ביוסטטיסטיקה למתקדמים

פרק 5 - מבחנים אפרמטריים למדגמים בלתי תלויים

תוכן העניינים

1. מבחן מאן וויטני - שימוש בפלטים 1
2. מבחן קרוסקל וליס -ניתוח פלטים 5
3. מבחן פישר 9
4. מבחן פישר - קירוב נורמלי 13

מבחנים אפרמטרים למדגמים בלתי תלויים

ניתוח פלטים במבחן מן-וויטני – רקע

מבחן מן-ויטני מיועד לבדוק האם לשתי אוכלוסיות התפלגות שווה.
המבחן בוחן באופן רוחבי את כל תחום הערכים ולא מתמקד בערך מרכזי אחד.
נשתמש במבחן זה כאשר יש שני מדגמים בלתי תלויים והמשתנה הכמותי הנחקר אינו מתפלג נורמלית או שמדובר במשתנה מסולם סדר.
המבחן מתבסס על דירוג כל התצפיות. בעצם, מבחן זה הוא המענה האפרמטרי למבחן הפרמטרי להפרש תוחלות במדגמים בלתי תלויים.

דוגמה (פתרון בהקלטה) :

להלן תוצאות הערכות שקיבלו שני מורים: ד"ר A ופרופסור B. סטודנטים נתנו משוב כללי על המורים בסקלה של 1 (גרוע) עד 5 (מצוין). הטענה היא שד"ר A הוא מרצה טוב יותר מאשר פרופסור B.

Mann-Whitney Test

Ranks

	teacher	N	Mean Rank	Sum of Ranks
grade	dr A	17	25.00	425.00
	prof B	20	13.90	278.00
	Total	37		

Test Statistics[a]

	grade
Mann-Whitney U	68.000
Wilcoxon W	278.000
Z	-3.249
Asymp. Sig. (2-tailed)	.001
Exact Sig. [2*(1-tailed Sig.)]	.001[b]

a. Grouping Variable: teacher
b. Not corrected for ties.

- הסבירו מדוע נעשה כאן מבחן מן-וויטני.
- מה המסקנה ברמת מובהקות של 1%?

© כל הזכויות שמורות לברק קנדל

שאלות

1) מחקרים טוענים שקיים הבדל בין שעות השינה של גברים לשעות השינה של נשים. כיוון שלא ניתן להוכיח ששעות שינה הינו משתנה המתפלג נורמלית ביצעו מבחן מן-וויטני בו לקחו נשים וגברים אקראיים ובדקו את שעות השינה שלהם.

Ranks

	gender	N	Mean Rank	Sum of Ranks
sleeptime	male	10	???	135.00
	female	???	7.50	75.00
	Total	20		

Mann-Whitney Test

Test Statistics[a]

	sleeptime
Mann-Whitney U	20.000
Wilcoxon W	75.000
Z	-2.319
Asymp. Sig. (2-tailed)	.020

א. השלימו את סימני השאלה החסרים בפלט.

ב. מהי המסקנה ברמת מובהקות של 5%?

ג. מה הייתה מובהקות התוצאה אם טענת המחקר הייתה שגברים ישנים יותר מנשים?

© כל הזכויות שמורות לברק קנדל

2) שני אנשים נתבקשו לבדוק את מספר תאונות הדרכים בשבוע בשני קטעי כביש שונים. כל אחד בחר את השבועות באופן אקראי ובלתי תלוי באחר וספר כמה תאונות היו בכל כביש בשבוע. הפלטים שהתקבלו:

מספר תאונות

	Frequency
.00	2
1.00	1
2.00	2
3.00	1
Total	6

a. road = 1.00

מספר תאונות

	Frequency
.00	2
1.00	1
2.00	1
3.00	1
4.00	1
5.00	1
Total	7

a. road = 2.00

Ranks

	road	N	Mean Rank	Sum of Ranks
מספר תאונות	1.00	A	B	C
	2.00	D	E	F
	Total	G		

Test Statistics[a]

	VAR00002
Asymp. Sig. (2-tailed)	.465

א. השלימו בטבלה השלישית את המספרים החסרים במקום האותיות.

ב. מהי מובהקות התוצאה לבדיקת ההשערה שאין הבדל בין הכבישים מבחינת התפלגות תאונות הדרכים?

ג. כיצד הייתה משתנה התשובה של הסעיף הקודם אם כל חוקר היה מוסיף נתונים על שבוע נוסף לכל כביש?

ד. מה המסקנה ברמת מובהקות של 5%?

