}^{k} f_i$	100%

רושמים את התצפיות בטבלה שבה עמודה אחת מבטאת את ערכי המשתנה והשנייה את השכיחות. יעיל עבור משתנה איכותי וכמותי בדיד וכשיש מספר רב של תצפיות. לא יעיל למשתנה כמותי רציף.

© כל הזכויות שמורות לברק קנדל

דוגמה:

להלן התפלגות הציונים בכיתה מסוימת:

הציון- X	מספר התלמידים – השכיחות - f	F_i	$\frac{f_i}{n}$
5	2	2	0.08=2/25
6	4	6	0.16=4/25
7	8	14	0.32=8/25
8	5	19	0.2=5/25
9	4	23	0.16=4/25
10	2	25	0.08=2/25

שכיחות מצטברת – צבירה של השכיחויות.

השכיחויות F_i – השכיחות המצטברת נותנת כמה תצפיות קטנות או שוות לערך.

שכיחות יחסית (פרופורציה) – השכיחות מחולקת לכמות התצפיות הכללי:

$\frac{f_i}{n}$ – איזה חלק מהתצפיות בקבוצה שוות לערך.

טבלת שכיחויות במחלקות:

משתמשים שהמשתנה כמותי רציף או כאשר יש מספר ערכים רב במשתנה הבדיד וטבלת שכיחויות תהיה ארוכה מידי.

דוגמה:

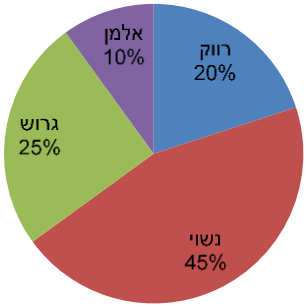
נתנו לקבוצת ילדים לבצע משימה, בדקו את התפלגות זמן הביצוע, בדקות.
להלן ההתפלגות שהתקבלה:

זמן בדקות	מספר הילדים
0.5-3.5	20
3.5-9.5	18
9.5-19.5	14
19.5-29.5	8

© כל הזכויות שמורות לברק קנדל

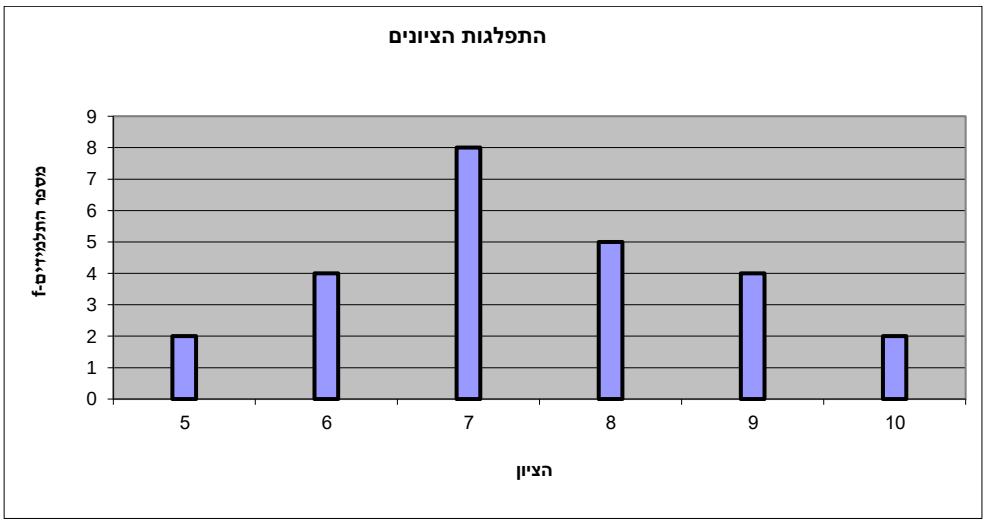
דיאגרמת עוגה:

זהו התיאור הגרפי של משתנה איכותי. בדיאגראמת עוגה כל ערך במשתנה מקבל "נתח", שהוא פרופורציונלי לשכיחות היחסית של ערך המשתנה בנתונים.



דיאגרמת מקלות:

הציר האופקי הוא הציר של המשתנה והציר האנכי של השכיחות, כך שהגובה של המקל מעיד על השכיחות. רלבנטי למשתנה כמותי בדיד. לא נהוג להשתמש בתיאור למשתנה איכותי וכמו כן לא למשתנה כמותי רציף, וכן בסולמות מדידה עבור משתנה מסולם סדר.

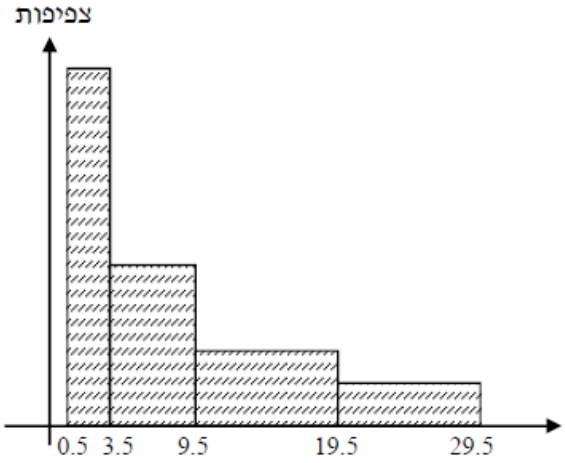


© כל הזכויות שמורות לברק קנדל

היסטוגרמה:

היסטוגרמה היא הדרך הגרפית כדי לתאר טבלת שכיחויות במחלקות, והיא רלוונטית למשתנה כמותי רציף.
בהיסטוגרמה הציר האופקי הוא הציר של המשתנה והציר האנכי הוא הציר של הצפיפות.
הצפיפות מחושבת בכל מחלקה על ידי חלוקת השכיחות ברוחב של כל המחלקה, והיא נותנת את מספר התצפיות הממוצע בכל מחלקה ליחדה. אם המחלקות הן שוות ברוחב, ניתן לשרטט את ההיסטוגרמה לפי השכיחות ואין צורך בצפיפות.

X	רוחב	אמצע	שכיחות	מצטברת	צפיפות
0.5 - 3.5	3	2	20	20	6.6667
3.5 - 9.5	6	6.5	18	38	3
9.5 - 19.5	10	14.5	14	52	1.4
19.5 - 29.5	10	24.5	8	60	0.8



פוליגון – מצולעון:

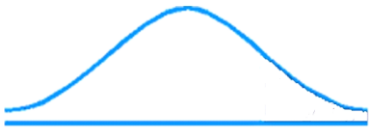
אם נחבר את אמצע קצה כל מלבן בקווים ישרים.
נותן מראה חזותי לצורה של התפלגות המשתנה.

© כל הזכויות שמורות לברק קנדל

צורות התפלגות נפוצות:

התפלגות סימטרית פעמונית

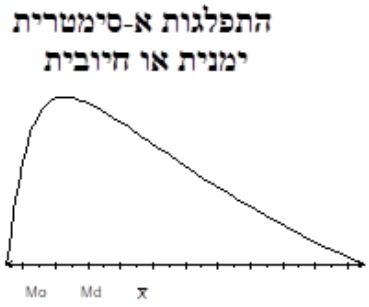
רוב התצפיות במרכז, וככל שנתרחק מהמרכז יהיו פחות תצפיות באופן סימטרי. לדוגמה, ציוני IQ.



ישנן התפלגויות סימטריות שאינן פעמוניות, כגון:

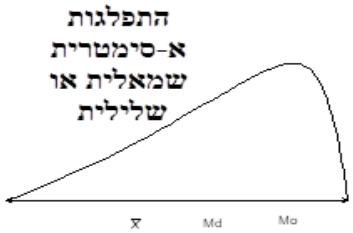
התפלגות אסימטרית ימנית (חיובית)

רוב התצפיות מקבלות ערכים נמוכים ויש מיעוט הולך וקטן של תצפיות שמקבלות ערכים גבוהים קיצוניים. לדוגמה, שכר במשק.



התפלגות אסימטרית שמאלית (שלילית)

רוב התצפיות מקבלות ערכים גבוהים ויש מיעוט הולך וקטן של תצפיות שמקבלות ערכים נמוכים קיצוניים. לדוגמה, אורך חיים.



© כל הזכויות שמורות לברק קנדל

שאלות:

1) בסקר צפייה בטלוויזיה התקבלו התוצאות הבאות: 25 צפו בערוץ הראשון, 25 צפו בערוץ 10, 75 צפו בערוץ השני, 50 צפו באחד מערוצי הכבלים ו-25 לא צפו בטלוויזיה בזמן הסקר.

א. רשמו את טבלת השכיחות ואת השכיחות היחסית.

ב. תארו את הנתונים באופן גרפי.

2) להלן נתונים על התפלגות המקצוע המועדף של תלמידי שכבה ו' בבית הספר "מעוף":

מספר התלמידים	המקצוע
44	מתמטיקה
20	תנ"ך
12	אנגלית
26	היסטוריה

א. מהו המשתנה הנחקר?

ב. מהי פרופורציית התלמידים שמעדיפים תנ"ך?

3) להלן התפלגות ההשכלה במקום עבודה מסוים:

מספר העובדים	השכלה
60	נמוכה
120	תיכונית
20	אקדמאית

א. מהו המשתנה הנחקר?
מאיזה סולם הוא?

ב. תארו את הנתונים באופן גרפי.

4) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
6, 5, 8, 7, 6, 9, 8, 6, 7, 6, 7, 8, 5, 4, 6, 10, 9, 8, 6, 7.

א. מהו המשתנה? האם הוא בדיד או רציף?

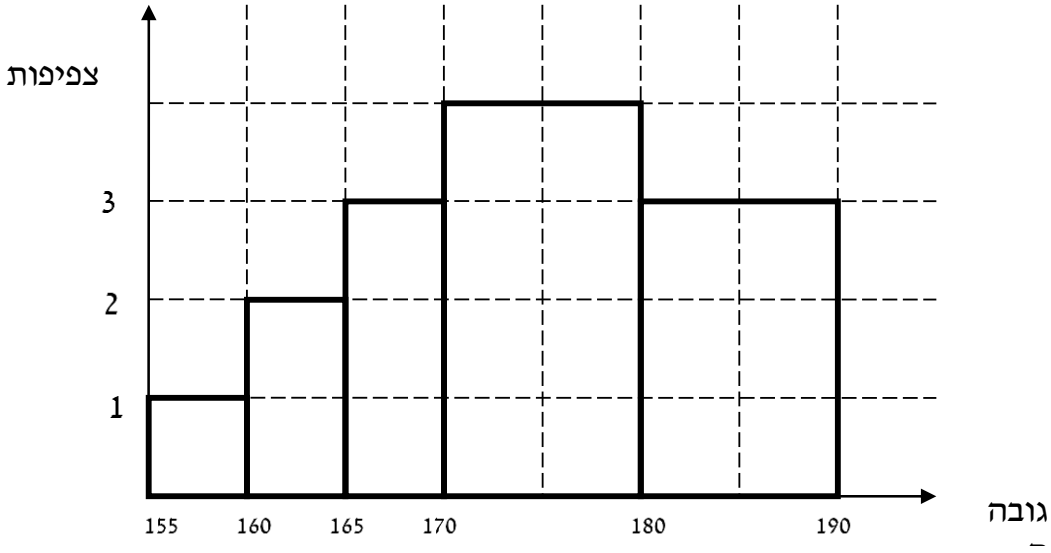
ב. תארו את הרשימה בטבלת שכיחויות.

ג. הוסיפו שכיחויות יחסיות לטבלה.

ד. תארו את הנתונים באופן גרפי.

© כל הזכויות שמורות לברק קנדל

5) להלן היסטוגרמה המתארת את התפלגות הגבהים בס״מ של קבוצה מסוימת:



א. מהו המשתנה הנחקר? האם הוא בדיד או רציף?
ב. תארו את הנתונים בטבלת שכיחויות במחלקות.
ג. הוסיפו שכיחות יחסית לטבלה.
ד. הוסיפו את הצפיפות של כל מחלקה לטבלה.
ה. מהי צורת ההתפלגות של הגבהים?

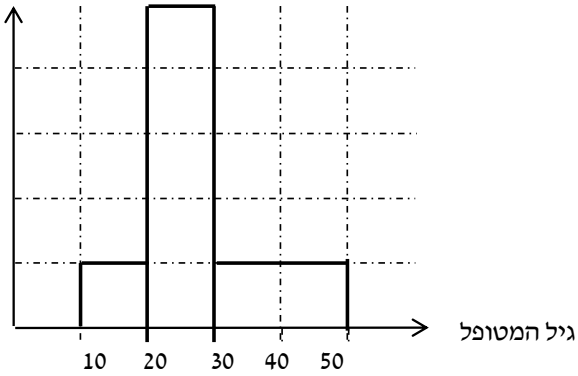
6) להלן התפלגות המשקל של קבוצה מסוימת בק״ג:

משקל	מספר מקרים
40-45	10
45-50	20
50-60	30
60-65	20
65-70	10

א. תארו את ההתפלגות באופן גרפי.
ב. מה ניתן להגיד על צורת ההתפלגות?

© כל הזכויות שמורות לברק קנדל

7) להלן גיל המטופלים של ד"ר שוורץ בשנים:
קנה מידה:



א. מה המשתנה הנחקר? האם הוא בדיד או רציף?
ב. מהי הקבוצה הנחקרת?
ג. תרגמו את ההסיטוגרמה לטבלת שכיחות.
ד. מהי הפרופורציה של המטופלים של ד"ר שוורץ בגילאים 20-30?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. להלן טבלה:

x	$f(x)$	$\frac{f(x)}{n}$	%
ערוץ 1	25	$\frac{25}{200}$	12.5%
ערוץ 10	25	$\frac{25}{200}$	12.5%
ערוץ 2	75	$\frac{75}{200}$	37.5%
כבלים	50	$\frac{50}{200}$	25%
לא צפו	25	$\frac{25}{200}$	12.5%
סה"כ	200	1	100%

ב. עיין גרף מלא בסרטון הוידאו.

2) א. מקצוע מועדף.

ב. 19.6%.

3) א. משתנה נחקר: השכלה, סוג: סדר.

ב. עיין גרף מלא בסרטון הוידאו.

© כל הזכויות שמורות לברק קנדל

4) א. המשתנה: ציון, משתנה בדיד.
ד. עיין גרף מלא בסרטון הוידאו.

ב+ג. להלן טבלה:

x	$f(x)$	$\frac{f(x)}{n}$	%
4	1	$\frac{1}{20}$	5%
5	2	$\frac{2}{20}$	10%
6	6	$\frac{6}{20}$	30%
7	4	$\frac{4}{20}$	20%
8	4	$\frac{4}{20}$	20%
9	2	$\frac{2}{20}$	10%
10	1	$\frac{1}{20}$	5%
סה"כ	20	20	100%

5) א. גובה בס"מ, רציף.

ב+ג+ד. להלן טבלה:

ה. אסימטרית.

x	$f(x)$	$\frac{f(x)}{n}$	%	d
155-160	5	$\frac{5}{100}$	5%	1
160-165	10	$\frac{10}{100}$	10%	2
165-170	15	$\frac{15}{100}$	15%	3
170-180	40	$\frac{40}{100}$	40%	4
180-190	30	$\frac{30}{100}$	30%	3

© כל הזכויות שמורות לברק קנדל

6) א. עיין גרף מלא בסרטון הוידאו.
ב. סימטרית.

7) א. המשתנה: גיל בשנים, משתנה רציף.
ב. המטופלים של ד"ר שוורץ.
ד. להלן טבלה:
ה. 62.5%.

x	$f(x)$
10-20	8
20-30	40
30-50	16

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 3 - סטטיסטיקה תיאורית-גבולות מדומים ואמיתיים

תוכן העניינים

1. כללי 15

סטטיסטיקה תיאורית – גבולות מדומים וגבולות אמיתיים:

רקע:

עבור משתנה רציף נהוג לתאר את הנתונים בטבלת שכיחויות במחלקות.
הנתונים שנאספים הם ברמת דיוק מסוימת.
לדוגמה: משקל של בני אדם ומשקל של יהלומים ישקלו ברמת דיוק שונה.

גבולות מדומים:

כאשר גבול עליון של מחלקה אחת שונה מגבול תחתון של המחלקה הבאה אז הגבולות הם גבולות מדומים. כשהגבולות מדומים, ההפרש בין גבול תחתון של מחלקה לבין גבול עליון של המחלקה הקודמת יהיה רמת הדיוק.

רמת הדיוק חייבת להיות קבועה - אין אפשרות שחלק מהאנשים נדייק ברמה אחת ואת השאר ברמה אחרת. בגלל שהמשתנה הוא משתנה רציף, כשננתח את הנתונים נעבור מגבולות מדומים לגבולות אמתיים. אם הנתונים יינתנו בגבולות מדומים נהפוך אותם תמיד לגבולות אמיתיים.

כיצד עוברים מגבולות מדומים לגבולות אמיתיים?

לוקחים את רמת הדיוק ומחלקים אותה ב-2, ואת התוצאה המתקבלת מוסיפים לגבולות העליונים ומפחיתים מהגבולות התחתונים. אם יתנו נתונים בגבולות מדומים אנחנו מוכרחים לעבור לגבולות אמיתיים על מנת להמשיך ולנתח, אך אם הנתונים כבר יינתנו בגבולות אמיתיים נשאיר אותם כמו שהם.

דוגמה (פתרון בהקלטה):

להלן התפלגות הגבהים בס״מ של תלמידי כיתה ח׳:
יש להעביר את הנתונים לגבולות אמתיים.

$f(x)$	X
20	130-139
25	140-149
30	150-159
20	160-169
10	170-189

© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן התפלגות של משתנה בהצגה של מחלקות.
יש להעביר את הנתונים לגבולות אמתיים:

$f(x)$	X
542	500-590
32	600-690
154	700-790
254	800-890

2) להלן התפלגות המשקלים בק"ג של קבוצת אנשים מסוימת.
יש לרשום את הנתונים בגבולות אמיתיים:

מספר אנשים	משקל בק"ג
18	60-64
24	65-69
52	70-79
19	80-89

תשובות סופיות:

1) להלן טבלה:

$f(x)$	x
542	495-595
32	595-695
154	695-795
254	795-895

2) להלן טבלה:

$f(x)$	x
18	59.5-64.5
24	64.5-69.5
52	69.5-79.5
19	79.5-89.5

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 4 - סטטיסטיקה תיאורית- סכימה

תוכן העניינים

1. כללי 17

סטטיסטיקה תיאורית – סכימה:

רקע:

בסטטיסטיקה ישנה צורת רישום מקובלת לסכום של תצפיות: $\sum_{i=1}^{n} X_i$.

נסביר את צורת הרישום על ידי הדוגמה הבאה:

i	X_i
1	5
2	0
3	1
4	3
5	2

(הסבר מלא מופיע בסרטונים באתר).

© כל הזכויות שמורות לברק קנדל

שאלות:

1) בבניין 5 דירות. לכל דירה רשמו את מספר החדרים שיש בדירה (X),
ומספר הנפשות החיות בדירה (Y). חשבו:

מספר דירה	X	Y
1	2	1
2	3	1
3	2	2
4	4	3
5	3	2

א. $\sum_{i=1}^{3} X_i$.

ב. $\sum_{i=1}^{5} Y_i$.

ג. $\sum_{i=1}^{4} X_i$.

ד. $\left(\sum_{i=1}^{4} X_i\right)^2$.

ה. $\sum X_i$.

ו. $\sum X_i Y_i$.

ז. $\sum(X_i)\sum(Y_i)$.

© כל הזכויות שמורות לברק קנדל

2) נתון לוח ערכי המשתנים X_i ו- Y_i, כאשר: $i=1,2,...,6$, ונתונים הקבועים: $a=2$, $b=5$. חשבו את הנוסחאות הבאות:

i	1	2	3	4	5	6
X_i	3	2	4	-2	1	4
Y_i	2	0	0	1	-5	2

א. $\sum_{i=1}^{4} y_i$.

ב. $\sum_{i=1}^{6} a$.

ג. $\sum_{i=1}^{6} x_i y_i$.

ד. $\sum_{i=1}^{6} (x_i + y_i)$.

ה. $\sum_{i=1}^{6} x_i + a$.

3) קבעו לכל זהות האם היא נכונה:

א. $\sum_{i=1}^{n} bX_i = b \cdot \sum_{i=1}^{n} X_i$.

ב. $\sum_{i=1}^{n} a = a \cdot n$.

ג. $\left(\sum_{i=1}^{n} X_i\right)^2 = \sum_{i=1}^{n} X_i^2$.

4) נתון: $\sum_{i=1}^{10} X_i^2 = 1640$, $\sum_{i=1}^{10} X_i = 80$.

חשבו: $\sum_{i=1}^{10} (X_i - 4)^2$.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. 7. ב. 9. ג. 11. ד. 121.
ה. 14. ו. 27. ז. 126.

2) א. 3. ב. 12. ג. 7.
ד. 12. ה. 14.

3) א. נכונה. ב. נכונה. ג. לא נכונה.

4) 1160.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 5 - סטטיסטיקה תיאורית -מדדי מרכז

תוכן העניינים

1. כללי ..... (ללא ספר)

מבוא לסטטיסטיקה בניהול

פרק 6 - סטטיסטיקה תיאורית - מדדי פיזור - הטווח, השונות וסטיית התקן

תוכן העניינים

1. כללי 21

סטטיסטיקה תיאורית – מדדי פיזור – הטווח, השונות וסטיית התקן:

רקע:

המטרה: למדוד את הפיזור של הנתונים, כלומר כמה הם רחוקים זה מזה ושונים זה מזה.

הטווח / תחום (RANGE):
ההפרש בין התצפית הגבוהה ביותר לנמוכה ביותר: $R = X_{\max} - X_{\min}$.

שונות וסטיית תקן:

שונות היא ממוצע ריבועי של הסטיות מהממוצע וסטיית התקן היא שורש של השונות.

עבור סדרת נתונים: $S_x^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2$.

דוגמאות:

1) נחשב את השונות של סדרת המספרים הבאה: 9, 4, 5.

עבור טבלת שכיחויות: $S_x^2 = \frac{\sum(x-\bar{x})^2 f}{n} = \frac{\sum x^2 \cdot f}{n} - \bar{x}^2$.

2) להלן התפלגות הציונים בכיתה מסוימת בה ממוצע הציונים הוא 7.44.

הציון X	השכיחות F	$x^2 \cdot F$
5	2	50
6	4	144
7	8	392
8	5	320
9	4	324
10	2	200
סה"כ		**1430**

$$S_x^2 = \frac{\sum_x x^2 f(x)}{n} - \bar{x}^2 = \frac{1430}{25} - 7.44^2 = 1.8464$$
$$S = \sqrt{S_x^2} = \sqrt{1.8464} = 1.3588$$

כשיש מחלקות נעזר באמצע המחלקה כדי לחשב את השונות.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן רשימת הציונים של 20 תלמידים שנבחנו במבחן הבנת הנקרא:
6, 5, 8, 7, 6, 9, 8, 6, 7, 6, 7, 8, 5, 4, 6, 10, 9, 8, 6, 7.
חשבו את השונות, סטיית התקן והטווח של הציונים.

2) להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן":

מספר מכוניות למשפחה	1	2	3	4	5
שכיחות	65	150	220	140	55

א. חשבו סטיית התקן.

ב. חשבו את הטווח של הנתונים.

הקפידו להסביר לגבי כל סעיף מה משמעות התוצאה שקיבלתם.

3) בחברה העוסקת בטלמרקטינג בדקו עבור כל עובד את מספר שנות הוותק שלו. התקבל שממוצע שנות הוותק הוא 4 שנים וסטיית התקן היא שנתיים.

א. האם הממוצע יגדל/יקטן/לא ישתנה וסטיית התקן תגדל/תקטן/לא תשנה כאשר יתווספו שני עובדים עם וותק של 4 שנים להתפלגות?

ב. האם הממוצע יגדל/יקטן/לא ישתנה וסטיית התקן תגדל/תקטן/לא תשנה כאשר יתווספו שני עובדים אשר אחד עם וותק של 0 שנים והשני עם וותק של 8 שנים להתפלגות?

4) נתונה רשימה של 5 תצפיות, אך רק עבור 4 מהן נרשמו הסטיות שלהן מהממוצע: 2, 3, 2, 1-. חשבו את השונות של חמש התצפיות.

5) בשכונה בדקו בכל דירה את מספר החדרים לדירה. בשכונה 200 דירות.

מספר חדרים	**פרופורציה**
1	0.1
2	0.2
3	0.4
4	0.15
5	

א. מה הממוצע של מספר החדרים לשכונה בדירה?

ב. חשבו את סטיית התקן של מספר החדרים לדירה.

ג. חלק מבעלי הדירות בנות 2 החדרים הפכו את דירתם לדירת חדר. כיצד הדבר ישפיע(יקטין, יגדל, לא ישנה)
על כל מדד שחישבתם בסעיפים הקודמים

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) שונות: 2.19, סטיית תקן: 1.48, טווח: 6.

2) א. סטיית תקן: 1.106. ב. טווח: 4.

3) א. ממוצע לא ישתנה, סטיית התקן תקטן.
ב. ממוצע לא ישתנה, סטיית התקן תגדל.

4) 10.8.

5) א. 3.05. ב. 1.16. ג. ממוצע: יקטן, סטיית התקן: תגדל.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 7 - סטטיסטיקה תיאורית- מדדי מיקום יחסי-ציון תקן

תוכן העניינים

1. כללי 24

סטטיסטיקה תיאורית – מדדי מיקום יחסי – ציון תקן:

רקע:

המטרה למדוד איך תצפית ממוקמת ביחס לשאר התצפיות בהתפלגות.

ציון תקן:

הנוסחה לציון תקן של תצפית היא: $Z = \frac{X - \overline{X}}{S}$.

ציון התקן נותן כמה סטיות תקן סוטה התצפית מהממוצע. כלומר, ציון התקן מעיד על כמה סטיות תקן התצפית מעל או מתחת לממוצע:

- ציון תקן חיובי אומר שהתצפית מעל הממוצע.
- ציון תקן שלילי אומר שהתצפית מתחת לממוצע.
- ציון תקן אפס אומר שהתצפית בדיוק בממוצע.

דוגמה (פתרון בהקלטה):

במקום עבודה מסוים, ממוצע המשכורות הוא 8 אלף ₪, עם סטית תקן של אלפיים ₪.
באותו מקום עבודה ההשכלה הממוצעת של העובדים הנה 14 שנים, עם סטית תקן של 1.5 שנים. ערן מרוויח במקום עבודה זה 11 אלף ₪ והשכלתו 16 שנים.
מה ערן יותר, באופן יחסי, משכיל או משתכר?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) תלמידי כיתה ח' ניגשו למבחן בלשון ולמבחן במתמטיקה.
להלן התוצאות שהתקבלו:

המקצוע	ממוצע	סטיית תקן
לשון	74	12
מתמטיקה	80	16

עודד קיבל: 68 בלשון ו-70 במתמטיקה.

א. באיזה מקצוע עודד טוב יותר באופן יחסי לשכבה שלו?

ב. איזה ציון עודד צריך לקבל במתמטיקה כדי שיהיה שקול לציונו בלשון?

2) במפעל לייצור מצברים לרכב בדקו במשך 40 ימים את התפוקה היומית (מספר מצברים במאות) ואת מספר הפועלים שעבדו באותו היום.
להלן טבלה המסכמת את המידע שנאסף על שני המשתנים:

	תפוקה	מספר פועלים
ממוצע	48	15
סטיית תקן	10	2

באחד הימים מתוך כלל הימים שנבדקו התפוקה הייתה 50 מאות מצברים ובאותו היום עבדו 13 פועלים.
מה יותר חריג באותו היום, יחסית לשאר הימים שנבדקו: נתוני התפוקה או כמות הפועלים?

א. התפוקה.

ב. כמות הפועלים.

ג. חריגים באותה מידה.

ד. חסרים נתונים כדי לדעת זאת.

3) הגובה הממוצע של המתגייסים לצבא הוא 175 סנטימטר עם סטיית תקן של 10 סנטימטר. המשקל הממוצע הוא 66 ק"ג עם סטיית תקן של 8 ק"ג.
ערן התגייס כשגובהו 180 ס"מ ומשקלו 59 ק"ג.

א. במה ערן חריג יותר ביחס לשאר המתגייסים, גובהו או משקלו?

ב. כמה ערן אמור לשקול כדי שמשקלו יהיה שקול לגובהו?

תשובות סופיות:

1) א. לשון. ב. 72.

2) ב'.

3) א. משקל. ב. 70.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 8 - סטטיסטיקה תיאורית-אחוזונים בטבלה בדידה

תוכן העניינים

1. כללי 26

סטטיסטיקה תיאורית – מדדי מיקום יחסי – אחוזונים בטבלה בדידה:

רקע:

האחוזון (המאון) ה- p הוא הערך בנתונים המחלק את הנתונים בצורה כזאת, שעד אליו (כולל) יש p % מהנתונים. מסמנים את האחוזון ה- p ב- X_p .

חישוב האחוזון מתוך נתונים בטבלת שכיחויות בדידה:

האחוזון הוא הערך שבו בפעם הראשונה השכיחות היחסית המצטברת (באחוזים) גדולה או שווה ל- p % .

דוגמה (פתרון בהקלטה):

בסניף בנק 250 לקוחות. ספרו לכל לקוח את מספר תוכניות החיסכון שלו:

# תוכניות החיסכון	$F(x)$	שכיחות מצטברת	שכיחות יחסית מצטברת
0	100		
1	75		
2	25		
3	25		
4	25		

א. מצאו את האחוזון ה-25.

ב. מצאו את הערך ש-20% מהמקרים מעליו.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן התפלגות של משתנה כלשהו:

X	$F(x)$
0	10
1	40
2	30
3	15
4	5

מצאו להתפלגות את:

א. האחוזון ה-60.

ב. המאון ה-40.

ג. העשירון העליון.

ד. הטווח בין הרבעונים.

2) להלן התפלגות מספר המכוניות למשפחה בישוב "הגורן":

מספר מכוניות למשפחה	1	2	3	4	5
שכיחות	65	150	220	140	55

חשבו את:

א. העשירון התחתון.

ב. האחוזון ה-30.

ג. הערך ש-20% מהתצפית גדולות ממנו.

ד. רבעון עליון.

תשובות סופיות:

1) א. 2 ב. 1. ג.3. ד. 1.

2) א. 1. ב. 2. ג. 4. ד. 4.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 9 - סטטיסטיקה תיאורית - טרנספורמציה לינארית

תוכן העניינים

1. כללי 28

סטטיסטיקה תיאורית – טרנספורמציה לינארית:

רקע:

מצב שבו מבצעים שינוי מסוג הוספה (או החסרה) של קבוע, והכפלה (או חילוק) של קבוע, לכל התצפיות: $y = a \cdot x + b$. כך יושפעו המדדים השונים:

מדדי המרכז:

$$MR_y = a \cdot MR + b$$
$$MO_y = a \cdot MO + b$$
$$\bar{y} = a \cdot \bar{x} + b$$
$$Md_y = a \cdot \leq Md_x + b$$

מדדי הפיזור:

$$R_y = |a| R_x$$
$$S_y = |a| S_x$$
$$S_y^{\ 2} = a^2 S_x^{\ 2}$$

מדדי המיקום היחסי:

$$Y_p = a \cdot X_p + b$$
$$Z_Y = \frac{a}{|a|} Z_x$$

שלבי העבודה:

1. נזהה שמדובר בטרנספורמציה לינארית (שינוי קבוע לכל התצפיות).
2. נרשום את כלל הטרנספורמציה לפי נתוני השאלה.
3. נפשט את הכלל ונזהה את ערכי a ו-b.
4. נציב בנוסחאות שלעיל בהתאם למדדים שנשאלים.

דוגמה (פתרון בהקלטה):

השכר הממוצע של עובדים הינו 9000 ₪ וטווח 6000 ₪. חשבו את המדדים הללו לאחר שהעלו את כל המשכורות ב-10% ואחר כך קנסו אותם ב-100 ₪.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) עבור סדרת נתונים התקבל: $\bar{x}=80, S=15, MO=70$.
הוחלט להכפיל את כל התצפיות ב-4 ולהחסיר מהתוצאה 5.
חשבו את המדדים הללו לאחר השינוי.

2) בחברה מסוימת השכר הממוצע הוא 40 ₪ לשעה עם סטיית תקן של 5 ₪ לשעה.
הוחלט להעלות את כל המשכורות ב-10%, אך זה לא סיפק את העובדים ולכן הם קיבלו לאחר מכן תוספת של 2 ₪ לשעה.
מה הממוצע ומהי השונות של השכר לשעה לאחר כל השינויים.

3) במבחן מסוים הציון החציוני היה 73, טווח הציונים היה 40 נקודות והעשירון העליון היה הציון 87. כיוון שהציונים בבחינה היו נמוכים, המורה החליט לתת פקטור של 4 נק' לכל התלמידים.
חשבו את המדדים לאחר הפקטור.

4) דגמו מקו ייצור 50 קופסאות של גפרורים. בדקו בכל קופסא בה יש 40 גפרורים את כמות הגפרורים הפגומים. התקבל שבממוצע יש 3 גפרורים פגומים בקופסא, עם סטיית תקן של 1.5 גפרורים.
מה יהיה הממוצע ומה תהיה סטיית התקן של מספר התקינים בקופסא?

5) חברת בזק הציעה את ההצעה הבאה: שלושים שקלים דמי מנוי חודשיים קבועים וכן 10 אגורות לכל דקה של שיחה יוצאת. אדם בדק במשך שנה את דקות השיחות היוצאות שלו, וקיבל שבממוצע חודשי יש לו 600 דקות שיחות יוצאות עם שונות של 2500 דקות רבועות, כמו כן בחודש ינואר ציון התקן היה 2.
חשבו את המדדים הללו עבור חשבון הטלפון החודשי של אותו אדם בשקלים אם היה משתמש בחבילה המוצעת לו על ידי בזק.

6) הוכיחו שאם כל התצפיות בהתפלגות עברו טרנספורמציה לינארית: $Y_i = a \cdot X_i + b$,
אזי הממוצע והשונות של כלל התצפיות לאחר הטרנספורמציה יהיו בהתאמה:
$\bar{y} = a \cdot \bar{x} + b, S_y^2 = a^2 S_x^2$.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) ממוצע: 315, סטיית תקן: 60, שכיח: 275.
2) ממוצע: 46, שונות: 30.25.
3) טווח: 40, חציון: 77, עשירון עליון: 91.
4) ממוצע: 37, סטיית תקן: 1.5.
5) ממוצע: 90, שונות: 25, ציון תקן: 2.
6) $a^2 \cdot S_x^2$.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 10 - סטטיסטיקה תיאורית-מקדם ההשתנות

תוכן העניינים

1. כללי 31

סטטיסטיקה תיאורית – מקדם ההשתנות:

רקע:

כאשר מחשבים סטיית תקן למספר קבוצות בעלי ממוצע שונה, השוואת מידת פיזור הנתונים אינה מתייחסת לערך מרכז הנתונים (לממוצע למשל).
על מנת לתת מדד פיזור המתחשב בממוצע הנתונים נחשב את מקדם ההשתנות – Coefficient of Variation: $CV = \frac{S(X)}{\overline{X}}$.
ככל שמקדם ההשתנות נמוך יותר, כך המשתנה מרוכז יותר סביב הממוצע, וככל שמקדם ההשתנות גבוה יותר, מידת הפיזור סביב הממוצע גבוהה יותר.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן נתונים לגבי ציונים במבחן באנגלית ב-3 כיתות מתוך שכבה י' בתיכון:

כיתה	ממוצע	מס' תלמידים	סטיית תקן
1	76	40	12
2	68	20	15
3	82	30	10

א. חשבו את מקדם ההשתנות בכל כיתה.

ב. מהי הכיתה הכי הטרוגנית?

2) נתונות שתי קבוצות: הממוצע בקבוצה א' הוא 100 והשונות 100.
הממוצע בקבוצה ב' הוא 500 והשונות 400.
באיזו קבוצה מידת הפיזור יחסית קטן יותר?

3) במפעל לייצור מצברים לרכב בדקו במשך 40 ימים את התפוקה היומית (מספר מצברים במאות) ואת מספר הפועלים שעבדו באותו היום.
להלן טבלה המסכמת את האינפורמציה שנאספה על שני המשתנים:

	תפוקה	מספר פועלים
ממוצע	48	15
סטיית תקן	10	2

לפי קריטריון CV:

א. הפיזור באופן יחסי שווה בין התפוקה היומית לכמות הפועלים העובדים ביום.

ב. הפיזור יחסית יותר גדול עבור התפוקה היומית מאשר עבור מספר הפועלים ביום.

ג. הפיזור יחסית יותר גדול עבור מספר הפועלים ביום מאשר עבור התפוקה היומית.

ד. אין מספיק נתונים כדי לחשב את CV.

תשובות סופיות:

1) א. $\frac{\sigma}{\bar{X}}$. ב. כיתה ב'.

2) קבוצה ב'.

3) ב'.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 11 - סטטיסטיקה תיאורית- תרשים קופסא

תוכן העניינים

1. כללי 33

סטטיסטיקה תיאורית – תרשים קופסא (Boxplot):

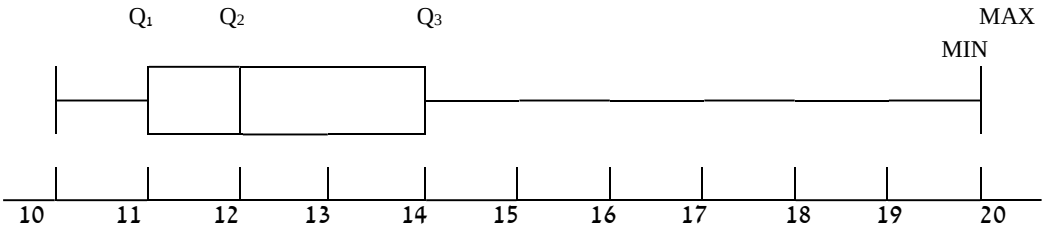
רקע:

תרשים קופסא הינו תרשים שבעזרתו ניתן לבחון:

1) את המרכז של ההתפלגות על ידי החציון (Q_2).

2) את הפיזור של הנתונים (הטווח והטווח הבין רבעוני).

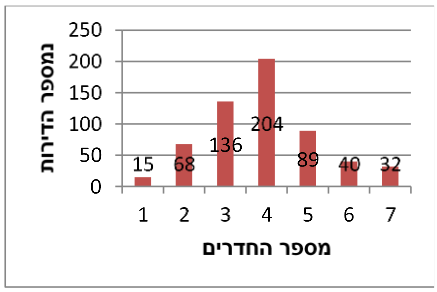
3) את צורת ההתפלגות (סימטרית \אסימטרית ימנית או אסימטרית שמאלית).



© כל הזכויות שמורות לברק קנדל

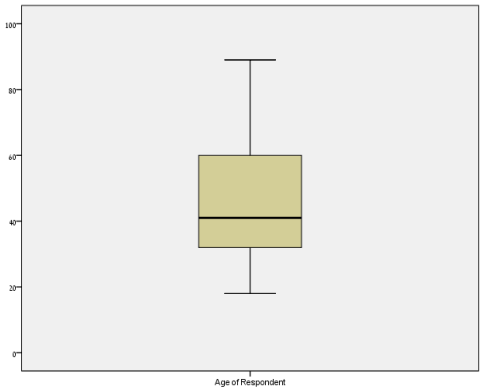
שאלות:

1) להלן התפלגות מספר החדרים לדירות שנבנו בשנת 2009 בעיר אשדוד:



א. מצאו את החציון, הרבעון התחתון והרבעון העליון של ההתפלגות.
ב. שרטטו דיאגרמת קופסא להתפלגות.
ג. מה ניתן לומר על צורת ההתפלגות?

2) להלן דיאגרמת קופסא המתארת את התפלגות הגיל (בשנים) באוכלוסייה מסוימת:



א. מה הגיל החציוני?
ב. מה בערך טווח הגילאים?
ג. מה ניתן להגיד על צורת ההתפלגות?

תשובות סופיות:

1) א. חציון: 4 , רבעון תחתון: 3, רבעון עליון: 5.
ב. ראה גרף מלא בסרטון וידאו. ג. כמעט סימטרית.

2) א. חציון: 40. ב. טווח: 70. ג. התפלגות אסימטרית ימנית.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 12 - סטטיסטיקה תיאורית- ניתוח פלטים

תוכן העניינים

1. כללי 35

סטטיסטיקה תיאורית – ניתוח פלטים:

שאלות:

1) להלן פלט על התפלגות הגילאים באוכלוסייה מסוימת:

		Statistic
Age of Respondent	Mean	45.63
	Median	41.00
	Variance	317.140
	Std. Deviation	a
	Minimum	18
	Maximum	b
	Range	71
	Interquartile Range	28

א. מצאו את הערכים בטבלה המסומנים ב- a ו- b.

ב. נתון שההתפלגות היא אסימטרית, האם היא נוטה ימינה או שמאלה?

2) להלן התפלגות ההשכלה של העובדים בחברת "מתאר":

		Statistic
years of education	Mean	?
	Median	12.0000
	Variance	?
	Std. Deviation	2.54786
	Minimum	?
	Maximum	?
	Range	?
	Interquartile Range	?

מלאו את הערכים המסומנים בסימני שאלה.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. $b=89$, $a=17.81$. ב. אסימטרית ימנית.

2) ממוצע: 11.909, שונות: 6.492, טווח: 10, טב"ר: 3.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 13 - סטטיסטיקה תיאורית - שאלות מסכמות

תוכן העניינים

1. כללי 37

סטטיסטיקה תיאורית – שאלות מסכמות:

שאלות:

1) בדקו עבור 5 תלמידים את המשקל שלהם:

מספר תלמיד	משקל בק"ג
1	58
2	62
3	48
4	34
5	58

א. מהו המשתנה הנחקר? בדיד או רציף?

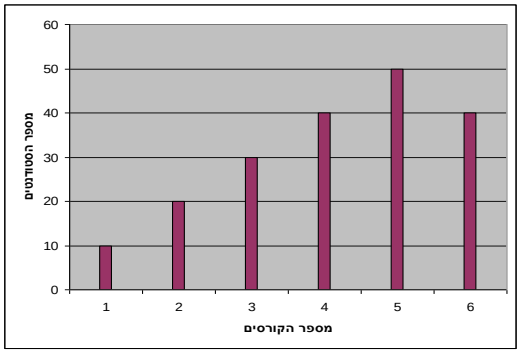
ב. מה המשקל החציוני, הממוצע והשכיח?

ג. מה הטווח וסטיית התקן של המשקל?

ד. לאותם תלמידים חישבו גם את הגובה בס"מ וקיבלו גובה ממוצע של 168 וסטיית תקן 6. במה תלמיד מספר 3, שגובהו 162, יותר חריג – במשקל או בגובה?

ה. הוסיפו עוד תלמיד השוקל 52 ק"ג בדיוק. הסבירו ללא חישוב כיצד הדבר ישפיע על הממוצע וסטיית התקן (יגדיל, יקטין או לא ישנה).

2) בפקולטה להנדסה אספה המזכירות נתונים לגבי מס' הקורסים שכל סטודנט סיים בשנה הראשונה ללימודיו בשנת 2008. להלן התוצאות שהתקבלו:



א. מה המשתנה הנחקר? בדיד או רציף?

ב. מהי צורת ההתפלגות?

ג. תארו את הנתונים בטבלת שכיחויות.

ד. חשבו את השכיח, החציון והטווח.

© כל הזכויות שמורות לברק קנדל

3) להלן התפלגות הציונים בבחינה בלשון שנעשתה עבור תלמידי כיתות ד'.
במחקר השתתפו 150 תלמידים. ממוצע הציונים שהתקבל: $\overline{X} = 7\frac{1}{15}$.

ציון	מספר התלמידים
4	12
5	16
6	
7	38
8	
9	14
10	10

א. השלימו את השכיחויות החסרות בטבלה.
ב. חשבו את הציון החציוני, השכיח.
ג. חשב שונות וסטיית תקן להתפלגות הציונים.

4) חברה סלולארית דגמה 200 אנשים. עבור כל אדם נבדקה מידת שביעות הרצון של הלקוח מהחברה (1 - שביעות רצון נמוכה ו-5 - שביעות רצון גבוהה).
להלן ההתפלגות שהתקבלה:

שביעות רצון	מספר האנשים
1	40
2	60
3	50
4	30
5	20

א. מה אחוז האנשים עם רמת שביעות רצון נמוכה?
ב. מה המשתנה הנחקר ומאיזה סוג הוא?
ג. מהי הדרך הגרפית המתאימה ביותר לתיאור הנתונים?
   i. היסטוגרמה.
   ii. דיאגרמת מקלות.
   iii. דיאגרמת עוגה.
ד. חשבו את המדדים הבאים:
   i. טווח.
   ii. שכיח.
   iii. חציון.

© כל הזכויות שמורות לברק קנדל

5) להלן מספר טענות, עבור כל טענה ציינו אם היא נכונה או לא נכונה ונמקו:

א. בסדרה שבה כל התצפיות שוות זו לזו השונות הינה 0.

ב. ציון התקן של החציון תמיד יהיה 0.

ג. ציון התקן של האחוזון ה-70 בהתפלגות אסימטרית ימנית (חיובית) תמיד יהיה חיובי.

ד. אם נוסיף תצפיות לסדרה של תצפיות, הדבר בהכרח יגדיל את הממוצע של הסדרה.

ה. בסדרה החציון הינו 80. הוספו שתי תצפיות אחת 79 ואחת 100 לכן החציון יגדל.

ו. אם נוסיף את הערך 4 לכל התצפיות אז סטיית התקן לא תשתנה.

ז. אם נחלק את כל התצפיות בהתפלגות ב-2 אז השונות תקטן פי 2.

ח. אם נגדיל את ממוצע המשכורות של עובדים בחברה אז גם השונות תגדל.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. המשתנה הנחקר: משקל תלמיד בק"ג, משתנה כמותי רציף.

ב. $\bar{X}=52$, $Md=X_{\frac{n+1}{2}}=X_3=58$, שכיח: 58.

ג. $S=10.12$, $R=28$.

ד. הוא חריג יותר בגובה כי שם ציון התקן בערך מוחלט יותר גבוה.

ה. הממוצע לא ישתנה אך סטיית התקן תקטן.

2) א. מספר הקורסים, בדיד. ב. התפלגות אסימטרית שמאלית.

ג. להלן טבלה: ד. שכיח: 5, טווח: 5.

x	$f(x)$
1	10
2	20
3	30
4	40
5	50
6	40
סה"כ	**190**

3) א. 20 תלמידים קיבלו ציון 6 ו-40 תלמידים קיבלו ציון 8.

ב. חציון: 7, שכיח: 8. ג. שונות: 2.533, סטיית תקן: 1.592.

4) א. 20%. ב. שביעות רצון (סדר).

ג. 2. ד. טווח: 4, שכיח: 2, חציון: 2.

5) א. נכון. ב. לא נכון. ג. לא נכון. ד. לא נכון. ה. לא נכון.

ו. נכון. ז. לא נכון. ח. לא נכון.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 14 - יסודות ההסתברות

תוכן העניינים

1. כללי 41

הגדרות יסודיות:

רקע:

ניסוי מקרי: תהליך לו כמה תוצאות אפשריות. התוצאה המתקבלת נודעת רק לאחר ביצוע התהליך. למשל: תוצאה בהטלת קובייה, מזג האוויר בעוד שבועיים.

מרחב מדגם: כלל התוצאות האפשריות בניסוי המקרי. לדוגמה, בהטלת קובייה: $\{1,2,3,4,5,6\}$, או: מזג האוויר בעוד שבועיים: {נאה, שרבי, מושלג, גשום, מעונן חלקית, אביך}.

מאורע: תת קבוצה מתוך מרחב במדגם. מסומן באותיות: A , B , C. בהטלת קובייה למשל, המאורע 'לקבל לפחות 5' יסומן: $A=\{5,6\}$. המאורע 'לקבל תוצאה זוגית' יסומן: $B=\{2,4,6\}$.

גודל מרחב המדגם: מספר התוצאות האפשריות במרחב המדגם. בהטלת קובייה למשל נקבל: $|\Omega|=6$.

גודל המאורע: מספר התוצאות האפשריות במאורע עצמו. למשל, בהטלת הקובייה האירועים הקודמים יסומנו: $|A|=2$, $|B|=3$.

מאורע משלים: מאורע המכיל את כל התוצאות האפשריות במרחב המדגם פרט לתוצאות במאורע אותו הוא משלים. למשל, בהטלת הקובייה: $\bar{A}=\{1,2,3,4\}$, $\bar{B}=\{1,3,5\}$.

מרחב מדגם אחיד (סימטרי): מרחב מדגם בו לכל התוצאות במרחב המדגם יש את אותה עדיפות , אותה סבירות למשל, קובייה הוגנת, אך לא כמו מזג האוויר בשבוע הבא.

© כל הזכויות שמורות לברק קנדל

הסתברות במרחב מדגם אחיד: במרחב מדגם אחיד הסיכוי למאורע יהיה: $P(A)=\frac{|A|}{|\Omega|}$.

דוגמה: מה הסיכוי בהטלת קובייה לקבל לפחות 5? $P(A)=\frac{|A|}{|\Omega|}=\frac{2}{6}$.

דוגמה: מה הסיכוי בהטלת קובייה לקבל תוצאה זוגית? $P(B)=\frac{|B|}{|\Omega|}=\frac{3}{6}$.

הסתברות במרחב לא אחיד: תחושב לפי השכיחות היחסית: $\frac{f}{n}$.

דוגמה:

להלן התפלגות הציונים בכיתה מסוימת:

x – הציון	f – מספר התלמידים – השכיחות
5	2
6	4
7	8
8	5
9	4
10	2

מה ההסתברות שתלמיד אקראי שניבחר בכיתה קיבל את הציון 8? $\frac{f}{n}=\frac{5}{25}=0.2$.

מה ההסתברות שתלמיד אקראי שניבחר בכיתה יכשל? $\frac{f}{n}=\frac{2}{25}=0.08$.

הסתברות למאורע משלים: הסתברות לקבוצת המשלים של המאורע ביחס למרחב המדגם: $P(\bar{A})=1-P(A)$. למשל, בדוגמה הקודמת הסיכוי לעבור את הבחינה יכול להיות מחושב לפי הסיכוי להיכשל: $P(A)=1-\frac{2}{25}=\frac{23}{25}$.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) מהאותיות E, F ו-G יש ליצור מילה בת 2 אותיות, לא בהכרח בת משמעות.

א. הרכיבו את כל המילים האפשריות.

ב. רשמו את המקרים למאורע:

i. במילה נמצאת האות E.

ii. במילה האותיות שונות.

ג. רשמו את המקרים למאורע $\bar{A}$.

2) מטילים זוג קוביות.

א. רשמו את מרחב המדגם של הניסוי. האם מרחב המדגם אחיד?

ב. רשמו את כל האפשרויות לאירועים הבאים:

i. סכום התוצאות 7.

ii. מכפלת התוצאות 12.

ג. חשבו את הסיכויים לאירועים שהוגדרו בסעיף ב'.

3) נבחר באקראי ספרה מבין הספרות 0-9.

א. מה ההסברות שהספרה שנבחרה גדולה מ-5?

ב. מה ההסתברות שהספרה שנבחרה היא לכל היותר 3?

ג. מה ההסתברות שהספרה שנבחרה היא אי זוגית?

4) להלן התפלגות מספר מקלטי הטלוויזיה עבור כל משפחה בישוב מסוים:

מספר משפחות	22	28	18	22	10
מספר מקלטים	0	1	2	3	4

נבחרה משפחה באקראי מהישוב.

א. מה ההסתברות שאין מקלטים למשפחה?

ב. מה ההסתברות שיש מקלטים למשפחה?

ג. מה ההסתברות שיש לפחות 3 מקלטים למשפחה?

5) להלן התפלגות מספר המכוניות למשפחה ביישוב "עדן":

מספר משפחות	20	40	100	30	10
מספר מכוניות	0	1	2	3	4

נבחרה משפחה אקראית מן הישוב.

א. מה ההסתברות שאין לה מכוניות?

ב. מה ההסתברות שבבעלות המשפחה לפחות 3 מכוניות?

ג. מה הסיכוי שבבעלותה פחות מ-3 מכוניות?

© כל הזכויות שמורות לברק קנדל

6) נטיל מטבע רגיל 3 פעמים. בצד אחד של המטבע מוטבע עץ ובצד השני פלי.

א. רשמו את מרחב המדגם של הניסוי. האם מרחב המדגם הוא אחיד?

ב. רשמו את כל האפשרויות לאירועים הבאים:

i. התקבל פעם אחת עץ.

ii. התקבל לפחות פלי אחד.

ג. מהו המאורע המשלים ל-D?

ד. חשבו את הסיכויים לאירועים שהוגדרו בסעיפים ב-ג.

תשובות סופיות:

1) א. $\Omega=\{EE,EF,EG,FE,FF,FG,GE,GF,GG\}$.

ב. $B\{EF,EG,FE,FG,GE,GF\}$, $A=\{EE,EF,EG,FE,GE\}$.

ג. $\bar{A}=\{FF,FG,GF,GG\}$.

2) א. $\Omega=\begin{Bmatrix}(1,1)&(2,1)&(3,1)&(5,1)&(4,1)&(6,1)\\(1,2)&(2,2)&(3,2)&(4,2)&(5,2)&(6,2)\\(1,3)&(2,3)&(3,3)&(4,3)&(5,3)&(6,3)\\(1,4)&(2,4)&(3,4)&(4,4)&(5,4)&(6,4)\\(1,5)&(2,5)&(3,5)&(4,5)&(5,5)&(6,5)\\(1,6)&(2,6)&(3,6)&(4,6)&(5,6)&(6,6)\end{Bmatrix}$.

ב. $C=\{(2,6),(3,4),(4,3),(6,2)\}$, $A=\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$.

ג. הסיכוי ל-A: $\frac{1}{6}$. הסיכוי ל-B: $\frac{1}{9}$.

3) א. 0.4 ב. 0.4 ג. 0.5.

4) א. 0.22 ב. 0.78 ג. 0.32.

5) א. 0.1 ב. 0.2 ג. 0.8.

6) א. $\Omega=\{PPP,PPE,PEP,EPP,PEE,EPE,EEP,EEE\}$.

ב. $D=\{PPP,PPE,PEP,EPP,PEE,EPE,EEP\}$, $A=\{PPE,PEP,EPP\}$.

ג. $\bar{D}=\{EEE\}$.

ד. $\frac{1}{8}$.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 15 - פעולות בין מאורעות (חיתוך ואיחוד) - מאורעות זרים ומכילים

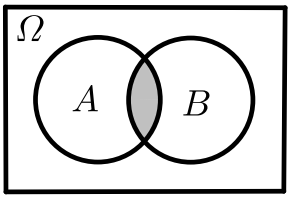
תוכן העניינים

1. כללי 45

פעולות בין מאורעות (חיתוך ואיחוד) – מאורעות זרים ומכילים:

רקע:

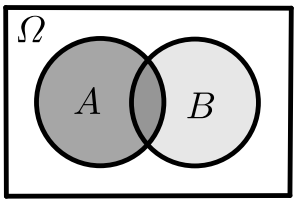
פעולת חיתוך:



נותנת את המשותף בין המאורעות הנחתכים.
חיתוך בין המאורע A למאורע B יסומן כך: $A\cap B$.
מדובר בתוצאות שנמצאות ב-A וגם ב-B.

דוגמה:
בהטלת קובייה, למשל, האפשרויות לקבל לפחות 5 הן: $A=\{5,6\}$.
האפשרויות לקבל תוצאה זוגית הן: $B=\{2,4,6\}$.
החיתוך שביניהם הוא: $A\cap B=\{6\}$.

פעולת איחוד:



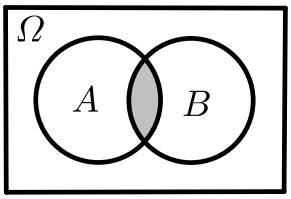
נותנת את כל האפשריות שנמצאות לפחות באחת מהמאורעות, ומסומנת: $A\cup B$.
הפעולה נותנת את אשר נמצא ב-A או B.
כלומר, לפחות אחד מהמאורעות קורה.

דוגמה:
בהטלת קובייה האפשרויות לקבל לפחות 5 הן: $A=\{5,6\}$.
האפשרויות לקבל תוצאה זוגית: $B=\{2,4,6\}$.
האפשרויות לקבל לפחות 5 וגם תוצאה זוגית: $A\cup B=\{2,4,5,6\}$.

דוגמה (הפתרון נמצא בהקלטה):
סטודנט ניגש בסמסטר לשני מבחנים. מבחן בסטטיסטיקה ומבחן בכלכלה. ההסתברות שלו לעבור את המבחן בסטטיסטיקה הוא 0.9, ההסתברות שלו לעבור את המבחן בכלכלה הוא 0.8 וההסתברות לעבור את המבחן בסטטיסטיקה ובכלכלה היא 0.75.
מה ההסתברות שלו לעבור את המבחן בסטטיסטיקה בלבד?
מה ההסתברות שלו להיכשל בשני המבחנים?
מה ההסתברות לעבור לפחות מבחן אחד?

© כל הזכויות שמורות לברק קנדל

נוסחת החיבור לשני מאורעות:



ההסתברות של איחוד מאורעות תחושב ע"י הקשר הבא:

$$P(A\cup B)=P(A)+P(B)-P(A\cap B)$$

חוקי דה מורגן לשני מאורעות:

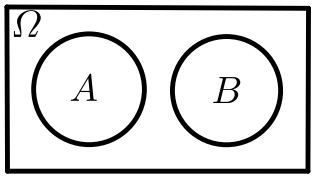
$$\overline{A\cup B}=\overline{A}\cap\overline{B}$$
$$\overline{A\cap B}=\overline{A}\cup\overline{B}$$
$$P(A\cap B)=1-P(\overline{A}\cup\overline{B})$$
$$P(A\cup B)=1-P(\overline{A}\cap\overline{B})$$

שיטת ריבוע הקסם:

השיטה רלבנטית רק אם יש שני מאורעות במקביל בדומה לתרגיל הקודם:

	$\overline{A}$	A	
B	$P(\overline{A}\cap B)$	$P(A\cap B)$	$P(B)$
$\overline{B}$	$P(\overline{A}\cap\overline{B})$	$P(A\cap\overline{B})$	$P(\overline{B})$
	$P(\overline{A})$	$P(A)$	1

מאורעות זרים:



מאורעות זרים הם כאשר אין להם אף איבר משותף:
$A\cap B=\{\ \}$. כלומר, הם לא יכולים להתרחש בו זמנית.

ההסתברות של חיתוך המאורעות היא אפס: $P(A\cap B)=0$.

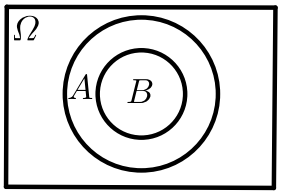
ההסתברות של איחוד המאורעות תחושב: $P(A\cup B)=P(A)+P(B)$.

דוגמה:

בהטלת קובייה, האפשרויות לקבל לפחות 5 הן: $A=\{5,6\}$ והאפשרות לקבל 3 היא: $B=\{3\}$, ולכן החיתוך ביניהם הוא אפס, כלומר: $A\cap B=\{\ \}$.

© כל הזכויות שמורות לברק קנדל

מאורעות מוכלים:



נתונים שני מאורעות A ו-B, השונים מאפס.
נאמר שהמאורע B מוכל במאורע A אם כל איברי המאורע B כלולים במאורע A ונרשום: $B\subset A$.

מאורע A מכיל את מאורע B כל התוצאות שנמצאות ב-B מוכלות בתוך מאורע A.

קשר זה מסומן באופן הבא: $B\subset A$.

$$A\cap B=B\ \ P(A\cap B)=P(B)$$
$$A\cup B=A\ \ P(A\cup B)=P(A)$$

למשל: $\begin{matrix}A=\{2,4,6\}\\B=\{2,4\}\end{matrix}$.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) מהאותיות E, F ו-G יוצרים מילה בת 2 אותיות – לא בהכרח בת משמעות.
נגדיר את המאורעות הבאים:
A - במילה נמצאת האות E.
B - במילה אותיות שונות.
א. רשמו את כל האפשרויות לחיתוך A עם B.
ב. רשמו את כל האפשרויות לאיחוד של A עם B.

2) תלמיד ניגש בסמסטר לשני מבחנים מבחן בכלכלה ומבחן בסטטיסטיקה.
נגדיר את המאורעות הבאים:
A - לעבור את המבחן בסטטיסטיקה.
B - לעבור את המבחן בכלכלה.
היעזרו בפעולות חיתוך, איחוד ומשלים בלבד כדי להגדיר את המאורעות הבאים וסמנו בדיאגראמת וון את השטח המתאים:
א. התלמיד עבר רק את המבחן בכלכלה.
ב. התלמיד עבר רק את המבחן בסטטיסטיקה.
ג. התלמיד עבר את שני המבחנים.
ד. התלמיד עבר לפחות מבחן אחד.
ה. התלמיד נכשל בשני המבחנים.
ו. התלמיד נכשל בכלכלה.

3) נתבקשתם לבחור ספרה באקראי. נגדיר את A להיות הספרה שנבחרה היא זוגית. נגדיר את B להיות הספרה שנבחרה קטנה מ-5.
א. רשמו את כל התוצאות למאורעות הבאים:
A , B , $\overline{B}$, $A\cap B$, $A\cup B$.
ב. חשבו את ההסתברויות לכל המאורעות מהסעיף הקודם.

4) נסמן ב-Ω את מרחב המדגם וב-ϕ קבוצה ריקה.
נתון כי A הינו מאורע בתוך מרחב המדגם.
להלן מוגדרים מאורעות שפתרונם הוא Ω או ϕ או A.
קבעו עבור כל מאורע מה הפתרון שלו:
$\overline{\overline{A}}$, $A\cap\phi$, $A\cup\phi$, $A\cap\Omega$, $A\cup\Omega$, $A\cap\overline{A}$, $\overline{\phi}$, $A\cup\overline{A}$.

© כל הזכויות שמורות לברק קנדל

5) הוגדרו המאורעות הבאים:
A - אדם שגובהו מעל 1.7 מטר
B - אדם שגובהו מתחת ל-1.8 מטר.
קבעו את גובהם של האנשים הבאים:
א. $A\cap B$.
ב. $A\cup B$.
ג. $\overline{A}\cap B$.
ד. $\overline{A}\cup\overline{B}$.
ה. $\overline{\overline{A}}$.

6) נגדיר את המאורעות הבאים:
A - אדם דובר עברית.
B - אדם דובר ערבית.
C - אדם דובר אנגלית.
השתמשו בפעולות איחוד, חיתוך והשלמה לתיאור המאורעות הבאים:
א. אדם דובר את כל שלוש השפות.
ב. אדם דובר רק עברית.
ג. אדם דובר לפחות שפה אחת מתוך השפות הללו.
ד. אדם אינו דובר אנגלית.
ה. קבוצת התלמידים שדוברים שתי שפות בדיוק (מהשפות הנ"ל).

7) שתי מפלגות רצות לכנסת הבאה. מפלגת "גדר" תעבור את אחוז החסימה בהסתברות של 0.08 ומפלגת "עתיד" תעבור את אחוז החסימה בהסתברות של 0.20. בהסתברות של 76% שתי המפלגות לא תעבורנה את אחוז החסימה.
א. מה ההסתברות שלפחות אחת מהמפלגות תעבור את אחוז החסימה?
ב. מה ההסתברות ששתי המפלגות תעבורנה את אחוז החסימה?
ג. מה ההסתברות שרק מפלגות "עתיד" תעבור את אחוז החסימה?

8) במקום עבודה מסוים 40% מהעובדים הם גברים. כמו כן, 20% מהעובדים הם אקדמאים. 10% מהעובדים הינן נשים אקדמאיות.
א. איזה אחוז מהעובדים הם גברים אקדמאיים?
ב. איזה אחוז מהעובדים הם גברים או אקדמאיים?
ג. איזה אחוז מהעובדים הם נשים לא אקדמאיות?

© כל הזכויות שמורות לברק קנדל

9) הסיכוי של מניה A לעלות הנו 0.5 ביום מסוים והסיכוי של מניה B לעלות ביום מסוים הנו 0.4. בסיכוי של 0.7 לפחות אחת מהמניות תעלה ביום מסוים.
חשבו את ההסתברויות הבאות לגבי שתי המניות הללו ביום מסוים:

א. ששתי המניות תעלנה.

ב. שאף אחת מהמניות לא תעלנה.

ג. שמניה A בלבד תעלה.

10) מטילים זוג קוביות, אדומה ושחורה. נגדיר את המאורעות הבאים:

A - בקובייה האדומה התקבלה התוצאה 4 ובשחורה 2.

B - סכום התוצאות משתי הקוביות הוא 6.

C - מכפלת התוצאות בשתי הקוביות היא 10.

א. האם A ו-B מאורעות זרים?

ב. האם המאורע B מכיל את המאורע A?

ג. האם A ו-C מאורעות זרים?

ד. האם A ו-C מאורעות משלימים?

11) עבור המאורעות A ו-B ידועות ההסתברויות הבאות: $P(A)=0.6$,
$P(B)=0.3$, $P(\bar{A}\cap\bar{B})=0.1$.

א. האם A ו-B מאורעות זרים?

ב. חשבו את $P(\bar{A}\cap B)$.

12) מטבע הוטל פעמיים. נגדיר את המאורעות הבאים:

A - קיבלנו עץ בהטלה הראשונה.

B - קיבלנו לפחות עץ אחד בשתי ההטלות.

איזו טענה נכונה?

א. A ו-B מאורעות זרים.

ב. A ו-B מאורעות משלימים.

ג. B מכיל את A.

ד. A מכיל את B.

13) בהגרלה חולקו 100 כרטיסים. על 3 מהם רשום חופשה ועל 2 מהם רשום מחשב שאר הכרטיסים ריקים. אדם קיבל כרטיס אקראי.

א. מה הסיכוי לזכות בחופשה או במחשב? האם המאורעות הללו זרים?

ב. מה ההסתברות לא לזכות בפרס?

© כל הזכויות שמורות לברק קנדל

14) נתון כי: $P(A)=0.3$, $P(B)=0.25$, $P(A\cup B)=0.49$

א. חשבו את הסיכוי ל-$P(A\cap B)$.

ב. האם A ו-B מאורעות זרים?

ג. מה ההסתברות שרק A יקרה או שרק B יקרה?

15) A ו-B מאורעות זרים. נתון ש: $P(\overline{A}\cap\overline{B})=P(A\cap\overline{B})=2\cdot P(B\cap\overline{A})$.

מה הסיכוי למאורע A ומה ההסתברות למאורע B?

16) קבעו אילו מהטענות הבאות נכונות:

א. $A\cap B=B\cap A$.

ב. $\overline{A\cup B}=A\cap B$.

ג. $A\cap B\cap C=A\cap B\cap(C\cup B)$.

ד. $\overline{A\cap B\cap C}=\overline{A}\cup\overline{B}\cup\overline{C}$.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. $A\cap B=\{EG,EF,FE,GE\}$.

ב. $A\cup B=\{EG,EF,EE,FE,GE,EG,GF\}$.

2) א. $B\cap\bar{A}$. ב. $A\cap\bar{B}$. ג. $A\cap B$. ד. $A\cup B$. ה. $\bar{A}\cap\bar{B}$. ו. $\bar{B}$.

3) א. $A=0,\ 2,\ 4,\ 6,\ 8$, $B=0,\ 1,\ 2,\ 3,\ 4$, $\bar{B}=5,\ 6,\ 7,\ 8,\ 9$,

$A\cap B=0,\ 2,\ 4$, $A\cup B=0,\ 2,\ 4,\ 6,\ 8,\ 1,\ 3$.

ב. $P(A)=0.5$, $P(B)=0.5$, $P(\bar{B})=0.5$, $P(A\cap B)=0.3$, $P(A\cup B)=0.7$.

4) $\bar{\bar{A}}=A$, $A\cap\phi=\phi$, $A\cup\phi=A$, $A\cap\Omega=A$, $A\cup\Omega=\Omega$,

$A\cap\bar{A}=\phi$, $\bar{\phi}=\Omega$, $A\cup\bar{A}=\Omega$.

5) א. $A\cap B$: גובה בין 1.7 ל-1.8. ב. $A\cup B$: כל גובה אפשרי .

ג. $\bar{A}=\bar{A}\cap B$: גובה לכל היותר 1.7. ד. $\bar{A}\cup\bar{B}$: לכל היותר 1.7 או לפחות 1.8.

ה. $\bar{\bar{A}}=A$: גובה מעל 1.7.

6) א. $A\cap B\cap C$. ב. $A\cap\bar{B}\cap\bar{C}$. ג. $A\cup B\cup C$.

ד. $\bar{C}$. ה. $(A\cap B\cap\bar{C})\cup(B\cap C\cap\bar{A})\cup(A\cap C\cap\bar{B})$.

7) א. $P(A\cup B)=0.24$. ב. $P(A\cap B)=0.04$. ג. $P(B\cap\bar{A})=0.16$.

8) א. $P(A\cap B)=10\%$. ב. $P(A\cup B)=50\%$. ג. $P(\bar{A}\cap\bar{B})=50\%$.

9) א. $P(A\cap B)=0.2$. ב. $P(\bar{A}\cap\bar{B})=0.3$. ג. $P(A\cup\bar{B})=0.3$.

10) א. לא. ב. כן. ג. כן. ד. לא.

11) א. כן. ב. $P(\bar{A}\cap B)=0.3$.

12) הטענה הנכונה היא ג'.

13) א. 0.05. ב. 0.95.

14) א. $P(A\cap B)=0.06$. ב. לא. ג. $P((A\cap\bar{B})\cup(B\cap\bar{A}))=0.43$.

15) $P(A)=\frac{2}{5}$, $P(B)=\frac{1}{5}$.

16) א. נכון. ב. לא נכון. ג. לא נכון. ד. נכון.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 16 - הסתברות מותנית-במרחב מדגם אחיד

תוכן העניינים

1. כללי 53

הסתברות מותנית – במרחב מדגם אחיד:

רקע:

לעיתים אנו נדרשים לחשב הסתברות למאורע כלשהו כאשר ברשותנו אינפורמציה לגבי מאורע אחר. הסתברות מותנית הינה סיכוי להתרחשות מאורע כלשהו כאשר ידוע שמאורע אחר התרחש/ לא התרחש.

ההסתברות של A בהינתן ש- B כבר קרה: $P(A|B)$.

כשמרחב המדגם אחיד: $P(A|B)=\frac{|A\cap B|}{|B|}$.

דוגמה (פתרון בהקלטה):

נטיל קובייה.

נגדיר:

A - התוצאה זוגית.

B - התוצאה גדולה מ-3.

נרצה לחשב את: $P(A|B)$.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) נבחרה ספרה זוגית באקראי. מה הסיכוי שהספרה גדולה מ-6?

2) יוסי הטיל קובייה.
מה הסיכוי שקיבל את התוצאה 4, אם ידוע שהתוצאה שהתקבלה זוגית?

3) הוטלו צמד קוביות. נגדיר:
A - סכום התוצאות בשתי ההטלות הינו 7.
B - מכפלת התוצאות 12.
חשבו את $P(A|B)$.

4) מטבע הוטל פעמיים.
ידוע שהתקבל לכל היותר ראש אחד, מה הסיכוי שהתקבלו שני ראשים?

5) זוג קוביות הוטלו והתקבל שהתוצאות זהות.
מה הסיכוי שלפחות אחת התוצאות 5?

6) זוג קוביות הוטלו והתקבל לפחות פעם אחת 4.
מה הסיכוי שאחת התוצאות 5?

7) נבחרה משפחה בת שני ילדים, שמהם אחד הוא בן.
מה ההסתברות שבמשפחה שני בנים בקרב הילדים?

8) נבחרה משפחה בת שלושה ילדים, ונתון שהילד האמצעי בן.
מה הסיכוי שיש בנות בקרב הילדים?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) 0.2.

2) $\frac{1}{3}$.

3) 0.5.

4) 0.

5) $\frac{1}{6}$.

6) $\frac{2}{11}$.

7) $\frac{1}{3}$.

8) $\frac{3}{4}$.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 17 - הסתברות מותנית - מרחב לא אחיד

תוכן העניינים

1. כללי 56

הסתברות מותנית – מרחב לא אחיד:

רקע:

הסיכוי שמאורע A יתרחש, בהינתן שמאורע B כבר קרה: $P(A|B)=\frac{P(A\cap B)}{P(B)}$.

במונה: הסיכוי לחיתוך של שני המאורעות, זה הנשאל וזה הנתון שהתרחש.
במכנה: הסיכוי למאורע נתון שהתרחש.

דוגמה (פתרון בהקלטה):

נבחרו משפחות שיש להם שתי מכוניות. ל-30% מהמשפחות הללו המכונית הישנה יותר היא מתוצרת אירופה ואצל 60% מהמשפחות הללו המכונית החדשה יותר מתוצרת אירופה. כמו כן, בקרב 15% מהמשפחות הללו שתי המכוניות הן מתוצרת אירופאית.
אם המכונית הישנה של המשפחה היא אירופאית, מה ההסתברות שגם החדשה אירופאית?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) תלמיד ניגש בסמסטר לשני מבחנים: מבחן בכלכלה ומבחן בסטטיסטיקה:
נגדיר את המאורעות הבאים:
A - לעבור את המבחן בסטטיסטיקה.
B - לעבור את המבחן בכלכלה.
כמו כן נתון שהסיכוי לעבור את המבחן בכלכלה הנו 0.8, הסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9 והסיכוי לעבור את שני המבחנים הנו 0.75.
חשבו את הסיכויים למאורעות הבאים:

א. התלמיד עבר בסטטיסטיקה, מה ההסתברות שהוא עבר בכלכלה?

ב. התלמיד עבר בכלכלה, מה ההסתברות שהוא עבר בסטטיסטיקה?

ג. התלמיד עבר בכלכלה, מה ההסתברות שהוא נכשל בסטטיסטיקה?

ד. התלמיד נכשל בסטטיסטיקה, מה ההסתברות שהוא נכשל בכלכלה?

ה. התלמיד עבר לפחות מבחן אחד, מה ההסתברות שהוא יעבור את שניהם?

2) במדינה שתי חברות טלפון סלולארי: "סופטי" ו"בלי". 30% מהתושבים הבוגרים רשומים אצל חברת "בלי", 60% מהתושבים הבוגרים רשומים אצל חברת "סופטי" ול-15% מהתושבים הבוגרים אין טלפון סלולארי כלל.

א. איזה אחוז מהתושבים הבוגרים רשומים אצל שתי החברות?

ב. נבחר אדם שרשום אצל חברת "סופטי", מה ההסתברות שהוא רשום גם אצל חברת "בלי"?

ג. אם אדם לא רשום אצל חברת "בלי", מה ההסתברות שהוא כן רשום בחברת "סופטי"?

ד. אם אדם רשום אצל חברה אחת בלבד, מה ההסתברות שהוא רשום בחברת "סופטי"?

3) במכללה שני חניונים: חניון קטן וחניון גדול. בשעה 08:00 יש סיכוי של 60% שבחניון הגדול יש מקום, סיכוי של 30% שבחניון הקטן יש מקום וסיכוי של 20% שבשני החניונים יש מקום.

א. מה ההסתברות שיש מקום בשעה 08:00 רק בחניון הגדול של המכללה?

ב. ידוע שבחניון הקטן יש מקום בשעה 08:00, מה הסיכוי שבחניון הגדול יש מקום?

ג. אם בשעה 08:00 בחניון הגדול אין מקום, מה ההסתברות שבחניון הקטן יהיה מקום?

ד. נתון שלפחות באחד מהחניונים יש מקום בשעה 08:00, מה ההסתברות שבחניון הגדול יש מקום?

© כל הזכויות שמורות לברק קנדל

4) נלקחו 200 שכירים ו-100 עצמאים. מתוך השכירים 20 הם אקדמאיים, ומתוך העצמאיים 30 הם אקדמאיים.

א. בנו טבלת שכיחות משותפת לנתונים.

ב. נבחר אדם אקראי מה ההסתברות שהוא שכיר?

ג. מה ההסתברות שהוא שכיר ולא אקדמאי?

ד. מה ההסתברות שהוא שכיר או אקדמאי?

ה. אם האדם שנבחר הוא עצמאי מהי ההסתברות שהוא אקדמאי?

ו. אם האדם שנבחר הוא לא אקדמאי, מה ההסתברות שהוא שכיר?

תשובות סופיות:

1) א. 0.833 ב. 0.9375 ג. 0.0625 ד. 0.5 ה. 0.789

2) א. 5% ב. 0.0833 ג. 0.786 ד. 0.6875

3) א. 0.4 ב. $\frac{2}{3}$ ג. 0.25 ד. $\frac{6}{7}$

4) א. להלן טבלה: ב. $\frac{2}{3}$ ג. 0.6 ד. $\frac{23}{30}$

	אקדמאי	לא אקדמאי	סה"כ
שכיר	20	180	200
עצמאי	30	70	100
סה"כ	50	250	300

ה. 0.3 ו. 0.72

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 18 - דיאגרמת עצים - נוסחת בייס ונוסחת ההסתברות השלמה

תוכן העניינים

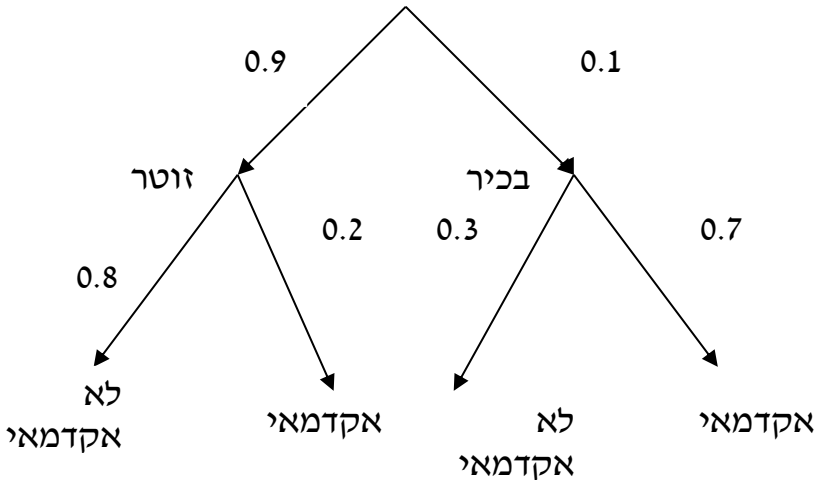
1. כללי 59

דיאגרמת עצים – נוסחת בייס ונוסחת ההסתברות השלמה:

נשתמש בשיטה זו כאשר יש תרגיל שבו התרחשות המאורעות היא בשלבים, כך שכל תוצאה של כל שלב תלויה בשלב הקודם, פרט לשלב הראשון:

דוגמה:

בחברה מסוימת 10% מוגדרים בכירים והיתר מוגדרים זוטרים. מבין הבכירים 70% הם אקדמאים ומבין הזוטרים 20% הם אקדמאים. נשרטט עץ שיתאר את הנתונים, השלב הראשון של העץ אינו מותנה בכלום ואילו השלב השני מותנה בשלב הראשון.



כדי לקבל את הסיכוי לענף מסוים נכפיל את כל ההסתברויות על אותו ענף.
נבחר אדם באקראי מאותה חברה.

1) מה הסיכוי שהוא בכיר אקדמאי? $0.1\cdot 0.7 = 0.07$.

2) מה הסיכוי שהוא זוטר לא אקדמאי? $0.9\cdot 0.8 = 0.72$.

כדי לקבל את הסיכוי לכמה ענפים נחבר את הסיכויים של כל ענף
(רק אחרי שבתוך הענף הכפלנו את ההסתברויות).

3) מה הסיכוי שהוא אקדמאי? $0.1\cdot 0.7 + 0.9\cdot 0.2 = 0.25$.

4) נבחר אקדמאי מה ההסתברות שהוא עובד זוטר?
מדובר כאן על שאלה בהסתברות מותנה ולכן נשתמש בעיקרון של הסתברות מותנה: $P(zutar \mid academay) = \frac{0.9\cdot 0.2}{0.25} = \frac{0.18}{0.25} = 0.72$.

© כל הזכויות שמורות לברק קנדל

נוסחת ההסתברות השלמה:

בהינתן B, מאורע כלשהו, וחלוקה של מרחב המדגם Ω ל- $A_1,\ldots,A_n$ כך ש- $\bigcup_i A_i = \Omega$,

אזי: $P(B)=\sum_{i=1}^{n} P(A_i)\cdot P\left(\frac{B}{A_i}\right)$.

נוסחת בייס:

$$P\left(\frac{A_j}{B}\right)=\frac{P(A_j)P\left(\frac{B}{A_j}\right)}{\sum^{n} P(A_i)\cdot P\left(\frac{B}{A_i}\right)}.$$

© כל הזכויות שמורות לברק קנדל

שאלות:

1) בשקית סוכריות 4 סוכריות תות ו-3 לימון. מוציאים באקראי סוכרייה. אם היא בטעם תות אוכלים אותה ומוציאם סוכרייה נוספת, ואם היא בטעם לימון מחזירים אותה לשקית ומוציאים סוכרייה נוספת.

א. מה ההסתברות שהסוכריה הראשונה שהוצאה בטעם תות והשנייה בטעם לימון?

ב. מה ההסתברות שהסוכריה השנייה בטעם לימון?

2) באוכלוסייה מסוימת 30% הם ילדים, 50% בוגרים והיתר קשישים. לפי נתוני משרד הבריאות הסיכוי שילד יחלה בשפעת במשך החורף הוא 80%, הסיכוי שמבוגר יחלה בשפעת במשך החורף הוא 40% והסיכוי שקשיש יחלה בשפעת במשך החורף הוא 70%.

א. איזה אחוז מהאוכלוסייה הינו קשישים שלא יחלו בשפעת במשך החורף?

ב. מה אחוז האנשים שיחלו בשפעת במשך החורף?

ג. נבחר אדם שחלה במשך החורף בשפעת, מה ההסתברות שהוא קשיש?

ד. נבחר ילד, מה ההסתברות שהוא לא יחלה בשפעת במשך החורף?

3) בכד א' 5 כדורים כחולים ו-5 כדורים אדומים. בכד ב' 6 כדורים כחולים ו-4 כדורים אדומים. בוחרים באקראי כד, מוציאים ממנו כדור ומבלי להחזירו מוציאים כדור נוסף.

א. מה ההסתברות ששני הכדורים שיוצאו יהיו בצבעים שונים?

ב. אם הכדורים שהוצאו הם בצבעים שונים, מה ההסתברות שהכדור השני שהוצא יהיה בצבע אדום?

4) חברת סלולר מסווגת את לקוחותיה לפי 3 קבוצות גיל: נוער, בוגרים ופנסיונרים. נתון כי: 10% מהלקוחות בני נוער, 70% מהלקוחות בוגרים והיתר פנסיונרים. מתוך בני הנוער 90% מחזיקים בסמארט-פון, מתוך האוכלוסייה הבוגרת ל-70% יש סמארט-פון ומתוך אוכלוסיית הפנסיונרים 30% מחזיקים בסמארט-פון.

א. איזה אחוז מלקוחות החברה הם בני נוער עם סמארט-פון?

ב. נבחר לקוח אקראי ונתון שיש לו סמארט-פון. מה ההסתברות שהוא פנסיונר?

ג. אם ללקוח אין סמארט-פון, מה ההסתברות שהוא לא בן נוער?

© כל הזכויות שמורות לברק קנדל

5) כדי להתקבל למקום עבודה יש לעבור שלושה מבחנים. המבחנים הם בשלבים, כלומר לאחר כישלון במבחן מסוים אין אפשרות לגשת למבחן הבא אחריו.
70% מהמועמדים עוברים את המבחן הראשון. מתוכם, 50% עוברים את המבחן השני. מבין אלה שעוברים את המבחן השני 40% עוברים את המבחן השלישי.

א. מה ההסתברות להתקבל לעבודה?

ב. מועמד לא התקבל לעבודה. מה ההסתברות שהוא נכשל במבחן הראשון?

ג. מועמד לא התקבל לעבודה. מה ההסתברות שהוא עבר את המבחן השני?

6) משרד הבריאות פרסם את הנתונים הבאים:
מתוך אוכלוסיית הילדים והנוער 80% חולים בשפעת בזמן החורף.
מתוך אוכלוסיית המבוגרים (עד גיל 65) 60% חולים בשפעת בזמן החורף.
30% מהתושבים הם ילדים ונוער. 50% הם מבוגרים. היתר קשישים.
כמו כן נתון ש68% מהאוכלוסייה תחלה בשפעת בחורף.

א. מה אחוז החולים בשפעת בקרב האוכלוסייה הקשישה?

ב. נבחר אדם שלא חלה בשפעת, מה ההסתברות שהוא לא קשיש?

7) רדאר שנמצא על החוף צריך לקלוט אנייה הנמצאת ב-1 מ-4 האזורים: A, B, C, D.
אם האנייה נמצאת באזור A הרדאר מזהה אותה בסיכוי 0.8, סיכוי זה פוחת ב-0.1 ככל שהאנייה מתקדמת באזור. כמו כן נתון שבהסתברות חצי האנייה נמצאת באזור D, בהסתברות 0.3 באזור C, באזור B היא נמצאת בסיכוי 0.2, אחרת היא נמצאת באזור A.

א. מה הסיכוי שהאנייה תתגלה ע"י הרדאר?

ב. אם האנייה התגלתה ע"י הרדאר, מה ההסתברות שהיא נמצאת באזור C?

ג. אם האנייה התגלתה ע"י הרדאר, מה הסיכוי שהיא לא נמצאת באזור B?

8) סימפטום X מופיע בהסתברות של 0.4 במחלה A, בהסתברות של 0.6 במחלה B ובהסתברות של 0.5 במחלה C. סימפטום X מופיע אך ורק במחלות הללו, אדם לא יכול לחלות ביותר ממחלה אחת מבין המחלות הללו. לקליניקה מגיעים אנשים כדלקמן: 8% חולים במחלה A, 10% במחלה B, 2% במחלה C והיתר בריאים.
כמו כן נתון שבמחלה A, סימפטום X מתגלה בסיכוי של 80%,
ובמחלות B, C הסימפטום מתגלה בסיכוי של 90% בכל מחלה.

א. מה ההסתברות שאדם הגיע לקליניקה וגילו אצלו את סימפטום X?

ב. אם התגלה אצל אדם סימפטום X, מה ההסתברות שהוא חולה במחלה A?

ג. אם לאדם יש את סימפטום X, מה ההסתברות שהוא חולה במחלה A?

ד. אם לא גילו אצל אדם את סימפטום X, מה ההסתברות שהוא בריא?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1)	א. $\frac{2}{7}$.	ב. $\frac{23}{49}$.		
2)	א. 6%.	ב. 58%.	ג. 0.241.	ד. 0.2.
3)	א. 0.544.	ב. 0.5.		
4)	א. 9%.	ב. 0.09375.	ג. 0.9722.	
5)	א. 0.14.	ב. 0.3488.	ג. 0.2442.	
6)	א. 70%.	ב. 0.8125.		
7)	א. 0.57.	ב. 0.3158.	ג. 0.7543.	
8)	א. 0.0886.	ב. 0.2889.	ג. 0.3137.	ד. 0.8778.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 19 - תלות ואי תלות בין מאורעות

תוכן העניינים

1. כללי 64

תלות ואי תלות בין מאורעות:

רקע:

אם מתקיים ש: $P(B|A)=P(B)$, נגיד שמאורע B בלתי תלוי ב-A.

הדבר גורר גם ההפך: $P(A|B)=P(A)$, כלומר, גם A אינו תלוי ב-B.

כשהמאורעות בלתי תלויים מתקיים ש: $P(A\cap B)=P(A)\cdot P(B)$.

הוכחה לכך: $P(A/B)=P(A)=\frac{P(A\cap B)}{P(B)}\Rightarrow P(A\cap B)=P(A)\cdot P(B)$.

נשתמש בנוסחאות של מאורעות בלתי תלויים רק אם נאמר במפורש שהמאורעות בלתי תלויים בתרגיל או שמההקשר אפשר להבין ללא צל של ספק שהמאורעות בלתי תלויים.

למשל,

חוקר מבצע שני ניסויים בלתי תלויים הסיכוי להצליח בניסוי הראשון הנו 0.7 והסיכוי להצליח בניסוי השני הוא 0.4.

א. מה הסיכוי להצליח בשני הניסויים יחדו?

כיוון שהמאורעות הללו בלתי תלויים:

$$P(A\cap B)=P(A)\cdot P(B)=0.7\cdot 0.4=0.28$$

ב. מה הסיכוי להיכשל בשני הניסויים?