ה. מהי מובהקות התוצאה לבדיקת ההשערה שכביש מספר 1 עם התפלגות תאונות גבוהה יותר מאשר כביש מספר 2?

© כל הזכויות שמורות לברק קנדל

תשובות סופיות

1) א. 10 ,13.5 ב. נדחה H_0. ג. 0.01.

2) א. 6, 6.17, 37, 7, 7.71, 54, 13 ב. 0.465.

ג. לא ניתן לדעת. ד. לא נדחה H_0. ה. 0.7675.

© כל הזכויות שמורות לברק קנדל

מבחן קרוסקל וואליס – ניתוח פלטים – רקע

כאשר לא ניתן לבצע מבחן ניתוח שונות חד כיווני (למשל, כאשר השונויות לא שוות, המשתנה התלוי אינו מתפלג נורמאלי או המשתנה התלוי נמדד בסולם סדר) נבחר במבחן לא-פרמטרי להשוואת k קבוצות בלתי תלויות.

מבחן Kruskal-Wallis הוא מבחן אפרמטרי להשוואת מספר אוכלוסיות בלתי תלויות. מבחן מן-וויטני הינו מקרה פרטי של מבחן Kruskal-Wallis כאשר יש שתי קבוצות בלתי תלויות.

השערות המבחן:

- השערת האפס: אין הבדל בין הקבוצות הנבדקות.
- השערת המחקר: קיים הבדל בין 2 קבוצות לפחות.

מבחן זה מתבסס על דירוג כלל התצפיות במחקר מכל המדגמים.

דוגמה (פתרון בהקלטה):

השוו את שביעות הרצון מהשירות של 4 חברות טלפון סלולרי שונות. שביעות הרצון הוא ציון שהלקוח היה צריך לתת לרמת השירות של החברה מ-1 ועד 10.

א. מהו המשתנה התלוי ומהו המשתנה הבלתי תלוי במחקר זה?
ב. מהן השערות המחקר?
ג. מהו המבחן הסטטיסטי במקרה זה?
ד. מה מסקנות המחקר ברמת מובהקות של 5%?

Count

		company				
		1	2	3	4	Total
שביעות רצון	4.00	0	2	0	0	2
	5.00	0	1	0	0	1
	6.00	2	1	0	1	4
	7.00	2	2	1	0	5
	8.00	1	2	2	2	7
	9.00	1	1	3	2	7
	10.00	1	1	0	0	2
Total		7	10	6	5	28

© כל הזכויות שמורות לברק קנדל

Kruskal-Wallis Test

Ranks

	company	N	Mean Rank
שביעות רצון	1	7	13.93
	2	10	11.40
	3	6	18.50
	4	5	16.70
	Total	28	

Test Statistics[a,b]

	שביעות רצון
Chi-Square	3.363
df	3
Asymp. Sig.	.339

a. Kruskal Wallis Test

b. Grouping Variable: company

© כל הזכויות שמורות לברק קנדל

שאלות

1) מחקר פסיכולוגי רצה להשוות 3 טיפול שונים לטיפול בחרדה. נלקחו נבדקים הסובלים מחרדה. הם חולקו באקראי ל-3 קבוצות וכל קבוצה קיבלה שיטת טיפול אחרת לחרדה. להלן תוצאות המבחן הסטטיסטי שבוצע:

Kruskal-Wallis Test

Ranks

	group	N	Mean Rank
anxiety	1.00	8	15.88
	2.00	4	6.88
	3.00	8	6.94

Test Statistics[a,b]

	anxiety
Chi-Square	11.149
df	2
Asymp. Sig.	.004

א. כמה נבדקים השתתפו במחקר?

ב. באיזו שיטת טיפול רמת החרדה הייתה הנמוכה ביותר במדגם?

ג. האם, עבור ר"מ 0.05, יש הבדל ברמת החרדה המתקבלת בשיטות הטיפול השונות?

© כל הזכויות שמורות לברק קנדל

2) חברת תרופות לקחה 15 אנשים ברמת בריאות דומה. החברה חילקה את האנשים לשלוש קבוצות שוות בגודלן. לכל קבוצה ניתנה אותה תרופה במינון שונה (dosage). המינונים שניתנו הם: 10 מ״ג, 20 מ״ג ו-30 מ״ג. לאחר שעה מזמן לקיחת התרופה ניבדק קצב פעימות הלב של כל אדם (pulse).
הנתונים שלהלן הוזנו לתוכנה סטטיסטית, והתקבלו התוצאות הבאות:

מינון	10	10	10	10	10	20	20	20	20	20	30	30	30	30	30
דופק	87	87	83	81	79	79	76	82	82	82	68	67	75	73	72

Kruskal-Wallis Test

Ranks

	dosage	N	Mean Rank
pulse	10.00	A	E
	20.00	B	F
	30.00	C	G
	Total	D	

Test Statistics[a,b]

	pulse
Chi-Square	10.204
df	2
Asymp. Sig.	.006

א. השלימו את המספרים החסרים במקום האותיות A עד G בטבלת הדירוגים.