באופן דומה:

$$P(\bar{A}\cap\bar{B})=P(\bar{A})\cdot P(\bar{B})=(1-0.7)(1-0.4)=0.18$$

הרחבה: אי תלות בין n מאורעות:

n מאורעות $A_1,\ldots,A_n$ הם בלתי תלויים אם ורק אם: $P\left(\bigcap_{i=1}^{n}A_i\right)=\prod_{i=1}^{n}P(A_i)$.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) נתון: $P(A\cup B)=0.6$, $P(B)=0.5$, $P(A)=0.2$.
האם המאורעות הללו בלתי תלויים?

2) תלמיד ניגש לשני מבחנים שהצלחתם לא תלויה זו בזו.
הסיכוי שלו להצליח במבחן הראשון הוא 0.7 והשני 0.4.
א. מה הסיכוי להצליח בשני המבחנים יחדו?
ב. מה הסיכוי שניכשל בשני המבחנים?

3) במדינה מסוימת יש 8% אבטלה, נבחרו באקראי שני אנשים מהמדינה.
א. מה ההסתברות ששניהם מובטלים?
ב. מה ההסתברות שלפחות אחד מהם מובטל?

4) מוצר צריך לעבור בהצלחה ארבע בדיקות בלתי תלויות לפני שיווקו, אחרת הוא נפסל ולא יוצא לשוק. הסיכוי לעבור בהצלחה כל אחת מהבדיקות הוא 0.8. בכל מקרה מבוצעות כל 4 הבדיקות.
א. מה הסיכוי שהמוצר יפסל?
ב. מה ההסתברות שהמוצר יעבור בהצלחה לפחות בדיקה אחת?

5) במדינה מסוימת יש 8% אבטלה, נבחרו באקראי חמישה אנשים מהמדינה.
א. מה ההסתברות שכולם מובטלים במדגם?
ב. מה ההסתברות שלפחות אחד מהם מובטל?

תשובות סופיות:

1) כן.
2) א. 0.28. ב. 0.18.
3) א. 0.0064. ב. 0.1536.
4) א. 0.5904. ב. 0.9984.
5) א. 0.08^5. ב. 0.3409.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 20 - שאלות מסכמות בהסתברות

תוכן העניינים

1. כללי 66

שאלות מסכמות בהסתברות:

שאלות:

1) נלקחו משפחות שיש להם שתי מכוניות. ל-30% מהמשפחות הללו המכונית הישנה יותר היא מתוצרת אירופה ואצל 60% מהמשפחות הללו המכונית החדשה יותר מתוצרת אירופה. כמו כן 15% מהמשפחות הללו שתי המכוניות הן מתוצרת אירופאית.

א. מה ההסתברות שמשפחה אקראית בת שתי מכוניות תהיה ללא מכוניות מתוצרת אירופה?

ב. מה ההסתברות שלפחות מכונית אחת תהיה אירופאית?

ג. ידוע שלמשפחה יש מכונית אירופאית.
מה ההסתברות שרק המכונית החדשה שלה היא מתוצרת אירופאית?

ד. אם המכונית הישנה של המשפחה היא אירופאית, מה ההסתברות שגם החדשה אירופאית?

2) במדינת "שומקום" 50% מהחלב במרכולים מיוצר במחלבה א', 40% במחלבה ב' והיתר במחלבה ג'. 3% מתוצרת מחלבה א' מגיעה חמוצה למרכולים ואילו במחלבה ב' 10%.
כמו כן ידוע שבמדינת "שומקום" בסך הכול 7.5% מהחלב חמוץ.

א. איזה אחוז מהחלב שמגיע למרכול ממחלבה ג' חמוץ?

ב. אם נרכש חלב חמוץ במרכול. מה הסיכוי שהוא יוצר במחלבה ג'?

ג. ברכישת חלב נמצא שהוא אינו חמוץ. מה הסיכוי שהוא יוצר במחלבה א'?

ד. האם המאורעות: "חלב חמוץ" ו-"יוצר במחלבה א'" בלתי תלויים?

3) רוני ורונה יצאו לבלות במרכז בילוים עם מספר אפשרויות בילוי:
בהסתברות של 0.3 הם ייצאו לבאולינג, בהסתברות של 0.5 הם ייצאו לבית קפה ובהסתברות של 0.7 הם יצאו לפחות לאחד מהם (באולינג/קפה).

א. מה ההסתברות שהם יצאו רק לבאולינג?

ב. האם המאורעות "לצאת לבאולינג" ו-"לצאת לבית קפה" זרים?

ג. האם המאורעות "לצאת לבאולינג" ו-"לצאת לבית קפה" תלויים?

ד. מה ההסתברות שיום אחד הם יצאו רק לבאולינג וביום למחרת לא יצאו לאף אחד מהמקומות?

© כל הזכויות שמורות לברק קנדל

4) 70% מהנבחנים בסטטיסטיקה עוברים את מועד א'. כל מי שלא עובר את מועד א' ניגש לעשות מועד ב', מתוכם 80% עוברים אותו. מבין אלה שנכשלים בשני המועדים 50% נרשמים לקורס מחדש, והיתר פורשים מהתואר.

א. מה הסיכוי שסטודנט אקראי עבר את הקורס?

ב. אם סטודנט אקראי עבר הקורס, מה הסיכוי שעבר במועד ב'?

ג. מה אחוז הסטודנטים שפורשים מהתואר?

ד. נבחרו 2 סטודנטים אקראיים רונית וינאי, מה ההסתברות שרונית עברה במועד א' ושינאי עבר במועד ב'?

5) באוכלוסייה מסוימת 40% הם גברים והיתר הן נשים. מבין הגברים 10% מובטלים. בסך הכול 13% מהאוכלוסייה מובטלת.

א. מה אחוז האבטלה בקרב הנשים?

ב. נבחר אדם מובטל, מה ההסתברות שזו אישה?

ג. נגדיר את המאורעות הבאים: A - נבחר אדם מובטל, B - נבחר גבר. האם המאורעות הללו זרים? והאם הם בלתי תלויים?

6) בתיבה 10 מטבעות, מתוכם 7 מטבעות רגילים (ראש, זנב) ו-3 מטבעות שבשני צדדיהם טבוע ראש. אדם בוחר באקראי מטבע ומטיל אותו פעמיים.
נסמן ב-A את ההטלה הראשונה ראש, וב-B את ההטלה השנייה ראש.

א. חשבו את הסיכויים למאורעות A ו-B.

ב. האם המאורע A ו-B בלתי תלויים?

ג. ידוע שבהטלה הראשונה התקבל ראש, מה ההסתברות שהמטבע שהוטל הוא מטבע הוגן?

7) ערן מעוניין למכור את רכבו והוא מפרסם מודעה באינטרנט ומודעה בעיתון. מבין אלה שמעוניינים לרכוש רכב משומש 30% יראו את המודעה באינטרנט, 50% יראו את המודעה בעיתון ו-72% יראו את המודעה בלפחות אחת מהמדיות.

א. מה אחוז האנשים, מאלה שמעוניינים לרכוש רכב משומש, שיראו את 2 המודעות?

ב. אם אדם ראה את המודעה באינטרנט, מה ההסתברות שהוא לא ראה את המודעה בעיתון?

ג. האם המאורעות: "לראות את המודעה באינטרנט" ו-"לראות את המודעה בעיתון" בלתי תלויים?

ד. אדם שראה את המודעה באינטרנט בלבד יתקשר לערן בהסתברות של 0.7, אם הוא ראה את המודעה בעיתון בלבד הוא יתקשר לערן בהסתברות של 0.6 ואם הוא ראה את שתי המודעות הוא יתקשר לערן בהסתברות של 0.9.

i. מה ההסתברות שאדם המעוניין לרכוש רכב משומש יתקשר לערן?

ii. אדם המעוניין לרכוש רכב משומש התקשר לערן. מה ההסתברות שהוא ראה את שתי המודעות?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. 0.25 ב. 0.75 ג. 0.6 ד. 0.5

2) א. 0.2 ב.0.267 ג. 0.524 ד. תלויים.

3) א. 0.2 ב. אינם זרים. ג. תלויים. ד. 0.06

4) א. 0.94 ב. 0.255 ג. 0.03 ד. 0.168

5) א. 15% ב. 0.692 ג. לא זרים ותלויים.

6) א. 0.65 ב. תלויים. ג. 0.5384

7) א. 8% ב. 0.733 ג. תלויים. ד. i. 0.478

ד. ii. 0.15

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 21 - המשתנה המקרי הבדיד - פונקציית ההסתברות

תוכן העניינים

1. כללי 69

המשתנה המקרי הבדיד – פונקציית ההסתברות:

רקע:

משתנה מקרי בדיד:

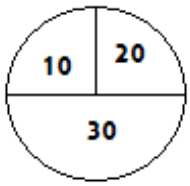
משתנה מקרי בדיד הינו משתנה היכול לקבל כמה ערכים בודדים בהסתברויות שונות.
מתארים את המשתנה המקרי על ידי פונקציית הסתברות.

פונקציית הסתברות:

פונקציה המתאימה לכל ערך אפשרי של המשתנה את ההסתברות שלה.
סכום ההסתברויות על פונקציית ההסתברות חייב להיות 1.

דוגמה (פתרון בהקלטה):

בקזינו יש רולטה כמתואר בשרטוט:



אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה ב-₪.
בנו את פונקציית ההסתברות של סכום הזכייה במשחק בודד.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) ידוע שביישוב מסוים התפלגות מספר המכוניות למשפחה היא:
50 משפחות אינן מחזיקות במכונית.
70 משפחות עם מכונית אחת.
60 משפחות עם 2 מכוניות.
20 משפחות עם 3 מכוניות .
בוחרים באקראי משפחה מהיישוב, נגדיר את X להיות מספר המכוניות של המשפחה שנבחרה. בנו את פונקציית ההסתברות של X.

2) מהאותיות: A , B , C יוצרים קוד דו תווי.
א. כמה קודים ניתן ליצור?
ב. רשמו את כל הקודים האפשריים.
ג. נגדיר את X להיות מספר הפעמים שהאות B מופיעה בקוד.
בנו את פונקציית ההסתברות של X.

3) תלמיד ניגש בסמסטר לשני מבחנים: מבחן בכלכלה ומבחן בסטטיסטיקה.
כמו כן, נתון שהסיכוי לעבור את המבחן בכלכלה הנו 0.8, הסיכוי לעבור את המבחן בסטטיסטיקה הנו 0.9 והסיכוי לעבור את שני המבחנים הנו 0.75.
יהי X מספר המבחנים שהסטודנט עבר. בנו את פונקציית ההסתברות של X.

4) הסיכוי לזכות במשחק מסוים הינו 0.3. אדם משחק את המשחק עד אשר הוא מנצח אך בכל מקרה הוא לא משחק את המשחק יותר מ-4 פעמים.
נגדיר את X להיות מספר הפעמים שהוא שיחק את המשחק.
בנו את פונקציית ההסתברות של X.

5) חברה לניהול פרויקטים מנהלת 3 פרויקטים במקביל. הסיכוי שפרויקט א' יצליח הינו 0.7, הסיכוי שפרויקט ב' יצליח הינו 0.8, והסיכוי שפרויקט ג' יצליח הינו 0.9. נתון שהצלחת כל פרויקט בלתי תלויה זו בזו. נגדיר את X להיות מספר הפרויקטים שיצליחו. בנו את פונקציית ההסתברות של X.

6) להלן פונקציית הסתברות של משתנה מקרי כלשהו: $k=1,2...4$, $P(X=k)=\frac{k}{A}$.
מצאו את ערכו של A.

7) בגן ילדים 8 ילדים, מתוכם 5 בנים ו-3 בנות.
בוחרים באקראי 3 ילדים להשתתף בהצגה.
נגדיר את X כמספר הבנים שנבחרו להצגה.
בנו את פונקציית ההסתברות של X.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) להלן טבלה:

X	0	1	2	3
$P(X)$	0.25	0.35	0.3	0.1

2) להלן טבלה:

X	0	1	2
$P(X)$	$\frac{4}{9}$	$\frac{4}{9}$	$\frac{1}{9}$

3) להלן טבלה:

X	0	1	2
$P(X)$	0.05	0.20	0.75

4) להלן טבלה:

X	1	2	3	4
$P(X)$	0.3	0.21	0.147	0.343

5) להלן טבלה:

X	0	1	2	3
$P(X)$	0.006	0.092	0.398	0.504

6) .10

7) להלן טבלה:

X	1	2	3	4
$P(X)$	$\frac{1}{56}$	$\frac{15}{56}$	$\frac{30}{56}$	$\frac{10}{56}$

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 22 - המשתנה המקרי הבדיד - תוחלת - שונות וסטיית תקן

תוכן העניינים

1. כללי 72

המשתנה המקרי הבדיד – תוחלת, שונות וסטיית תקן:

רקע:

תוחלת:

ממוצע של פונקציית ההסתברות, אם נבצע את התהליך אינסוף פעמים כמה בממוצע נקבל. התוחלת היא צפי של המשתנה המקרי.

מגדירים תוחלת באופן הבא: $E(X)=\sum_i x_i P(x_i)=\mu$.

שונות:

תוחלת ריבועי הסטיות מהתוחלת – נותן אינדיקציה על הפיזור והסיכון של פונקציית ההסתברות.

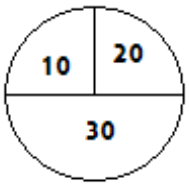
מגדירים שונות באופן הבא: $V(X)=\sum_i (x_i-\mu)^2 P(x_i)=\sum_i x_i^2 P(x_i)-\mu^2=\sigma^2$.

סטיית תקן:

שורש של השונות – הפיזור הממוצע הצפוי סביב התוחלת. מסמנים: $STD=\sigma$.

דוגמה:

בקזינו רולטה כמוראה בשרטוט. אדם מסובב את הרולטה וזוכה בסכום הרשום על הרולטה ב-₪. הסתברות לקבלת הסכומים השונים:



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

$$E(X)=10\cdot 0.25+20\cdot 0.25+30\cdot 0.5=22.5=\mu$$

$$V(X)=\sum_i (x_i-\mu)^2 P(x_i)=$$

$$=(10-22.5)^2\cdot 0.25+(20-22.5)^2\cdot 0.25+(30-22.5)^2\cdot 0.5=68.75=\sigma^2$$

כדי לחשב את סטיית התקן נוציא שורש לשונות: $\sigma_x=\sqrt{V(X)}=\sqrt{68.75}=8.29$.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) אדם משחק במשחק מזל.
נגדיר את X להיות סכום הזכייה.
להלן פונקציית ההסתברות של X:

X	-30	0	20	40
$P(X)$	0.4	0.1	0.3	0.2

מהי התוחלת, השונות וסטית התקן של X?

2) בישוב מסוים שני סניפי בנק: בנק פועלים ובנק לאומי. מתוך האוכלוסייה הבוגרת בישוב, ל-50% חשבון בנק בסניף הפועלים, ל-40% חשבון בנק בסניף לאומי ול-20% מהתושבים הבוגרים אין חשבון באף אחד מהסניפים.
יהי X מס' סניפי הבנק שלבוגר בישוב יש בהם חשבון.
חשבו את: $E(X)$.

3) ידוע של- 20% מהמשפחות יש חיבור לווייני בביתם. בסקר אדם מחפש לראיין משפחה המחוברת ללוויין. הוא מטלפן באקראי למשפחה וממשיך עד אשר הוא מגיע למשפחה המחוברת ללוויין. בכל מקרה הסוקר לא יתקשר ליותר מ-5 משפחות. נגדיר את X להיות מספר המשפחות שאליהן האדם יתקשר.

א. בנו את פונקציית ההסתברות של X.

ב. חשבו את התוחלת וסטיית התקן של X.

4) לאדם צרור מפתחות. בצרור 5 מפתחות אשר רק אחד מתאים לדלת של ביתו. האדם מנסה את המפתחות באופן מקרי. לאחר שניסה מפתח מסוים הוא מוציא אותו מהצרור כדי שלא ישתמש בו שוב.
נסמן ב-X את מספר הניסיונות עד שהדלת תפתח.

א. בנו את פונקציית ההסתברות של X.

ב. חשבו את התוחלת והשונות של X.

© כל הזכויות שמורות לברק קנדל

5) נתונה פונקציית ההסתברות של המשתנה המקרי X:

X	2	4	6	8
$E(X)$		0.3		0.2

כמו כן נתון ש: $E(X)=4.2$.

א. מצאו את ההסתברויות החסרות בטבלה.

ב. חשבו את: $V(X)$.

6) משתנה מקרי בדיד מקבל את הערכים 0-5 ו-5-.
נתון שהתוחלת של המשתנה 0 ושהשונות היא 10.
מצא את פונקציית ההסתברות.

תשובות סופיות:

1) תוחלת: 2 , שונות: 796.

2) 0.9.

3) א. ראו סרטון. ב. תוחלת: 3.36 , סטיית תקן: 1.603.

4) א. ראו טבלה: ב. תוחלת: 3 , שונות: 2.

X	1	2	3	4	5
$P(X)$	0.2	0.2	0.2	0.2	0.2

5) א. ראו טבלה: ב. 5.16.

X	2	4	6	8
$P(X)$	0.4	0.3	0.1	0.2

6) ראו טבלה:

X	-5	0	5
$P(X)$	0.2	0.6	0.2

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 23 - המשתנה המקרי הבדיד- טרנספורמציה לינארית

תוכן העניינים

1. כללי 75

המשתנה המקרי הבדיד – טרנספורמציה לינארית:

רקע:

טרנספורמציה לינארית היא מצב שבו מבצעים הכפלה של קבוע ו/או הוספה של קבוע על המשתנה המקורי (כולל גם חלוקה של קבוע והחסרה של קבוע).

בניסוח מתמטי נאמר כי אם משתנה אקראי Y מיוצג ע"י משתנה אקראי X כאשר a,b הם קבועים כלשהם: $Y=aX+b$, אזי מתקיימים:

1) $E(Y)=aE(X)+b$.
2) $V(Y)=a^2\cdot V(X)$.
3) $\sigma_Y=|a|\sigma_x$.

שלבי העבודה:

1) נזהה שמדובר בטרנספורמציה ליניארית (שינוי קבוע לכל התצפיות).
2) נרשום את כלל הטרנספורמציה לפי נתוני השאלה.
3) נפשט את הכלל ונזהה את ערכי a ו- b.
4) נציב בנוסחאות שלעיל בהתאם למדדים שנשאלים.

דוגמה – הרולטה:

בהמשך לנתוני שאלת הרולטה נתון שעלות השתתפות במשחק 15 ₪.
מהי התוחלת והשונות של הרווח במשחק?

פתרון (בהקלטה):

חישבנו קודם ש: $V(X)=68.75=\sigma^2$, $E(X)=22.5=\mu$.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) סטודנט ניגש ל-5 קורסים הסמסטר. נניח שכל קורס שסטודנט מסיים מזכה אותו ב-4 נקודות אקדמאיות. חשבו את התוחלת והשונות של סך הנקודות שיצבור הסטודנט כאשר נתון שתוחלת מספר הקורסים שיסיים היא 3.5 עם שונות 2.

2) תוחלת סכום הזכייה במשחק מזל הינה 10 עם שונות 3. הוחלט להכפיל את סכום הזכייה במשחק. עלות השתתפות במשחק הינה 12.
מה התוחלת ומהי השונות של הרווח במשחק?

3) תוחלת של משתנה מקרי הינה 10 וסטית התקן 5. הוחלט להוסיף 2 למשתנה ולאחר מכן להעלות אותו ב-10%.
מהי התוחלת ומהי סטיית התקן לאחר השינוי?

4) X הינו משתנה מקרי. כמו כן נתון ש- $E(X)=4$ ו- $V(X)=3$.
Y הינו משתנה מקרי חדש, עבורו: $Y=7-X$. חשבו את: $E(Y)$ ו- $V(Y)$.

5) אדם החליט לבטח את רכבו; שווי הרכב 100,000 ₪ להלן התביעות האפשריות והסתברותן: בהסתברות של 0.001 תהיה תביעה טוטאלוסט (כל שווי הרכב).
בהסתברות של 0.02 תהיה תביעה בשווי מחצית משווי הרכב.
בהסתברות של 5% תהיה תביעה בשווי רבע משווי הרכב.
אחרת אין תביעה בכלל. החברה מאפשרת תביעה אחת בשנה.
נסמן ב-X את גובה התביעה השנתית, באלפי₪.

א. בנו את פונקציית ההסתברות של X.
ב. חשבו את התוחלת והשונות של גובה התביעה.
ג. פרמיית הביטוח היא 4,000 ₪.

מהי התוחלת ומהי השונות של רווח חברת הביטוח לביטוח הרכב הנ"ל?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) תוחלת: 14, שונות: 32.

2) תוחלת: 8, שונות: 12.

3) תוחלת: 13.2, סטיית תקן: 5.5.

4) תוחלת: 3, שונות: 3.

5) א. להלן טבלה: ב. תוחלת: 2350, שונות: $85,727.5^2$.

X	100	50	25	0
$P(X)$	0.001	0.02	0.05	0.929

ג. תוחלת: 1650, שונות: $85,727.5^2$.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 24 - תוחלת ושונות של סכום משתנים מקריים

תוכן העניינים

1. כללי 78

תוחלת ושונות של סכום משתנים מקריים:

רקע:

אם: $X_1, X_2, ..., X_n$ משתנים מקרים אזי:

$$E(T) = E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n).$$

אם: $X_1, X_2, ..., X_n$ משתנים מקריים בלתי תלויים בזוגות, אזי:

$$V(T) = V(X_1 + X_2 + ... + X_n) = V(X_1) + V(X_2) + ... + V(X_n).$$

דוגמה:

אדם משחק בשני משחקי מזל בלתי תלויים. תוחלת סכות הזכייה של המשחק הראשון היא 7 עם סטיית תקן 3. תוחלת סכום הזכייה של המשחק השני היא 2- עם סטיית תקן 4.
מה התוחלת ומהי השונות של סכום הזכייה הכולל של שני המשחקים יחד?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) הרווח ממניה א' הוא עם תוחלת של 5 ושונות 10.
הרווח ממניה ב' הוא עם תוחלת של 4 ושונות.
ידוע שההשקעות של שתי המניות בלתי תלויות זו בזו.
מה התוחלת והשונות של הרווח הכולל מהשקעה בשתי המניות יחד?

2) X ו-Y הם משתנים בלתי תלויים, סטיית התקן של X היא 3.
סטיית התקן של Y היא 4. מהי סטיית התקן של X+Y?

3) אדם משחק בשני משחקי מזל בלתי תלויים זה בזה:
X - סכום הזכייה במשחק הראשון.
Y - סכום הזכייה במשחק השני.
נתון:

$$E(x)=10 \quad , \quad \sigma(X)=3$$

$$E(y)=12 \quad , \quad \sigma(Y)=4$$

מהי התוחלת ומהי סטיית התקן של סכום הזכייה בשני המשחקים?

4) ברולטה הסיכוי לזכות ב- 30 ₪ הוא חצי, ב-10 ₪ רבע וכך גם ב-20 ₪.
מה היא התוחלת והשונות של סכום הזכייה הכולל לאדם המשחק ברולטה 4 פעמים?

5) נתון משתנה מקרי בעל פונקציית ההסתברות הבאה:

$$P(X=K)=\frac{A}{K-1} \quad , \quad K=2,3,4,5$$

ת	אחר
	0

מצאו את ערכו של A.

א. חשבו את התוחלת והשונות של X.

ב. נלקחו n משתנים מקריים בלתי תלויים מההתפלגות הנ"ל.
בטאו באמצעות n את תוחלת והשונות של סכום המשתנים.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) תוחלת: 9, שונות: 15.

2) 5.

3) תוחלת: 22, שונות: 5.

4) תוחלת: 90, שונות: 275.

5) א. $A=\frac{12}{25}=0.48$. ב. תוחלת: 2.92, שונות: 1.1136.

ג. תוחלת: 2.92, שונות: $1.1136n$.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 25 - התפלגויות בדידות מיוחדות -התפלגות בינומית

תוכן העניינים

1. כללי 81

התפלגויות בדידות מיוחדות – התפלגות בינומית:

רקע:

נגדיר את המושג ניסוי ברנולי:
ניסוי ברנולי הנו ניסוי שיש לו שתי תוצאות אפשריות: "הצלחה" ו"כישלון".
למשל מוצר פגום או תקין, אדם עובד או מובטל, עץ או פלי בהטלת מטבע וכדומה.
בהתפלגות בינומית חוזרים על אותו ניסוי ברנולי n פעמים באופן בלתי תלוי זה בזה.
מגדירים את X להיות מספר ההצלחות שהתקבלו בסך הכול. נסמן ב- P את הסיכוי להצלחה בניסוי בודד, וב- Q את הסיכוי לכישלון בניסוי בודד.
ואז נגיד ש: $X \sim B(n,p)$.

פונקציית ההסתברות של X:

לכל $k=0,1,2,....,n$ $\quad P(X=K)=\binom{n}{k}p^k(1-p)^{n-k}$

כאשר: $0!=1$; $n!=n\cdot(n-1)\cdot(n-2)\cdot....\cdot 1$; $\binom{n}{k}=\frac{n!}{k!(n-k)!}$

לגודל: $\binom{n}{k}$ ניתן לחשב באמצעות המחשבון.

תוחלת: $E(X)=np$.

שונות: $V(X)=npq$.

שימו לב, כדי לזהות שמדובר בהתפלגות בינומית צריכים להתקיים כל התנאים הבאים:

1) חוזרים על <u>אותו ניסוי ברנולי</u> באופן <u>בלתי תלוי</u> זה בזה.
2) חוזרים על הניסוי n פעמים.
3) X – מוגדר <u>כמספר ההצלחות המתקבלות</u> בסך הכול.

דוגמה (פתרון בהקלטה):

במדינה מסוימת ל-80% מהתושבים יש רישיון נהיגה.
נבחרו 10 תושבים אקראיים מהמדינה.

א. מה ההסתברות שבדיוק ל-9 מהם יש רישיון נהיגה?
ב. מה ההסתברות שלפחות ל-9 מהם יש רישיון נהיגה?
ג. מהי התוחלת ומהי סטיית התקן של מספר התושבים שנדגמו ושיש להם רישיון נהיגה?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) במדינה 10% מהאוכלוסייה מובטלת. נבחרו 5 אנשים באקראי מאותה אוכלוסייה. נגדיר את X להיות מספר המובטלים שהתקבלו במדגם.

א. מהי ההתפלגות של X ?

ב. מה ההסתברות שיהיה בדיוק מובטל אחד?

ג. מה ההסתברות שכולם יעבדו במדגם?

ד. מה ההסתברות ששלושה יעבדו במדגם?

ה. מה ההסתברות שלפחות אחד יהיה מובטל?

ו. מה תוחלת ומהי השונות של מספר המובטלים במדגם?

2) על פי נתוני משרד התקשורת ל-70% מהאוכלוסייה יש סמארטפון. נבחרו 10 אנשים באקראי. נגדיר את X כמספר האנשים שנדגמו עם סמארטפון.

א. מהי ההתפלגות של X ? הסבירו.

ב. מה ההסתברות שבמדגם ל-8 אנשים יש סמארט-פון?

ג. מה ההסתברות שבמדגם לפחות ל-9 יהיו סמארט-פון?

ד. מה התוחלת ומה סטיית התקן של מספר האנשים שנדגמו ולהם סמארט-פון?

3) בבית הימורים יש שורה של 6 מכונות מזל מאותו סוג. משחק במכונת מזל כזו עולה 5 ₪. ההסתברות לזכות ב-20₪ בכל אחת מהמכונות היא 0.1 וההסתברות להפסיד את ההשקעה היא 0.9 בכל מכונה. מהמר נכנס לבית ההימורים ומכניס 5 ₪ לכל אחת מ-6 המכונות.

א. מה ההסתברות שיפסיד בכל המכונות?

ב. מה ההסתברות שיזכה בדיוק בשתי מכונות?

ג. מה ההסתברות שיזכה ביותר כסף מה-30 ₪ שהשקיע?

ד. מהן התוחלת וסטיית התקן של הרווח נטו של המהמר (הזכיות בניכוי ההשקעה)?

4) במדינה מסוימת התפלגות ההשכלה בקרב האוכלוסייה מעל גיל 30 היא כזו:

השכלה	נמוכה	תיכונית	תואר I	תואר II ומעלה
פרופורציה	0.1	0.6	0.2	0.1

נבחרו 20 אנשים אקראיים מעל גיל 30.

א. מה ההסתברות ש-5 מהם אקדמאים?

ב. מה התוחלת של מס' בעלי ההשכלה הנמוכה?

© כל הזכויות שמורות לברק קנדל

5) במכללה מסוימת 20% מהסטודנטים גרים בת"א. מבין הסטודנטים שגרים בת"א 30% מגיעים ברכבם, ומבין הסטודנטים שלא גרים בת"א 50% מגיעים ברכבם למכללה.

א. השומר בשער המכללה בודק לכל סטודנט את תיקו בהיכנסו למכללה. מה ההסתברות שבקרב 5 סטודנטים שנבדקו ע"י השומר רק 1 מתוכם הגיע למכללה ברכבו?

ב. בהמשך לסעיף הקודם מה ההסתברות שרוב הסטודנטים בקרב ה-5 הגיעו למכללה ברכבם?

6) במבחן אמריקאי 20 שאלות. סטודנט ניגש למבחן והסיכוי שהוא יודע שאלה כלשהי הוא 0.8. אם הוא לא יודע הוא מנחש את התשובה.
לכל שאלה 4 תשובות אפשריות שרק אחת מהן נכונה.

א. מה הסיכוי לענות על שאלה מסוימת נכון?

ב. מה הסיכוי שיענה נכונה על בדיוק 16 שאלות?

ג. על כל שאלה שענה נכון התלמיד מקבל 5 נקודות, על כל שאלה ששגה מופחתת נקודה, מה התוחלת ומהי השונות של ציון התלמיד?

7) 5% מקו היצור פגום. המוצרים נארזים בתוך קופסת קרטון. בכל קופסא 10 מוצרים שונים. הקופסאות נארזות בתוך מכולה. בכל מכולה 20 קופסאות.

א. מה ההסתברות שבקופסא אקראית לפחות מוצר פגום אחד?

ב. מה התוחלת ומהי סטיית התקן של מספר הקופסאות במכולה בהן לפחות מוצר פגום אחד?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. $X \sim B(n=5, p=0.1)$. ב. 0.32805. ג. 0.59049.
ד. 0.0729. ה. 0.40954. ו. תוחלת: 0.5, שונות: 0.45.

2) ב. 0.2335. ג. 0.1493. ד. תוחלת: 7, סטיית תקן: 1.449.

3) א. 0.5314. ב. 0.0984. ג. 0.1143.
ד. תוחלת: 18-, סטיית תקן: 14.697.

4) א. 0.1789. ב. 2.

5) א. 0.1956. ב. 0.4253.

6) א. 0.85. ב. 0.182. ג. תוחלת: 82, שונות: 91.8.

7) א. 0.401. ב. תוחלת: 8.025, סטיית תקן: 2.193.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 26 - התפלגויות רציפות מיוחדות - התפלגות נורמלית

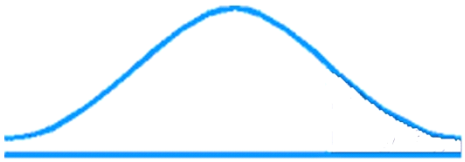
תוכן העניינים

1. כללי 85

התפלגויות רציפות מיוחדות – התפלגות נורמלית:

רקע:

התפלגות נורמלית הינה התפלגות של משתנה רציף. ישנם משתנים רציפים מסוימים שנהוג להתייחס אליהם כנורמליים כגון: זמן ייצור, משקל תינוק ביום היוולדו ועוד. פונקציית הצפיפות של ההתפלגות הנורמלית נראית כמו פעמון:



לעקומה זו קוראים גם עקומת גאוס ועקומה אחת נבדלת מהשנייה באמצעות הממוצע וסטיית התקן שלה.

אלה הם הפרמטרים שמאפיינים את ההתפלגות: $X \sim N\left(\mu,\sigma^2\right)$.

נוסחת פונקציית הצפיפות: $f\left(x\right)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

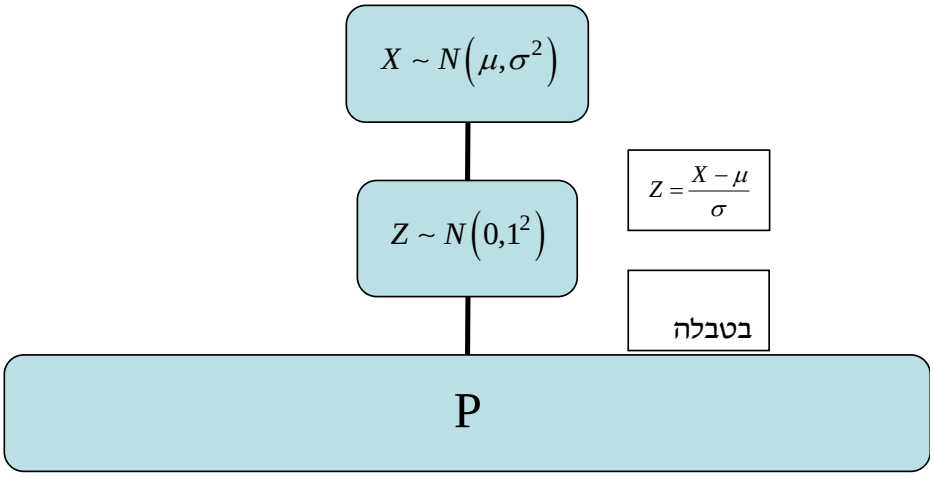
כדי לחשב הסתברויות בהתפלגות נורמלית יש לחשב את השטחים הרלוונטיים שמתחת לעקומה. כדי לחשב שטחים אלה נמיר כל התפלגות נורמלית להתפלגות נורמלית סטנדרטית על ידי תהליך הנקרא תקנון. התפלגות נורמלית סטנדרטית היא התפלגות נורמלית שהממוצע שלה הוא אפס וסטיית התקן היא אחת, והיא תסומן באות Z: $Z \sim N\left(0,1^2\right)$.

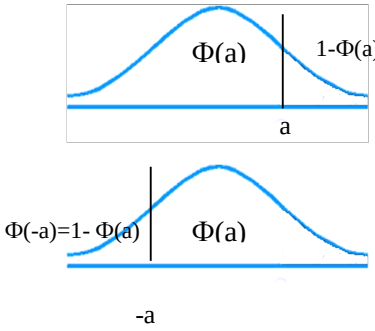
תהליך התקנון מבוצע על ידי הנוסחה הבאה: $Z=\frac{X-\mu}{\sigma}$.

אחרי תקנון מקבלים ערך הנקרא ציון תקן.
ציון התקן משמעו בכמה סטיות תקן הערך סוטה מהממוצע.

לאחר חישוב ציון התקן של ערך מסוים נעזרים בטבלה של ההתפלגות הנורמלית הסטנדרטית לחישוב השטח הרצוי, ובאופן כללי נתאר את הסכמה הבאה:

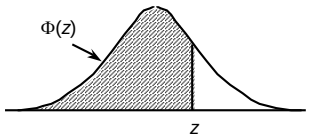
© כל הזכויות שמורות לברק קנדל





© כל הזכויות שמורות לברק קנדל

טבלת ההתפלגות המצטברת הנורמלית סטנדרטית – ערכי $\Phi(z)$:



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

z	1.282	1.645	1.960	2.326	2.576	3.090	3.291	3.891	4.417
Φ(z)	0.90	0.95	0.975	0.99	0.995	0.999	0.9995	0.99995	0.999995

© כל הזכויות שמורות לברק קנדל

דוגמה (הפתרון בהקלטה):

משקל חפיסות שוקולד המיוצרות בחברה מתפלג נורמלית עם ממוצע 100 גרם בסטיית תקן של 8 גרם.

1) מה אחוז חפיסות השוקולד ששוקלות מתחת ל-110 גרם?
2) מה אחוז חפיסות השוקולד השוקלות מעל 110 גרם?
3) מה אחוז חפיסות השוקולד השוקלות מתחת ל 92 גרם?
4) מהו המשקל ש-90% מהחפיסות בקו הייצור שוקלים פחות מהם?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) הגובה של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 170 ס״מ וסטית תקן של 10 ס״מ.

א. מה אחוז האנשים שגובהם מתחת ל-182.4 ס״מ?

ב. מה אחוז האנשים שגובהם מעל 190 ס״מ?

ג. מה אחוז האנשים שגובהם בדיוק 173.6 ס״מ?

ד. מה אחוז האנשים שגובהם מתחת ל-170 ס״מ?

ה. מה אחוז האנשים שגובהם לכל היותר 170 ס״מ?

2) נתון שהזמן שלוקח לתרופה מסוימת להשפיע מתפלג נורמלית עם ממוצע של 30 דקות ושונות של 9 דקות רבועות.

א. מהי פרופורציית המקרים בהן התרופה תעזור אחרי יותר משעה?

ב. מה אחוז מהמקרים שבהן התרופה תעזור בין 35 ל-37 דקות?

ג. מה הסיכוי שהתרופה תעזור בדיוק תוך 36 דקות?

ד. מה שיעור המקרים שבהן ההשפעה של התרופה תסטה מ-30 דקות בפחות מ-3 דקות?

3) המשקל של אנשים באוכלוסייה מסוימת מתפלג נורמלית עם ממוצע של 60 ק״ג וסטיית תקן של 8 ק״ג.

א. מה אחוז האנשים שמשקלם נמוך מ-55 ק״ג?

ב. מהי פרופורציית האנשים באוכלוסייה שמשקלם לפחות 50 ק״ג?

ג. מהי השכיחות היחסית של האנשים באוכלוסייה שמשקלם בין 60 ל-70 ק״ג?

ד. לאיזה חלק מהאוכלוסייה משקל הסוטה מהמשקל הממוצע בלא יותר מ-4 ק״ג?

ה. מה הסיכוי שאדם אקראי ישקול מתחת ל-140 ק״ג?

4) משקל תינוקות ביום היוולדם מתפלג נורמלית עם ממוצע של 3300 גרם וסטיית תקן 400 גרם.

א. מצאו את העשירון העליון.

ב. מצאו את האחוזון ה-95.

ג. מצאו את העשירון התחתון.

© כל הזכויות שמורות לברק קנדל

5) ציוני מבחן אינטליגנציה מתפלגים נורמלית עם ממוצע 100 ושונות 225.

א. מה העשירון העליון של הציונים במבחן האינטיליגנציה?

ב. מה העשירון התחתון של ההתפלגות?

ג. מהו הציון ש-20% מהנבחנים מקבלים מעליו?

ד. מהו האחוזון ה-20?

ה. מהו הציון ש-5% מהנבחנים מקבלים מתחתיו?

6) נפח משקה בבקבוק מתפלג נורמלית עם סטיית תקן של 20 מ״ל, ונתון ש-33% מהבקבוקים בעלי נפח שעולה על 508.8 מ״ל.

א. מה ממוצע נפח משקה בבקבוק?

ב. 5% מהבקבוקים המיוצרים עם הנפח הגבוה ביותר נשלחים לבדיקה, החל מאיזה נפח שולחים בקבוק לבדיקה?

ג. 1% מהבקבוקים עם הנפח הקטן ביותר נתרמים לצדקה, מהו הנפח המקסימלי לצדקה?

7) אורך חיים של מכשיר מתפלג נורמלית. ידוע שמחצית מהמכשירים חיים פחות מ-500 שעות, כמו כן ידוע ש-67% מהמכשירים חיים פחות מ-544 שעות.

א. מהו ממוצע אורך חיי מכשיר?

ב. מהי סטית בתקן של אורך חיי מכשיר?

ג. מה הסיכוי שמכשיר אקראי יחיה פחות מ-460 שעות?

ד. מהו המאיון העליון של אורח חיי מכשיר?

ה. 1% מהמכשירים בעלי אורך החיים הקצר ביותר נשלח למעבדה לבדיקה מעמיקה. מהו אורך החיים המקסימלי לשליחת מכשיר למעבדה?

8) להלן שלוש התפלגויות נורמליות של שלוש קבוצות שונות ששורטטו באותה מערכת צירים. ההתפלגויות מוספרו כדי להבדיל ביניהן.

א. לאיזו התפלגות הממוצע הגבוה ביותר?

ב. במה מבין המדדים הבאים התפלגות 1 ו-2 זהות?

i. בעשירון העליון.

ii. בממוצע.

iii. בשונות.

ג. לאיזו התפלגות סטיית התקן הקטנה ביותר?

i. 1.

ii. 2.

iii. 3.

iv. אין לדעת.

© כל הזכויות שמורות לברק קנדל

9) הזמן שלוקח לאדם להגיע לעבודתו מתפלג נורמלית עם ממוצע של 40 דקות וסטית תקן של 5 דקות.

א. מה ההסתברות שמשך הנסיעה של האדם לעבודתו יהיה לפחות שלושת רבעי השעה?

ב. אדם יצא לעבודתו בשעה 08:10 מביתו. הוא צריך להגיע לעבודתו בשעה 09:00. מה הסיכוי שיאחר לעבודתו?

ג. אם ידוע שזמן נסיעתו לעבודה היה יותר משלושת רבעי השעה. מה ההסתברות שזמן הנסיעה הכולל יהיה פחות מ-50 דקות?

ד. מה הסיכוי שבשבוע (חמישה ימי עבודה) בדיוק פעם אחת יהיה זמן הנסיעה לפחות שלושת רבעי השעה?

10) ההוצאה החודשית לבית אב בעיר "טרירה" מתפלגת נורמלית עם ממוצע של 2000 דולר וסטית תקן של 300 דולר. בחרו באקראי 5 בתי אב. ההסתברות שלפחות אחד מהם מוציא בחודש מעל ל- T דולר היא 0.98976.

א. מה ערכו של T?

ב. מה הסיכוי שההוצאה החודשית של בית אב בעיר תהיה לפחות סטיית תקן אחת מעל T?

ג. מסתבר שנפלה טעות בנתונים, ויש להוסיף 100 דולר להוצאות החודשית של כל בתי האב בעיר. לאור זאת, מה ההסתברות שההוצאה החודשית של בית אב נמוכה מ-1800 דולר?

11) אורך שיר אקראי המשודר ברדיו מתפלג נורמלית עם תוחלת של 3.5 דקות וסטיית תקן של שלושים שניות.

א. מה ההסתברות שאורך של שיר אקראי המנוגן ברדיו יהיה בין 3 ל-2.5 דקות?

ב. מהו הטווח הבין רבעוני של אורך שיר המשודר ברדיו?

ג. ביום מסוים מנוגנים 200 שירים ברדיו. כמה שירים מתוכם תצפה שיהיו באורך הנמוך מ-3.5 דקות?

ד. בשעה מסוימת שודרו 8 שירים. מה ההסתברות שרבע מהם בדיוק היו ארוכים מ-4 דקות והיתר לא?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1)	א. 89.25%	ב. 2.28%	ג. 0	ד. 50% ה. 50%
2)	א. 0%	ב. 3.76%	ג. 0%	ד. 68.26%
3)	א. 26.43%	ב. 89.44%	ג. 39.44%	ד. 0.383
	ה. 100%			
4)	א. 3812.8	ב. 3958	ג. 2787.2	
5)	א. 119.2	ב. 80.8	ג. 112.6	ד. 87.4
6)	א. 500	ב. 532.9	ג. 453.48	
7)	א. 500	ב. 100	ג. 0.3446	ד. 733
	ה. 267			
8)	א. 3	ב. בממוצע.	ג. 1	
9)	א. 0.1587	ב. 0.0228	ג. 0.8563	ד. 0.3975
10)	א. 1925	ב. 0.2266	ג. 0.1587	
11)	א. 0.1359	ב. 0.675	ג. 100	ד. 0.25

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

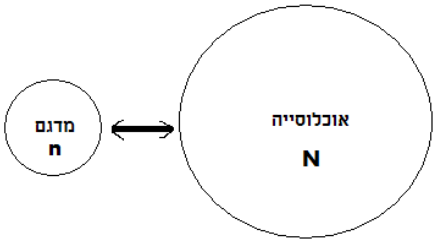
פרק 27 - הסקה סטטיסטית - הקדמה

תוכן העניינים

1. כללי 93

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, אנו מוגבלים בזמן ובאמצעים טכניים ולכן מבצעים מדגם במטרה לבצע הסקה סטטיסטית מהמדגם לאוכלוסייה.
הדגימה בקורס תהיה דגימה מקרית - הכוונה לדגימה שבה לכל תצפית באוכלוסייה יש את אותו סיכוי להיכלל במדגם.

סטטיסטי:

גודל המחושב על המדגם.

פרמטר:

גודל המתאר את האוכלוסייה.

© כל הזכויות שמורות לברק קנדל

הסימונים לפרמטר וסטטיסטי הם שונים:

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

פרמטר הוא גודל קבוע גם אם אנו 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?

© כל הזכויות שמורות לברק קנדל

שאלות:

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?

2) להלן התפלגות מספר מקלטי הטלוויזיה למשפחה בישוב "העוגן".
נגדיר את X להיות מספר המקלטים של משפחה אקראית. מתכננים לדגום מאוכלוסייה זו 4 משפחות ולהתבונן בממוצע מספר מקלטי הטלוויזיה במדגם.

א. מיהי האוכלוסייה ומהו המשתנה הנחקר?

ב. מהו הסטטיסטי שיילקח מהמדגם ומה סימונו?

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

3) נתון כי 20% מהשכירים במדינה הם אקדמאיים. נבחרו באקראי 10 שכירים באותה אוכלוסייה ומתכננים לפרסם את מספר האקדמאיים שנדגמו.

א. מהי האוכלוסייה?

ב. מה המשתנה באוכלוסייה?

ג. מהם הפרמטרים?

ד. מהו הסטטיסטי?

תשובות סופיות:

1) א. כלל הסטודנטים במכללה שסיימו סטטיסטיקה א. ב. ציון.
ג. ממוצע: 78, סטיית תקן: 15. ד. 2.

2) א. האוכלוסייה: 1000 משפחות בישוב העוגן, המשתנה הנחקר: מס' מקלטים.
ב. $\bar{X}$ = ממוצע מדגם.

3) א. השכירים במדינה. ב. השכלה: אקדמאי, לא אקדמאי.
ג. שיעור ההצלחות באוכלוסייה: 0.2. ג. מס' האקדמאים במדגם.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 28 - התפלגות הדגימה ומשפט הגבול המרכזי

תוכן העניינים

1. התפלגות ממוצע המדגם ומשפט הגבול המרכזי 96
2. התפלגות מספר ההצלחות במדגם - קירוב נורמלי להתפלגות הבינומית 103
3. התפלגות פרופורציית ההצלחות במדגם 108

התפלגות 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.

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. מיהי אוכלוסיית המחקר? מהו המשתנה הנחקר?
ב. מהם הפרמטרים של האוכלוסייה?
ג. מה התוחלת ומהי סטית התקן של ממוצע המדגם?

© כל הזכויות שמורות לברק קנדל

דגימה מהתפלגות נורמאלית:

אם נדגום מתוך אוכלוסייה שהמשתנה בה מתפלג נורמאלית עם ממוצע μ ושונות σ^2.

ממוצע המדגם גם יתפלג נורמאלית: $Z_{\bar{x}}=\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$, $\bar{x}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$.

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?

משפט הגבול המרכזי:

אם אוכלוסייה מתפלגת כלשהו עם ממוצע μ ושונות σ^2 אזי עבור מדגם מספיק גדול $(n\geq 30)$ ממוצע המדגם מתפלג בקירוב נורמאלי: $\bar{x}\overset{\cdot}{\sim} N\left(\mu,\frac{\sigma^2}{n}\right)$.

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?

© כל הזכויות שמורות לברק קנדל

שאלות:

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. מיהי 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 הטלוויזיה למשפחה בישוב מסוים:

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

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?

© כל הזכויות שמורות לברק קנדל

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 מתחולת 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 הפתרון 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. בקבוקיי היין שבארגז נמזגים לקערה עם קיבולת של שלושה ליטר. מה ההסתברות שהיין יגלוש מהקערה?

8) משתנה מתפלג נורמאלית עם תוחלת 80 וסטיית תקן 4.

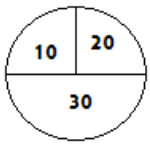
א. מה ההסתברות שממוצע המדגם יסטה מתוחלתו בלא יותר מיחידה כאשר גודל המדגם הוא 9?

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

ג. הסבר את ההבדל בתשובות של שני הסעיפים.

© כל הזכויות שמורות לברק קנדל

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 בחמשת 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) קובייה הוטלה 50 פעמים.
מה ההסתברות שהממוצע של התוצאות יהיה לפחות 3.72?

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. היעזרו במשפט הגבול המרכזי.

13) נתון משתנה מקרי בדיד בעל פונקציית ההסתברות הבאה:

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

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

© כל הזכויות שמורות לברק קנדל

14) נתון ש-$X \sim N\left(\mu,\sigma^2\right)$. דגמו 5 תצפיות מאותה התפלגות והתבוננו בממוצע המדגם $\bar{X}$ לכן: $P\left(\bar{X}>\mu\right)$ יהיה (בחרו 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 (בחרו בתשובה הנכונה):

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

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

ג. $\bar{X} \sim N\left(\mu,\sigma^2\right)$.

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

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

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

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. כלל הסטודנטים במכללה שסיימו סטטיסטיקה א. ב. ציון.
ג. ממוצע: 78, סטיית תקן: 15. ד. 2.
ה. 78. ו. 10.6.

2) א. להלן טבלה:

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

ב. $\mu = 2.05$, $\sigma^2 = 0.9475$, $\sigma = 0.973$

ג. $\mu_{\bar{x}} = 2.05$, $\sigma_{\bar{x}}^2 = 0.2369$, $\sigma\left(\overline{X}\right) = 0.486$.

3) $\mu_{\bar{x}} = 3.5$, $\sigma\left(\overline{X}\right) = 1.21$.

4) א. 0.8413. ב. 0.0013. ג. 0. ד. 0.1974.

5) א. 0. ב. 0. ג. 0.9544. ד. 178.205.

6) א. 0.0465. ב. 27.71. ג. 0.2628. ד. התשובה הייתה קטנה.

7) א. 0. ב. 0.1587. ג. 0.1587. ד. 0.5.

8) א. 0.5468. ב. 0.6826.

9) א. להלן טבלה:

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

ב. התוחלת: 22.5, השונות: 68.75. ג. התוחלת: 22.5, השונות: 13.75.
ד. 0.8997.

10) 0.0475.

11) 0.1814.

12) א. 0.9772. ב. 0.0228. ג. 271.

13) 0.5.

14) ב'.

15) ד'.

16) ג'.

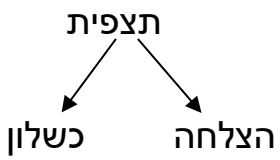
© כל הזכויות שמורות לברק קנדל

התפלגות מספר ההצלחות במדגם – קירוב נורמלי להתפלגות הבינומית:

רקע:

תזכורת על התפלגות בינומית:

בפרק זה נדון בהתפלגות מספר ההצלחות במדגם אקראי (תצפיות בלתי תלויות זו בזו).
את 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: $p(y=k)=\binom{n}{k}p^k q^{n-k}$,

תוחלת: $E(y)=np$.

שונות: $V(y)=npq$.

קירוב נורמלי עבור התפלגות בינומית:

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

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

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

אז: $Z_y = \dfrac{y-np}{\sqrt{npq}}$.

© כל הזכויות שמורות לברק קנדל

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. $p(Y=a) \cong p\left(a-\frac{1}{2} \le Y \le a+\frac{1}{2}\right)$.
2. $P(Y \le a) \cong P(Y \le a+0.5)$.
3. $P(Y \ge a) \cong P(Y \ge a-0.5)$.

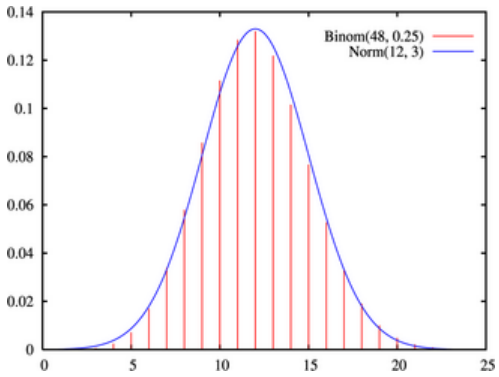
הערות:

- התנאים למעבר מבינומי לנורמלי הם נזילים, כלומר משתנים ממרצה אחד לשני. התנאי שהצגתי כאן הוא הפופולרי ביותר:
  1. $n \cdot p \ge 5$.
  2. $n \cdot (1-p) \ge 5$.
- ישנם מרצים שנותנים את התנאי המחמיר הבא:
  1. $n \cdot p \ge 10$.
  2. $n \cdot (1-p) \ge 10$.
- וישנם מרצים שהתנאי שהם נותנים הוא: $(n \ge 30)$.
- 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) בפתרונות שאציג תמיד אבצע תיקון רציפות במעבר מבינומי לנורמלי כיוון שכך הפתרון יהיה יותר מדויק (בכל מקרה שהמדגמים גדולים העניין זניח).

© כל הזכויות שמורות לברק קנדל

דוגמה (הפתרון בהקלטה):
נתון שבקרב אוכלוסיית הנוער 25% זקוקים למשקפיים. נדגמו באקראי 48 בני נוער.

1. מה הסיכוי שבדיוק 14 מתוכם יהיו זקוקים למשקפיים?
2. מה הסיכוי שלכל היותר 13 מתוכם זקוקים למשקפיים?



© כל הזכויות שמורות לברק קנדל

שאלות:

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 שהסיכוי 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?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. 0.201. ב. 0.3758. ג. התוחלת: 2, סטיית התקן: 1.2649.
2) א. 0.9332. ב. 0.6915.
3) 0.1611.
4) א. 0.9406. ב. בינומית - 0.0788, קירוב לנורמלית - 0.0778.
5) א. 0.015. ב. 398.
6) 0.9996.
7) 0.8643.

© כל הזכויות שמורות לברק קנדל

התפלגות פרופורציית ההצלחות במדגם:

רקע:

בפרק זה נדון בהתפלגות הדגימה של פרופורציית המדגם.
Y - מספר ההצלחות במדגם (למשל, מספר המובטלים במדגם).

$\hat{p}=\frac{y}{n}$ - פרופורציית ההצלחות במדגם.

למשל, שיעור המובטלים במדגם - $n=200$:
מספר המובטלים : $Y=20$.

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

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:

$$E\left(\hat{p}\right)=p\ \ ,\ V\left(\hat{P}\right)=\frac{pq}{n}\ .$$

משפט הגבול המרכזי עבור הפרופורציה המדגמית:

אם : $np\geq 5\ \&\ nq\geq 5$, אזי : $\hat{p}\sim N\left(p,\frac{pq}{n}\right)$. $Z_{\hat{p}}=\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}}$.

הערות:

- התנאים לקרוב הנורמאלי הם נזילים, כלומר משתנים ממרצה אחד לשני.
  התנאי שהצגתי כאן הוא הפופולרי ביותר:
  1. $n\cdot p\geq 5$.
  2. $n\cdot(1-p)\geq 5$.
- ישנם מרצים שנותנים את התנאי המחמיר הבא:
  1. $n\cdot p\geq 10$.
  2. $n\cdot(1-p)\geq 10$.
- וישנם מרצים המשתמשים בתנאי : $\left(n\geq 30\right)$.
- תאלצו לבדוק מהו התנאי שנתנו לכם בכיתה כדי לעבור לנורמלית.

© כל הזכויות שמורות לברק קנדל

- 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.

א. מה ההסתברות שהשכיחות היחסית $\left(\hat{p}\right)$ של הזכאים לבגרות במדגם תעלה על 60%?

ב. מה ההסתברות שפרופורציית הזכאים לבגרות במדגם תעלה על 70%?

© כל הזכויות שמורות לברק קנדל

שאלות:

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 היות 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 סמאטרפון?

ג. מה ההסתברות שפרופורציית בעלי הסמרטפון 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 האמיתית ביותר מ-4%?

ג. כמה בתי אב יש לדגום כדי שהסטייה בין הפרופורציה המדגמית לפרופורציה האמיתית 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.
ב"מועדון ה-700" נכללים נבחנים שמקבלים ציון מעל 700 בפסיכומטרי.
מה הסיכוי שבמועד בו נבחנו 2000 נבחנים אקראיים יהיו לפחות 3% המשתייכים למועדון?

© כל הזכויות שמורות לברק קנדל

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

א. הוכיחו ש: $V(\hat{P}) = \frac{P(1-P)}{n}$, $E(\hat{P}) = p$.

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

תשובות סופיות:

1) א. התוחלת: 0.1, השונות: 0.00064. ב. 0.5. ג. 0.3446.

2) א. 0.0618. ב. 0.0618. ג. 0.8238.

3) א. 0.5. ב. 0. ג. 0.8968. ד. גדלה.

4) א. 0.0062. ב. 0.0456. ג. 481. ד. 0.77436.

5) 0.0154.

6) א. שאלת הוכחה. ב. 0.5.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 29 - מושגי יסוד באמידה

תוכן העניינים

1. כללי 112

מושגי יסוד 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-$\hat{\theta}$. $\hat{\theta}$ הוא סטטיסטי המחושב על המדגם ובאמצעותו נאמוד את θ.
- שגיאת אמידה: $\left|\hat{\theta}-\theta\right|$ - 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.

א. מה הפרמטר בדוגמה זו?

ב. מהי טעות האמידה של ערוץ 10?

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

© כל הזכויות שמורות לברק קנדל

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 אומר חסר הטיה 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 אומר חסר הטיה 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 הטיה 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: 3, 1, 2, 5, 4, 1, 2, 3, 1, 2.
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.

© כל הזכויות שמורות לברק קנדל

שאלות:

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 האומד 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 האומד?

ג. מהו האומדן לפרמטר?

ד. מהי טעות האמידה?

3) נדגמו עשרה מתגייסים לצה"ל. גובהם נמדד בס"מ. להלן התוצאות שהתקבלו: 168, 184, 192, 171, 180, 177, 187, 168, 177 ו-175.

א. מצאו אומדן חסר הטיה לגובה הממוצע של מתגייסי צה"ל.

ב. מצאו אומדן חסר הטיה לשונות הגבהים של מתגייסי צה"ל.

ג. מצאו 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: $\sum_{i=1}^{20} X_i = 162$, $\sum_{i=1}^{20} X^2_i = 1502.2$.

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.

© כל הזכויות שמורות לברק קנדל

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 תהייה:

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 באומד חסר הטיה לשונות האוכלוסייה ממנה הוצא המדגם.
ה. כאשר מעוניינים לחשב את שונות התפלגות הדגימה של ממוצע המדגם.

8) $X_1, X_2, ..., X_{16}$ מדגם מקרי מתוך אוכלוסייה בעלת ממוצע μ לא ידוע ושונות: $\sigma^2 = 64$. טעות התקן של האומד 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.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. 0.25 ב. 0.019 ג. 0.24 ד. 0.01
2) א. אוכלוסייה: מקררים של יצרן, תוחלת: 2400, סטיית תקן: 500.
ב. 100 ג. 2342 ד. 58
3) א. 177.9 ב. 64.1 ג. 0.4
4) א. 8.1 ב. 3.16
5) ד'.
6) ג'.
7) ד'.
8) ד'.
9) ג'.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 30 - רווח סמך לתוחלת (ממוצע)

תוכן העניינים

1. רווח סמך כששונות האוכלוסיה ידועה 117
2. קביעת גודל מדגם 122
3. רווח סמך כששונות האוכלוסיה לא ידועה 124

רווח סמך כששונות האוכלוסייה ידועה:

רקע:

ממוצע המדגם הוא אומד לממוצע האוכלוסייה, אך לא באמת ניתן להבין ממנו על גודלו של ממוצע האוכלוסייה. ההסתברות שממוצע המדגם יהיה בדיוק כמו הממוצע האמתי הוא אפסי.
מה שנהוג לעשות כדי לאמוד את ממוצע האוכלוסייה, זה לבנות רווח סמך.
נבנה מרווח בטחון שהסיכוי שהפרמטר μ ייכלל בתוכו הוא: $1-\alpha$.
$1-\alpha$: נקרא רמת בטחון או רמת סמך. כך ש: $P\left(A\le\mu\le B\right)=1-\alpha$.
A - גבול התחתון של רווח הסמך.
B - הגבול העליון של רווח הסמך.
$L=B-A$ - אורך רווח הסמך.

דוגמה (פתרון בהקלטה):

חוקר דגם 25 חיילים שנבחנו במבחן הפסיכומטרי. הוא בנה רווח סמך לממוצע הציונים במבחן הפסיכומטרי בקרב אוכלוסיית החיילים וקיבל בין 510 ל-590.
רווח הסמך נבנה ברמת סמך של 95%.

1. מהי אוכלוסיית המחקר?
2. מה המשתנה באוכלוסייה?
3. מה הפרמטר שהחוקר רצה לאמוד?
4. מהו רווח הסמך?
5. מה אורך רווח הסמך?
6. מהי רמת הביטחון של רווח הסמך?

© כל הזכויות שמורות לברק קנדל

בפרק זה נרצה לבנות רווח סמך לתוחלת (μ) במקרה ש-σ^2 (שונות האוכלוסייה) ידועה.

פרמטר אותו נרצה לאמוד: μ.

אומד נקודתי: $\bar{x}$.

תנאים לבניית רווח הסמך: $X \sim N$ או $n \geq 30$.

σ^2 (שונות האוכלוסייה) ידועה.

נוסחה לרווח הסמך: $\bar{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$.

דוגמה (פתרון בהקלטה):

על פי נתוני היצרן אורך חיי סוללה מתפלג נורמאלית עם סטיית תקן של 1 שעה. מעוניינים לאמוד את תוחלת חיי סוללה. נדגמו באקראי 4 סוללות, אורך החיים הממוצע שהתקבל הוא 13.5 שעות. בנו רווח סמך ברמת סמך של 95% לתוחלת אורך חיי סוללה.

שגיאת האמידה המקסימלית: $\varepsilon = Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$.

ε - נותן את שגיאת האמידה המקסימלית, דבר שנקרא גם טעות סטטיסטית, טעות דגימה.

דוגמה (פתרון בהקלטה):

בהמשך לשאלה עם הסוללות. מה ניתן להגיד בביטחון של 95% על שגיאת האמידה?

קשרים מתמטיים ברווח הסמך:

- אורך רווח הסמך הוא פעמיים שגיאת האמידה המקסימלית: $L = 2\varepsilon$.
- ממוצע המדגם נופל תמיד באמצע רווח הסמך: $\bar{X} = \frac{A+B}{2}$.
- ככל שמספר התצפיות (n) גבוה יותר, כך יש יותר אינפורמציה ולכן האומד יותר מדויק, ולכן נקבל רווח סמך יותר קצר.
- ככל שרמת הביטחון ($1-\alpha$) גבוהה יותר, כך: $z_{1-\frac{\alpha}{2}}$ יותר גבוה, ורווח הסמך יותר ארוך.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) חוקר התעניין לאמוד את השכר הממוצע במשק. על סמך מדגם הוא קבע שבביטחון של 95% כי השכר הממוצע במשק נע בין 9200 ל-9800 ₪.

א. מי האוכלוסייה במחקר?

ב. מה המשתנה הנחקר?

ג. מה הפרמטר שאותו רוצים לאמוד?

ד. מה רווח הסמך לפרמטר?

ה. מהי רמת הסמך לפרמטר?

ו. מה אורך רווח הסמך?

ז. מה הסיכוי שטעות הדגימה תעלה על 300 ₪?

2) מעוניינים לאמוד את התפוקה היומית הממוצעת של מפעל מסוים ברמת סמך של 95%. במדגם אקראי של 100 ימים התקבלה תפוקה ממוצעת 4950 מוצרים ביום. לצורך פתרון הנח שסטיית התקן האמתית ידועה ושווה 150 מוצרים ביום. בנו את רווח הסמך.

3) 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. הסבירו כיצד ומדוע השתנה רווח הסמך.

4) דגמו 200 עובדים מהמשק הישראלי. השכר הממוצע שלהם היה 9700 ₪. נניח שסטיית התקן של השכר במשק היא 3000 ₪.

א. בנו רווח סמך ברמת סמך של 95% לתוחלת השכר במשק.

ב. מה ניתן לומר בביטחון של 95% על הסטייה המרבית בין ממוצע המדגם לתוחלת השכר?

ג. מה היה צריך להיות גודל המדגם אם היינו רוצים להקטין את רווח הסמך ב-50%?

ד. אם היינו מגדילים את גודל המדגם ובונים רווח סמך באותה רמת סמך האם היה ניתן לטעון בביטחון רב יותר שרווח הסמך מכיל את הפרמטר?

© כל הזכויות שמורות לברק קנדל

5) בנו רווח סמך לממוצע הציונים של מבחן אינטליגנציה. ידוע שסטיית התקן היא 15 והמדגם מתבסס על 100 תצפיות. רווח הסמך שהתקבל הוא (105,99). שחזרו את:

א. ממוצע המדגם.

ב. שגיאת האמידה המקסימאלית.

ג. רמת הסמך.

6) 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.

7) חוקר בנה רווח סמך לממוצע וקיבל את רווח הסמך הבא: $82 < \mu < 92$.
נתון שסטיית התקן בהתפלגות שווה ל-10 ושהמדגם מתבסס על 16 תצפיות. התפלגות המשתנה היא נורמאלית.

א. מהו ממוצע המדגם?

ב. מהי רמת הסמך של רווח הסמך שנבנה?

ג. מה הסיכוי ששגיאת האמידה באמידת ממוצע האוכלוסייה תעלה על 5?

8) חוקר בנה רווח סמך לתוחלת כאשר השונות בהתפלגות ידועה ברמת סמך של 95%. אם החוקר כעת יבנה על סמך אותם נתונים רווח סמך ברמת סמך קטנה מ-95%, איזה מהמשפטים הבאים לא יהיה נכון.

א. אורך רווח הסמך החדש יהיה קטן יותר.

ב. גודל המדגם יהיה כעת קטן יותר.

ג. המרחק בין ממוצע המדגם לקצות רווח הסמך יהיו קטנים יותר ברווח הסמך החדש.

ד. רמת הביטחון לבנות רווח הסמך החדש תהיה קטנה יותר.

© כל הזכויות שמורות לברק קנדל

9) חוקר בנה רווח סמך ל- μ וקיבל: $48 < \mu < 54$. מה נכון בהכרח:

א. $\mu = 51$.

ב. $\overline{X} = 6$.

ג. $\overline{X} = 51$.

ד. אורך רווח הסמך הינו 3.

10) איזה מהגורמים הבאים אינו משפיע על גודלו של רווח בר סמך, כאשר שונות האוכלוסייה ידועה (בחרו בתשובה הנכונה):

א. רמת הביטחון.

ב. סטיית התקן באוכלוסייה.

ג. מספר המשתתפים.

ד. סטיית התקן במדגם.

תשובות סופיות:

1) א. העובדים במשק. ב. שכר ב-₪. ג. μ. ד. $9200 < \mu < 9800$.
ה. 0.95. ו. 600. ז. 0.05.

2) $4920.6 < \mu < 4979.4$.

3) א. $223.42 < \mu < 236.58$. ב. $222.16 < \mu < 237.84$.
ג. ראה סרטון.

4) א. $9284 < \mu > 10,116$. ב. הסטיה המירבית בין $\bar{x}$ ל- μ היא 416 ₪ בבטחון של 95%.
ג. 800. ד. לא.

5) א. 102. ב. 3. ג. 0.9544.

6) א. $83.5 < \mu < 4.42$. ב. יקטן פי 2. ג. גדל.

7) א. 87. ב. 5. ג. 0.9544.

8) ב'.

9) ג'.

10) ד'.

© כל הזכויות שמורות לברק קנדל

קביעת גודל מדגם:

רקע:

אם מעוניינים לאמוד את ממוצע האוכלוסייה כאשר סטיית התקן של האוכלוסייה ידועה: σ ברמת סמך של $1-\alpha$ ושגיאת אמידה שלא תעלה על ε מסוים, נציב בנוסחה הבאה: $n \geq \left(\frac{z_{1-\frac{\alpha}{2}} \cdot \sigma}{\varepsilon} \right)^2$.

כדי להציב בנוסחה צריך שהמשתנה הנחקר יתפלג נורמלית או שהמדגם ייצא בגודל של לפחות 30 תצפיות.

דוגמה (פתרון בהקלטה):

חברת תעופה מעוניינת לאמוד את תוחלת משקל המטען של נוסע. נניח שמשקל מטען של נוסע מתפלג נורמאלית עם סטיית תקן של 2 ק"ג. כמה נוסעים יש לדגום אם מעוניינים שבביטחון של 98% הסטייה המרבית בין ממוצע המדגם לממוצע האמתי לא יעלה על 0.5 ק"ג? (תשובה: 87).

© כל הזכויות שמורות לברק קנדל

שאלות:

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

2) מעוניינים לאמוד את הדופק הממוצע של מתגייסים לצבא.
מעוניינים שבביטחון של 95% שגיאת האמידה המרבית תהיה 0.5.
נניח שהדופק מתפלג נורמאלית על סטיית תקן של 3 פעימות לדקה.

א. כמה מתגייסים יש לדגום?

ב. אם ניקח מדגם הגדול פי 4 מהמדגם של סעיף א ונאמוד את הממוצע באותה רמת סמך כיצד הדבר ישפיע על שגיאת האמידה?

3) יהי X משתנה מקרי עם ממוצע μ וסטיית תקן σ. חוקר רוצה לבנות רווח בר סמך ל- μ ברמת ביטחון של 0.95, כך שהאורך של הרווח יהיה 0.5σ.
מהו גודל המדגם הנדרש?

תשובות סופיות:

1) 780.

2) א. 139. ב. הדבר יקטין את ε פי 2.

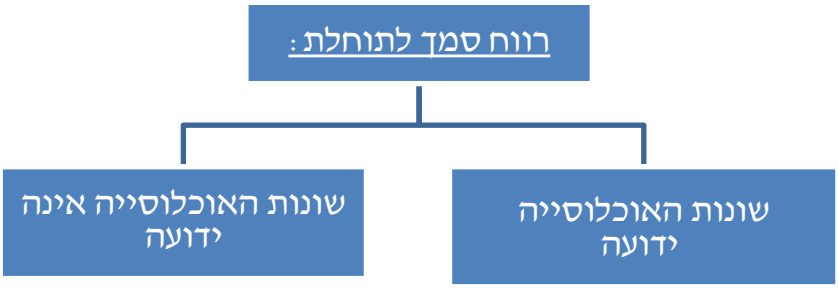
3) $n=62$.

© כל הזכויות שמורות לברק קנדל

רווח סמך כששונות האוכלוסייה לא ידועה:

רקע:

בבואנו לבנות רווח סמך לתוחלת אנו צריכים להתמקד בשני המצבים הבאים:

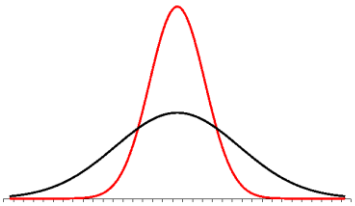


בפרק זה נעסוק במקרה ששונות האוכלוסייה $\left(\sigma^2\right)$ אינה ידועה לנו.

מקרה יותר פרקטי.

התנאי: $X \sim N$ או שהמדגם גדול.

רווח סמך: $\bar{X} \pm t^{(n-1)}_{1-\frac{\alpha}{2}} \cdot \frac{S}{\sqrt{n}}$.



האומד לשונות: $S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1} = \frac{\sum_{i=1}^{n} X^2{}_i - n\bar{X}^2}{n-1}$.

התפלגות T:

הינה התפלגות סימטרית פעמונית שהתוחלת שלה היא 0. ההתפלגות דומה להתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים.
התפלגות T תלויה במושג שנקרא דרגות חופש. דרגות החופש הן: $df = n-1$.
ככל שדרגות החופש עולות ההתפלגות הופכת להיות יותר גבוהה וצרה.
כשדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דוגמה (פתרון בהקלטה):

הזמן שלוקח לפתור שאלה מסוימת בחשבון מתפלג אצל תלמידי כיתות ח' נורמאלית.
במטרה לאמוד את תוחלת זמן הפתרון נדגמו 4 תלמידים בכיתה ח'. להלן התוצאות שהתקבלו בדקות: 5.3, 4.6, 5.2, 4.7.
בנו רווח סמך ברמת סמך של 95% לממוצע זמן הפתרון לשאלה בקרב תלמידי כיתה ח'.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) מחקר מעוניין לדעת כיצד תרופה מסוימת משפיעה על קצב פעימות הלב.
ל-5 אנשים שנטלו את התרופה מדדו את הדופק והתקבל מספר פעימות לדקה: 84, 88, 84, 79, 89.
הערה: לצורך פתרון הנח שקצב פעימות הלב מתפלג נורמאלית בקירוב.

א. בנו רווח סמך ברמת סמך של 95% לתוחלת הדופק של נוטלי התרופה הנ"ל.

ב. נתון שהדופק הממוצע ללא לקיחת התרופה הינו 70. לאור זאת, האם בביטחון של 95% התרופה משפיעה על הדופק?

ג. בהמשך לסעיף א', אם היינו בונים את רווח הסמך ברמת ביטחון של 99%, כיצד הדבר היה משפיע על רווח הסמך?

2) במדגם שנעשה על 25 מתגייסים לצבא האמריקאי התקבל כי גובה ממוצע של חייל הינו 178 ס"מ עם סטיית תקן $S=13$ ס"מ.
בנו רווח סמך ברמת סמך של 90% לתוחלת גובה המתגייסים לצבא האמריקאי. מה יש להניח לצורך פתרון?

3) אדם מעוניין לאמוד את זמן הנסיעה הממוצע שלו לעבודה. לצורך כך הוא דוגם 5 ימים שזמן הנסיעה בהם בדקות הוא: 30, 40, 32, 34, 27.

א. ברמת ביטחון של 95% אמוד את זמן הנסיעה הממוצע. מהי ההנחה הדרושה לצורך פתרון?

ב. איך גודל רווח הסמך היה משתנה אם היו דוגמים עוד ימים?

4) ציוני מבחן אינטליגנציה מתפלגים נורמאלית. נדגמו 25 מבחנים והתקבל ממוצע ציונים 102 וסטיית תקן מדגמית 13.

א. בנו רווח סמך לממוצע הציונים באוכלוסייה ברמת ביטחון של 95%.

ב. חזרו על סעיף א' אם סטיית התקן הינה סטיית התקן האמיתית של כלל הנבחנים.

ג. הסבירו את ההבדלים בין שני הסעיפים הנ"ל.

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

© כל הזכויות שמורות לברק קנדל

6) נדגמו 120 אנשים אקראיים מעל גיל 50. עבור כל אדם נבדק מספר שנות השכלתו. להלן התוצאות שהתקבלו: $\bar{x}=13.8$, $S=2$.
בנו רווח סמך ברמת סמך של 96% לממוצע ההשכלה של אזרחים מעל גיל 50.

7) שני סטטיסטיקאים בנו רווח בר-סמך לאותו פרמטר μ.
לכל אחד מהסטטיסטיקאים מדגם אחר, אך באותו גודל 10.
שניהם קבעו אותה רמת סמך.
סטטיסטיקאי א: הניח $\sigma=20$ 2.
סטטיסטיקאי ב: חישב לפי המדגם וקיבל $S=20$.
למי משני הסטטיסטיקאים יהיה רווח סמך ארוך יותר?

א. סטטיסטיקאי א.
ב. סטטיסטיקאי ב.
ג. אותו אורך רווח סמך לשני הסטטיסטיקאים.
ד. תלוי בתוצאות המדגם של כל סטטיסטיקאי.

תשובות סופיות:

1) א. $79.88 < \mu < 89.72$ ב. כן. ג. הוא היה גדל.

2) ראה בסרטון.

3) א. צריך להניח שהמשתנה מתפלג נורמלית. ב. לא ניתן לדעת.

4) א. $96.63 < \mu < 107.37$ ב. $96.90 < \mu < 107.10$ ג. ראה בסרטון.

5) $3.149 < \mu < 3.351$.

6) $13.42 < \mu < 14.18$.

7) ב'.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 31 - מבוא לבדיקת השערות על פרמטרים

תוכן העניינים

1. הקדמה 127
2. סוגי טעויות 131

הקדמה:

רקע:

תהליך של בדיקת השערות הוא תהליך מאד נפוץ בעולם הסטטיסטיקה.
בבדיקת השערות על פרמטרים נעבוד לפי השלבים הבאים:

שלב א: נזהה את הפרמטר הנחקר.

שלב ב: נרשום את השערות המחקר.
השערת האפס המסומנות ב- H_0.
בדרך כלל השערת האפס מסמלת את אשר היה מקובל עד עכשיו, את השגרה, הנורמה.
השערה אלטרנטיבית (השערת המחקר) המסומנת ב- H_1.
ההשערה האלטרנטיבית מסמלת את החדשנות בעצם ההשערה האלטרנטיבית מדברת על הסיבה שהמחקר נעשה היא שאלת המחקר.

שלב ג: נבדוק האם התנאים לביצוע התהליך מתקיימים ונניח הנחות במידת הצורך.

שלב ד: נרשום את כלל ההכרעה. בתהליך של בדיקת השערות יוצרים כלל שנקרא כלל הכרעה. הכלל יוצר אזורים שנקראים:

1. **אזור דחייה:**
   דחייה של השערת האפס כלומר קבלה של האלטרנטיבה.
2. **אזור קבלה:**
   קבלה של השערת האפס ודחייה של האלטרנטיבה. כלל ההכרעה מתבסס על איזשהו סטטיסטי. אזור הדחיה מוכתב על ידי סיכון שלוקח החוקר מראש שנקרא רמת מובהקות ומסומן ב- α.

שלב ה: בתהליך יש ללכת לתוצאות המדגם ולחשב את הסטטיסטי המתאים ולבדוק האם התוצאות נופלות באזור הדחייה או הקבלה.

שלב ו: להסיק מסקנה בהתאם לתוצאות המדגם.

© כל הזכויות שמורות לברק קנדל

דוגמה (פתרון בהקלטה):

משרד הבריאות פרסם שמשקל ממוצע של תינוקות ביום לידתם בישראל 3300 גרם.
משרד הבריאות רוצה לחקור את הטענה שנשים מעשנות בזמן ההיריון יולדות תינוקות במשקל נמוך מהממוצע. במחקר השתתפו 20 נשים מעשנות בהריון.
להלן תוצאות המדגם שבדק את המשקל של התינוקות בעת הלידה:

$n=20$, $\overline{X}=3120$, $S=280$.

א. מהי אוכלוסיית המחקר?

ב. מה המשתנה הנחקר?

ג. מה הפרמטר הנחקר?

ד. מהן השערות המחקר?

© כל הזכויות שמורות לברק קנדל

שאלות:

בשאלות הבאות, ענו על הסעיפים הבאים:

א. מהי אוכלוסיית המחקר?

ב. מה המשתנה הנחקר?

ג. מה הפרמטר הנחקר?

ד. מהן השערות המחקר?

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.

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.

3) במשך שנים אחוז המועמדים שהתקבל לפקולטה למשפטים היה 25%.
השנה מתוך מדגם של 120 מועמדים התקבלו 22.
מחקר מעוניין לבדוק האם השנה מקשים על הקבלה לפקולטה למשפטים.

4) בחודש ינואר השנה פורסם שאחוז האבטלה במשק הוא 8% במדגם עכשווי התקבל שמתוך 200 אנשים 6.5% מובטלים. רוצים לבדוק ברמת מובהקות של 5% האם כיום אחוז האבטלה הוא כמו בתחילת השנה.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. נבחנים בבגרות באנגלית.
ב. ציון.
ג. ממוצע הציונים בשיטת לימוד חדשה.
ד. $\begin{matrix} H_0:\mu=72 \\ H_1:\mu>72 \end{matrix}$.

2) א. משקאות בבקבוק של חברה מסוימת.
ב. נפח משקה בסמ"ק.
ג. ממוצע נפח המשקה בבקבוק.
ד. $\begin{matrix} H_0:\mu=500 \\ H_1:\mu<500 \end{matrix}$.

3) א. מועמדים לפקולטה למשפטים.
ב. משתנה דיכוטומי (התקבל, לא התקבל).
ג. אחוז הקבלה.
ד. $\begin{matrix} H_0:p=0.25 \\ H_1:p<0.25 \end{matrix}$.

4) א. אזרחים בוגרים במשק.
ב. משתנה דיכוטומי (מובטל, עובד).
ג. אחוז האבטלה כיום.
ד. $\begin{matrix} H_0:p=0.08 \\ H_1:p\neq 0.08 \end{matrix}$.

© כל הזכויות שמורות לברק קנדל

סוגי טעויות:

רקע:

בתהליך של בדיקת השערות יוצרים כלל שניקרא כלל הכרעה.
הכלל יוצר אזורים שנקראים:

1. אזור דחייה – דחייה של השערת האפס כלומר קבלה של האלטרנטיבה.
2. אזור קבלה – קבלה של השערת האפס ודחייה של האלטרנטיבה.

כלל ההכרעה מתבסס על איזשהו סטטיסטי .
בתהליך יש ללכת לתוצאות המדגם ולבדוק האם התוצאות נופלות באזור הדחייה או הקבלה וכך להגיע למסקנה – המסקנה היא בעירבון מוגבל כיוון שהיא תלויה בכלל ההכרעה ובתוצאות המדגם. אם נשנה את כלל ההכרעה אז אנחנו יכולים לקבל מסקנה אחרת. אם נבצע מדגם חדש אז אנחנו עלולים לקבל תוצאה אחרת.
לכן יתכנו טעויות במסקנות שלנו:

		הכרעה	
		H_0	H_1
מציאות	H_0	אין טעות	טעות מסוג 1
	H_1	טעות מסוג 2	אין טעות

הגדרת הטעויות:

טעות מסוג ראשון: להכריע לדחות את H_0 למרות שבמציאות H_0 נכונה.

טעות מסוג שני: להכריע לקבל את H_0 למרות שבמציאות H_1 נכונה.

דוגמה (פתרון בהקלטה):

אדם חשוד בביצוע עבירה ונתבע בבית המשפט.
אילו סוגי טעויות אפשריות בהכרעת הדין?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) 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. בסופו של דבר הוחלט להכריע לטובת חברת המשקאות.

א. רשמו את השערות המחקר.

ב. מה מסקנת המחקר?

ג. איזו סוג טעות יתכן וביצעו במחקר?

2) במחקר על פרמטר מסוים הוחלט בסופו של דבר לדחות את השערת האפס.

א. האם ניתן לדעת אם בוצע טעות במחקר?

ב. מה סוג הטעות האפשרית?

3) לפי נתוני משרד הפנים בשנת 1980 למשפחה ממוצעת היה 2.3 ילדים למשפחה עם סטיית תקן 0.4. ישנה טענה שכיום ממוצע מספר הילדים במשפחה קטן יותר. לצורך כך הוחלט לדגום 121 משפחות. במדגם התקבל ממוצע 2.17 ילדים למשפחה. על סמך תוצאות המדגם נקבע שלא ניתן לקבוע שבאופן מובהק תוחלת מספר הילדים למשפחה קטנה כיום.

א. מהי אוכלוסיית המחקר?

ב. מה המשתנה הנחקר?

ג. מה הפרמטר הנחקר?

ד. מה השערות המחקר?

ה. מה מסקנת המחקר?

ו. מהי סוג הטעות האפשרית במחקר?

תשובות סופיות:

1) א. $\begin{matrix} H_0:\mu=500 \\ H_1:\mu<500 \end{matrix}$. ב. לא דחינו את H_0. ג. טעות מסוג שני.

2) א. לא ניתן לדעת. ב. טעות מסוג ראשון.

3) א. משפחות כיום. ב. מס' הילדים.

ג. תוחלת מספר הילדים למשפחה כיום. ד. $\begin{matrix} H_0:\mu=2.3 \\ H_1:\mu<2.3 \end{matrix}$.

ה. לא לדחות את H_0. ו. טעות מסוג שני.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

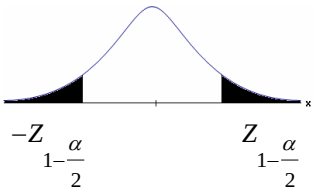
פרק 32 - בדיקת השערות על תוחלת (ממוצע)

תוכן העניינים

1. בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסיה ידועה 133
2. סיכוי לטעויות ועוצמה (ששונות האוכלוסיה ידועה) 137
3. קביעת גודל מדגם (ששונות האוכלוסיה ידועה) 143
4. מובהקות תוצאה - אלפא מינימלית (ששונות האוכלוסיה ידועה) 146
5. בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסיה לא ידועה 151
6. מובהקות תוצאה - אלפא מינימלית (ששונות האוכלוסיה לא ידועה) 155
7. הקשר בין רווח סמך לבדיקת השערות על תוחלת (ממוצע) 158

בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסייה ידועה:

רקע:

השערת האפס: השערה אלטרנטיבה:	$H_0:\mu=\mu_0$ $H_1:\mu\neq\mu_0$	$H_0:\mu\geq\mu_0$ $H_1:\mu<\mu_0$	$H_0:\mu\leq\mu_0$ $H_1:\mu>\mu_0$
תנאים:	1. σ ידועה 2. $X\sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0:	$Z_{\bar{x}}<-Z_{1-\frac{\alpha}{2}}$ או $Z_{\bar{x}}>Z_{1-\frac{\alpha}{2}}$  $-Z_{1-\frac{\alpha}{2}}$ $Z_{1-\frac{\alpha}{2}}$ H_0 דוחים את	$Z_{\bar{x}}<-Z_{1-\alpha}$ $-Z_{1-\alpha}$ H_0 דוחים את	$Z_{\bar{x}}>Z_{1-\alpha}$ $Z_{1-\alpha}$ H_0 דוחים את

סטטיסטי המבחן: $Z_{\bar{x}}=\dfrac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$.

חלופה אחרת לכלל הכרעה:

נדחה H_0 אם מתקיים:	$\bar{X}>\mu_0+Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}$ או $\bar{X}<\mu_0-Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}$	$\bar{X}<\mu_0-Z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}$	$\bar{X}>\mu_0+Z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}$

© כל הזכויות שמורות לברק קנדל

דוגמה:

יבול העגבניות מתפלג נורמלית עם תוחלת של 10 טון לדונם וסטיית תקן של 2.5 טון לדונם בעונה. משערים ששיטת זיבול חדשה תעלה את תוחלת היבול לעונה מבלי לשנות את סטיית התקן. נדגמו 4 חלקות שזובלו בשיטה החדשה. היבול הממוצע שהתקבל היה 12.5 טון לדונם. בדקו את ההשערה ברמת מובהקות של 1%.

פיתרון:

<u>אוכלוסייה</u>: עגבניות.

<u>המשתנה</u>: X = יבול העגבניות בטון לעונה.

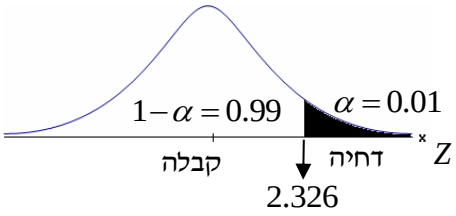
<u>הפרמטר</u>: μ = תוחלת היבול בשיטת הזיבול החדשה.

<u>השערות</u>: $H_0 : \mu = 10$
$H_1 : \mu > 10$

<u>תנאים</u>:

1. $X \sim N$.
2. $\sigma = 2.5$.

<u>כלל הכרעה</u>:



נדחה את H_0 אם $Z_{\bar{x}} > 2.326$

<u>תוצאות</u>: $n = 4$, $\bar{x} = 12.5$

<u>סטטיסטי המבחן</u>: $Z_{\bar{x}} = \dfrac{\bar{X} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}}$.

<u>נציב</u>: $Z_{\bar{X}} = \dfrac{1.25 - 10}{\dfrac{2.5}{\sqrt{4}}} = 2 < 2.326$

<u>מסקנה</u>:

לא נדחה H_0 (נקבל H_0).

ברמת מובהקות של 1% לא נוכל לקבל את הטענה ששיטת הזיבול החדשה מעלה את תוחלת היבול של העגבניות.

© כל הזכויות שמורות לברק קנדל

שאלות:

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 תהיה 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%?

4) המשקל הממוצע של הספורטאים בתחום ספורט מסוים הוא 90 ק"ג, עם סטיית תקן 8 ק"ג. לפי דעת מומחים בתחום יש צורך בהורדת המשקל ובשימוש בדיאטה מסוימת שצריכה להביא להורדת המשקל. לשם בדיקת יעילות הדיאטה נלקח מדגם מקרי של 50 ספורטאים ובתום שנה של שימוש בדיאטה התברר שהמשקל הממוצע במדגם זה היה 84 ק"ג.
יש לבדוק בר"מ של 10%, האם הדיאטה גורמת להורדת המשקל.