ב. בדוק ברמת מובהקות של 5% האם קיים הבדל בין המינונים השונים מבחינת הדופק של האנשים?

ג. הסבירו ללא חישוב כיצד הייתה משתנה התשובה לסעיף הקודם אם הינו מעלים את הדופק של כל התצפיות במחקר ב-2.

תשובות סופיות

1) א. 20 ב. מדגם 2 ג. נדחה H_0.

2) א. 3 ,9.3 ,11.7 ,5,5,5,15 ב. נדחה H_0. ג. לא תשתנה.

© כל הזכויות שמורות לברק קנדל

מבחן פישר – רקע

מבחן זה הוא מבחן הנכנס לקטגוריית המבחנים האפרמטריים.
משתמשים במבחן כאשר מתעניינים להשוות בין שתי אוכלוסיות והמשתנה התלוי הוא דיכוטומי, כלומר משתנה שיש לו שני ערכים אפשריים.
במבחן זה יוצרים מדגם אחד שמקבל טיפול כלשהו ומדגם אחר המהווה קבוצת ביקורת ואינו מקבל את הטיפול. מבחן זה הוא החלופה האפרמטרית למבחן הפרמטרי להשוואת שתי פרופורציות על סמך שני מדגמים בלתי תלויים.
במבחן הפרמטרי דורשים שבכל מדגם מספר ההצלחות וגם מספר הכישלונות יהיה לפחות 10 (הפחות מחמירים דורשים לפחות 5).
מבחן פישר, אותו נעשה בפרק זה, נקרא בספרות המקצועית:
"Fisher exact probability test".

דוגמה:

מעוניינים לבדוק האם שיעורי עזר יעילים בשיפור ההישגים בקורס סטטיסטיקה.
נלקחו 2 כיתות הלומדות סטטיסטיקה בנות 15 תלמידים כל אחת. בכיתה אחת ניתנו שיעורי עזר ובכיתה השנייה לא ניתנו שיעורי עזר. בכיתה בה ניתנו שיעורי עזר 1 נכשל בקורס ובכיתה שבה לא ניתנו שיעורי עזר 3 נכשלו בקורס.
מהן השערות המחקר ומהו המבחן הסטטיסטי המתאים?

הטכניקה הנוחה ביותר למבחן פישר היא לחשב את מובהקות התוצאה ולדחות את השערת האפס אם $PV \le \alpha$.
מובהקות התוצאה היא הסיכוי לתוצאות של המדגם וקיצוני יותר בהנחת השערת האפס.
כדי לחשב את מובהקות התוצאה נבנה טבלת שכיחות משותפת במבנה הבא:

	"כישלון"	"הצלחה"	סה"כ
קבוצת טיפול	A	B	A+B
קבוצת ביקורת	C	D	C+D
סה"כ	A+C	B+D	N

בנו את טבלת השכיחות המשותפת המתאימה לדוגמה:

			סה"כ
קבוצת טיפול	A	B	A+B
קבוצת ביקורת	C	D	C+D
סה"כ	$A+C$	$B+D$	N

© כל הזכויות שמורות לברק קנדל

דוגמה:

בהנחת השערת האפס התפלגות של A הינו משתנה מקרי היפר גיאומטרי שבו יש אוכלוסייה בגודל N , מתוכם $C+A$ "מיוחדים" ואנו דוגמים מתוכם מדגם בגודל $B+A$.

פונקציית ההסתברות של ההתפלגות ההיפר גיאומטרית במקרה זה,

כאשר X מייצג את השכיחות A תהיה: $\frac{\binom{A+C}{x}\binom{B+D}{A+B-x}}{\binom{N}{A+B}}$.

חשבו את מובהקות התוצאה בדוגמה ומה תהיה המסקנה ברמת מובהקות של 5%?