5) לפי מפרט נתון, על עובי בורג להיות 4 מ"מ עם סטיית תקן של 0.2 מ"מ.
במדגם של 25 ברגים העובי הממוצע היה 4.07 מ"מ.
קבעו ברמת מובהקות 0.05, האם עובי הברגים מתאים למפרט.
הניחו כי עובי של בורג מתפלג נורמלית וסטיית התקן של עובי בורג היא אכן 0.2 מ"מ.

6) במחקר נמצא שתוצאה היא מובהקת ברמת מובהקות של 5% מה תמיד נכון?
בחרו בתשובה הנכונה.

א. הגדלת רמת המובהקות לא תשנה את מסקנת המחקר.

ב. הגדלת רמת המובהקות תשנה את מסקנת המחקר.

ג. הקטנת רמת המובהקות לא תשנה את מסקנת המחקר.

ד. הקטנת רמת המובהקות תשנה את מסקנת המחקר.

© כל הזכויות שמורות לברק קנדל

7) חוקר ערך מבחן דו צדדי ברמת מובהקות של α והחליט לדחות את השערת האפס.
אם החוקר היה עורך מבחן צדדי ברמת מובהקות של $\frac{\alpha}{2}$ אזי בהכרח:

א. השערת האפס הייתה נדחית.

ב. השערת האפס הייתה לא נדחית.

ג. לא ניתן לדעת מה תהיה מסקנתו במקרה זה.

8) שני סטטיסטיקאים בדקו השערות: $H_0:\mu=\mu_0$ כנגד $H_1:\mu>\mu_0$,
עבור שונות ידועה ובאותה רמת מובהקות.
שני החוקרים קבלו אותו ממוצע במדגם אך לחוקר א' היה מדגם בגודל 100 ולחוקר ב' מדגם בגודל 200.

א. אם חוקר א' החליט לדחות את H_0, מה יחליט חוקר ב'? נמקו.

ב. אם חוקר א' יחליט לא לדחות את H_0, מה יחליט חוקר ב'? נמקו.

תשובות סופיות:

1) נקבל H_0, ברי"מ של 5% לא נקבל את הטענה של המורה ששיטת הלימוד שלו מעלה את ממוצע הציונים.

2) א. נדחה H_0, ברי"מ של 2.5% נקבל את תלונת אגודת הצרכנים בדבר הפחתת נפח המשקה בבקבוק.

ב. הגדלנו את רמת המובהקות לכן אנחנו נשארים בדחייה של H_0 והמסקנה לא תשתנה.

3) נדחה H_0, ברי"מ של 5% נקבע שהמכונה לא מאופסת.

4) נדחה H_0, ברי"מ של 0.1 נקבל את הטענה שהדיאטה יעילה ומפחיתה את המשקל הממוצע.

5) נקבל H_0, ברי"מ של 0.05 נכריע שתוחלת עובי הבורג מתים למפרט.

6) א'.

7) ג'.

8) א. לדחות. ב. לא ניתן לדעת.

© כל הזכויות שמורות לברק קנדל

סיכוי לטעויות ועוצמה (ששונות האוכלוסייה ידועה):

רקע:

		הכרעה	
		H_0	H_1
מציאות	H_0	אין טעות	טעות מסוג 1
	H_1	טעות מסוג 2	אין טעות

הגדרת הסתברויות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות):

(לדחות H_0 | H_0 נכונה) $= P_{H_0}$ (לדחות את H_0) $= \alpha$.

הסיכוי לבצע טעות מסוג 2:

(לקבל H_0 | H_1 נכונה) $= P_{H_1}$ (לקבל את H_0) $= \beta$.

רמת בטחון:

(לקבל H_0 | H_0 נכונה) $= P_{H_0}$ (לקבל את H_0) $= (\alpha - 1)$.

עוצמה:

(לדחות H_0 | H_1 נכונה) $= P_{H_1}$ (לדחות את H_0) $= \pi = (\beta - 1)$.

© כל הזכויות שמורות לברק קנדל

התהליך לחישוב סיכוי לטעות מסוג שני:

השערת האפס: **השערה אלטרנטיבה:**	$H_0:\mu=\mu_0$ $H_1:\mu\neq\mu_0$	$H_0:\mu=\mu_0$ $H_1:\mu<\mu_0$	$H_0:\mu=\mu_0$ $H_1:\mu>\mu_0$
תנאים:	1. σ ידועה 2. $X\sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0:	$\bar{X}>\mu_0+Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}$ או $\bar{X}<\mu_0-Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}$	$\bar{X}<\mu_0-Z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}$	$\bar{X}>\mu_0+Z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}$
חישוב β:	$P_{\mu_1}\left(\mu_0-Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}<\bar{X}<\mu_0+Z_{1-\frac{\alpha}{2}}\cdot\frac{\sigma}{\sqrt{n}}\right)$	$P_{\mu_1}\left(\bar{X}>\mu_0-Z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}\right)$	$P_{\mu_1}\left(\bar{X}<\mu_0+Z_{1-\alpha}\cdot\frac{\sigma}{\sqrt{n}}\right)$

התפלגות ממוצע המדגם: $\bar{X}\sim N\left(\mu,\frac{\sigma^2}{n}\right)$.

התקנון: $Z=\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$.

דוגמה:

בתחילת השנה חשבון הטלפון הסלולארי הממוצע לאדם היה 200₪ עם סטיית תקן של 80 ₪ לחודש. בעקבות כניסתן של חברות טלפון סלולארית חדשות מעוניינים לבדוק האם כיום ממוצע חשבון הטלפון הסלולארי פחת. לצורך בדיקה דגמו באקראי 36 אנשים וחשבון הטלפון הסלולארי שלהם היה 150 ₪ בממוצע לחודש.

א. רשמו את השערות המחקר ובנו כלל הכרעה במונחי חשבון ממוצע מדגמי ברמת מובהקות של 5%.

ב. מה מסקנתכם? איזה סוג טעות אפשרית במסקנה?

ג. נניח שבמציאות כיום החשבון הממוצע הוא 160 ₪. מה הסיכוי לבצע טעות מסוג שני?

ד. אם נקטין את רמת המובהקות מסעיף א', כיצד הדבר ישפיע על התשובה מסעיף ג'?

© כל הזכויות שמורות לברק קנדל

פתרון:

א. אוכלוסייה: משלמי חשבון טלפון סלולאר כיום.

המשתנה: X = חשבון הטלפון החודשי בשקלים.

הפרמטר: μ.

השערות: $\begin{matrix} H_0: \mu = 200 \\ H_1: \mu < 200 \end{matrix}$.

תנאים:

1. $\mu = 200$.

2. $n = 36$.

$K = \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$, $\bar{X} < \mu_0 - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$.

$$\alpha = 0.05$$
$$Z_{1-\alpha} = Z_{0.95} = 1.645$$

נציב: $K = 200 - 1.645 \cdot \frac{80}{\sqrt{36}} = 178.07$ שקלים.

כלל ההכרעה: דחה את H_0 אם $\bar{X} < 178.07$ שקלים.

ב. ברמת מובהקות של 5% נכריע שאכן ממוצע חשבון הטלפון הסלולרי פחת מתחילת השנה.

ג. השערות: $\begin{matrix} H_0: \mu_0 = 200 \\ H_1: \mu < 200 \end{matrix}$.

כלל ההכרעה: נדחה את H_0 אם $\bar{X} < 178.07$.

$H_1: \bar{X} \sim N\left(160, \frac{80^2}{36}\right)$.

$$Z = \frac{178.07 - 160}{\frac{80}{\sqrt{36}}} = 1.36$$

$$\beta = P_{H_1}\left(\begin{matrix}\text{לקבל את} \\ H_0\end{matrix}\right) = P_{H_1}\left(\bar{X} > 178.07\right) = 1 - \phi(136) = 1 - 0.9131 = 0.0869$$

ד. הקטנת α מגדילה את β.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) נתון ש: $X \sim N\left(\mu, \sigma^2 = 1\right)$.

להלן השערות של חוקר לגבי הפרמטר μ: $H_1: \mu = 7$, $H_0: \mu = 5$.

מעוניינים ליצור כלל הכרעה המתבסס על הסמך תצפית בודדת כך שרמת המובהקות תהיה 5%.

א. עבור אילו ערכים של X שידגם נדחית השערת H_0?

ב. מה הסיכוי לבצע טעות מסוג שני?

ג. אם במדגם התקבל ש-$X = 6.9$ מה תהיה המסקנה ומה הטעות האפשרית?

2) 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 וברמת 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 העצמה של הכלל מסעיף א'?

3) להלן נתונים על תהליך של בדיקת השערות על תוחלת:

$n = 225$, $\sigma = 30$, $H_1: \mu \neq 200$, $H_0: \mu = 200$.

א. רשמו כלל הכרעה במונחי ממוצע מדגם קריטי וברמת מובהקות של 10%.

ב. בהמשך לסעיף א', מהי העצמה אם התוחלת שווה ל-195?

ג. הסבירו, ללא חישוב, איך העצמה תשתנה אם רמת המובהקות תהיה 5%?

4) 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. רשמו את ההשערות ואת כלל 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. הסבירו ללא חישוב כיצד התשובה לסעיף ב תשתנה אם רמת המובהקות תגדל.

ד. הסבירו ללא חישוב כיצד התשובה לסעיף ב תשתנה אם התוחלת האמיתית היא 47 ולא 48 מ"מ.

© כל הזכויות שמורות לברק קנדל

5) להלן השערות של מחקר: $H_1:\mu=58$, $H_0:\mu=50$.
מעוניינים לדגום 100 תצפיות. ידוע שסטיית התקן של ההתפלגות הינה 20.

א. בנו כלל הכרעה שהסיכוי לטעות מסוג שני בו הוא 10%.
מהי רמת המובהקות?

ב. כיצד הייתה משתנה רמת המובהקות אם (כל סעיף בפני עצמו)?

i. סטיית התקן הייתה יותר גדולה.

ii. הסיכוי לטעות מסוג שני גדול יותר.

השאלות שלהלן הן שאלות רב-ברירה, בחרו בתשובה הנכונה ביותר:

6) אם חוקר החליט להגדיל את רמת המובהקות במחקר שלו אזי:

א. הסיכוי לטעות מסוג ראשון גדל.

ב. העוצמה של המבחן גדלה.

ג. הסיכוי לטעות מסוג שני גדל.

ד. תשובות א' ו-ב' נכונות.

7) חוקר ביצע מחקר ובו עשה טעות מסוג שני לכן:

א. השערת האפס נכונה.

ב. השערת האפס נדחתה.

ג. השערת האפס לא נדחתה.

ד. אף אחת מהתושבות לא נכונה בהכרח.

8) מה המצב הרצוי לחוקר המבצע בדיקת השערה:

	α	$1-\beta$
א.	גדולה	גדולה
ב.	גדולה	קטנה
ג.	קטנה	גדולה
ד.	קטנה	קטנה

9) נערך שינוי בכלל ההחלטה של בדיקת השערה מסוימת ובעקבותיו אזור דחיית H_0 קטן. כל שאר הגורמים נשארו ללא שינוי. כתוצאה מכך:

א. הן α, והן $1-\beta$, יקטנו.

ב. α יישאר ללא שינוי ואילו $1-\beta$ יגדל.

ג. α יגדל ואילו $1-\beta$ יקטן.

ד. הן α והן $1-\beta$ יגדלו.

© כל הזכויות שמורות לברק קנדל

10) ידוע כי לחץ דם תקין באוכלוסייה הוא 120. רופא מניח שלחץ הדם בקרב עיתונאים גבוה יותר מהממוצע באוכלוסייה. הוא לקח מדגם של 60 עיתונאים וקיבל ממוצע 137. על סמך המדגם, הוא בודק טענתו ברמת מובהקות 0.02 ומסיק שלחץ הדם בקרב העיתונאים אינו גבוה יותר.
מה הטעות האפשרית שהרופא עושה?

א. טעות מסוג ראשון.

ב. טעות מסוג שני.

ג. טעות מסוג שלישי.

ד. אין טעות במסקנתו.

תשובות סופיות:

1) א. מעל 6.645. ב. 0.3594.

ג. דחינו את H_0, תתכן טעות מסוג ראשון.

2) א. נדחה H_0 אם $\bar{X} < 2.24$. ב. נדחה H_0. ג. 1.

3) א. נדחה H_0 אם $\bar{X} > 203.29$ או $\bar{X} < 196.71$. ב. 0.8051. ג. תקטן.

4) א. נדחה H_0 אם $\bar{X} < 48.9$. ב. 0.0885. ג. תקטן. ד. תקטן.

5) א. 0.0033. ב. i. רמת המובהקות הייתה קטנה.

ב. ii. רמת המובהקות הייתה גדלה.

6) ד'.

7) ג'.

8) ג'.

9) א'.

10) ב'.

© כל הזכויות שמורות לברק קנדל

קביעת גודל מדגם (ששונות האוכלוסייה ידועה):

רקע:

השערות המחקר הן: $H_1:\mu=\mu_1$, $H_0:\mu=\mu_0$.

סטיית התקן של האוכלוסייה ידועה σ ומעוניינים לבצע מחקר שרמת המובהקות לא תעלה על α והסיכוי לטעות מסוג שני לא יעלה על β.

הנוסחה הבאה נותנת את גודל המדגם הרצוי: $n\geq\left(\frac{\left(Z_{1-\alpha}+Z_{1-\beta}\right)\times\sigma}{\mu_0-\mu_1}\right)^2$.

דוגמה:

משרד החינוך מפעיל בגן חובה שיטת חינוך שפותחה בשנת 1995. לפי שיטת חינוך זו תוחלת הציון במבחן אוצר מילים לגיל הרך הוא 70. אנשי חינוך החליטו לבדוק שיטת חינוך שפותחה בהולנד הנותנת שם תוחלת ציון אוצר מילים של 80.
נניח שציוני מבחן זה מתפלגים נורמאלית עם $\sigma=17$.
כדי לבדוק האם גם בישראל הפעלת שיטת החינוך ההולנדית תעבוד בגנים, רוצים לבנות מחקר ברמת מובהקות של 5%. כמו כן, מעוניינים שאם בהפעלת השיטה ההולנדית תוחלת הציונים תעלה לכדי 80, המחקר יגלה זאת בסיכוי של 90%.
כמה ילדי גן חובה דרושים למחקר?

פתרון:

<u>האוכלוסייה</u>: ילדי גן חובה.

<u>המשתנה</u>: X = ציון במבחן אוצר מילים.

<u>הפרמטר</u>: μ.

<u>השערות</u>: $\begin{matrix}H_0:\mu=70\\H_1:\mu=80\end{matrix}$.

$X\sim N\left(\mu,\sigma^2=17^2\right)$.

אם בהפעלת השיטה ההולנדית התוחלת תעלה ל80, נגלה זאת בסיכוי 90%.

© כל הזכויות שמורות לברק קנדל

$$n \geq \left(\frac{\left(Z_{1-\alpha} + Z_{1-\beta}\right) \times \sigma}{\mu_0 - \mu_1} \right)^2$$

$$\alpha = 0.05$$

$$1-\beta = 0.9$$

$$\mu_0 = 70$$

$$\mu_1 = 80$$

$$\sigma = 17$$

$$Z_{1-\alpha} = Z_{0.95} = 1.645$$
$$Z_{1-\beta} = Z_{0.9} = 1.282$$

נציב: $n \geq \left(\frac{(1.645+1.282) \times 17}{70-80} \right)^2 = 24.76$

לכן, $n_{\min} = 25$.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) במבחן אינטליגנציה הציונים מתפלגים נורמאלית עם סטיית תקן 8 וממוצע 100. פסיכולוג מעוניין לבדוק את הטענה שבאוכלוסיות במצב סוציו אקונומי נמוך תוחלת הציונים היא 95.
אם מעוניינים לגלות את הטענה בהסתברות של לפחות 99% כשרמת המובהקות היא 5% מהו גודל המדגם הדרוש?

2) משרד התקשורת טוענים שאדם מדבר בממוצע 180 דקות בחודש בטלפון הסלולרי. חברות הטלפון הסלולרי טוענות שאינפורמציה זו אינה נכונה ואדם מדבר בממוצע פחות: כ-160 דקות. לצורך פתרון נניח שסטיית התקן של זמן השיחה החודשי ידוע ושווה ל-60 דקות. כמה אנשים יש לדגום כך שאם טענת משרד התקשורת נכונה נדחה אותה בסיכוי של 5% (איך קוראים להסתברות זאת?) כמו כן אם טענת חברות הטלפון הסלולרית נכונה המחקר יגלה זאת בסיכוי של 90% (איך קוראים להסתברות זאת?).

3) השערות המחקר הן: $H_1:\mu=\mu_1$, $H_0:\mu=\mu_0$.
כמו כן נתון שהמשתנה מתפלג נורמלית עם סטיית התקן ידועה σ מעוניינים לבצע מחקר שרמת המובהקות לא תעלה על α והסיכוי לטעות מסוג שני לא יעלה על β. הוכיחו שגודל המדגם הרצוי לכך יהיה: $n\geq\left(\frac{\left(Z_{1-\alpha}+Z_{1-\beta}\right)\times\sigma}{\mu_0-\mu_1}\right)^2$.

תשובות סופיות:

1) 41.

2) 78.

3) שאלת הוכחה.

© כל הזכויות שמורות לברק קנדל

מובהקות תוצאה – אלפא מינימלית (ששונות האוכלוסייה ידועה):

רקע:

דרך נוספת להגיע להכרעות שלא דרך כלל הכרעה, היא דרך חישוב מובהקות התוצאה:

באמצעות תוצאות המדגם מחשבים את מובהקות התוצאה שמסומן ב- p_v.

את רמת המובהקות החוקר קובע מראש לעומת זאת, את מובהקות התוצאה החוקר יוכל לחשב רק אחרי שיהיו לו את התוצאות.

המסקנה של המחקר תקבע לפי העיקרון הבא: אם $p_v \le \alpha$, דוחים את H_0.

מובהקות התוצאה זה הסיכוי לקבלת תוצאות המדגם וקיצוני מתוצאות אלה בהנחת השערת האפס.

(לקבל את תוצאות המדגם וקיצוני) $p_v = P_{H_0}$.

אם ההשערה היא דו צדדית:

(לקבל את תוצאות המדגם וקיצוני) $p_v = 2P_{H_0}$.

מובהקות התוצאה היא גם האלפא המינימלית לדחיית השערת האפס.

השערת האפס: **השערה אלטרנטיבה:**	$H_0:\mu=\mu_0$ $H_1:\mu\neq\mu_0$	$H_0:\mu=\mu_0$ $H_1:\mu<\mu_0$	$H_0:\mu=\mu_0$ $H_1:\mu>\mu_0$
תנאים:	1. σ ידועה 2. $X\sim N$ או מדגם מספיק גדול		
p-value	אם $\bar{x}>\mu_0 \Rightarrow 2\cdot P_{H_0}\left(\bar{X}\ge\bar{x}\right)$ אם $\bar{x}<\mu_0 \Rightarrow 2\cdot P_{H_0}\left(\bar{X}\le\bar{x}\right)$	$P_{H_0}(\bar{X}\le\bar{x})$	$P_{H_0}(\bar{X}\ge\bar{x})$

כאשר בהנחת השערת האפס: $\bar{X}\sim N\left(\mu_0,\frac{\sigma^2}{n}\right)$, $Z_{\bar{x}}=\frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$.

© כל הזכויות שמורות לברק קנדל

דוגמה:

המשקל הממוצע של מתגייסים לצבא לפני 20 שנה היה 65 ק"ג. מחקר מעוניין לבדוק האם כיום המשקל הממוצע של מתגייסים גבוה יותר. נניח שמשקל המתגייסים מתפלג נורמאלית עם סטיית תקן של 12 ק"ג. במדגם של 16 מתגייסים התקבל משקל ממוצע של 71 ק"ג.

א. מהי מובהקות התוצאה?

ב. מה המסקנה אם רמת המובהקות היא 5% ואם רמת המובהקות היא 1%?

פתרון:

א. <u>אוכלוסייה</u>: המתגייסים לצבא כיום.

<u>משתנה</u>: X = משקל בק"ג.

<u>פרמטר</u>: μ.

<u>השערות</u>: $\begin{matrix} H_0:\mu=65 \\ H_1:\mu>65 \end{matrix}$.

<u>תנאים</u>:

1. $X\sim N$.
2. $\sigma=12$.

<u>תוצאות מדגם</u>:

$n=16$

$\bar{X}=71$

$$P_V=P_{H_0}\left(\begin{matrix}\text{לתוצאות}\\ \text{המדגם}\\ \text{וקיצוני}\end{matrix}\right)=P_{H_0}\left(\bar{X}\geq 71\right)=1-\phi\left(2\right)=1-0.9772=0.0228$$

$$Z_{\bar{x}}=\frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}=\frac{71-65}{\frac{12}{\sqrt{16}}}=2$$

$\alpha_{\min}=0.0228$

© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן השערות של מחקר: $H_1:\mu>70$, $H_0:\mu=70$.
המשתנה הנחקר מתפלג נורמלית עם סטיית תקן 20.
במדגם מאותה אוכלוסייה התקבלו התוצאות הבאות: $\bar{x}=74$, $n=100$.
מהי מובהקות התוצאה?

2) השכר הממוצע במשק בשנת 2012 היה 8800 ₪ עם סטיית תקן 2000. במדגם שנעשה אתמול על 100 עובדים התקבל שכר ממוצע 9500 ₪. מטרת המחקר היא לבדוק האם כיום חלה עליה בשכר. עבור אילו רמות מובהקות שיבחר החוקר יוחלט שחלה עליה בשכר הממוצע במשק?

3) אדם חושד שחברת ממתקים לא עומדת בהתחייבויותיה, ומשקלו של חטיף מסוים אותו הוא קונה מדי בוקר נמוך מ-100 גרם.
חברת הממתקים טוענת מצידה שהיא אכן עומדת בהתחייבויותיה. ידוע כי סטיית התקן של משקל החטיף היא 12 גרם. האדם מתכוון לשקול 100 חפיסות חטיפים ולאחר מכן להגיע להחלטה.
לאחר הבדיקה הוא קיבל משקל הממוצע של 98.5 גרם.

א. רשמו את השערות המחקר.

ב. מהי רמת המובהקות המינימלית עבורה דוחים את השערת האפס?

ג. מהי רמת המובהקות המקסימלית עבורה נקבל את השערת האפס?

ד. מה המסקנה ברמת מובהקות של 5%?

4) מכונה לחיתוך מוטות במפעל חותכת מוטות באורך שמתפלג נורמאלית עם תוחלת אליה כוונה המכונה וסטיית תקן 2 ס"מ. ביום מסוים כוונה המכונה לחתוך מוטות באורך 80 ס"מ. אחראי האיכות מעוניין לבדוק האם המכונה מכוילת. לצורך כך נדגמו מקו הייצור 16 מוטות שנחתכו אורכן הממוצע היה 81.7 ס"מ.

א. מהי רמת המובהקות המינימלית עבורה נכריע שהמכונה לא מכוילת?

ב. אם נוסיף עוד תצפית שערכה יהיה 82 ס"מ, כיצד הדבר ישפיע על התשובה של הסעיף הקודם?

ג. הכרע ברמת מובהקות של 5% האם המכונה מכוילת.

5) אם מקבלים בחישובים אלפא מינימלית (P value) קטנה מאד, סביר להניח כי החוקר ידחה את השערת האפס בקלות. נכון/לא נכון? נמק.

© כל הזכויות שמורות לברק קנדל

6) בבדיקת השערות התקבל שה- $p-value=0.02$.
מה תהיה מסקנת חוקר המשתמש ברמת מובהקות 1%?
בחרו בתשובה הנכונה.

א. יקבל את השערת האפס בכל מקרה.

ב. ידחה את השערת האפס מקרה.

ג. ידחה את השערת האפס רק אם המבחן הנו דו צדדי.

ד. לא ניתן לדעת כי אין מספיק נתונים.

7) מובהקות התוצאה (PV) היא גם (בחרו בתשובה הנכונה):

א. רמת המובהקות המינימאלית לדחות השערת האפס.

ב. רמת המובהקות המקסימאלית לדחיית השערת האפס.

ג. רמת המובהקות שנקבעת מראש על ידי החוקר שטרם קיבל את תוצאות המחקר.

ד. רמת המובהקות המינימאלית לאי דחיית השערת האפס.

8) בבדיקת השערות מסוימת התקבל: $p\ value=0.0254$ לכן (בחרו בתשובה הנכונה):

א. ברמת מובהקות של 0.01 אך לא של 0.05 נדחה את H_0 .

ב. ברמת מובהקות של 0.01 ושל 0.05 לא נדחה את H_0 .

ג. ברמת מובהקות של 0.05 אך לא של 0.01 נדחה את H_0 .

ד. ברמת מובהקות של 0.01 ושל 0.05 נדחה את H_0 .

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) 0.0228.

2) עבור כל רמת מובהקות סבירה.

3) א. $\begin{matrix} H_0:\mu=100 \\ H_1:\mu<100 \end{matrix}$. ב. 0.1056. ג. 0.1056.

ד. נכריע שיש עמידה בהתחייבות של החברה.

4) א. 0.0006. ב. יקטן. ג. נכריע שאין כיול.

5) נכון.

6) א'.

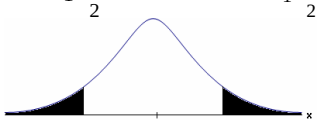
7) א'.

8) ג'.

© כל הזכויות שמורות לברק קנדל

בדיקת השערות על תוחלת (ממוצע) כששונות האוכלוסייה לא ידועה:

רקע:

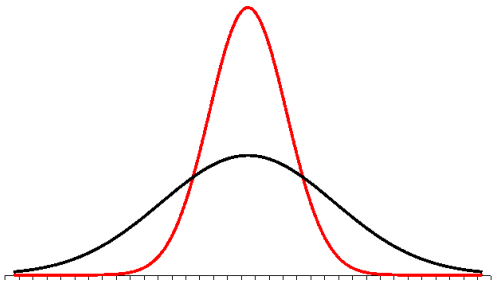
השערת האפס: השערה אלטרנטיבה:	$H_0:\mu=\mu_0$ $H_1:\mu\neq\mu_0$	$H_0:\mu=\mu_0$ $H_1:\mu<\mu_0$	$H_0:\mu=\mu_0$ $H_1:\mu>\mu_0$
תנאים:	1. σ אינה ידועה 2. $X\sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0:	$t_{\bar{x}}<-t_{1-\frac{\alpha}{2}}^{(n-1)}$ או $t_{\bar{x}}>t_{1-\frac{\alpha}{2}}^{(n-1)}$  $-t_{1-\frac{\alpha}{2},n-1}$ $t_{1-\frac{\alpha}{2},n-1}$ ■ - דוחים את H_0	$t_{\bar{x}}<-t_{1-\alpha}^{(n-1)}$ $-t_{1-\alpha,n-1}$ ■ - דוחים את H_0	$t_{\bar{x}}>t_{1-\alpha}^{(n-1)}$ $t_{1-\alpha,n-1}$ ■ - דוחים את H_0
חלופה לכלל הכרעה: נדחה H_0 אם מתקיים:	$\bar{X}>\mu_0+t_{1-\frac{\alpha}{2}}^{n-1}\cdot\frac{S}{\sqrt{n}}$ או $\bar{X}<\mu_0-t_{1-\frac{\alpha}{2}}^{n-1}\cdot\frac{S}{\sqrt{n}}$	$\bar{X}<\mu_0-t_{1-\alpha}^{n-1}\cdot\frac{S}{\sqrt{n}}$	$\bar{X}>\mu_0+t_{1-\alpha}^{n-1}\cdot\frac{S}{\sqrt{n}}$

סטטיסטי המבחן: $t_{\bar{x}}=\dfrac{\bar{x}-\mu_0}{\frac{S}{\sqrt{n}}}$.

$S^2=\dfrac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}{n-1}=\dfrac{\sum_{i=1}^{n}X^2{}_i-n\bar{X}^2}{n-1}$.

© כל הזכויות שמורות לברק קנדל

התפלגות T:



הינה התפלגות סימטרית פעמונית שהתוחלת שלה היא 0. ההתפלגות דומה להתפלגות Z רק שהיא יותר רחבה ולכן הערכים שלה יהיו יותר גבוהים. התפלגות T תלויה במושג שנקרא דרגות חופש.

דרגות החופש הן: $df = n-1$.

ככל שדרגות החופש עולות ההתפלגות הופכת להיות יותר גבוהה וצרה.

כשדרגות החופש שואפות לאינסוף התפלגות T שואפת להיות כמו התפלגות Z.

דוגמה (פתרון בהקלטה):

מפעל קיבל הזמנה לייצור משטחים בעובי של 0.1 ס״מ.

כדי לבדוק האם המפעל עומד בדרישה נדגמו 10 משטחים ונמצא שהעובי הממוצע הוא 0.104 עם אומדן לסטיית תקן 0.002 ס״מ.

א. מהן השערות המחקר?

ב. מה ההנחה הדרושה לצורך פתרון?

ג. בדוק ברמת מובהקות של 5%.

© כל הזכויות שמורות לברק קנדל

שאלות:

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: 125, 80, 100, 95, 90 שעות. מה מסקנתכם ברמת מובהקות של 5%.
מהי ההנחה הדרושה לצורך הפתרון?

2) משרד הבריאות פרסם שמשקל ממוצע של תינוקות ביום היוולדם בישראל 3300 גר'. משרד הבריאות רוצה לחקור את הטענה שנשים מעשנות בזמן ההיריון יולדות תינוקות במשקל נמוך מהממוצע.
במחקר השתתפו 20 נשים מעשנות בהריון.
להלן תוצאות המדגם שבדק את המשקל של התינוקות בעת הלידה:

$$n = 20$$
$$\bar{x} = 3120$$
$$S = 280$$

מה מסקנתכם ברמת מובהקות של 5% מה יש להניח לצורך פתרון?

3) ציוני מבחן אינטליגנציה מתפלגים נורמלית. בארה"ב ממוצע הציונים הוא 100. במדגם שנעשה על 23 נבחנים ישראלים, התקבל ממוצע ציונים 104.5 וסטיית התקן המדגמית 16. האם בישראל ממוצע הציונים שונה מארה"ב? הסיקו ברמת מובהקות של 5%.

4) באוכלוסייה מסוימת נדגמו 10 תצפיות והתקבלו התוצאות הבאות:

$$\sum_{i=1}^{10} X_i = 750$$
$$\sum_{i=1}^{10} \left(X_i - \bar{X} \right)^2 = 900$$

נתון שההתפלגות היא נורמלית.
בדוק ברמת מובהקות של 5% האם התוחלת של ההתפלגות שונה מ-80 .

© כל הזכויות שמורות לברק קנדל

5) ליאור ורוני העלו את אותן השערות על ממוצע האוכלוסייה.
כמו כן הם התבססו על אותן תוצאות של מדגם.
ליאור השתמש בטבלה של התפלגות Z.
רוני השתמשה בטבלה של התפלגות t.
מה נוכל לומר בנוגע להחלטת המחקר שלהם? בחר בתשובה הנכונה.

א. אם ליאור ידחה את השערת האפס אז גם בהכרח רוני.

ב. אם רוני תדחה את השערת האפס אז גם בהכרח ליאור.

ג. שני החוקרים בהכרח יגיעו לאותה מסקנה.

ד. לא ניתן לדעת על היחס בין דחיית השערת האפס של שני החוקרים.

6) נתון ש: $X \sim N(\mu,\sigma^2)$ כמו כן נתונות ההשערות הבאות: $\begin{matrix} H_0:\mu=\mu_0 \\ H_1:\mu<\mu_0 \end{matrix}$.
חוקר בדק את ההשערות הללו על סמך מדגם שכלל 10 תצפיות.
σ^2 לא הייתה ידועה לחוקר. החוקר החליט לדחות את השערת האפס ברמת מובהקות של 5% לאחר מכן כדי לחזק את קביעתו הוא דגם עוד 5 תצפיות ושקלל את תוצאות אלה גם למדגם כך שכלל עכשיו 15 תצפיות.
בחר בתשובה הנכונה:

א. כעת בברור הוא ידחה את השערת האפס.

ב. כעת הוא דווקא יקבל את השערת האפס.

ג. כעת לא ניתן לדעת מה תהיה מסקנתו.

תשובות סופיות:

1) נדחה H_0.

2) נדחה H_0.

3) נקבל H_0.

4) נקבל H_0.

5) ב'.

6) ג'.

© כל הזכויות שמורות לברק קנדל

מובהקות תוצאה – אלפא מינימלית (ששונות האוכלוסייה לא ידועה):

רקע:

נזכיר שהמסקנה של המחקר תיקבע לפי העיקרון הבא: אם $p_v \le \alpha$ דוחים את H_0.

מובהקות התוצאה היא הסיכוי לקבלת תוצאות המדגם וקיצוני מתוצאות אלה בהנחת השערת האפס.

(לקבל את תוצאות המדגם וקיצוני) $p_v = P_{H_0}$.

אם ההשערה היא דו צדדית:

(לקבל את תוצאות המדגם וקיצוני) $p_v = 2P_{H_0}$.

מובהקות התוצאה היא גם האלפא המינימלית לדחיית השערת האפס.

השערת האפס: השערה אלטרנטיבה:	$H_0 : \mu = \mu_0$ $H_1 : \mu \ne \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$	$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$
תנאים:	1. σ אינה ידועה או 2. מדגם מספיק גדול $X \sim N$		
p-value	אם $\bar{x} > \mu_0 \Rightarrow 2 \cdot P_{H_0}\left(\bar{X} \ge \bar{x}\right)$ אם $\bar{x} < \mu_0 \Rightarrow 2 \cdot P_{H_0}\left(\bar{X} \le \bar{x}\right)$	$P_{H_0}\left(\bar{X} \le \bar{x}\right)$	$P_{H_0}\left(\bar{X} \ge \bar{x}\right)$

$$t_{\bar{x}} = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$$

$$S^2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1} = \frac{\sum_{i=1}^{n} X^2{}_i - n\bar{X}^2}{n-1}$$

$$d.f = n-1$$

© כל הזכויות שמורות לברק קנדל

דוגמה:

ממוצע זמן הנסיעה של אדם לעבודה הינו 40 דקות. הוא מעוניין לבדוק דרך חלופית שאמורה להיות יותר מהירה. לצורך כך הוא דוגם 5 ימים שבהם הוא נוסע בדרך החלופית. זמני הנסיעה שקיבל בדקות הם: 27, 32, 30, 40, 34.
הניחו שזמן הנסיעה מתפלג נורמלית.

א. רשמו את השערות המחקר.

ב. מצאו חסמים למובהקות התוצאה.

ג. מה המסקנה ברמת מובהקות של 5%?

פתרון:

<u>אוכלוסייה</u>: כלל הנסיעות לעבודה בדרך החלופית.

<u>משתנה</u>: X = זמן נסיעה בדקות.

<u>תנאים</u>: $X \sim N$.

<u>פרמטר</u>: μ.

א. <u>השערות</u>:
$$\begin{matrix} H_0: \mu = 40 \\ H_1: \mu < 40 \end{matrix}.$$

ב. <u>תוצאות המדגם</u>:

$$n=5\ ,\ \bar{X}\ \frac{\sum X_i}{n} = \frac{34+40+...}{5} = 32.6$$

$$S^2 = \frac{\sum X_i{}^2 - n\cdot \bar{X}^2}{n-1} = \frac{34^2+40^2-...-5\cdot 32.6^2}{5-1} = 23.4$$

$$S = \sqrt{23.4}$$

$$t_{\bar{X}} = \frac{\bar{X}-\mu_0}{\frac{S}{\sqrt{n}}} = \frac{32.6-40}{\frac{4.88}{\sqrt{5}}} = -3.39$$

$$P_V = P_{H_0} = \left(\bar{X} \le 32.6\right) = P\left(t \le -3.39\right)$$

$$d.f = 5-1 = 4$$

$$1\% < P_V < 2.5\%$$

$P_V < \alpha = 0.05$, לכן דוחים את H_0.

<u>מסקנה</u>: ברי"מ של 5% נכריע שהדרך החלופית מהירה יותר.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) קו ייצור אריזות סוכר נארזות כך שהמשקל הממוצע של אריזות הסוכר צריך להיות אחד קילוגרם. בכל יום דוגמים מקו הייצור 5 אריזות במטרה לבדוק האם קו הייצור תקין. בבדיקה דגמו 5 אריזות סוכר ולהלן משקלן בגרמים:
997, 996, 1005, 1008, 1024.

א. רשמו את השערות המחקר.

ב. מהי מובהקות התוצאה? הצג חסמים.

ג. מה המסקנה ברמת מובהקות של 5%?

2) חוקר בדק את הטענה כי פועלים העובדים במשמרת לילה איטיים יותר מפועלים העובדים ביום. ידוע כי משך הזמן הממוצע הדרוש לייצר מוצר מסוים ביום הוא 6 שעות. במדגם מיקרי של 25 פועלים שעבדו במשמרת לילה נמצא כי הזמן הממוצע לייצר אותו מוצר הוא 7 שעות עם סטית תקן של 3 שעות.
מהי ה- α המינימלית שלפיה ניתן להחליט שאכן העובדים במשמרת לילה איטיים יותר?

3) הגובה של מתגייסים לצה"ל מתפלג נורמלית. במדגם של 25 מתגייסים מדדו את הגבהים שלהם בס"מ והתקבלו התוצאות הבאות:
$\sum(x_i-\bar{x})^2=2832$, $\bar{x}=176.2$.
מטרת המחקר היא לבדוק האם תוחלת הגבהים של המתגייסים גבוה מ-174 ס"מ באופן מובהק.
מהי בקרוב מובהקות התוצאה ועל פיה מה תהיה המסקנה ברמת מובהקות של 6%?

תשובות סופיות:

1) א. $H_0:\mu=1000$, $H_1:\mu\neq1000$. ב. $20\%\le P_v\le 50\%$.

ג. ברמת מובהקות של 5% לא נוכל לקבוע שקו הייצור אינו תקין.

2) 10%.

3) 1.01, נקבל את H_0.

© כל הזכויות שמורות לברק קנדל

הקשר בין רווח סמך לבדיקת השערות על תוחלת (ממוצע):

רקע:

ניתן לבצע בדיקת השערות דו צדדית ברמת מובהקות α על μ :

$H_1:\mu\neq\mu_0$, $H_0:\mu=\mu_0$.

על ידי בניית רווח סמך ברמת סמך של $1-\alpha$ ל- μ :

אם μ_0 נופל ברווח ← נקבל את H_0 .

אם μ_0 לא נופל ברווח ← נדחה את H_0 .

דוגמה:

חוקר ביצע בדיקת השערות לתוחלת. להלן השערותיו:

$H_1:\mu\neq 80$, $\alpha=5\%$, $H_0:\mu=80$.

החוקר בנה רווח סמך ברמה של 90% וקיבל: $79<\mu<84$.

האם אפשר לדעת מה מסקנתו, ואם כן מהי?

פתרון (פתרון מלא בהקלטה):

רווח הסמך ברמת סמך של 90% מכיל "80".

ברמת סמך של 95% רווח הסמך יגדל ויכיל "80".

לכן, ברמת מובהקות של 5% נקבל H_0 .

© כל הזכויות שמורות לברק קנדל

שאלות:

1) חוקר רצה לבדוק את ההשערות הבאות: $H_1:\mu\neq 90$, $H_0:\mu=90$.
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.

2) 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 10.4 מ"ג. נלקח מדגם של 60 נבדקים שניזונו מדיאטה זו. נמצא כי ממוצע מספר המיליגרם סוכר היה 115.5 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 90 מ"ג לסמ"ק. האם לדעתך ניתן להסיק על סמך תוצאת סעיף א שהדיאטה משפיעה על רמת הסוכר בדם? הסבירו.

3) יצרן אנטיביוטיקה רושם על גבי התרופות שכמות הפנצלין היא 200 מ"ג לקפסולה.משרד הבריאות ביצע מדגם של 8 קפסולות אקראיות מקו הייצור ומצא שבממוצע יש 196 מ"ג פנצילין לקפסולה עם סטיית תקן מדגמית של 5 מ"ג. בהנחה וכמות הפנצלין בקפסולה מתפלגת נורמלית.

א. בנו רווח סמך ברמת סמך של 95% לממוצע כמות הפנצלין לקפסולה המיוצרת על ידי יצרן האנטיביוטיקה.

ב. בדקו ברמת מובהקות של 5% האם יש אמת באינפורמציה המסופקת על ידי היצרן.

תשובות סופיות:

1) נקבל השערת.

2) א. $112.87\leq\mu\leq 118.13$.

ב. נכריע שהדיאטה משפיעה על תוחלת רמת הסוכר בדם.

3) א. $191.8\leq\mu\leq 200.2$. ב. נכריע שיש אמת בפרסום.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 33 - רווח סמך להפרש תוחלות (ממוצעים) במדגמים בלתי תלויים

תוכן העניינים

1. כששונויות האוכלוסיה ידועות..... 160
2. כששונויות האוכלוסיה לא ידועות ובהנחת שוויון שונויות..... 162

בששונויות האוכלוסייה ידועות:

רקע:

המטרה היא לאמוד את פער התוחלות: $\mu_1 - \mu_2$, כלומר ההבדלים של הממוצעים בין שתי האוכלוסיות.

האומד נקודתי: $\bar{x}_1 - \bar{x}_2$.

התנאים לבניית רווח הסמך:

1. σ^2_1 , σ^2_2 ידועות.
2. X_1 , $X_2 \sim N$ או $n_1, n_2 > 30$.
3. שני מדגמים בלתי תלויים.

רווח סמך: $\left(\bar{x}_1 - \bar{x}_2\right) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2}}$.

אם הערך אפס נופל בגבולות רווח הסמך נגיד שבביטחון של $1-\alpha$, לא קיים הבדל בין התוחלות.

דוגמה (פתרון בהקלטה):

נדגמו 100 תושבים מאזור A והמשכורת הממוצעת הייתה שם 9200 ₪.
כמו כן נדגמו 120 תושבים מאזור B וממוצע המשכורות שהתקבל שם 8700 ₪.
לצורך פתרון נניח שסטיית התקן של המשכורות באוכלוסיית שני האזורים היא 1800 ₪. אמדו ברמת סמך של 90% את הפרש השכר הממוצע בין אזור A לאזור B.

© כל הזכויות שמורות לברק קנדל

שאלות:

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) ציוני IQ מתוכננים כך שיתפלגו נורמאלית עם סטיית תקן של 15. במדגם של 20 נבחנים ישראלים התקבל ממוצע ציונים 104. במדגם של 23 נבחנים אמריקאיים התקבל ממוצע ציונים 99.

א. בנו רווח סמך ברמת סמך של 95% לפער בין ישראל לארה"ב בממוצע הציונים במבחן ה-IQ.

ב. האם קיים הבדל בין ישראלים לאמריקאים מבחינת ממוצע הציונים?

3) חברה להנדסת בניין מעוניינת להשוות ברמת הקשיות של שני סוגי ברגים. ידוע שרמת הקשיות של ברגים מתפלגת נורמלית עם סטיית תקן של 4 יחידות. במדגם של 15 ברגים מסוג א' התקבל רמת קשיות ממוצעת של 28 יחידות ובמדגם של 12 ברגים מסוג ב' התקבל רמת קשיות ממוצעת של 25. עבור אילו רמות בטחון יקבע שאין הבדל בין שני סוגי הברגים מבחינת ממוצע רמת הקשיות שלהם?

תשובות סופיות:

1) (-20,90).

2) א. $-3.99 < \mu_1 - \mu_2 < 13.99$.

ב. לא נוכל לטעון בביטחון של 95% שקיים הבדל בין ישראל לארה"ב.

3) רמות בטחון הגבוהות מ-0.9476.

© כל הזכויות שמורות לברק קנדל

כששונויות האוכלוסייה לא ידועות ובהנחת שוויון שונויות:

רקע:

המטרה היא לאמוד את פער התוחלות: $\mu_1 - \mu_2$, כלומר ההבדלים של הממוצעים בין שתי האוכלוסיות.

האומד נקודתי: $\bar{x}_1 - \bar{x}_2$.

התנאים לבניית רווח הסמך:

1. $\sigma^2_1 = \sigma^2_2$.
2. $X_1, X_2 \sim N$.
3. מדגמים בלתי תלויים.

השונות המשוקללת: כיוון שאנו מניחים שבין שתי האוכלוסיות השונויות שוות אנו אומדים את השונות הזו על ידי שקלול שתי השונויות של שני המדגמים על ידי הנוסחה הבאה: $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$.

דרגות החופש: $d.f = n_1 + n_2 - 2$.

רווח סמך: $(\bar{x}_1 - \bar{x}_2) \pm t^{n_1+n_2-2}_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{S^2_p}{n_1} + \frac{S^2_p}{n_2}}$.

אם הערך אפס נופל בגבולות רווח הסמך נגיד שבביטחון של $1-\alpha$, לא קיים הבדל בין התוחלות.

דוגמה (פתרון בהקלטה):

מחקר מעוניין לבדוק האם קיים הבדל בין תל אביב לבאר שבע מבחינת ההכנסה הממוצעת של אקדמאים.להלן תוצאות המדגם שנעשה:

	תל אביב	באר שבע
מספר האקדמאים	20	10
ממוצע הכנסות של אקדמאים	11,000	9500
סטיית התקן של הכנסות אקדמאים	200	250

בנו רווח סמך ברמת ביטחון של 90% להפרש תוחלות ההכנסה בשני האזורים.
הניחו שהשכר מתפלג נורמלית עם אותה שונות בכל אחד מהאזורים.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) נדגמו 15 ישראלים ו-15 אמריקאים. כל הנדגמים נגשו למבחן IQ.
להלן תוצאות המדגם:

המדינה	ישראל	ארה"ב
גודל המדגם	15	15
סכום הציונים	1560	1470
סכום ריבועי הציונים	165,390	147,560

מצאו רווח סמך ברמת סמך של 95% לסטייה בין ממוצע הציונים בישראל לממוצע הציונים בארה"ב. רשמו את כל ההנחות הדרושות לצורך פתרון התרגיל.

2) להלן 4 תצפיות על משתנה X שמתפלג: $N\left(\mu_x,\sigma^2\right)$, ומשתנה Y שמתפלג: $N\left(\mu_y,\sigma^2\right)$.

X	22	20	21	25
Y	18	25	17	12

חשבו רווח סמך ל- $\mu_y-\mu_x$ ברמת הסמך 90%, בהנחה ששני המדגמים בלתי תלויים.

תשובות סופיות:

1) הנחות:
1. השונות שווה.
2. שהציונים מתפלגים נורמלית.
3. המדגמים אינם תלויים זה בזה.

$-5.52<\mu_1-\mu_2<17.52$.

2) $-9.6<\mu_y-\mu_x<1.6$.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

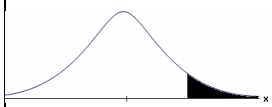
פרק 34 - בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

תוכן העניינים

1. כששונויות האוכלוסיה לא ידועות ומניחים שהן שוות 164
2. כששונויות האוכלוסייה ידועות 168

בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

כששונויות האוכלוסייה לא ידועות ומניחים שהן שוות – רקע

השערת האפס: השערה אלטרנטיבית:	$H_0 \quad \mu_1-\mu_2=c$ $H_1 \quad \mu_1-\mu_2\neq c$	$H_0 \quad \mu_1-\mu_2=c$ $H_1 \quad \mu_1-\mu_2<c$	$H_0 \quad \mu_1-\mu_2=c$ $H_1 \quad \mu_1-\mu_2>c$
תנאים:	1. מדגמים בלתי תלויים 2. σ_1, σ_2 לא ידועות אך שוות 3. המשתנים בכל אוכלוסייה מתפלגים נורמלית		
אזור הדחייה של H_0:	$t_{\bar{x}_1-\bar{x}_2} < -t^{(n_1+n_2-2)}_{1-\frac{\alpha}{2}}$ או $t_{\bar{x}_1-\bar{x}_2} > t^{(n_1+n_2-2)}_{1-\frac{\alpha}{2}}$ H_0 דוחים את	$t_{\bar{x}_1-\bar{x}_2} < -t^{(n_1+n_2-2)}_{1-\alpha}$ H_0 דוחים את	$t_{\bar{x}_1-\bar{x}_2} > t^{(n_1+n_2-2)}_{1-\alpha}$  H_0 דוחים את

סטטיסטי המבחן:

$$t_{\bar{x}_1-\bar{x}_2}=\frac{(\bar{x}_1-\bar{x}_2)-c}{\sqrt{\frac{S_p^{\,2}}{n_1}+\frac{S_p^{\,2}}{n_2}}}$$

השונות המשוקללת:

$$S_p^{\,2}=\frac{(n_1-1)S_1^{\,2}+(n_2-1)S_2^{\,2}}{n_1+n_2-2}$$

© כל הזכויות שמורות לברק קנדל

חלופה אחרת לכלל הכרעה:

נדחה H_0 אם מתקיים:	
$\bar{x}_1-\bar{x}_2 > c+t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}\cdot\sqrt{\frac{S_p^2}{n_1}+\frac{S_p^2}{n_2}}$ **או** $\bar{x}_1-\bar{x}_2 < c-t_{1-\frac{\alpha}{2}}^{(n_1+n_2-2)}\cdot\sqrt{\frac{S_p^2}{n_1}+\frac{S_p^2}{n_2}}$	$\bar{x}_1-\bar{x}_2 < c-t_{1-\alpha}^{(n_1+n_2-2)}\cdot\sqrt{\frac{S_p^2}{n_1}+\frac{S_p^2}{n_2}}$
$\bar{x}_1-\bar{x}_2 > c+t_{1-\alpha}^{(n_1+n_2-2)}\cdot\sqrt{\frac{S_p^2}{n_1}+\frac{S_p^2}{n_2}}$	

דוגמה (פתרון בהקלטה):

חברה המייצרת מוצרי בנייה טוענת שפיתחה סגסוגת (תערובת מתכות) שטמפרטורת ההתכה שלה גבוהה משמעותית מטמפרטורת ההתכה של הסגסוגת לבנייה שמשתמשים בה כיום לבניית בניינים. לצורך בדיקת טענת המחקר נדגמו 10 יחידות של מתכות מהסוג הישן ו-12 יחידות של מתכות מהסוג החדש. להלן תוצאות המדגם:

טמפרטורת ההתכה הממוצעת במתכת הישנה 1170 מעלות עם אומד חסר הטיה לשונות $S^2=200$.

טמפרטורת ההתכה הממוצעת במתכת החדשה 1317 מעלות עם אומד חסר הטיה לשונות $S^2=260$.

נניח לצורך פתרון שטמפרטורת ההתכה מתפלגת נורמאלית עם אותה שונות במתכות השונות. בדקו ברמת מובהקות של 5%.

© כל הזכויות שמורות לברק קנדל

שאלות

1) להלן נתונים של שטחי דירות מתוך דירות שנבנו בשנת 2012 ובשנת 2013 (במ"ר):

2012	120	94	90	130	95	112	120
2013	100	82	91	105	74	69	

בדקו שבשנת 2013 הייתה ירידה משמעותית בשטחי הדירות לעומת שנת 2012 עבור רמת מובהקות של 5%.
הניחו ששטחי הדירות בכל שנה מתפלגים נורמלית עם אותה שונות.

2) נדגמו 15 ישראלים ו-15 אמריקאים. כל הנדגמים נגשו למבחן IQ. להלן תוצאות המדגם:

המדינה	ישראל	ארה"ב
גודל המדגם	15	15
סכום הציונים	1560	1470
סכום ריבועי הציונים	165,390	147,560

בדקו ברמת מובהקות של 5% האם קיים הבדל של נקודה בין ישראלים לאמריקאים מבחינת ממוצע הציונים במבחן ה-IQ לטובת ישראל.
רשמו את כל ההנחות הדרושות לצורך פתרון התרגיל.

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

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

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

ב. עבור איזו רמת מובהקות ניתן לקבוע שנורות מסוג W 60 דולקות בממוצע יותר מאשר נורות מסוג 100?

ג. בדקו ברמת מובהקות של 5% האם נורות מסוג W 60 דולקות יותר מ 1000 שעות. רשמו את כל ההנחות הדרושות.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות

1) נדחה את H_0.

2) הנחות:

1. סטיות התקן שוות.
2. המשתנים מתפלגים נורמלית.

נקבל את H_0.

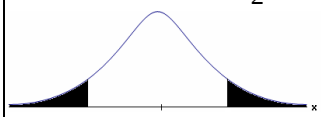
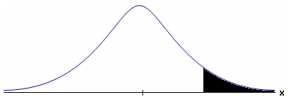
3) א. נדחה את H_0. ב. רמת מובהקות של לפחות 5%.

ג. לא נדחה את H_0.

© כל הזכויות שמורות לברק קנדל

בדיקת השערות על הפרש תוחלות במדגמים בלתי תלויים

כשהשונויות של האוכלוסייה ידועות – רקע

השערת האפס: השערה אלטרנטיבית:	$H_0 \quad \mu_1-\mu_2=c$ $H_1 \quad \mu_1-\mu_2\neq c$	$H_0 \quad \mu_1-\mu_2=c$ $H_1 \quad \mu_1-\mu_2<c$	$H_0 \quad \mu_1-\mu_2=c$ $H_1 \quad \mu_1-\mu_2>c$
תנאים:	מדגמים בלתי תלויים ידועות σ_1,σ_2 $X_1,X_2 \sim N$ או מדגמים מספיק גדולים		
כלל ההכרעה: אזור הדחייה של H_0:	$Z_{\bar{x}_1-\bar{x}_2} < -Z_{1-\frac{\alpha}{2}}$ או $Z_{\bar{x}_1-\bar{x}_2} > Z_{1-\frac{\alpha}{2}}$  $-Z_{1-\frac{\alpha}{2}}$ $Z_{1-\frac{\alpha}{2}}$ H_0 דוחים את ■	$Z_{\bar{x}_1-\bar{x}_2} < -Z_{1-\alpha}$ $-Z_{1-\alpha}$ H_0 דוחים את ■	$Z_{\bar{x}_1-\bar{x}_2} > Z_{1-\alpha}$  $Z_{1-\alpha}$ H_0 דוחים את ■

סטטיסטי המבחן: $$Z_{\bar{x}_1-\bar{x}_2}=\frac{\bar{x}_1-\bar{x}_2-c}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

© כל הזכויות שמורות לברק קנדל

חלופה אחרת לכלל הכרעה:

נדחה H_0 אם מתקיים:	
$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ **או** $\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$\bar{x}_1 - \bar{x}_2 > c + Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$
$\bar{x}_1 - \bar{x}_2 < c - Z_{1-\alpha} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	

התפלגות הפרש הממוצעים: $\bar{x}_1 - \bar{x}_2 \sim N(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2})$

התקנון: $Z_{\bar{x}_1 - \bar{x}_2} = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

דוגמה (פתרון בהקלטה):

בשנת 2004 הפער בין השכר הממוצע של הגברים לנשים היה 3000₪ לטובת הגברים.
מעוניינים לבדוק האם כיום הצטמצם הפער בין הגברים לנשים מבחינת השכר הממוצע.
נדגמו 100 עובדים גברים. שכרם הממוצע היה 9,072₪. נדגמו 80 עובדות, שכרן הממוצע היה 7809₪. לצורך פתרון נניח שסטיות התקן של השכר ידועות ושוות ל-2000₪ באוכלוסיית הנשים ו-3000₪ באוכלוסיית הגברים.
מה המסקנה ברמת מובהקות של 5%?

© כל הזכויות שמורות לברק קנדל

שאלות

1) מחקר טוען שאנשים החיים במרכז הארץ צופים בממוצע בטלוויזיה יותר מאנשים שלא חיים במרכז. נדגמו 100 אנשים מהמרכז ו-107 אנשים לא מהמרכז. אנשים אלה נשאלו כמה שעות ביום הם נוהגים לצפות בטלוויזיה.
במדגם של מרכז הארץ התקבל ממוצע 2.7 שעות.
במדגם של מחוץ למרכז הארץ התקבל ממוצע 1.8 שעות.
לצורך פתרון הניחו שבכל אזור, סטיית התקן היא שעה 1 ביום.
בדקו את טענת המחקר ברמת מובהקות של 1%.

2) 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%.

3) 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.

א. מהי מובהקות התוצאה לבדיקה האם התפוקה הממוצעת היומית גבוהה מהתפוקה הממוצעת הלילית.

ב. מה תהיה המסקנה ברמת מובהקות של 8%?

4) במחקר מקיף שנעשה באירופה נקבע שגברים גבוהים מנשים ב-8 ס״מ בממוצע. מחקר ישראלי מתעניין לבדוק האם בישראל הפער גדול יותר. לצורך המחקר נדגמו 40 גברים ו 40 נשים באקראי. כמו כן, נניח שסטיות התקן של הגברים והנשים ידועות ושוות ל-6 ס״מ אצל הנשים ו-12 ס״מ אצל הגברים.

א. מהן השערות המחקר ומהו כלל ההכרעה ברמת מובהקות של 10%?

ב. אם בישראל הפער בין גברים לנשים מבחינת הגובה הממוצע הוא 11 ס״מ, מה ההסתברות שהמחקר לא יגלה זאת? איך קוראים להסתברות הזאת?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות

1) נדחה H_0.

2) לא נדחה את H_0.

3) א. 0 ב. נדחה את H_0.

4) א. נדחה את H_0, אם במדגם הגברים יהיו גבוהים בממוצע מהנשים ביותרמ-10.72 ס״מ.

ב. 0.6331

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 35 - רווח סמך לתוחלת (ממוצע) ההפרשים במדגמים מזווגים

תוכן העניינים

1. רווח סמך לתוחלת (ממוצע) ההפרשים במדגמים מזווגים 172

רווח סמך לתוחלת (ממוצע) ההפרשים במדגמים מזווגים:

רקע:

מדגם מזווג: מדגם אחד שבו יש n צמדים. כל תצפית במדגם תנפק זוג ערכים: X ו- Y.
ניצור משתנה חדש: $D = x - y$.

הפרמטר שנרצה לאמוד: μ_D.

התנאים לבניית רווח הסמך:

1. $x, y \sim N$.
2. המדגם מזווג.

נוסחת רווח הסמך: $\bar{D} \pm t^{n-1}_{1-\frac{\alpha}{2}} \frac{S_D}{\sqrt{n}}$.

כאשר דרגות החופש: $df = n-1$.

דוגמה (פתרון בהקלטה):

מעוניינים לבדוק האם יש הבדל בין מהירות הריצות של שתי תוכנות מחשב.
לקחו 5 קבצים אקראיים והריצו אותם בשתי התוכנות:

הקובץ	1	2	3	4	5
הזמן בתוכנה הראשונה	25	48	49	46	38
הזמן בתוכנה השנייה	27	46	42	40	48

הניחו כי זמני הריצות מתפלגים נורמלית.
מִצאו רווח סמך של 95% להפרש תוחלת הזמן בין שתי התוכנות.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) נדגמו 5 סטודנטים שסיימו את הקורס סטטיסטיקה ב'.
להלן הציונים בסמסטר א' ו-ב':

סמסטר א'	74	68	90	75	82
סמסטר ב'	80	84	87	76	100

נניח שהציונים מתפלגים נורמאלית.

א. בנו רווח סמך ברמת סמך של 95% לתוחלת פער הציונים בין סמסטר א' לבין סמסטר ב'.

ב. האם על סמך רווח הסמך קיים הבדל בין הסמסטרים מבחינת תוחלת הציונים?

ג. מה צריך לשנות בנתונים כדי שהמדגמים יהיו בלתי תלויים?

2) במטרה לבדוק האם קיים הבדל בין קווי זהב לבזק מבחינת ממוצע המחירים לשיחות בינ"ל. נדגמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקת שיחה.
להלן התוצאות:

חברה/ מדינה	**ארה"ב**	**קנדה**	**הולנד**	**פולין**	**מצרים**	**סין**	**יפן**
בזק - X	1.5	2.1	2.2	3	3.5	3.2	4.2
קווי זהב - Y	1.4	2	1.9	3.1	3.3	3.2	4.2

בהנחה והמחירים מתפלגים נורמלית עבור כל חברה,
בנו רווח סמך ברמת סמך של 90% לתוחלת הפרש המחירים של שתי החברות.

תשובות סופיות:

1) א. $-19 < \mu_0 < 38$. ב. בביטחון של 95% לא קיים הבדל.
ג. ראה הסבר בסרטון.

2) $-0.013 < \mu < 0.185$.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

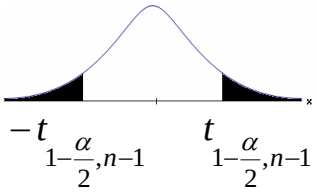
פרק 36 - בדיקת השערות לתוחלת ההפרש במדגמים מזווגים

תוכן העניינים

1. בדיקת השערות למדגמים מזווגים 174

בדיקת השערות על תוחלת ההפרשים במדגמים מזווגים (תלויים)

בדיקת השערות למדגמים מזווגים – רקע

השערת האפס: השערה אלטרנטיבית:	$H_0: \mu_D = C$ $H_1: \mu_D \neq C$	$H_0: \mu_D = C$ $H_1: \mu_D < C$	$H_0: \mu_D = C$ $H_1: \mu_D > C$
תנאים:	1. σ_D אינה ידועה 2. $D \sim N$ או מדגם מספיק גדול		
כלל ההכרעה: אזור הדחייה של H_0:	$t_{\bar{D}} > t^{(n-1)}_{1-\frac{\alpha}{2}}$ או $t_{\bar{D}} < -t^{(n-1)}_{1-\frac{\alpha}{2}}$  ■ - דוחים את H_0	$t_{\bar{D}} < -t^{(n-1)}_{1-\alpha}$ ■ - דוחים את H_0	$t_{\bar{D}} > t^{(n-1)}_{1-\alpha}$ ■ - דוחים את H_0
חלופה לכלל הכרעה: נדחה H_0 אם מתקיים:	$\bar{D} > C + t^{n-1}_{1-\frac{\alpha}{2}} \cdot \frac{S_D}{\sqrt{n}}$ או $\bar{D} < C - t^{n-1}_{1-\frac{\alpha}{2}} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} < C - t^{n-1}_{1-\alpha} \cdot \frac{S_D}{\sqrt{n}}$	$\bar{D} > C + t^{n-1}_{1-\alpha} \cdot \frac{S_D}{\sqrt{n}}$

סטטיסטי המבחן: $t_{\bar{D}} = \dfrac{\bar{D} - \mu_D}{S_D / \sqrt{n}}$, $S_D^2 = \dfrac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1} = \dfrac{\sum_{i=1}^{n} D_i^2 - n\bar{D}^2}{n-1}$

דוגמה (פתרון בהקלטה):

© כל הזכויות שמורות לברק קנדל

חברה שיווקית מעוניינת לבדוק את טענת רשת השיווק "מגה בעיר" הטוענת שמחיריה נמוכים מהמחירים מרשת השיווק "שופרסל".
לצורך הבדיקה נבחרו באקראי 4 מוצרים שונים. המחירים נבדקו בשתי הרשתות.
להלן המחירים:

המוצר / רשת	מגה בעיר	שופרסל
שמפו	17	18
ג'ל כביסה	48	57
עוגת גבינה	35	35
לחם	12	10
קפה נמס	49	47
בקבוק יין	113	142
גבינה בולגרית	20	26

בהנחה והמחירים מתפלגים נורמאלית, בדקו ברמת מובהקות של 5% את טענת רשת "מגה בעיר".

© כל הזכויות שמורות לברק קנדל

שאלות

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

חברה/ מדינה	ארה"ב	קנדה	הולנד	פולין	מצרים	סין	יפן
X	1.5	2.1	2.2	3	3.5	3.2	4.2
Y	1.4	2	1.9	3.1	3.2	3.2	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 מתפלגים נורמלית.

3) נדגמו 5 סטודנטים שסיימו את הקורס סטטיסטיקה ב'. להלן הציונים שלהם בסמסטר א' ו- ב':

סטטיסטיקה א'	74	68	90	75	82
סטטיסטיקה ב'	80	84	87	76	100

פורסם שתלמידים שמסיימים את סמסטר ב משפרים בממוצע את הציונים ב-5 נקודות לעומת סמסטר א'. הניחו שהציונים מתפלגים נורמלית.

א. מהי מובהקות התוצאה לבדיקת הטענה שהשיפור הוא יותר מ 5 נקודות?

ב. על סמך הסעיף הקודם, מהי רמת המובהקות המינימלית להכרעה שהשיפור הוא יותר מ- 5 נקודות?

ג. לאור זאת, מה המסקנה ברמת מובהקות של 10%?

4) לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים. אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני ללא היפנוזה. לאחר מכן נערך לשניהם מבחן באנגלית. נניח שציוני המבחן מתפלגים נורמאלית ללא ידיעת השונות האמתית. המבחן שיש לבצע כאן הוא:

א. מבחן Z למדגם יחיד.

ב. מבחן T למדגם יחיד.

ג. מבחן T למדגמים בלתי תלויים.

ד. מבחן T למדגמים מזווגים.

© כל הזכויות שמורות לברק קנדל

5) בתחנת טיפת חלב מסוימת יש שני מכשירי שקילה. על מנת להשוות בין שני המשקלים נדגמו 4 תינוקות. כל תינוק בן חודשיים נשקל בכל אחד מהמשקלים. להלן תוצאות השקילה (בק"ג):

משקל במכשיר 1	4.5	9.6	0.7	2.5
משקל במכשיר 2	3.5	6.9	1.7	0.5

נניח שהמשקלים מתפלגים נורמלית, המבחן שיש לבצע כאן הוא:

א. מבחן Z למדגם יחיד.

ב. מבחן T למדגם יחיד.

ג. מבחן T למדגמים בלתי תלויים.

ד. מבחן T למדגמים מזווגים.

6) כדי להשוות בין שני אצנים נדגמו 5 תוצאות מריצת 100 מטר של כל אצן. זמני הריצה נרשמו ויש להניח שמתפלגים נורמלית. המטרה להשוות בין האצנים. המבחן שיש לבצע כאן הוא:

א. מבחן Z למדגם יחיד.

ב. מבחן T למדגם יחיד.

ג. מבחן T למדגמים בלתי תלויים.

ד. מבחן T למדגמים מזווגים.

תשובות סופיות

1) לא נדחה H_0.

2) לא נדחה H_0.

3) א. $0.25 \le pv \le 0.5$ ב. 0.5 ג. לא נדחה H_0.

4) ד'.

5) ד'.

6) ג'.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 37 - הקשר בין רווח סמך לבדיקת השערות להפרש תוחלות

תוכן העניינים

1. הקשר בין רווח סמך לבדיקת השערות להפרש תוחלות 178

הקשר בין רווח סמך לבדיקת השערות על הפרש תוחלות

רקע

ניתן לבצע בדיקת השערות דו צדדית ברמת מובהקות α על $\mu_1-\mu_2$:

$H_0:\mu_1-\mu_2=C,\qquad H_1:\mu_1-\mu_2\neq C$.

על ידי בניית רווח סמך ברמת סמך של $1-\alpha$ ל- $\mu_1-\mu_2$:

אם C נופל ברווח ← נקבל את H_0 .

אם C לא נופל ברווח ← נדחה את H_0 .

דוגמה (פתרון בהקלטה) :

חוקר ביצע בדיקת השערות לתוחלת ההפרש במדגם מזווג.

להלן השערותיו : $\alpha=5\%$, $H_1:\mu_D\neq 80,\qquad H_0:\mu_D=80$.

החוקר בנה רווח סמך ברמה של 90%, $78<\mu_D<83$.

האם אפשר לדעת מה מסקנתו, ואם כן מהי?

© כל הזכויות שמורות לברק קנדל

שאלות

1) נדגמו 5 סטודנטים שסיימו את הקורס סטטיסטיקה ב'.
להלן ציוניהם בסמסטר א' ו- ב':

סמסטר א	סמסטר ב
74	80
68	84
90	87
75	76
82	100

א. בנו רווח סמך ברמת סמך של 95% לתוחלת פער הציונים בין סמסטר א' לבין סמסטר ב'.

ב. פורסם שתלמידים שמסיימים את סמסטר ב משפרים בממוצע את הציונים ב-5 נק' לעומת סמסטר א'. האם יש אמת בפרסום?

2) הוחלט להשוות הציונים אצל מרצה X ואצל מרצה Y. נבחרו באקראי 6 סטודנטים, 3 סטודנטים של מרצה X ו-3 סטודנטים של מרצה Y, עבורם התקבלו הציונים הבאים:

מרצה X	82	90	68
מרצה Y	68	81	64

א. חשבו רווח סמך ברמת סמך 90% להפרש בין התוחלות של הציונים אצל שני המרצים.

ב. האם ברמת מובהקות של 10% נכריע שיש הבדל בין תוחלות הציונים אצל שני המרצים?

שאלות רב-ברירה:

3) סטטיסטיקאי נתבקש לאמוד את הפרש הממוצעים של שני טיפולים לפי שני מדגמים מקריים בלתי תלויים.
הוא חישב רווח סמך להפרש ברמת סמך 0.98, וקיבל את הרווח $-2 < \mu_1 - \mu_2 < 4.5$.
אילו יתבקש החוקר לבדוק לפי אותם נתונים את ההשערות:
$H_0 : \mu_1 - \mu_2 = 0$; $H_1 : \mu_1 - \mu_2 \neq 0$ ברמת מובהקות 0.05, מסקנתו תהיה:

א. לדחות את השערת האפס.

ב. לא לדחות את השערת האפס.

ג. שלא ניתן לדעת את המסקנה עבור רמת מובהקות 0.05 .

ד. שלא נתונות בשאלה סטיות התקן של האוכלוסיות, ולכן לא ניתן להסיק דבר.

© כל הזכויות שמורות לברק קנדל

4) במטרה לבדוק האם קיים הבדל בין קווי זהב לבזק 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 וקיבלו : $-0.0293 < \mu_D < 0.2145$, רווח הסמך הוא ברמת סמך של 95%.
לכן מסקנת המחקר היא :

א. ברמת מובהקות של 5% לא נוכל לקבוע שקיים הבדל בין החברות.

ב. ברמת מובהקות של 5% נקבע שקיים הבדל מובהק בין החברות.

ג. לא ניתן לדעת מה המסקנה ברמת מובהקות של 5% כיוון שלא נאמר מה ההגדרה של D.

תשובות סופיות

1) א. $-3.8 \le \mu_D \le 19$ ב. נכריע שיש אמת בפרסום.

2) א. $-8.5 \le \mu_X - \mu_Y \le 26.5$ ב. נכריע שאין הבדל.

3) ג.

4) א.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 38 - רווח סמך לשונות וסטיית תקן

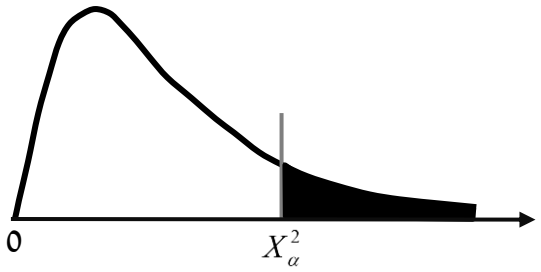
תוכן העניינים

1. רווח סמך לשונות וסטיית תקן גרסה 2 181

רווח סמך לשונות וסטיית תקן:

רקע:

בפרק זה נדון על בניית רווח סמך לשונות האוכלוסייה.
התנאי לבניית רווח הסמך : המשתנה הנחקר מתפלג נורמלית,
למרות שנהוג לא לדרוש את התנאי הזה אם המדגם מספיק גדול.
רווח הסמך יתבסס על התפלגות הנקראת חי בריבוע. התפלגות זו היא התפלגות אסימטרית חיובית המתחילה מהערך אפס ותלויה בדרגות חופש.
דרגות החופש במקרה זה יהיו : $n-1$.



רווח הסמך לשונות : $\frac{(n-1)\hat{S}^2}{\chi^2_{1-\frac{\alpha}{2},n-1}} \le \sigma^2 \le \frac{(n-1)\hat{S}^2}{\chi^2_{\frac{\alpha}{2},n-1}}$.

כאשר : $\hat{S}^2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1} = \frac{\sum_{i=1}^{n}X^2{}_i - n\cdot\bar{X}^2}{n-1}$ אומד לשונות הלא-ידועה.

אם נרצה לבנות רווח סמך לסטיית תקן אז נוציא שורש לרווח סמך לשונות.

© כל הזכויות שמורות לברק קנדל

דוגמה:

זמן התגובה מתפלג נורמאלית. במטרה לאמוד את שונות זמן התגובה נדגמו 4 תצפיות.
להלן התוצאות בשניות: 5.3, 4.6, 5.2, 4.7.
בנו רווח סמך, ברמת סמך של 95%, לשונות זמן התגובה באוכלוסייה.

פתרון:

פרמטר: σ^2.

$X \sim N\left(\mu,\sigma^2\right)$ = זמן תגובה (בשניות).

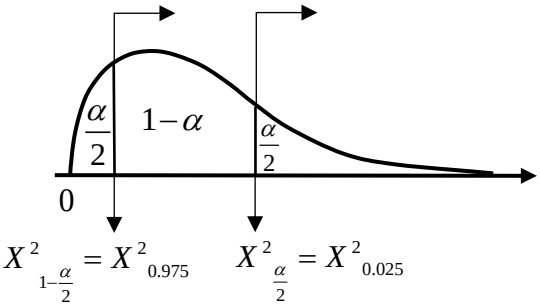
תוצאות מדגם: $n=4$.

$$\bar{X}=\frac{4.7+5.2+4.6+5.3}{4}=4.95$$

$$d.f=n-1=4-1=3$$

נציב: $$S^2=\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}{n-1}=\frac{\sum_{i=1}^{n}X^2{}_i-n\cdot\bar{X}^2}{n-1}$$

$$S^2=\frac{4.7^2+5.2^2...-4\cdot4.95^2}{4-1}=0.123$$



$$1-\alpha=0.95$$
$$\alpha=0.05$$
$$\frac{\alpha}{2}=0.025$$

(טבלת התפלגות חי-בריבוע מופיעה בעמוד האחרון).

נציב: $$\frac{(n-1)S^2}{\chi^2_{\frac{1-\alpha}{2},n-1}}\le\sigma^2\le\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2},n-1}}$$

$$\frac{(4-1)\cdot0.123}{9.35}<\sigma^2<\frac{(4-1)\cdot0.123}{0.216}$$

$$0.039<\sigma^2<1.708$$

© כל הזכויות שמורות לברק קנדל

שאלות:

1) חמישה מטופלים קבלו תרופה מסוימת. בדקו לכל מטופל את זמני התגובה שלו. להלן הזמנים שהתקבלו בדקות: 28, 26, 21, 17, 18.
בהנחה וזמני התגובה מתפלגים נורמאלית, בנו רווח סמך ברמת סמך של 95% לשונות זמן התגובה.

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 נורמאלית:

א. בנו רווח סמך לתוחלת הטמפ' בחודשים אלה בת"א ברמת סמך של 95%.

ב. בנו רווח סמך לסטיית התקן של הטמפ' בחודשים אלה בת"א ברמת סמך של 95%.

3) ציוני IQ בארה"ב מתפלגים נורמאלית עם ממוצע 100 וסטיית תקן 5.

נבחנו 20 נבחנים ישראלים במבחן ה-IQ.

להלן התוצאות שהתקבלו: $\sum_{i=1}^{20} X_i^2 = 218,220$, $\sum_{i=1}^{20} X_i = 2080$.

נניח שגם בישראל הציונים מתפלגים נורמאלית.

א. מצאו אומדנים לממוצע הציונים בישראל ולשונות הציונים בישראל באמצעות אומדנים חסרי הטיה.

ב. אמדו ברמת ביטחון של 95% את תוחלת הציונים של נבחנים בישראל.

ג. אמדו ברמת סמך של 90% את סטיית התקן של הציונים של נבחנים ישראלים.

ד. על סמך הסעיפים הקודמים, האם בישראל ממוצע הציונים וסטיית התקן של הציונים שונה מבארה"ב? הסבירו.

4) באוכלוסייה מסוימת נדגמו 10 תצפיות והתקבלו התוצאות הבאות:

$\sum_{i=1}^{10} X_i = 750$, $\sum_{i=1}^{10} \left(X_i - \bar{X}\right)^2 = 900$.

נתון ש: $X_i \sim N\left(\mu, \sigma^2\right)$.

א. בנו רווח סמך ל- μ ברמת סמך של 95%.

ב. בנו רווח סמך ל- σ^2 ברמת סמך של 95%.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) $8.4 < \sigma^2 < 194.2$.

2) א. $30.285 < \mu < 31.315$ ב. $0.836 < \sigma < 1.606$.

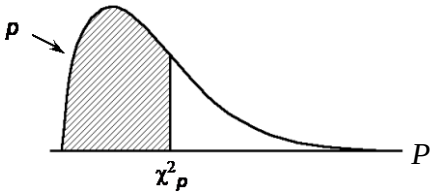
3) א. ממוצע: 104, שונות: 100. ב. $99.32 \le \mu \le 108.68$ ג. $7.94 < \sigma < 13.7$.

ד. בבטחון של 95% תוחלת הציונים בישראל אינה שונה משל ארה"ב.
בבטחון של 90% סטית התקן של הציונים בישראל שונה משל ארה"ב.

4) א. $68.75 < \mu < 82.15$. ב. $47.4 < \sigma^2 < 333.3$.

© כל הזכויות שמורות לברק קנדל

נספח - טבלת התפלגות חי-בריבוע – ערכי החלוקה $\chi^2 p$:



df	.005	.01	.025	.05	.10	.25	.50	.75	.90	.95	.975	.99	.995
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	0.0158	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	0.0717	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 39 - רווח סמך ליחס שונויות

תוכן העניינים

1. רווח סמך ליחס שונויות..... 186

רווח סמך ליחס שונויות:

רקע:

נרצה לאמוד את ההבדל בין שתי שונויות משתי אוכלוסיות שונות.

הפרמטר יהיה: $\frac{\sigma_1^2}{\sigma_2^2}$: כלומר היחס בין השונויות.

התנאים:

- $X_1, X_2 \sim N$ או מדגמים גדולים.
- מדגמים בלתי תלויים.

רווח הסמך יבנה על סמך התפלגות הנקראת התפלגות F. התפלגות זו היא אסימטרית חיובית ומושפעת משתי דרגות החופש, זו של המונה וזו של המכנה.

$$df_1 = n_1 - 1$$
$$df_2 = n_2 - 1$$

רווח הסמך יהיה: $\frac{s_1^2}{s_2^2}\frac{1}{F_{1-\frac{\alpha}{2}}(n_1-1,n_2-1)} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2}{s_2^2}F_{1-\frac{\alpha}{2}}(n_2-1,n_1-1)$.

דוגמה (פתרון בהקלטה):

מחקר סוציולוגי מעוניין לחקור את הרגלי הבילויים בקבוצות גיל שונות:
במדגם שנעשה על סטודנטים בגילאי 21-26 התקבל אומד חוסר הטיה לשונות ההוצאה החודשית על בילויים 10,000. כמות הסטודנטים שנדגמו 16.
במדגם שנעשה על 11 מבוגרים בשנות השלושים התקבל אומד חסר הטיה לשונות ההוצאה החודשית על בילויים 490,000.
נניח שההוצאה החודשית לבילוי בכל קבוצת גיל מתפלג נורמאלית.
בנו רווח סמך ברמת סמך של 95% ליחס בין השונויות.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) בתחום הבינוי משתמשים בשני סוגי מתכות: מתכת A ומתכת B. מחקר מעוניין לבדוק האם קיים הבדל בין שני סוגי המתכות מבחינת שונות החוזק שלהן. דגמו מספר יחידות מתכת מכל סוג והתקבלו התוצאות הבאות:

סוג המתכת	A	B
N	8	10
$\sum X_i$	16	30
$\sum X_i^2$	60	198

יש להניח שרמת החוזק של המתכות מתפלגת נורמאלית.

א. בנו רווח סמך ליחס השונויות של רמות החוזק בין שני סוגי המתכות ברמת סמך של 90%.

ב. בנו רווח סמך ליחס סטיות התקן של רמות החוזק בין שני סוגי המתכות ברמת סמך של 90%.

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

אמוד ברמת ביטחון של 95% פי כמה גדולה השונות של הגברים באוכלוסייה מהשונות של הנשים.

מה יש להניח לצורך פתרון?

תשובות סופיות:

1) א. $0.1013<\frac{\sigma_A^2}{\sigma_B^2}<1.2267$. ב. $0.318<\frac{\sigma_A}{\sigma_B}<1.108$.

2) $0.39\le\frac{\sigma_m^2}{\sigma_F^2}\le 5.26$.

© כל הזכויות שמורות לברק קנדל

טבלת התפלגות F לפי זנב ימני של α

		$d.f.1.$											
α	$d.f.2$	1	2	3	4	5	6	7	8	9	10	12	15
.05	1	161	200	216	225	230	234	237	239	241	242	244	246
.025		648	800	864	900	922	937	948	957	963	969	977	985
.01		4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6106	6157
.005		16211	20000	21615	22500	23056	23437	23715	23925	24091	24224	24426	24630
.05	2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43
.025		38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40	39.41	39.43
.01		98.50	99.00	99.16	99.25	99.30	99.33	99.36	99.38	99.39	99.40	99.42	99.43
.005		198.50	199.01	199.16	199.24	199.30	199.33	199.36	199.38	199.39	199.39	199.42	199.43
.05	3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70
.025		17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25
.01		34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.05	26.87
.005		55.55	49.80	47.47	46.20	45.39	44.84	44.43	44.13	43.88	43.68	43.39	43.08
.05	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86
.025		12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66
.01		21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20
.005		31.33	26.28	24.26	23.15	22.46	21.98	21.62	21.35	21.14	20.97	20.70	20.44
.05	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62
.025		10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43
.01		16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72
.005		22.78	18.31	16.53	15.56	14.94	14.51	14.20	13.96	13.77	13.62	13.38	13.15
.05	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94
.025		8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27
.01		13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56
.005		18.63	14.54	12.92	12.03	11.46	11.07	10.79	10.57	10.39	10.25	10.03	9.81
.05	7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51
.025		8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57
.01		12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31
.005		16.24	12.40	10.88	10.05	9.52	9.16	8.89	8.68	8.51	8.38	8.18	7.97
.05	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22
.025		7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.10
.01		11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52
.005		14.69	11.04	9.60	8.81	8.30	7.95	7.69	7.50	7.34	7.21	7.01	6.81

© כל הזכויות שמורות לברק קנדל

d.f.1.

α	*d.f.2.*	1	2	3	4	5	6	7	8	9	10	12	15
.05	9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01
.025		7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77
.01		10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96
.005		13.61	10.11	8.72	7.96	7.47	7.13	6.88	6.69	6.54	6.42	6.23	6.03
.05	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85
.025		6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52
.01		10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56
.005		12.83	9.43	8.08	7.34	6.87	6.54	6.30	6.12	5.97	5.85	5.66	5.47
.05	12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62
.025		6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18
.01		9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01
.005		11.75	8.51	7.23	6.52	6.07	5.76	5.52	5.35	5.20	5.09	4.91	4.72
.05	15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40
.025		6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86
.01		8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52
.005		10.80	7.70	6.48	5.80	5.37	5.07	4.85	4.67	4.54	4.42	4.25	4.07
.05	20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20
.025		5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57
.01		8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09
.005		9.94	6.99	5.82	5.17	4.76	4.47	4.26	4.09	3.96	3.85	3.68	3.50
.05	30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01
.025		5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31
.01		7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70
.005		9.18	6.35	5.24	4.62	4.23	3.95	3.74	3.58	3.45	3.34	3.18	3.01
.05	60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84
.025		5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06
.01		7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35
.005		8.49	5.79	4.73	4.14	3.76	3.49	3.29	3.13	3.01	2.90	2.74	2.57
.05	120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75
.025		5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	2.05	1.94
.01		6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19
.005		8.18	5.54	4.50	3.92	3.55	3.28	3.09	2.93	2.81	2.71	2.54	2.37
.05	∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67
.025		5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.94	1.83
.01		6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04
.005		7.88	5.30	4.28	3.72	3.35	3.09	2.90	2.74	2.62	2.52	2.36	2.19

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 40 - בדיקת השערות

תוכן העניינים

1. כללי ..... (ללא ספר)

מבוא לסטטיסטיקה בניהול

פרק 41 - רווח סמך לפרופורציה

תוכן העניינים

1. רווח הסמך לפרופורציה 190

רווח הסמך לפרופורציה:

רקע:

המטרה היא לאמוד את P – פרופורציה באוכלוסייה.

האומד הנקודתי:

$\hat{p}=\frac{y}{n}$ (Y - מספר ההצלחות שבמדגם).

רווח הסמך ל- p : $\hat{p}\pm Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

תנאי לבניית רווח הסמך:

מדגם של לפחות 30 תצפיות (לעיתים נותנים תנאי של מספר הצלחות ומספר כשלונות לפחות 5 או לפחות 10).

האומד לטעות התקן: $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

מתקיים ש: $\hat{P}=\frac{A+B}{2}$, $L=2\varepsilon$.

דוגמה (פתרון בהקלטה):

במטרה לאמוד את אחוז המובטלים במשק נדגמו 200 אזרחים, מתוכם התקבל ש-24 היו מובטלים.

א. בנו רווח סמך לאחוז המובטלים באוכלוסייה ברמת סמך של 95%.

ב. מהו האומד לטעות התקן?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) נדגמו 200 דירות בעיר חיפה. 48 מתוכן נמצאו כבעלות ממ"ד.

א. בנו רווח סמך ברמת סמך של 95% לאחוז הדירות בחיפה עם ממ"ד.

ב. על סמך סעיף א' מה ניתן לומר על שגיאת האמידה המקסימאלית?

ג. בהנחה ובחיפה 80 אלף דירות, בנו רווח סמך ברמת סמך של 95% למספר הדירות בחיפה עם ממ"ד בפועל.

2) במדגם של 300 אנשי היי-טק התקבל ש-180 מהם אקדמאים.

א. בנו רווח סמך לפרופורציית אקדמאים ברמת סמך של 95% (בקרב אנשי הייטק).

ב. כיצד רווח הסמך של סעיף א' היה משתנה אם היינו מקטינים את רמת הסמך?

ג. כיצד רווח הסמך היה משתנה אם הינו מגדילים את גודל המדגם?

3) במדגם של 400 נהגים התקבל רווח סמך לפרופורציית הנהגים החדשים: $0.08 < p < 0.18$.

א. כמה נהגים במדגם היו נהגים חדשים?

ב. מהי רמת הסמך של רווח הסמך שנבנה?

4) במסגרת מערכת הבחירות בארה"ב נשאלו 840 אנשים עבור איזה מועמד יצביעו. 510 אנשים ענו כי יצביעו בעד ברק אובמה. בסקר פורסם שתתכן סטייה של $\pm 3\%$ מתוצאות האמת. באיזו רמת ביטחון הסקר השתמש?

5) 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 באוכלוסייה הנחקרת.

ב. מצאו רווח סמך ברמה של 99% לסיכוי שבאוכלוסייה הנחקרת תמצא אישה לא נשואה?

6) ביצעו מדגם באוכלוסייה. שיעור ההצלחות במדגם היה 10% ורווח הסמך ניבנה ברמת סמך של 95%. אורכו הינו 8.3156%. מהו גודל המדגם שנלקח?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. $18.1\% < p < 29.9\%$.
ב. בביטחון של 95% שגיאת האמידה היא לכל היותר 0.059.
ג. $14,480 < \mu < 23,920$.

2) א. $0.545 \le p \le 0.655$. ב. האורך שלו היה קטן. ג. לא ניתן לדעת.

3) א. 52. ב. 0.997.

4) 0.925.

5) א. $22.5\% < p < 30.9\%$. ב. $45.91\% < p < 60.72\%$.

6) 200.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 42 - רווח סמך להפרש פרופורציות

תוכן העניינים

1. רווח סמך להפרש פרופורציות 193

רווח סמך להפרש פרופורציות:

רקע:

המטרה: לאמוד את $p_1 - p_2$: הפרש פרופורציות בין שתי אוכלוסיות שונות.

האומד הנקודתי: $\hat{p}_1 - \hat{p}_2$.

התנאי לבניית רווח הסמך: כל מדגם מעל 30 או לבדוק שמספר ההצלחות ומספר הכישלונות בכל מדגם לפחות 5 בכל מדגם (יש כאלה שבודקים לפחות 10).

רווח סמך: $\left(\hat{p}_1 - \hat{p}_2\right) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1\left(1-\hat{p}_1\right)}{n_1} + \frac{\hat{p}_2\left(1-\hat{p}_2\right)}{n_2}}$.

רק שאפס נופל בתחומי רווח הסמך להפרש הפרופורציה נאמר שלא ניתן לקבוע שקיים הבדל מובהק בין הפרופורציות באוכלוסיות.

דוגמה (פתרון בהקלטה):

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?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) 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 .

2) במחקר רפואי השתתפו 200 אנשים הסובלים מכאבים כרוניים. הם חולקו באקראי ל-2 קבוצות שוות בגודלן. קבוצה 1 קיבלה את תרופה A וקבוצה שנייה קיבלה את תרופה B. בקרב לוקחי תרופה A 90 טענו שמצבם השתפר. בקרב לוקחי תרופה B 70 טענו שמצבם השתפר.

א. בנו רווח סמך ברמת סמך של 95% להפרש בין שיעורי ההצלחה של שתי התרופות.

ב. האם על סמך סעיף א' ניתן לקבוע שקיים הבדל בין התרופות מבחינת שיעורי ההצלחה?

3) נדגמו 200 משפחות מגוש דן. ל-70% מתוכן מכשיר DVD בבית.
נדגמו 300 משפחות מאזור הצפון ל-65% מתוכן מכשיר DVD בבית.

א. בנו רווח סמך ברמת סמך של 98% לפרופורציות המשפחות בגוש דן עם DVD בבית.

ב. בנו רווח סמך ברמת סמך של 95% להפרש בין פרופורציות המשפחות בגוש דן עם DVD לבין פרופורציות המשפחות בצפון עם DVD.

תשובות סופיות:

1) א. $-4.9\% < P_F - P_M < 14.9\%$.

ב. $22.5\% < p < 31.8\%$.

2) א. $0.093 < P_A - P_B < 0.307$.

ב. כן.

3) $0.625 < p < 0.7754$.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

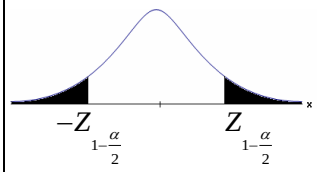
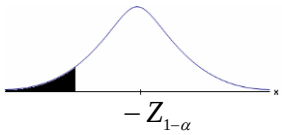
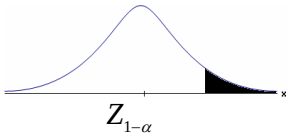
פרק 43 - בדיקת השערות על פרופורציה

תוכן העניינים

1. התהליך 195
2. סיכוי לטעויות ועוצמה 198
3. מובהקות התוצאה - אלפא מינימלית 202

התהליך:

רקע:

השערת האפס: השערה אלטרנטיבית:	$H_0: p = p_0$ $H_1: p \neq p_0$	$H_0: p = p_0$ $H_1: p < p_0$	$H_0: p = p_0$ $H_1: p > p_0$
תנאים:	$np_0 \geq 5 \;\&\; n(1-p_0) \geq 5$		
כלל ההכרעה: אזור הדחייה של H_0:	$Z_{\hat{p}} < -Z_{1-\frac{\alpha}{2}}$ או $Z_{\hat{p}} > Z_{1-\frac{\alpha}{2}}$  H_0 דוחים את	$Z_{\hat{p}} < -Z_{1-\alpha}$  H_0 דוחים את	$Z_{\hat{p}} > Z_{1-\alpha}$  H_0 דוחים את

סטטיסטי המבחן: $Z_{\hat{p}} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$.

חלופה אחרת לכלל הכרעה:

כלל ההכרעה – אזור הדחייה של H_0:		
$\hat{p} > p_0 + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\dfrac{p_0(1-p_0)}{n}}$ $\hat{p} < p_0 - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\dfrac{p_0(1-p_0)}{n}}$	$\hat{p} < p_0 - Z_{1-\alpha} \cdot \sqrt{\dfrac{p_0(1-p_0)}{n}}$	$\hat{p} > p_0 + Z_{1-\alpha} \cdot \sqrt{\dfrac{p_0(1-p_0)}{n}}$

דוגמה (פתרון בהקלטה):

בחודש ינואר השנה פורסם שאחוז האבטלה במשק הוא 8% במדגם עכשווי התקבל שמתוך 200 אנשים 6.5% מובטלים.
בדקו ברמת מובהקות של 5% האם כיום אחוז האבטלה הוא כמו בתחילת השנה.

© כל הזכויות שמורות לברק קנדל

שאלות:

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%?

4) קפיטריה במכללה מסוימת מעריכה כי אחוז הסטודנטים שקונים קפה בקפיטריה הינו 20%. נערך סקר אשר כלל 200 סטודנטים. התברר כי 33 מהם רוכשים קפה בקפיטריה. מטרת הסקר הייתה לבדוק את אמיתות הערכה של הקפיטריה.

א. רשמו את ההשערות.

ב. בדקו את ההשערות ברמת מובהקות של 10%.

ג. מה תהיה המסקנה אם נקטין את רמת המובהקות?

5) חבר כנסת רוצה להעביר חוק. לצורך כך הוא דוגם 400 אזרחים במטרה לבדוק האם רוב האזרחים תומכים בחוק. במדגם התקבל ש-276 אזרחים תומכים בחוק.

א. מה מסקנתכם ברמת מובהקות של 5%?

ב. האם ניתן לדעת מה תהיה המסקנה אם רמת המובהקות תהיה גדולה יותר? הסבירו.

6) שני חוקרים בדקו את ההשערות הבאות: $H_0: p = p_0$, $H_1: p > p_0$.
חוקר א' השתמש ברמת מובהקות α_1 וחוקר ב' ברמת מובהקות α_2 החוקר הראשון דחה את H_0 ואילו החוקר השני קיבל את H_0.
שניהם התבססו על אותם תוצאות של מדגם. בחר בתשובה הנכונה:

א. $\alpha_1 = \alpha_2$.

ב. $\alpha_1 > \alpha_2$.

ג. $\alpha_1 < \alpha_2$.

ד. המצב המתואר לא אפשרי.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) נדחה H_0.

2) נדחה H_0.

3) נקבל H_0.

4) א. $\begin{matrix} H_0 : p = 0.2 \\ H_1 : p \neq 0.2 \end{matrix}$. ב. נקבל H_0. ג. המסקנה לא תשתנה.

5) א. נדחה H_0. ב. המסקנה לא תשתנה.

6) ג'.

© כל הזכויות שמורות לברק קנדל

סיכוי לטעויות ועוצמה:

רקע:

הגדרת הסתברויות:

הסיכוי לבצע טעות מסוג 1 (רמת מובהקות):

(לדחות H_0) $= P_{H_0}$ (H_0 נכונה $\mid$ לדחות את H_0) $= P(H_0$ לדחות את $\mid H_0$ נכונה$) = \alpha$.

הסיכוי לבצע טעות מסוג 2:

(לקבל H_0) $= P_{H_1}$ (H_1 נכונה $\mid$ לקבל את H_0) $= P(H_0$ לקבל את $\mid H_0$ נכונה$) = \beta$.

רמת בטחון:

(לקבל H_0) $= P_{H_0}$ (H_0 נכונה $\mid$ לקבל את H_0) $= P(H_0$ לקבל את $\mid H_0$ נכונה$) = (1-\alpha)$.

עוצמה:

(לדחות H_0) $= P_{H_1}$ (H_1 נכונה $\mid$ לדחות את H_0) $= P(H_0$ לדחות את $\mid H_0$ נכונה$) = \pi = (1-\beta)$.

	הכרעה		
		H_0	H_1
מציאות	H_0	אין טעות	טעות מסוג 1
	H_1	טעות מסוג 2	אין טעות

התהליך לחישוב סיכוי לטעות מסוג שני:

השערת האפס: השערה אלטרנטיבית:	$H_0: p = p_0$ $H_1: p \neq p_0$	$H_0: p = p_0$ $H_1: p < p_0$	$H_0: p = p_0$ $H_1: p > p_0$
תנאים:	$np_0 \geq 5 \;\&\; n(1-p_0) \geq 5$		
כלל ההכרעה: אזור הדחייה של H_0:	$\hat{p} > p_0 + Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$ או $\hat{p} < p_0 - Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	$\hat{p} < p_0 - Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$	$\hat{p} > p_0 + Z_{1-\alpha} \cdot \sqrt{\frac{p_0(1-p_0)}{n}}$

© כל הזכויות שמורות לברק קנדל

חישוב β :
$P_{H_1}\left(\hat{p} < p_0 + Z_{1-\alpha}\cdot\sqrt{\frac{p_0(1-p_0)}{n}}\right)$
$P_{H_1}\left(p_0 - Z_{1-\frac{\alpha}{2}}\cdot\sqrt{\frac{p_0(1-p_0)}{n}} < \hat{p} < p_0 + Z_{1-\frac{\alpha}{2}}\cdot\sqrt{\frac{p_0(1-p_0)}{n}}\right)$
$P_{H_1}\left(\hat{p} > p_0 - Z_{1-\alpha}\cdot\sqrt{\frac{p_0(1-p_0)}{n}}\right)$

כאשר: $\hat{P} \sim N\left(p\ , \frac{p(1-p)}{n}\right)$.

והתקנון: $Z_{\hat{p}} = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$.

דוגמה (פתרון בהקלטה):

רופאי שיניים טוענים שיותר ממחצית האוכלוסייה הבוגרת בארץ אינם מבקרים אצל רופא שיניים באופן קבוע, כנדרש. כדי לבדוק טענה זו, נערך סקר בקרב 150 אנשים בוגרים.

א. רשמו את ההשערות וכלל הכרעה ברמת מובהקות של 10%.

ב. מהי עוצמת המבחן אם מסתבר ש 60% מהאוכלוסייה אינם מבקרים אצל רופא שיניים באופן קבוע.

© כל הזכויות שמורות לברק קנדל

שאלות:

1) משרד הבריאות פרסם ש-10% מתושבי המדינה סובלים ממחלת האסטמה. מחקר דורש לבדוק האם בחיפה, בגלל זיהום האוויר, שיעור הסובלים מאסטמה גבוה יותר. לצורך המחקר נבדקו 260 מתושבי חיפה.

א. רשמו את השערות המחקר, וצרו מבחן ברמת מובהקות של 5% לבדיקתן.

ב. מהי עצמת המבחן של סעיף א' בהנחה ובחיפה 16% מהתושבים סובלים מאסטמה?

ג. כיצד תשנה התשובה לסעיף ב' אם מסתבר שבחיפה 18% סובלים מאסטמה?

ד. בהמשך לסעיף א' האם נכון לומר שבהסתברות של 5% ההשערה שבחיפה 10% מהתושבים סובלים מאסטמה אינה נכונה?

2) 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.
נניח שהתרופה נבדקת אכן מורידה את פרופורציות הסובלים מתופעות הלוואי ל-10%, מהי עצמת המבחן עבור רמת מובהקות של 5%?

3) 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%.

א. חשבו את הסיכוי לבצע טעות מסוג שני בהנחה והיום יש 28% אקדמאים.

ב. כיצד התשובה לסעיף הקודם תשתנה אם נגדיל את רמת המובהקות?

4) מעוניינים לבדוק האם בפקולטה מסוימת ישנה העדפה לגברים. הוחלט לדגום 200 מתקבלים ועל סמך מספר הבנים לקבוע אם טענת המחקר מתקבלת. חוקר א' קבע רמת מובהקות של 5% וחוקר ב' החליט לקבל את טענת המחקר אם במדגם יהיו לפחות 120 בנים. למי מבין החוקרים רמת מובהקות גדולה יותר?

5) חוקר ביצע מחקר ובו עשה טעות מסוג שני לכן (בחרו בתשובה הנכונה):

א. השערת האפס נכונה.

ב. השערת האפס נדחתה.

ג. השערת האפס לא נדחתה.

ד. אף אחת מהתושבות לא נכונה בהכרח.

6) קבעו אם הטענה הבאה נכונה: בבדיקת השערות לא ניתן לבצע בו זמנית טעות מסוג ראשון וטעות מסוג שני.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. $\begin{matrix} H_0 : p = 0.1 \\ H_1 : p > 0.1 \end{matrix}$. ב. 0.9015 ג. תגדל. ד. טענה לא נכונה.

2) 0.4404.

3) א. 0.1446 ב. תקטן.

4) חוקר א'.

5) ג'.

6) נכונה.

© כל הזכויות שמורות לברק קנדל

מובהקות התוצאה – אלפא מינימלית:

רקע:

דרך נוספת להגיע להכרעות שלא דרך כלל הכרעה, היא דרך חישוב מובהקות התוצאה:
באמצעות תוצאות המדגם מחשבים את מובהקות התוצאה שמסומן ב- p_v.
את רמת המובהקות החוקר קובע מראש לעומת זאת, את מובהקות התוצאה החוקר יוכל לחשב רק אחרי שיהיו לו את התוצאות.
המסקנה של המחקר תקבע לפי העיקרון הבא:

אם $p_v \le \alpha$ דוחים את H_0.

מובהקות התוצאה זה הסיכוי לקבלת תוצאות המדגם וקיצוני מתוצאות אלה בהנחת השערת האפס.
לקבל את תוצאות המדגם וקיצוני) $p_v = P_{H_0}$.

אם ההשערה היא דו צדדית:
לקבל את תוצאות המדגם וקיצוני) $p_v = 2P_{H_0}$.

מובהקות התוצאה היא גם האלפא המינימלית לדחיית השערת האפס.

השערת האפס: השערה אלטרנטיבית:	$H_0 : p = p_0$ $H_1 : p \neq p_0$	$H_0 : p = p_0$ $H_1 : p < p_0$	$H_0 : p = p_0$ $H_1 : p > p_0$
תנאים:	$np_0 \ge 5 \,\&\, n(1-p_0) \ge 5$		
p-value	אם $\hat{p} > p_0 \Rightarrow 2 \cdot P_{H_0}(\hat{P} \ge \hat{p})$ אם $\hat{p} < p_0 \Rightarrow 2 \cdot P_{H_0}(\hat{P} \le \hat{p})$	$P_{H_0}(\hat{P} \le \hat{p})$	$P_{H_0}(\hat{P} \ge \hat{p})$

כאשר בהנחת השערת האפס: $\hat{P} \sim N\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.

התקנון: $Z_{\hat{p}} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$.

© כל הזכויות שמורות לברק קנדל

דוגמה (פתרון בהקלטה):

ישנה טענה שיש הבדל בין אחוז הבנים ואחוז הבנות הפונים ללמוד להנדסאי מחשבים. לשם כך נלקח מדגם מקרי של 200 תלמידים הלומדים מחשבים והתברר כי 112 מהם בנים.

א. מהי מובהקות התוצאה?

ב. מה המסקנה ברמת מובהקות של 5%?

© כל הזכויות שמורות לברק קנדל

שאלות:

1) 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%?

2) נהוג לחשוב ש-60% מהילדים בגיל שלוש קמים מהמיטה במהלך הלילה לפחות פעם אחת. ישנה טענה שללא שנת צהריים פחות מ-60% מהילדים בגיל זה יקומו לפחות פעם אחת במהלך הלילה. נדגמו 80 ילדים בגיל 3 אשר אינם ישנים בצהריים מתוכם התקבל ש-41 קמו במהלך הלילה.

א. מהי רמת המובהקות המינימלית עבורה תתקבל הטענה במחקר?

ב. מהי רמת המובהקות המקסימלית עבורה לא תתקבל טענת המחקר?

ג. עבור אילו רמות מובהקות נקבל את טענת המחקר?

ד. מה תהיה מסקנת המחקר ברמת מובהקות של 6%?

3) במטרה לבדוק האם מטבע הוא הוגן מטילים אותו 80 פעמים. התקבל ש-60 מההטלות הראו עץ. רשמו את השערות המחקר, חשבו את מובהקות התוצאה והסיקו מסקנה ברמת מובהקות של 5%.

4) בבדיקת השערות על פרופורציה התקבל שה- $p\text{-}value = 0.02$.
מה תהיה מסקנת חוקר המשתמש ברמת מובהקות 5% :
(בחרו בתשובה הנכונה)

א. יקבל את השערת האפס

ב. ידחה את השערת האפס.

ג. לא ניתן לדעת כי אין מספיק נתונים.

5) קבעו אם הטענה הבאה נכונה :
"במבחן לבדיקת השערות חד-צדדי התקבל ערך $p-value$ של 3%,
לכן אם היינו מבצעים מבחן דו-צדדי (כאשר יתר הנתונים ללא שינוי),
היינו מקבלים ערך $p-value$ של 6%".

6) במפעל 10% מהעובדים נפגעים לפחות פעם אחת בשנה מתאונות עבודה.
לאור זאת, המפעל החליט לצאת בתוכנית לצמצום שיעור הנפגעים. תכנית זו נוסתה על 100 עובדים. מתוכם 12 נפגעו בתאונות עבודה במשך השנה.
מהי רמת המובהקות הקטנה ביותר עבורה יוחלט שהתכנית יעילה?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. 0.0455.

ב. ברמת מובהקות של 1% : לא דוחים את H_0.

ברמת מובהקות של 5% : נדחה את H_0.

2) א. 0.0548. ב.0.0548. ג. מעל 0.0548.

ד. נכריע לטובת טענת המחקר.

3) $p_v = 0$, נדחה את H_0.

4) ב'.

5) הטענה נכונה.

6) 0.7486.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 44 - בדיקת השערות על הפרש פרופורציות

תוכן העניינים

1. כללי 206

בדיקת השערות על הפרש פרופורציות

רקע

השערת האפס : **השערה אלטרנטיבית:**	$H_0: p_1 - p_2 = 0$ $H_1: p_1 - p_2 \neq 0$	$H_0: p_1 - p_2 = 0$ $H_1: p_1 - p_2 < 0$	$H_0: p_1 - p_2 = 0$ $H_1: p_1 - p_2 > 0$
תנאים:	1.מדגמים בלתי תלויים	2. מדגמים גדולים	
כלל ההכרעה: אזור הדחייה של :	$Z_{\hat{p}_1-\hat{p}_2} < -Z_{1-\frac{\alpha}{2}}$ או $Z_{\hat{p}_1-\hat{p}_2} > Z_{1-\frac{\alpha}{2}}$ $-Z_{1-\frac{\alpha}{2}}$ $Z_{1-\frac{\alpha}{2}}$ H_0 דוחים את - ■	$Z_{\hat{p}_1-\hat{p}_2} < -Z_{1-\alpha}$ $-Z_{1-\alpha}$ H_0 דוחים את - ■	$Z_{\hat{p}_1-\hat{p}_2} > Z_{1-\alpha}$ $Z_{1-\alpha}$ H_0 דוחים את - ■

סטטיסטי המבחן: $$Z_{\hat{p}_1-\hat{p}_2} \underset{H_0}{=} \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$$

כאשר הפרופורציה המשוקללת : $$\hat{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

© כל הזכויות שמורות לברק קנדל

חלופה אחרת לכלל הכרעה:

כלל ההכרעה: אזור הדחייה של H_0	
$\hat{p}_1-\hat{p}_2>0+Z_{1-\alpha/2}\cdot\sqrt{\frac{\hat{p}\ \hat{q}}{n_1}+\frac{\hat{p}\ \hat{q}}{n_2}}$ או $\hat{p}_1-\hat{p}_2<0-Z_{1-\alpha/2}\cdot\sqrt{\frac{\hat{p}\hat{q}}{n_1}+\frac{\hat{p}\ \hat{q}}{n_2}}$	$\hat{p}_1-\hat{p}_2<0-Z_{1-\alpha}\cdot\sqrt{\frac{\hat{p}\hat{q}}{n_1}+\frac{\hat{p}\ \hat{q}}{n_2}}$
$\hat{p}_1-\hat{p}_2>0+Z_{1-\alpha}\cdot\sqrt{\frac{\hat{p}\hat{q}}{n_1}+\frac{\hat{p}\ \hat{q}}{n_2}}$	

התפלגות של $\hat{p}_1-\hat{p}_2$:

$$\hat{p}_1-\hat{p}_2\sim N(p_1-p_2,\frac{p_1\cdot q_1}{n_1}+\frac{p_2\cdot q_2}{n_2})$$

תקנון:

$$Z_{\hat{p}_1-\hat{p}_2}=\frac{\hat{p}_1-\hat{p}_2-(p_1-p_2)}{\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1}+\frac{\hat{p}_2\hat{q}_2}{n_2}}}$$

$$Z_{\hat{p}_1-\hat{p}_2}\underset{H_0}{=}\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\frac{\hat{p}\hat{q}}{n_1}+\frac{\hat{p}\hat{q}}{n_2}}}$$

דוגמה (פתרון בהקלטה) :

נדגמו 80 סטודנטים שנבחנו במיקרו-כלכלה. מתוכם 60 עברו את הבחינה.
נדגמו 100 סטודנטים שנבחנו בסטטיסטיקה א'. מתוכם 82 עברו את הבחינה.
האם שיעור העוברים את הבחינה בסטטיסטיקה גבוה מאשר מהבחינה במיקרו כלכלה? בדקו ברמת מבוהקות של 10%.

© כל הזכויות שמורות לברק קנדל

שאלות

1) במדגם של 200 גברים, 8% היו מובטלים. במדגם של 180 נשים, 10% מהן היו מובטלות.
האם קיים הבדל מובהק בין פרופורציית המובטלים לפרופורציית המובטלות?
בדקו 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?

3) נדגמו 500 בוגרים מתוכם 200 גברים והיתר נשים. במדגם התקבל: מתוך הגברים ל-48% תעודת בגרות. מתוך הנשים ל-58% תעודת בגרות. מטרת המחקר היא לבדוק האם שיעור הזכאיות לבגרות גבוה משיעור הזכאים.

א. מהי מובהקות התוצאה?

ב. מה תהיה המסקנה ברמת מובהקות של 8%?

4) במדגם שנערך על 100 פרות מחוות בדרום הארץ התקבל כי 20 פרות נושאות וירוס מסוים. במדגם שנערך על 200 פרות מחוות בצפון הארץ התקבל כי 10 מתוכן נושאות וירוס גם כן.

א. בנו מבחן ברמת מובהקות של 5% לבדיקת הטענה כי הוויורוס תקף את פרות הדרום באופן משמעותי יותר מאשר את הפרות בצפון הארץ.

ב. מהי המסקנה לבדיקת הטענה של סעיף א ומהי הטעות האפשרית במסקנה?

ג. מהי עוצמת המבחן אם שיעור הפרות בדרום עם הוויורוס גבוה ב-10% משיעור הפרות בצפון עם הוויורוס?

ד. כיצד העוצמה תשתנה אם נגדיל את רמת המובהקות?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות

1) לא נדחה את H_0.

2) א. נדחה H_0. ב. נדחה H_0.

3) א. 0.0139 ב. נדחה H_0.

4) א. ראה סרטון. ב. נדחה H_0. ג. 0.8238

ד. תגדל.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 45 - שאלות מסכמות על רווחי סמך

תוכן העניינים

1. שאלות מסכמות על רווחי סמך 210

שאלות מסכמות על רווחי סמך:

שאלות:

1) מהירות הגלישה באינטרנט במקום מסוים מתפלגת נורמאלית. בדקו את מהירות הגלישה ב-30 זמנים אקראיים. מהירות הגלישה נמדדה ב-Mbps.
מהירות מתחת ל-10Mbps מוגדרת על ידי החברה כנמוכה.
התוצאות שהתקבלו במדגם: ממוצע היה 87 עם סטיית תקן 17 ו-12 פעמים המהירות הייתה נמוכה.
בנו רווחי סמך ברמת סמך של 95% לפרמטרים הבאים:

א. תוחלת מהירות הגלישה.

ב. סטיית תקן של מהירות הגלישה.

ג. הסיכוי שמהירות הגלישה תהיה נמוכה.

2) 200 אנשים נשאלו כמה פעמים ביום הם שותים כוס קפה.
להלן התפלגות התשובות:

מספר פעמים	0	1	2	3	4	5
מספר אנשים	86	34	28	22	20	10

א. תנו רווח סמך לממוצע מספר כוסות הקפה שאנשים נוהגים לשתות ביום. $\alpha = 0.05$.

ב. אדם השותה לפחות 4 כוסות קפה ביום נקרא "מכור לקפה". בנו רווח סמך לאחוז "המכורים לקפה". $\alpha = 0.1$.

3) חוקר בנה רווח סמך לאחוז האנשים שהתקררו לפחות פעם אחת בשנה. רווח הסמך שהתקבל הוא: $81 < p < 91$. רווח הסמך הנ"ל התבסס על מדגם של 500 איש.

א. כמה אנשים במדגם טענו שכלל לא התקררו השנה?

ב. באיזו רמת סמך נבנה רווח הסמך?

ג. בנו רווח סמך לאחוז האנשים שהתקררו לפחות פעם אחת השנה ברמת סמך של 96% על סמך תוצאות המדגם.

© כל הזכויות שמורות לברק קנדל

4) ציוני IQ בארה"ב מתפלגים נורמאלית עם תוחלת 100. במדגם של 20 ישראלים שנבחנו במבחן ה-IQ התקבלו התוצאות הבאות:

$$\sum_{i=1}^{20} x_i = 2040 \text{ , } \sum_{i=1}^{20} x_i^2 = 210740 .$$

א. אמדו ברמת ביטחון של 90% את ממוצע ציוני בחינת ה-IQ בישראל – מהי ההנחה הדרושה לפתרון?

ב. על סמך רווח הסמך של סעיף א' האם תקבלו את הטענה שבישראל ממוצע הציונים שונה מארה"ב?

ג. מה היה קורה לרווח הסמך אם היינו מגדילים את רמת הסמך שלו?

5) להלן תוצאות מדגם שבדק עבור כל משפחה האם יש לה בבית מכשיר טאבלט:

אזור מגורים	גוש דן	שאר הארץ
גודל המדגם	200	240
מספר משפחות בעלי טאבלט	160	168

א. בנו רווח סמך להבדל בין אחוז המשפחות עם טאבלט בגוש דן ואחוז המשפחות בעלי טאבלט בשאר חלקי הארץ. ברמת סמך של 98%.

ב. בנו רווח סמך לפרופורציות משפחות בעלות טאבלט בכלל הארץ ברמת סמך של 95%.

6) הגובה של מתגייסים לצה"ל מתפלג נורמלית במדגם של 25 מתגייסים.
התקבלו התוצאות הבאות: $\bar{x} = 176.2\text{cm}$, $\sum (x_i - \bar{x})^2 = 2832\text{cm}^2$.

א. אמדו את הגובה הממוצע של המתגייסים ברמת סמך של 98%.

ב. אמדו ברמת סמך של 90% את סטיית התקן של הגובה של מתגייסים של צה"ל.

7) בנק מתלבט האם לפתוח סניף באזור A או באזור B. לצורך פתרון נניח שסטיית 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%.

© כל הזכויות שמורות לברק קנדל

8) להלן מדגם של שכר הדירה ב-₪ של 5 דירות שלושה חדרים בשכונת בבלי בתל אביב:

שנת 2012	8000	7500	7000	6500	7500
שנת 2013	8000	8200	7800	6800	7700

בנו רווח סמך ברמת סמך של 95% לתוחלת עליית שכר הדירה משנת 2012 לשנת 2013 בשכונת בבלי. ניתן להניח ששכר הדירה בשכונה מתפלג נורמלית.

תשובות סופיות:

1) א. $80.65 \le \mu \le 93.35$. ב. $13.5 < \sigma < 22.9$. ג. $0.225 \le p \le 0.575$.

2) א. $1.21 \le \mu \le 1.65$. ב. $10.85\% \le p \le 19.15\%$.

3) א. 70. ב. 0.9988. ג. $83\% < p < 89\%$.

4) א. $97.4 \le \mu \le 106.6$. ב. לא. ג. יגדל.

5) א. $0.5\% \le p_1 - p_2 \le 19.5\%$. ב. $0.704 \le p \le 0.786$.

6) א. $170.8 \le \mu \le 181.6$. ב. $8.8 \le \sigma \le 14.3$.

7) א. $-372 \le \mu_A - \mu_B \le 772$, לא. ב. $6467 \le \mu_A \le 7133$.

8) $-21 \le \mu_D \le 821$.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

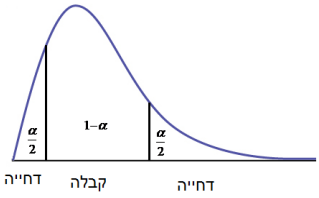
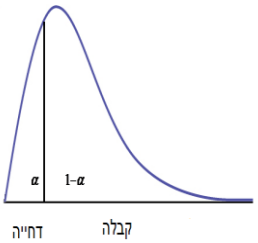
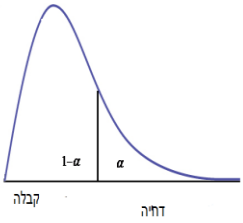
פרק 46 - בדיקת השערות על שונויות

תוכן העניינים

1. שונות וסטיית תקן 213
2. שתי שונויות 218

שונות וסטיית תקן:

רקע:

השערת האפס : השערה אלטרנטיבית:	$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$	$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$	$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$
תנאים :	$X \sim N$		
נדחה את השערת האפס אם:	 $\chi^2 < \chi^{2(n-1)}_{1-\frac{\alpha}{2}}$ או $\chi^2 > \chi^{2(n-1)}_{\frac{\alpha}{2}}$	 $\chi^2 < \chi^{2(n-1)}_{1-\alpha}$	 $\chi^2 > \chi^{2(n-1)}_{\alpha}$

סטטיסטי המבחן: $\chi^2 = \frac{(n-1)\hat{S}^2}{\sigma_0^2}$

התפלגות חי בריבוע:

אם $X_i \sim N(\mu,\sigma^2)$ והפרמטר μ אינו ידוע מתקיים ש: $\frac{(n-1)\hat{S}^2}{\sigma_0^2} \sim \chi^{2(n-1)}$

התפלגות זו היא התפלגות אסימטרית חיובית המתחילה מהערך אפס וערכיה שואפים לאינסוף. התפלגות זו תלויה בדרגות החופש.

אם μ אינו ידוע אז: $d.f = n-1$

© כל הזכויות שמורות לברק קנדל

דוגמה:

ציוני IQ לפי סטנדרטים אמריקאים מתפלגים נורמאלית עם $\sigma=15$.
מעוניינים לבדוק האם שונות הציונים של נבחנים ישראלים שונה מאמריקה.
במדגם של 20 ישראלים התקבל: $\sum_{i=1}^{20}(x_i-\bar{x})^2=3420$. מה המסקנה ברמת מובהקות של 5%?

פתרון:

<u>האוכלוסייה</u>: נבחנים ישראלים במבחן I.Q

<u>המשתנה</u>: X = ציון I.Q

<u>פרמטר</u>: σ^2

<u>השערות</u>:
$$H_0:\sigma^2=15^2=225$$
$$H_1:\sigma^2\neq 225$$

<u>הנחה</u>: $X\sim N$

<u>כלל הכרעה</u>: $d.f=n-1=20-1=19$

נדחה את H_0 אם $X^2>32.9$ או $X^2<8.91$

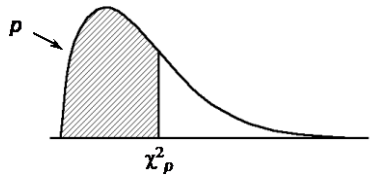
<u>תוצאות המדגם</u>: $n=20$

$$\sum(X_i-\bar{X})^2=3420$$
$$S^2=\frac{\sum(X_i-\bar{X})^2}{n-1}=\frac{3420}{20-1}=180$$
$$\chi^2=\frac{(n-1)S^2}{\sigma_0^2}=\frac{19\cdot 180}{225}=15.2$$

<u>מסקנה</u>: לא נדחה את H_0, לא נסיק ששונות הציונים של נבחנים ישראלים במבחן I.Q שונה מזו של אמריקאים. ($\alpha=5\%$)

© כל הזכויות שמורות לברק קנדל

טבלת התפלגות חי-בריבוע – ערכי החלוקה $\chi^2 p$



df	p												
	.005	.01	.025	.05	.10	.25	.50	.75	.90	.95	.975	.99	.995
1	0.0^4393	0.0^3157	0.0^3982	0.0^2393	0.0158	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	0.0717	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	37.2
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	38.6
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	40.0
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	41.4
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	42.8
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	44.2
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	45.6
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	46.9
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	48.3
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	49.6
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	51.0
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	52.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	53.7

© כל הזכויות שמורות לברק קנדל

שאלות:

1) חברה אורזת סוכר במשקל עם סטיית תקן 20 גרם. משקל הסוכר באריזה מתפלג נורמאלית. החברה החליפה את מכונות האריזה במטרה לדייק יותר במשקל הנארז. (רוצים שסטיית התקן תהיה קטנה יותר).
לצורך בדיקה דגמו 5 אריזות סוכר ולהלן משקלן (בגרם):
997, 996, 1005, 1024, 1008.
מה המסקנה ברמת מובהקות של 5%?

2) 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, 50, 72, 90, 110.

א. ברמת מובהקות של 5% בדקו האם סטיית התקן של זמן החלמה של התרופה החדשה נמוכה מהתרופה המקורית?

ב. האם ניתן לדעת מה תהיה התשובה לסעיף א', אם נגדיל את רמת המובהקות?

ג. האם ניתן לדעת מה תהיה התשובה לסעיף א אם נקטין את רמת המובהקות?

ד. האם ניתן לדעת מה תהיה התשובה לסעיף א אם נוסיף תצפית שערכה 70?

3) הגובה של אוכלוסייה מסוימת נחשב כמתפלג נורמלית עם 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, בבחירת המבחן המתאים הסתמך על המסקנה מסעיף א'.

4) השערות המחקר הן: $H_0:\sigma=2$, $H_1:\sigma<2$
במדגם של 21 תצפיות התקבלה סטיית תקן 1.143.
תן הערכה למובהקות התוצאה.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) לא נדחה H_o.

2) א. נדחה את H_o. ב. לא תשתנה. ג. לא ניתן לדעת.
ד. לא תשתנה.

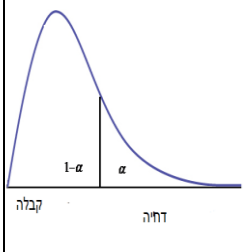
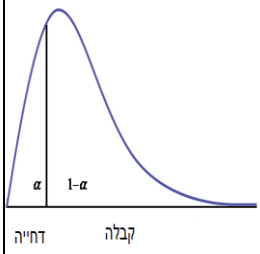
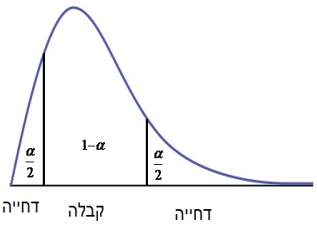
3) א. נדחה את H_o. ב. לא נדחה את H_o.

4) $0 \le P_V \le 0.005$

© כל הזכויות שמורות לברק קנדל

בדיקת השערות על שתי שונויות:

רקע:

השערת האפס: השערה אלטרנטיבית:	$H_0:\frac{\sigma_1^2}{\sigma_2^2}=1$ $H_1:\frac{\sigma_1^2}{\sigma_2^2}\neq 1$	$H_0:\frac{\sigma_1^2}{\sigma_2^2}=1$ $H_1:\frac{\sigma_1^2}{\sigma_2^2}<1$	$H_0:\frac{\sigma_1^2}{\sigma_2^2}=1$ $H_1:\frac{\sigma_1^2}{\sigma_2^2}>1$
תנאים:	1.מדגמים בלתי תלויים 2. $X_1, X_2 \sim N$		
נדחה את השערת האפס אם:	 או $F \geq f_{1-\alpha/2}^{(n_1-1,n_2-1)}$ $F \leq \frac{1}{f_{1-\alpha/2}^{(n_2-1,n_1-1)}}$	 $F \leq \frac{1}{f_{1-\alpha}^{(n_2-1,n_1-1)}}$	 $F \geq f_{1-\alpha}^{(n_1-1,n_2-1)}$

סטטיסטי המבחן: $F=\frac{S_1^2}{S_2^2}$

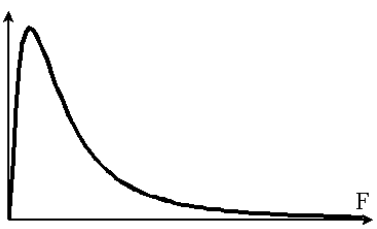
התפלגות F:

אם $X_1 \sim N(\mu_1,\sigma^2)$ ו- $X_2 \sim N(\mu_2,\sigma^2)$ אזי: $\frac{S_1^2}{S_2^2} \sim F(n_1-1,n_2-1)$.

התפלגות F הינה התפלגות אסימטרית חיובית התלויה בדרגות חופש של המונה ושל המכנה.

כמו כן בהתפלגות F מתקיימת התכונה הבאה: $F_\alpha(n_1-1,n_2-1)=\frac{1}{F_{1-\alpha}^{(n_2-1,n_1-1)}}$.

$df_1 = n_1 - 1$
$df_2 = n_2 - 1$



© כל הזכויות שמורות לברק קנדל

דוגמה:

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?

פתרון:

<u>האוכלוסיות:</u> נשים מול גברים.

<u>משתנה:</u> y = זמן ביצוע משימה של גבר , x = זמן ביצוע משימה של אישה

<u>פרמטר:</u> $\frac{\sigma^2_x}{\sigma^2_y}$

<u>השערות:</u>

$$H_0:\frac{\sigma^2_x}{\sigma^2_y}=1$$

$$H_1:\frac{\sigma^2_x}{\sigma^2_y}\neq 1$$

<u>הנחות:</u> 1. מדגימים ב"ת 2. $x,y\sim N$

<u>כלל הכרעה:</u>

$\alpha=2\%$

$n_1=13, d.f_1=n_1-1=12$

$n_2=10, d.f_2=n_2-1=9$

<u>נדחה את H_0</u> אם $F>5.11$ או $F<0.23$

<u>תוצאות המדגם:</u>

$$S^2_y=\frac{204}{10-1}=22\frac{2}{3}$$

$$S^2_x=\frac{200}{13-1}=16\frac{2}{3}$$

<u>סטטיסטי המבחן:</u>

$$F=\frac{S_1^2}{S_2^2}=\frac{16\frac{2}{3}}{22\frac{2}{3}}$$

<u>מסקנה:</u> ברמת מובהקות של 2% נקבל את H_0.

לא קיים הבדל מובהק בין גברים לנשים מבחינת השונויות שלהם.

© כל הזכויות שמורות לברק קנדל

טבלת ערכים קריטיים לפי התפלגות F

$\alpha = 0.05$

ד"ח מונה \ ד"ח מכנה	1	2	3	4	5	6	7	8	9	10	12	16	20	24	60	120	∞
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	246.46	248.01	249.05	252.20	253.25	254.31
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.69	8.66	8.64	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.84	5.80	5.77	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.60	4.56	4.53	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.92	3.87	3.84	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.49	3.44	3.41	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.20	3.15	3.12	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	2.99	2.94	2.90	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.83	2.77	2.74	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.70	2.65	2.61	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.60	2.54	2.51	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.51	2.46	2.42	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.44	2.39	2.35	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.38	2.33	2.29	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.33	2.28	2.24	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.29	2.23	2.19	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.25	2.19	2.15	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.21	2.16	2.11	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.18	2.12	2.08	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.16	2.10	2.05	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.13	2.07	2.03	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.11	2.05	2.01	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.09	2.03	1.98	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.07	2.01	1.96	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.05	1.99	1.95	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.04	1.97	1.93	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.02	1.96	1.91	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.01	1.94	1.90	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	1.99	1.93	1.89	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.90	1.84	1.79	1.64	1.58	1.51
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.95	1.85	1.78	1.74	1.58	1.51	1.44
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.82	1.75	1.70	1.53	1.47	1.39
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94	1.86	1.76	1.69	1.64	1.46	1.39	1.30
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.73	1.66	1.61	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.64	1.57	1.52	1.32	1.22	1.00

© כל הזכויות שמורות לברק קנדל

טבלת ערכים קריטיים לפי התפלגות F $\alpha = 0.01$ ראה איור מטה.

ד"ח מונה/ד"ח מכנה	1	2	3	4	5	6	7	8	9	10	12	16	20	24	60	120	∞
1	4052.18	4999.50	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6022.47	6055.85	6106.32	6170.10	6208.73	6234.63	6313.03	6339.39	6365.86
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.44	99.45	99.46	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.83	26.69	26.60	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.15	14.02	13.93	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.68	9.55	9.47	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.52	7.40	7.31	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.28	6.16	6.07	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.48	5.36	5.28	5.03	4.95	4.86
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.92	4.81	4.73	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.52	4.41	4.33	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.21	4.10	4.02	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	3.97	3.86	3.78	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.78	3.66	3.59	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.62	3.51	3.43	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.49	3.37	3.29	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.37	3.26	3.18	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.27	3.16	3.08	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.19	3.08	3.00	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.12	3.00	2.92	2.67	2.58	2.49
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.05	2.94	2.86	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	2.99	2.88	2.80	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.94	2.83	2.75	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.89	2.78	2.70	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.85	2.74	2.66	2.40	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.81	2.70	2.62	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.78	2.66	2.58	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.75	2.63	2.55	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.72	2.60	2.52	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.69	2.57	2.49	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.66	2.55	2.47	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.48	2.37	2.29	2.02	1.92	1.80
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70	2.56	2.38	2.27	2.18	1.91	1.80	1.68
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.31	2.20	2.12	1.84	1.73	1.60
90	6.93	4.85	4.01	3.53	3.23	3.01	2.84	2.72	2.61	2.52	2.39	2.21	2.09	2.00	1.72	1.60	1.46
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.15	2.03	1.95	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.00	1.88	1.79	1.47	1.32	1.00

© כל הזכויות שמורות לברק קנדל

שאלות:

1) להלן נתונים על שטחי דירות במ״ר עבור דירות חדשות שנבנו בשנת 2012 ובשנת 2013:

2012	120	100	82	112	95	130	90	94	120
2013		100	82	112	91	105	74	69	

א. בדקו ברמת מובהקות של 10% את ההשערה ששונויות שטחי הדירות החדשות בשנת 2012 ובשנת 2013 שוות. מה הן ההנחות הדרושות לביצוע הבדיקה?

ב. האם וכיצד הייתה משתנה המסקנה מהסעיף הקודם אם מסתבר שחלה טעות ברישום ויש להפחית 10 מ״ר מכל הדירות שמופיעות במדגם?

2) בתחום הבינוי משתמשים בשני סוגי מתכות: מתכת A ומתכת B. מחקר מעוניין לבדוק האם קיים הבדל בין שני סוגי המתכות מבחינת החוזק שלהן. דגמו מספר יחידות מתכת מכל סוג והתקבלו התוצאות הבאות:
יש להניח שרמת החוזק של המתכות מתפלגת נורמאלית.

סוג המתכת	A	B
n	8	10
$\sum X_i$	16	30
$\sum X_i^2$	60	198

א. האם קיים הבדל בין שונויות החוזק של מתכות?

ב. האם קיים הבדל בין תוחלות החוזק של מתכות? בכל סעיף רמת מובהקות של 10% .

3) מחקר סוציולוגי מעוניין לחקור את הרגלי הבילויים בקבוצות גיל שונות. ידוע כי בקרב האוכלוסייה הבוגרת (מעל 18) ההוצאה החודשית על בילויים מתפלגת נורמאלית עם תוחלת של 500₪ וסטיית תקן של 300₪.
במדגם שנעשה על סטודנטים בגילאי 21-26 התקבל אומד חוסר הטיה לשונות ההוצאה החודשית על בילויים 10,000₪. כמות הסטודנטים שנדגמה 16.
במדגם שנעשה על 11 מבוגרים בשנות השלושים התקבל אומד חסר הטיה לשונות ההוצאה החודשית על בילויים 490,000₪.

א. בדקו ברמת מובהקות של 5% האם שונות ההוצאה על בילויים בקרב סטודנטים בקבוצת גילאי 21-26 נמוכה מהשונות אצל כלל המבוגרים.

ב. בדקו ברמת מובהקות של 1% האם הפיזור של ההוצאה החודשית לבילויים גדולה יותר בקבוצת גיל ה-30 מאשר בקבוצת גיל 21-26.

תשובות סופיות:

1) א. לא נדחה את H_0. ב. מסקנה לא תשתנה.

2) א. לא נדחה את H_0. ב. לא נדחה את H_0.

3) א. נדחה את H_0. ב. נדחה את H_0.

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 47 - שאלות מסכמות בבדיקת השערות

תוכן העניינים

1. שאלות פתוחות מסכמות 223
2. שאלות רב ברירה (אמריקאיות) 227

שאלות מסכמות בבדיקת השערות על פרמטרים

שאלות

1) שני חוקרים נתבקשו לבדוק את ההשערות הבאות: $H_1:\mu>520$, $H_0:\mu=520$.
כל חוקר בדק מדגם של 225 נחקרים. ידוע ש- $\sigma=20$.
חוקר א' קבע את כלל ההכרעה לפי $\alpha=0.05$.
חוקר ב' מחליט לדחות H_0 אם $\bar{X}>522$.

א. למי מהחוקרים הסתברות לטעות מסוג ראשון קטנה יותר?

ב. מהי ההסתברות לטעות מסוג שני של חוקר ב' עבור $\mu_1=525$.

ג. הסבר ללא חישוב נוסף, האם ההסתברות לטעות מסוג שני עבור $\mu_1=525$, של חוקר א' שווה/קטנה/גדולה לזו של חוקר ב'.

ד. חוקר א' קיבל במדגם שלו $\bar{X}=523$. מהי מסקנתו?

2) ידוע כי תוחלת מספר הלייקים היומי של דנה היא 12 עם סטיית תקן 5.
דני טוען שהוא יותר פופולארי מדנה בכך שהוא מקבל יותר לייקים מדנה ביום. על-מנת לבדוק זאת ספר דני כמה לייקים הוא קיבל בכל יום במהלך 7 שבועות (כלומר, ב $-$ 49 ימים) וקיבל סך-הכול 637 לייקים. נניח כי סטיית התקן של מספר הלייקים שדני מקבל ביום זהה לסטיית התקן של דנה.

א. מהי רמת המובהקות שכדאי לדני לדרוש, כדי שדנה תשתכנע בצדקת טענתו (שדני פופולרי יותר בכך שהוא מקבל יותר לייקים מדנה ביום).

ב. אם דני משער שתוחלת מספר ה"לייק"ים שהוא מקבל ביום היא 14 וקובע רמת מובהקות 2.5%, מהי עוצמת המבחן של דני?

3) ברצוננו להשוות בין רשתות A לבין B. לשם כך בחרנו 4 מוצרים, ובדקנו את מחיריהם בשתי הרשתות. להלן התוצאות:

מוצר / רשת	A	B
1	5	5
2	4	5
3	5	3
4	7	4

הניחו כי המחירים מתפלגים נורמלית.
אם יש הנחות נוספות כדי לבצע את המבחן הפרמטרי רשמו אותן.

א. בדקו האם קיים הבדל בין הרשתות מבחינת תוחלת המחירים. רמת מובהקות של 5%.

ב. חזרו על הסעיף הקודם בהנחה ונבחרו בכל רשת מוצרים באקראי ולא בהכרח אותם מוצרים.

© כל הזכויות שמורות לברק קנדל

4) במדגם של 10 ישראלים שנבחנו במבחן ה-IQ נתקבלו התוצאות הבאות:

$$n=10,\quad \sum X_i=1020,\quad \sum X_i^2=105120.$$

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

$$n=14,\quad \sum X_i=1386,\quad \sum X_i^2=138644.$$

נתון שציוני הבחינה מתפלגים נורמלית בכל מדינה.

א. בדקו ברמת מובהקות של 10% האם קיים שוויון שונויות בין אוכלוסיית אמריקה לאוכלוסיית ישראל?

ב. בדקו האם קיים הבדל בממוצע הציונים בבחינת ה-IQ בין ישראל לארה"ב. ברמת מובהקות של 5%?

5) במטרה לבדוק האם סטודנטים הלומדים במכללות משקיעים יותר זמן ללימודים מאשר סטודנטים באוניברסיטאות נדגמו 12 סטודנטים ובדקו לכל סטודנט את הזמן שהוא משקיע ביום ללימודים. הזמנים נמדדו בדקות:

סטודנטים באוניברסיטאות	176	156	189	171	140	180
סטודנטים במכללות	180	190	191	186	204	150

א. נסחו את ההשערות ובדוק אותן ברמת מובהקות של 5%. רשום את כלל ההכרעה ואת ההנחות הדרושות לביצוע המבחן הפרמטרי.

ב. חשבו את p-value.

ג. ישנה טענה שממוצע זמן ההשקעה בלימודים במכללות הוא 3.5 שעות ביום. בדוק את הטענה כאשר רמת המובהקות הינה 5%.

6) במדינת טרפפו המשכורות במשק מתפלגות נורמלית עם ממוצע של 1 אלף דולר וסטיית תקן של 0.2 אלף דולר. בוצע מדגם מקרי בו השתתפו 5 נשים ו 5 גברים במדינת שומקום שבה המשכורות מתפלגות נורמאלית גם כן. להלן משכורותיהם באלפי דולר:

גברים	2	0.9	0.7	1.2	1.1
נשים	1.4	1.1	1.9	1.8	1.2

א. בדקו את הטענה שממוצע משכורותיהם של אזרחי שומקום גבוה מאשר ממוצע משכורותיהם של אזרחי טרפפו ברמת מובהקות של 5%. בהנחה שסטיית התקן זהה בשתי המדינות.

ב. חזרו על הסעיף הקודם ללא ההנחה הנ"ל.

ג. ישנה טענה שסטיית התקן במדינת שומקום גבוהה מזו של טרפפו. בדקו ברמת מובהקות של 5%.

© כל הזכויות שמורות לברק קנדל

7) במטרה להשוות בין אחוזי הצפייה של גברים ונשים בתוכנית טלוויזיה מסוימת בוצע סקר ובו התקבלו תוצאות הבאות:

	צופים	לא צופים
נשים	320	42
גברים	72	120

א. האם יש הבדל בין אחוזי הצפייה של גברים ונשים ברמת מובהקות של 1%?

ב. עבור רמת מובהקות של 5% בדוק טענה שמבין הצופים בתוכנית הטלוויזיה אחוז הנשים גדול פי 2 מאחוז הגברים.

8) בשנת 2000 ל-60% היה מדיח כלים בבית. מחקר רוצה לבדוק האם כיום פרופורציית המשפחות עם מדיח כלים עלה. הוחלט לבצע מדגם אקראי של 150 משפחות.

א. רשמו את השערות המחקר.

ב. מה היא מסקנת המחקר ברמת מובהקות של 5% אם במדגם ל-102 משפחות היה מדיח כלים.

ג. מהי הטעות האפשרית במסקנה מהסעיף הקודם.
האם ניתן לדעת את הסתברותה?

9) נערך מחקר על הקשר בין עישון ויתר לחץ דם. נבדק מדגם מקרי של 200 מעשנים ונמצא כי 30 סבלו מיתר לחץ דם.
ידוע שבאוכלוסייה 18% סובלים מיתר לחץ דם.

א. בדקו ברמת מובהקות 0.1 את ההשערה כי אחוז הסובלים מיתר לחץ דם בקרב המעשנים גדול מאשר כלל האוכלוסייה.

ב. מהי רמת המובהקות המינימלית לקבלת הטענה שאחוז הסובלים מיתר לחץ דם בקרב המעשנים גדול מאשר כלל האוכלוסייה.

ג. מהי עצמת המבחן, אם אחוז הסובלים מיתר לחץ דם בקרב אוכלוסיית המעשנים היא בפועל 25%.

10) להלן התפלגות מספר הנסיעות לחופשה השנתית במדגם של משפחות ישראליות. בדקו ברמת מובהקות של 5%:

מספר הנסיעות	0	1	2	3	4
מספר המשפחות	84	102	26	20	12

א. באיטליה משפחות נוסעות בממוצע פעמיים בשנה לחופשה. האם בישראל משפחות נוסעות פחות מאשר באיטליה?

ב. בהולנד 80% מהמשפחות נוסעות לפחות פעם אחת בשנה לחופשה, האם בישראל אחוז המשפחות שנוסעות לפחות פעם אחת בשנה לחופשה נמוך מאשר בהולנד?

© כל הזכויות שמורות לברק קנדל

11) נתון כי: $X \sim N(\mu, \sigma^2 = 10^2)$.
מעוניינים לבדוק את ההשערות: $H_1: \mu > 40, \quad H_0: \mu = 40$.
דגמו 25 תצפיות מהאוכלוסייה והתקבל $\bar{X} = 45$.
א. חשבו את p-value (מובהקות התוצאה).
ב. חזרו על סעיף א אם ההשערה האלטרנטיבית הייתה: $H_1: \mu < 40$
ג. חזרו על סעיף א אם ההשערה האלטרנטיבית הייתה: $H_1: \mu \neq 40$

12) ציוני בחינת הבגרות במתמטיקה מתפלגים נורמלית עם שונות 150. במדגם של 16 נבחנים מתל אביב התקבלה שונות מדגמית-190. במדגם של 25 ירושלמים התקבלה שונות מדגמית 118.
א. בדקו ברמת מובהקות של 2.5% 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.

תשובות סופיות

1) א. חוקר א' ב. 0.0122 ג. גדלה. ד. נדחה H_0.
2) א. לפחות 0.0808 ב. 0.7995.
3) א. לא נדחה H_0. ב. לא נדחה H_0.
4) א. לא נדחה H_0. ב. לא נדחה H_0.
5) א. לא נדחה H_0. ב. בין 5% ל-10%. ג. נדחה H_0.
6) א. נדחה H_0. ב. נדחה H_0. ג. נדחה H_0.
7) א. נדחה H_0. ב. נדחה H_0.
8) א. $H_0: p = 0.6$, $H_1: p > 0.6$ ב. נדחה H_0. ג. טעות מסוג ראשון בסיכוי של 0.05.
9) א. לא נדחה H_0. ב. 0.8643 ג. 0.8749.
10) א. נדחה H_0. ב. נדחה H_0.
11) א. 0.0062 ב. 0.9938 ג. 0.0124.
12) א. לא נדחה H_0. ב. לא נדחה H_0.

© כל הזכויות שמורות לברק קנדל

שאלות סיכום – שאלות רב ברירה על בדיקת השערות

1) בבדיקת השערה חד-צדדית ימנית ברמת מובהקות $\alpha = 0.01$, נדחתה השערת האפס. מה הייתה המסקנה לו נבדקה אותה ההשערה באמצעות אותם נתונים ברמת מובהקות $\alpha = 0.05$?

א. השערת האפס הייתה נדחית.

ב. השערת האפס לא הייתה נדחית.

ג. ההשערה המחקרית הייתה נדחית.

ד. בהעדר נתונים נוספים, לא ניתן לדעת.

2) על מנת לבדוק האם ההסתברות ללידת בן הינה חצי, נבחר מדגם מקרי של 200 ילדים, ונמצא שישנם 120 בנים.
מהן ההשערות האלטרנטיבית להשערת האפס?

א. $H_1 : p = 0.5$

ב. $H_1 : p = 0.6$

ג. $H_1 : p > 0.5$

ד. $H_1 : p \neq 0.5$

3) לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים. אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני ללא היפנוזה. לאחר מכן נערך לשניהם מבחן באנגלית. נניח שציוני המבחן מתפלגים נורמאלית ללא ידיעת השונות האמיתית. המבחן שיש לבצע כאן הוא:

א. מבחן Z למדגם יחיד.

ב. מבחן Tלמדגם יחיד.

ג. מבחן T למדגמים בלתי תלויים.

ד. מבחן T למדגמים מזווגים.

4) כדי לבדוק את הטענה שגברים רווקים שוקלים פחות מגברים נשואים לקח חוקר מדגם מקרי של 4 גברים ומדד את משקלם לפני נישואיהם ולאחר נישואיהם. הנה התוצאות:
מהן ההשערות הנבדקות?(ההפרש חושב $X - Y$)

לפני הנישואין- X	69	93	82	68
לאחר הנישואין- Y	80	88	84	71

א. $H_1 : \mu_d < 0, \quad H_0 : \mu_d = 0$

ב. $H_1 : \mu_X - \mu_Y < 0, \quad H_0 : \mu_X - \mu_Y = 0$

ג. $H_1 : \mu_X - \mu_Y < 0, \quad H_0 : \mu_X - \mu_Y = 0$

ד. $H_1 : \mu_d > 0, \quad H_0 : \mu_d = 0$

© כל הזכויות שמורות לברק קנדל

5) חוקר ביצע מחקר ובו עשה טעות מסוג שני לכן:

א. השערת האפס נכונה.

ב. השערת האפס נדחתה.

ג. השערת האפס לא נדחתה.

ד. אף אחת מהתושבות לא נכונה בהכרח.

6) ידוע כי ילד בגיל שנתיים ישן בממוצע 9 שעות בלילה. במדגם של 20 תינוקות בני שנתיים המתגוררים בצפון נמצא, כי ממוצע שעות השינה בלילה הינו 10 עם סטיית תקן של 1.1 במדגם של 10 תינוקות בדרום נמצא, כי ממוצע שעות השינה בלילה הינו 7.9 עם סטיית תקן של 1.1. על מנת להשוות בין ממוצע שעות השינה של ילדים מהצפון לבין זה של כלל הילדים יש לערוך _____, ועל מנת להשוות בין ממוצע שעות השינה של ילדים מהדרום לזה של ילדים המתגוררים בצפון יש לערוך _____.
יש להניח שההנחות הדרושות מתקיימות.

א. מבחן Z למדגם יחיד; מבחן T למדגם יחיד.

ב. מבחן T למדגם יחיד; מבחן T למדגמים תלויים.

ג. מבחן T למדגם יחיד; מבחן T למדגמים בלתי תלויים.

ד. מבחן T למדגמים בלתי תלויים; מבחן T לממוצע יחיד.

7) מובהקות התוצאה (PV) היא גם:

א. רמת המובהקות המינימאלית לדחות השערת האפס.

ב. רמת המובהקות המקסימאלית לדחיית השערת האפס.

ג. רמת המובהקות שנקבעת מראש על ידי החוקר טרם קיבל את תוצאות המחקר.

ד. רמת המובהקות המינימאלית לאי דחיית השערת האפס.

8) כדי לבדוק את הטענה שגברים רווקים שוקלים פחות מגברים נשואים לקח חוקר מדגם מקרי של 4 גברים ומדד את משקלם לפני נישואיהם ולאחר נישואיהם. הנה התוצאות:

לפני הנישואין	69	93	82	68
לאחר הנישואין	80	88	84	71

באיזה התפלגות משתמשים לבדיקת ההשערות, ובכמה דרגות חופש:

א. ההתפלגות Z ללא דרגות חופש.

ב. ההתפלגות T ו-3 דרגות חופש.

ג. ההתפלגות T ו-6 דרגות חופש.

ד. ההתפלגות χ^2 ו-3 דרגות חופש.

© כל הזכויות שמורות לברק קנדל

9) שני סטטיסטיקאים בודקים השערות ברמת מובהקות $\alpha = 0.05$ על סמך אותו מדגם. סטטיסטיקאי א' בודק את ההשערה: $H_0 : \mu = 20$ כנגד האלטרנטיבה $H_1 : \mu \neq 20$ ומחליט לא לדחות את השערת האפס.

סטטיסטיקאי ב' בודק את ההשערה $H_0 : \mu \leq 20$ כנגד האלטרנטיבה $H_1 : \mu > 20$ מה יחליט סטטיסטיקאי ב'?

א. לדחות את השערת האפס.

ב. לא לדחות את השערת האפס.

ג. ללא נתונים נוספים אי אפשר לדעת מה יחליט.

10) חוקר בדק השערה מסוימת והחליט לדחות את השערת האפס ברמת מובהקות 5%. מה נכון לומר?

א. הוא בוודאות ידחה את השערת האפס ברמת מובהקות 9% ואילו ברמת מובהקות 2% יש לבדוק מחדש.

ב. הוא בוודאות לא ידחה את השערת האפס ברמת מובהקות 9% ואילו ברמת מובהקות 2% יש לבדוק מחדש.

ג. הוא בוודאות ידחה את השערת האפס ברמת מובהקות 9% וברמת מובהקות 2%.

ד. הוא בוודאות לא ידחה את השערת האפס ברמת מובהקות 9% ואילו ברמת מובהקות 2% יש לבדוק מחדש.

11) רמת הכולסטרול בדמם של אנשים מתפלג נורמאלית עם תוחלת של 180 מ"ג (ל 100 סמ"ק דם). וסטיית תקן של 10 מ"ג. מעוניינים לבדוק את הטענה שצמחונים הם בעלי רמת כולסטרול נמוכה יותר.
נניח שסטיית התקן אצל צמחונים זהה לסטיית התקן של כלל האנשים.
במדגם של 20 צמחונים התקבל ממוצע רמת כולסטרול 174.5 מ"ג.
אם הוחלט לקבל את הטענה שצמחונים הם בעלי רמת כולסטרול נמוכה יותר איזה סוג טעות אפשרית במסקנה?

א. טעות מסוג ראשון.

ב. טעות מסוג שני.

ג. טעות מסוג שלישי.

ד. לא ניתן לדעת כיוון שאנו לא יודעים מה התוחלת האמתית אצל הצמחוניים.

© כל הזכויות שמורות לברק קנדל

12) בסקר שנערך התקבל ש 60% מתוך 220 נשאלים מבקרים אצל השיננית לפחות פעם אחת בשנה. עבור אילו רמות מובהקות ניתן יהיה לקבוע שרוב האוכלוסייה מבקרת אצל השיננית לפחות פעם בשנה?

א. רמת מובהקות הגדולה מ-5%.

ב. רמת מובהקות הקטנה מ-5%.

ג. רמת מובהקות הגדלה מ-0.0015.

ד. רמת מובהקות הקטנה מ-0.0015.

13) שני חוקרים העוסקים בתחום מחקרי משותף החליטו להסתמך על נתונים של מדגם שפורסם על ידי הלשכה המרכזית לסטטיסטיקה.
חוקר א' ניסח השערה דו צדדית ואילו חוקר ב' ניסח השערה חד צדדית.
מסקנתו של איזה מבין המשפטים הבאים הוא הנכון בנוגע למסקנות החוקרים?

א. אם חוקר א' ידחה את השערת האפס לא ניתן לדעת מה יחליט חוקר ב' באותה רמת מובהקות.

ב. אם חוקר א' יקבל את השערת האפס גם חוקר ב' יקבל את השערת האפס באותה רמת מובהקות.

ג. אם חוקר ב' ידחה את השערת האפס גם חוקר א' ידחה את השערת האפס באותה רמת מובהקות.

ד. אם חוקר א' ידחה את השערת האפס גם חוקר ב' ידחה את השערת האפס בתנאי שרמת המובהקות כפולה בגודלה.

14) ידוע מנתוני העבר כי תוחלת הציונים בבחינה בפסיכולוגיה היא 79. הועלתה השערה כי תוחלת הציונים בקרב העולים החדשים נמוכה יותר. לצורך בדיקת הטענה נלקח מדגם מקרי של 47 סטודנטים עולים ונמצא ממוצע של 75.
מה משמעות הפרמטר בניסוח ההשערות?

א. תוחלת ציוני העולים באוכלוסייה.

ב. ממוצע ציוני העולים במדגם.

ג. תוחלת ציוני האוכלוסייה מנתוני העבר.

ד. ממוצע ציוני שאר האוכלוסייה במדגם.

15) חוקר ביצע מחקר וידוע כי עשה טעות מסוג 1. מה מהבאים נכון?

א. החוקר דחה את השערת H_0 כאשר היא הייתה נכונה.

ב. החוקר דחה את השערת H_1 כאשר היא הייתה נכונה.

ג. החוקר לא דחה את השערת H_0 כאשר היא הייתה לא נכונה.

ד. המדגם של החוקר שייך בפועל להתפלגות הדגימה של H_1.

© כל הזכויות שמורות לברק קנדל

16) חוקר ביקש לבחון האם תאומים זהים אשר הופרדו בילדותם שונים מתאומים זהים אשר גדלו יחדיו מבחינת מידת הפער בין התאומים בלחץ הדם. הוא דגם 20 זוגות תאומים מכל אוכלוסייה ומדד את הפרש בין לחץ הדם בכל זוג תאומים. מהו המבחן הסטטיסטי המתאים?

א. מבחן T למדגמים בלתי תלויים עם 38 דרגות חופש.

ב. מבחן T למדגמים מזווגים, עם 39 דרגות חופש.

ג. מבחן T למדגמים בלתי תלויים עם 39 דרגות חופש.

ד. מבחן T למדגמים מזווגים עם 38 דרגות חופש.

17) בינואר השנה פורסם שהשכר הממוצע במשק הוא ₪8,900. במדגם שנעשה בחודש יוני על 60 עובדים נרשם עבור כל עובד במדגם האם השכר שלו נמוך או לא נמוך מהשכר הממוצע שפורסם בחודש ינואר. מהו המבחן המתאים כדי לבדוק שרוב העובדים בחודש יוני קיבלו שכר הנמוך מהשכר הממוצע שפורסם בחודש ינואר?

א. מבחן Z על פרופורציה.

ב. מבחן T על תוחלת אחת.

ג. מבחן T על שתי תוחלות במדגמים בלתי תלויים.

ד. מבחן T על שתי תוחלות במדגמים תלויים.

18) שלושה חוקרים רצו לבדוק את השפעתו של שידור פרסומות נגד תאונות דרכים על מהירות הנהיגה של נהגים בישראל (השונות של מהירות הנהיגה בישראל אינה ידועה). עידו השווה את מהירות הנהיגה של קבוצת נהגים אחת, חודש לפני שידור הפרסומות וחודש לאחר שידור הפרסומות.
רון השווה את מהירות הנהיגה של קבוצת נהגים, שראו את הפרסומות, למהירות הנהיגה של קבוצת נהגים, שלא ראו את הפרסומות.
יואב השווה את מהירות הנהיגה של קבוצת נהגים בחודש בו שודרו הפרסומות, למהירות הנהיגה הממוצעת בישראל על פי נתוני משרד התחבורה.
המבחנים בהם צריכים החוקרים להשתמש הם:

א. שלושתם במבחן T למדגמים בלתי תלויים.

ב. עידו במבחן T למדגמים מזווגים, ורון ויואב במבחן T למדגמים בלתי תלויים.

ג. עידו במבחן T למדגמים מזווגים, רון במבחן T למדגמים בלתי תלויים ויואב במבחן T למדגם יחיד.

ד. עידו במבחן T למדגמים מזווגים, רון ויואב במבחן T למדגם יחיד.

© כל הזכויות שמורות לברק קנדל

19) במחקר נמצא שתוצאה היא מובהקת ברמת מובהקות של 5%.
מה תמיד נכון?

א. הגדלת רמת המובהקות לא תשתנה את מסקנת המחקר.
ב. הגדלת רמת המובהקות תשנה את מסקנת המחקר.
ג. הקטנת רמת המובהקות לא תשנה את מסקנת המחקר.
ד. הקטנת רמת המובהקות תשנה את מסקנת המחקר.

20) חוקר ערך מבחן דו צדדי ברמת מובהקות של α והחליט לדחות את השערת האפס. אם החוקר היה עורך מבחן חד צדדי ברמת מובהקות של $\frac{\alpha}{2}$ אזי בהכרח:

א. השערת האפס הייתה נדחית.
ב. השערת האפס הייתה לא נדחית.
ג. לא ניתן לדעת מה תהיה מסקנתו במקרה זה.

21) ליאור ורוני העלו את אותן השערות על ממוצע האוכלוסייה. כמו כן הם התבססו על אותן תוצאות של מדגם.
ליאור השתמש בטבלה של התפלגות Z.
רוני השתמשה בטבלה של התפלגות T.
מה נוכל לומר בנוגע להחלטת המחקר שלהם?

א. אם ליאור ידחה את השערת האפס אז גם בהכרח רוני.
ב. אם רוני תדחה את השערת האפס אז גם בהכרח ליאור.
ג. שני החוקרים בהכרח יגיעו לאותה מסקנה.
ד. לא ניתן לדעת על היחס בין דחיית השערת האפס של שני החוקרים.

22) נתון ש $X \sim N(\mu,\sigma^2)$ כמו כן נתונות ההשערות הבאות: $H_1: \mu < \mu_0$, $H_0: \mu = \mu_0$.
חוקר בדק את ההשערות הללו על סמך מדגם שכלל 10 תצפיות. σ^2 לא הייתה ידועה לחוקר. החוקר החליט לדחות את השערת האפס ברמת מובהקות של 5%
לאחר מכן כדי לחזק את קביעתו הוא דגם עוד 5 תצפיות ושקלל את תוצאות אלה גם למדגם כך שכלל עכשיו 15 תצפיות.

א. כעת בברור הוא ידחה את השערת האפס.
ב. כעת הוא דווקא יקבל את השערת האפס.
ג. כעת לא ניתן לדעת מה תהיה מסקנתו.

23) אם חוקר החליט להגדיל את רמת המובהקות במחקר שלו אזי:

א. הסיכוי לטעות מסוג ראשון גדל.
ב. העוצמה של המבחן גדלה.
ג. הסיכוי לטעות מסוג שני גדל.
ד. תשובות א ו-ב נכונות.

© כל הזכויות שמורות לברק קנדל

24) חוקר ביצע מחקר ובו עשה טעות מסוג שני לכן:

א. השערת האפס נכונה.

ב. השערת האפס נדחתה.

ג. השערת האפס לא נדחתה.

ד. אף אחת מהתושבות לא נכונה בהכרח.

25) מה המצב הרצוי לחוקר המבצע בדיקת השערה:

	α	$1-\beta$
א.	גדולה	גדולה
ב.	גדולה	קטנה
ג.	קטנה	גדולה
ד.	קטנה	קטנה

26) נערך שינוי בכלל ההחלטה של בדיקת השערה מסוימת ובעקבותיו אזור דחיית H_0 קטן. כל שאר הגורמים נשארו ללא שינוי. כתוצאה מכך:

א. הן α, והן $(1-\beta)$, יקטנו.

ב. α יישאר ללא שינוי ואילו $(1-\beta)$ יגדל.

ג. α יגדל ואילו $(1-\beta)$ יקטן.

ד. הן α והן $(1-\beta)$ יגדלו.

27) ידוע כי לחץ דם תקין באוכלוסייה הוא 120. רופא מניח שלחץ הדם בקרב עיתונאים גבוה יותר מהממוצע באוכלוסייה. הוא לקח מדגם של 60 עיתונאים וקיבל ממוצע 137. על סמך המדגם, הוא בודק טענתו ברמת מובהקות 0.02 ומסיק שלחץ הדם בקרב העיתונאים אינו גבוה יותר. מה הטעות האפשרית שהרופא עושה?

א. טעות מסוג ראשון.

ב. טעות מסוג שני.

ג. טעות מסוג שלישי.

ד. אין טעות במסקנתו.

28) בבדיקת השערות התקבל שה- $p-value=0.02$.

מה תהיה מסקנת חוקר המשתמש ברמת מובהקות 1%?

בחר בתשובה הנכונה:

א. יקבל את השערת האפס בכל מקרה.

ב. ידחה את השערת האפס מקרה.

ג. ידחה את השערת האפס רק אם המבחן הנו דו צדדי.

ד. לא ניתן לדעת כי אין מספיק נתונים.

© כל הזכויות שמורות לברק קנדל

29) מובהקות התוצאה (PV) היא גם:

א. רמת המובהקות המינימאלית לדחות השערת האפס.

ב. רמת המובהקות המקסימאלית לדחיית השערת האפס.

ג. רמת המובהקות שנקבעת מראש על ידי החוקר טרם קיבל את תוצאות המחקר.

ד. רמת המובהקות המינימאלית לאי דחיית השערת האפס.

30) בבדיקת השערות מסוימת התקבל $p\ value = 0.0254$, לכן:

א. ברמת מובהקות של 0.01 אך לא של 0.05 נדחה את H_0.

ב. ברמת מובהקות של 0.01 ושל 0.05 לא נדחה את H_0.

ג. ברמת מובהקות של 0.05 אך לא של 0.01 נדחה את H_0.

ד. ברמת מובהקות של 0.01 ושל 0.05 נדחה את H_0.

31) רמת המובהקות במחקר הייתה 2% לכן.

א. בסיכוי של 2% נדחה את השערת האפס.

ב. בסיכוי של 2% לא נדחה את השערת האפס.

ג. בסיכוי של 2% השערת האפס לא נכונה.

ד. אף תשובה לא נכונה.

32) נתון ש: $X \sim N(\mu,\sigma^2)$. כמו כן נתונות ההשערות הבאות: $H_1: \mu < \mu_0, \quad H_0: \mu = \mu_0$.
חוקר בדק את ההשערות הללו על סמך מדגם שכלל 10 תצפיות.
σ^2 לא הייתה ידועה לחוקר. החוקר החליט לדחות את השערת האפס ברמת מובהקות של 5%. אם הוא היה מגדיל את רמת המובהקות ל-10% אזי:

א. כעת בברור הוא ידחה את השערת האפס.

ב. כעת הוא דווקא יקבל את השערת האפס.

ג. כעת לא ניתן לדעת מה תהיה מסקנתו.

33) לצורך בדיקת השפעת היפנוזה על לימוד אנגלית, נבחרו 10 זוגות תאומים זהים. אחד התאומים למד אנגלית בהשפעת היפנוזה, והשני ללא היפנוזה. לאחר מכן נערך לשניהם מבחן באנגלית. נניח שציוני המבחן מתפלגים נורמאלית ללא ידיעת השונות האמתית. מספר דרגות החופש במבחן הוא:

א. 9

ב. 19

ג. 18

ד. 8

© כל הזכויות שמורות לברק קנדל

34) בתחנת טיפת חלב מסוימת יש שני מכשירי שקילה. על מנת להשוות בין שני המשקלים נדגמו 4 תינוקות. כל תינוק בן חודשיים נשקל בכל אחד מהמשקלים. להלן תוצאות השקילה (בק"ג):

משקל במכשיר 1	4.5	9.6	0.7	2.5
משקל במכשיר 2	3.5	6.9	1.7	0.5

נניח שהמשקלים מתפלגים נורמלית.
המבחן שיש לבצע כאן הוא:

א. מבחן Z למדגם יחיד.
ב. מבחן T למדגם יחיד.
ג. מבחן T למדגמים בלתי תלויים.
ד. מבחן T למדגמים מזווגים.

35) כדי להשוות בין שני אצים נדגמו 5 תוצאות מריצת 100 מטר של כל אצן. זמני הריצה נרשמו ויש להניח שמתפלגים נורמלית. המטרה להשוות בין האצנים.
המבחן שיש לבצע כאן הוא:

א. מבחן Z למדגם יחיד.
ב. מבחן T למדגם יחיד.
ג. מבחן T למדגמים בלתי תלויים.
ד. מבחן T למדגמים מזווגים.

36) סטטיסטיקאי ערך מבחן סטטיסטי. הוא חישב את עוצמת המבחן וקיבל 0.
המשמעות של תוצאה זו היא:

א. לעולם לא לדחות את השערת האפס כאשר היא לא נכונה.
ב. תמיד לדחות את השערת האפס כאשר היא נכונה.
ג. לעולם לא לדחות את השערת האפס כאשר היא נכונה.
ד. תמיד לדחות את השערת האפס כאשר היא לא נכונה.

37) סטטיסטיקאי נתבקש לאמוד את הפרש הממוצעים של שני טיפולים לפי שני מדגמים מקריים בלתי תלויים. הוא חישב רווח סמך להפרש ברמת סמך 0.98, וקיבל את הרווח $-2 < \mu_1 - \mu_2 < 4.5$. אילו יתבקש החוקר לבדוק לפי אותם נתונים את ההשערות: $H_0 : \mu_1 - \mu_2 = 0$; $H_1 : \mu_1 - \mu_2 \neq 0$,
ברמת מובהקות 0.05 מסקנתו תהיה:

א. לדחות את השערת האפס.
ב. לא לדחות את השערת האפס.
ג. שלא ניתן לדעת את המסקנה עבור רמת מובהקות 0.05.
ד. שלא נתונות בשאלה סטיות התקן של האוכלוסיות, ולכן לא ניתן להסיק דבר.

© כל הזכויות שמורות לברק קנדל

38) במטרה לבדוק האם קיים הבדל בין קווי זהב לבזק מבחינת ממוצע המחירים לשיחות בינ"ל. נגדמו באקראי 7 מדינות ועבור כל מדינה נבדקה עלות דקת שיחה. בהנחה והמחירים מתפלגים נורמלית בנו רווח סמך למוצע ההפרשים וקיבלו: $-0.0293 < \mu_D < 0.2145$ רווח הסמך הוא ברמת סמך של 95%.
לכן מסקנת המחקר היא:

א. ברמת מובהקות של 5% לא נוכל לקבוע שקיים הבדל בין החברות.

ב. ברמת מובהקות של 5% נקבע שקיים הבדל מובהק בין החברות.

ג. לא ניתן לדעת מה המסקנה ברמת מובהקות של 5% כיוון שלא נאמר מה ההגדרה של D.

39) אם רמת מובהקות של מבחן סטטיסטי הינה 0, הכוונה היא:

א. תמיד נדחה H_0 כאשר היא נכונה, אך לא תמיד נדחה אותה כאשר היא לא נכונה.

ב. לא נדחה את H_0 אף פעם.

ג. לא נדחה את H_0 כאשר היא נכונה אך יתכן ונדחה אותה כאשר היא לא נכונה.

ד. כל התשובות לא נכונות.

40) חוקר ביצע ניסוי. הוא ניסח את ההשערות הבאות: $H_1: \mu \neq 10, \quad H_0: \mu = 10$.
לצורך בדיקה הוא לקח מדגם מקרי בגודל 5 מתוך אוכלוסייה המתפלגת נורמאלית עם שונות לא ידועה. על סמך תוצאות המדגם הוא חישב וקיבל: $t_{\bar{x}} = -2.63$. לכן המסקנה היא:

א. הוא ידחה H_0 ברמת מובהקות 0.1 אך לא כן ברמת מובהקות 0.05.

ב. הוא ידחה H_0 ברמת מובהקות 0.05 אך לא כן ברמת מובהקות 0.025.

ג. הוא ידחה H_0 ברמת מובהקות 0.025 אך לא כן ברמת מובהקות 0.01

ד. הוא לא ידחה H_0 ברמת מובהקות 0.1.

41) האיגוד האמריקני לרפואת ילדים מפרסם הנחיות חדשות הקובעות כי יש ליטול תוספת יוד במהלך תקופת ההיריון וההנקה. מחסור במינרל זה עלול לגרום לפגיעה מוחית אצל העובר והתינוק. החלטה זו נקבעה על סמך מחקר בו השתתפו 1050 נשים שנטלו יוד במהלך תקופת ההיריון וההנקה.
מתוך הנשים שהשתתפו במחקר, רק ל-21 נמצאו ילדים בעלי פגיעה מוחית לעומת 3% באוכלוסייה הכללית. בנוסף, פורסם שהאיגוד האמריקאי מגיע למסקנותיו על סמך רמת מובהקות של 0.5%. מה הסיכוי לבצע טעות מסוג ראשון במחקר?

א. 0.005

ב. 0.03

ג. 0.0287

ד. 0.05

© כל הזכויות שמורות לברק קנדל

42) חוקרת שיערה, כי משקלן של נשים כשנה לאחר החתונה גבוה ממשקלן בעת החתונה. החוקרת דגמה 15 נשים, ובדקה את משקלן בשתי נקודות הזמן (בעת החתונה, ושנה לאחריה), אך לא מצאה הבדל מובהק ברמת מובהקות 0.01. בהנחה, כי **במציאות** השערתה של החוקרת נכונה, סביר כי אם היא תגדיל את גודל המדגם, אזי:

א. יקטן הסיכוי לטעות מסוג שני (β).

ב. תגדל רמת הביטחון ($1-\alpha$).

ג. אף תשובה לא נכונה.

ד. כל התשובות נכונות.

43) איזה מהמשפטים הבאים נכון תמיד?

א. $\text{POWER}+\alpha+\beta=1$

ב. $\text{POWER}=0.5-\beta$

ג. $\text{POWER}+\alpha=1$

ד. $\beta+\alpha=1$

ה. הכול לא נכון.

44) מה נכון לומר לגבי הנחת שיוויון השונויות במבחן T למדגמים בלתי תלויים?

א. היא אומרת שהשונויות המדגמיות שוות.

ב. בלעדיה אין שום דרך לבדוק השערה על הפרש בין תוחלות.

ג. היא חשובה הן עבור מדגמים מזווגים והן עבור מדגמים בלתי תלויים.

ד. אף תשובה אינה נכונה.

45) חוקר החליט לא לדחות השערה ברמת מובהקות של α. במידה וחוקר זה היה בודק השערה זו ברמת מובהקות של 2α על סמך אותם נתונים, האם ההשערה תדחה?

א. ההשערה תדחה.

ב. ההשערה לא תדחה.

ג. התשובה תלויה בעוצמת המבחן.

ד. לא ניתן לדעת בוודאות אם ההשערה תדחה או לא.

46) חוקרת שיערה, כי בגילאי הגן בנות יותר תקשורתיות מבנים. אם החוקרת תדגום אקראית 30 בנים ו-30 בנות, ובמדגם יתקבל אותו ממוצע של ציון תקשורת. סטטיסטי המבחן יהיה:

א. אפס

ב. חיובי

ג. שלילי

ד. לא ניתן לדעת

© כל הזכויות שמורות לברק קנדל

47) עוצמה שווה ל-1 פרושה:

א. לעולם לא לדחות את השערת האפס כאשר היא נכונה.

ב. תמיד לדחות את השערת האפס כאשר היא נכונה.

ג. לעולם לא לדחות את השערת האפס כאשר היא לא נכונה.

48) מה מהבאים **<u>נכון</u>** לגבי מבחן T מדגמים מזווגים?

א. כל התצפיות במחקר אינן תלויות זו בזו.

ב. כל התצפיות במחקר תלויות זו בזו.

ג. כל הצמדים של תצפיות במחקר אינם תלויים זה בזה.

ד. התצפיות בתוך כל צמד אינן תלויות זו בזו.

49) לבדיקת ההשערה החד צדדית על התוחלת של התפלגות נורמלית $H_0:\mu\geq 10$, $H_1:\mu<10$. נלקח מדגם והתקבלה רמת מובהקות מינימאלית לדחיית השערת האפס 0.058. לו רצינו לבדוק את ההשערה הדו צדדית $H_0:\mu=10,\quad H_1:\mu\neq 10$, אז על סמך תוצאת אותו המדגם ברמת מובהקות 0.05:

א. ניתן להכריע בין ההשערות רק אם שונות האוכלוסייה נתונה.

ב. מקבלים את השערת האפס.

ג. דוחים את השערת האפס.

ד. לא ניתן להכריע בין ההשערות שכן חסרים נתונים.

50) לבדיקת ההשערה החד צדדית ימנית $H_0:\mu=55,\quad H_1:\mu=65$, נלקח מדגם מקרי בגודל n מאוכלוסייה בעלת התפלגות נורמלית ושונות σ^2. רמת המובהקות היא 5%. נמצא שהעוצמה היא 0.9. להלן 3 טענות:

- עבור מדגם בגודל n וברמת מובהקות 5% לבדיקת ההשערות: $H_0:\mu=55,\quad H_1:\mu=60$ העוצמה תהיה גדולה מ-0.9.
- עבור מדגם בגודל $2n$ ורמת מובהקות 5% לבדיקת ההשערות: $H_0:\mu=55,\quad H_1:\mu=65$ העוצמה תהיה גדולה מ-0.9.
- עבור מדגם בגודל n ורמת מובהקות 10% לבדיקת ההשערות: $H_0:\mu=55,\quad H_1:\mu=65$ העוצמה תהיה קטנה מ-0.9.

א. שלושת הטענות אינן נכונות.

ב. טענות 2 ו-3 אינן נכונות וטענה 1 נכונה.

ג. טענות 1 ו-2 נכונות וטענה 3 אינה נכונה.

ד. טענות 1 ו-3 אינן נכונות וטענה 2 נכונה.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

שאלה	תשובה		שאלה	תשובה
1	א		26	א
2	ד		27	ב
3	ד		28	א
4	א		29	א
5	ג		30	ג
6	ג		31	ד
7	א		32	א
8	ב		33	א
9	ג		34	ד
10	א		35	ג
11	א		36	א
12	ג		37	ג
13	א		38	א
14	א		39	ג
15	א		40	א
16	א		41	א
17	א		42	א
18	ג		43	ה
19	א		44	ד
20	ג		45	ד
21	ב		46	א
22	ג		47	ד
23	ד		48	ג
24	ג		49	ב
25	ג		50	ד

© כל הזכויות שמורות לברק קנדל

מבוא לסטטיסטיקה בניהול

פרק 48 - מבחני חי בריבוע

תוכן העניינים

1. מבחן טיב התאמה..... (ללא ספר)
2. מבחן טיב התאמה והקשר שלו לבדיקת השערות על פרופורציה אחת..... 240
3. מבחן לאי תלות..... (ללא ספר)
4. קשר בין מבחן אי תלות לבדיקת השערות להפרש פרופרוציות..... 242

הקשר בין מבחן טיב התאמה לבדיקת השערות על הפרופורציה – רקע

אם אנו רוצים לבצע מבחן טיב התאמה על משתנה שיש לו שתי קטגוריות בלבד (משתנה דיכוטומי), הדבר זהה לתהליך של בדיקת השערות דו צדדית על פרופורציה בודדת.

דוגמה (פתרון בהקלטה) :

הטילו מטבע 80 פעמים וקיבלו 48 פעמים את התוצאה "ראש".
בדקו האם המטבע הוא הוגן ברמת מובהקות של 5%.

א. באמצעות מבחן טיב התאמה.

ב. באמצעות מבחן Z לפרופורציה בודדת.

© כל הזכויות שמורות לברק קנדל

שאלות

1) בסקר שנעשה על 320 נשאלים, 43.75% טענו שהחיה המועדפת עליהם היא כלב. עד היום היה נהוג לחשוב ש40% מהאנשים מעדיפים כלבים.
בדקו ברמת מובהקות של 5% האם הסקר יישנה את הסברה שהייתה נהוגה עד היום לגבי העדפת כלב.

א. באמצעות מבחן טיב התאמה.

ב. באמצעות מבחן על פרופורציה.

2) לסוכנות מכוניות שלושה סניפים ברחבי הארץ. המכוניות נמכרות בסניפים השונים. מתוך 100 מכוניות נמצא ש-65 נמכרו בסניף תל-אביב, 23 בסניף ירושלים והיתר בסניף חיפה.

א. בדקו ברמת מובהקות של 5% האם שיעור המכוניות שנמכרות בסניף ת"א גדול פי 2 מכל סניף אחר.

ב. בדקו באמצעות מבחן טיב התאמה האם 60% מהמכוניות נהוגות להימכר בסניף תל אביב. האם יש דרך אחרת לבדוק את ההשערה?

3) בתחרות ריצה בית ספרית שלושה מסלולי ריצה.
ב-50 תחרויות בדקו באיזה מסלול היה הניצחון.
התוצאות שהתקבלו מסוכמות בטבלה הבאה:

המסלול	1	2	3
מספר הניצחונות	20	15	15

א. בדקו ברמת מובהקות של 5% האם יש מסלול מועדף לניצחון.

ב. בדקו ברמת מובהקות של 5% האם הסיכוי לנצח במסלול מספר 1 גבוה מ-$\frac{1}{3}$.

תשובות סופיות

1) לא נדחה H_0.

2) א. נדחה H_0. ב. לא נדחה H_0.

3) א. לא נדחה H_0. ב. לא נדחה H_0.

© כל הזכויות שמורות לברק קנדל

הקשר בין מבחן לאי תלות ובדיקת השערות להפרש פרופורציות – רקע

מבחן לאי תלות שבו לכל משתנה יש שתי קטגוריות שקול לבדיקת השערות דו צדדית על הפרש פרופורציות כאשר השערת האפס היא שהפרופורציות שוות. כל זאת, כמובן, אם התנאים למבחנים מתקיימים.

דוגמה (פתרון בהקלטה) :

בקרב מדגם של 200 נשים 120 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? האם אפשר לבדוק זאת גם על ידי מבחן לאי תלות וגם על ידי בדיקת השערות לשתי פרופורציות?

שאלות

1) בקרב מדגם של 200 נשים 120 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%.

א. על ידי מבחן לאי תלות.

ב. על ידי בדיקת השערות לשתי פרופורציות.

2) נלקחו 200 אנשים שמתוכם 60 הצהירו שהם עוסקים בפעילות גופנית סדירה. מתוך אלו שעוסקים בפעילות גופנית סדירה 50 נמצאו במצב בריאותי תקין. מתוך אלו שלא עוסקים בפעילות גופנית סדירה 90 נמצאו במצב בריאותי תקין.

א. בנו טבלת שכיחות משותפת לנתונים שהוצגו בשאלה.

ב. האם ניתן להגיד שהסיכוי להימצא במצב בריאותי תקין גבוה יותר כאשר עוסקים בפעילות גופנית סדירה לעומת המצב שלא עוסקים בפעילות גופנית סדירה? בדוק ברמת בטחון של 90%.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות

1) נדחה את השערת האפס לפירוט הדרך ראה וידאו.

2) א. להלן טבלה: ב. נדחה את השערת האפס.

פעילות/מצב בריאותי	תקין	לא תקין	
סדירה	50	10	60
לא סדירה	90	50	140
	140	60	200

© כל הזכויות שמורות לברק קנדל