© כל הזכויות שמורות לברק קנדל

שאלות

1) פסיכיאטרים נשאלו האם תרופה אנטי דיכאונית מסוימת אכן משפיעה על מצב הרוח. נלקחו 28 אנשים שהתלוננו על דיכאון ברמה דומה והם חולקו באקראי לשתי קבוצות: 16 נטלו את התרופה האנטי דיכאונית הנחקרת והיתר היוו קבוצת ביקורת ונטלו פלסיבו. כעבור 3 חודשים נבדק מצבם הנפשי של כל משתתפי המחקר. בקרב אלו שנטלו את התרופה רק 2 התלוננו על דיכאון ובקרב אלו שנטלו הפלסבו 6 התלוננו על דכאון.
מה המסקנה ברמת מובהקות של 10%?

2) חנות פרחים מעוניינת לבדוק את הטענה שתאורה אולטרה סגולה מגדילה את אורך החיים של הפרחים. נדגמו 20 פרחים מאותו סוג. הם חולקו באקראי ל-2 קבוצות: 10 פרחים יהיו בקבוצת הניסוי, כלומר בתאורה אולטרה סגולה והפרחים הנותרים יהיו בקבוצת הביקורת – באותם התנאים בדיוק אך ללא תאורה אולטרה סגולה.
כעבור 5 ימים נבדקו כלל הפרחים.

תאורה / מצב הפרח לאחר 5 ימים	נבול	תקין
אולטרה סגולה	1	9
רגילה	7	3

א. מה היא רמת המובהקות המינימלית עבורה יוסק שתאורה אולטרה סגולה מגדילה את אורך החיים של הפרחים?

ב. כיצד התשובה לסעיף הקודם הייתה משתנה אם בקבוצת הביקורת היו נמצאים פחות פרחים נבולים?

3) משרד החינוך הזמין מחקר שמטרתו היה לבדוק האם שנת צהריים קבועה בזמן לימודי התיכון משפיעה על הזכאות לבגרות.
נדגמו בוגרי תיכון אקראיים שנשאלו שתי שאלות:
$Q1$ - האם בזמן התיכון נהגת לישון צהריים באופן קבוע?
$Q2$ - האם את/ה זכאי/ת לתעודת בגרות?
להלן התוצאות שהתקבלו.

א. רשמו את השערות המבחן. מהו המבחן הסטטיסטי המתאים? נמקו.

ב. מהי מובהקות התוצאה?

ג. בדקו את השערות המחקר ברמת מובהקות של 6%.

שאלון/ מספר נשאל	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Q1	כן	לא	לא	לא	כן	כן	לא	לא	לא	לא	לא	כן	לא	לא
Q2	כן	כן	כן	לא	לא	כן	לא	לא	כן	כן	לא	כן	כן	לא

© כל הזכויות שמורות לברק קנדל

תשובות סופיות

1) נדחה את H_0.

2) א. 0.0099 ב. תקטן.

3) א. מבחן פישר. ב. 0.8112 ג. לא נדחה H_0.

© כל הזכויות שמורות לברק קנדל

קירוב נורמלי למבחן פישר – רקע

כאשר המדגמים גדולים דיו ניתן לבצע קירוב נורמלי למבחן פישר.
אחד מכללי האצבע לקרוב הנורמלי אם כל ארבעת הביטויים הבאים גדולים מ-7.5:

$$\frac{(A+C)^2}{N} \quad \frac{(B+D)^2}{N} \quad \frac{(A+B)^2}{N} \quad \frac{(D+C)^2}{N}$$

אך יתכן שחוקרים שונים ישתמשו בכללי אצבע אחרים.
נתייחס ל-A בתור משתנה מקרי המתפלג נורמלית עם התוחלת והשונות הבאות:

$$E(A)=\frac{(A+B)(A+C)}{N} \quad V(A)$$
$$=\frac{(A+B)(A+C)(B+D)(C+D)}{N^2(N-1)}$$

דוגמה:

מעוניינים לבדוק האם שיעורי עזר יעילים בשיפור ההישגים. נלקחו 2 כיתות בנות 40 תלמידים כל אחת. בכיתה אחת נתנו שיעורי עזר ובכיתה שנייה לא נתנו שיעורי עזר. בכיתה שנתנו שיעורי עזר 10 נכשלו ובכיתה ללא שיעורי עזר 16 נכשלו.
מה המסקנה ברמת מובהקות של 5%?

© כל הזכויות שמורות לברק קנדל

שאלות

1) כדי לבדוק אם חיסון חדש לכלבים מסייע בהפחתת ההדבקות במחלה מסוימת, נערך ניסוי שבו 50 כלבים קיבלו חיסון ו- 50לא קיבלו חיסון. מבין הכלבים שחוסנו 10 כלבים נדבקו במחלה, ומבין הכלבים שלא חוסנו 18 נדבקו במחלה. מהי המסקנה ברמת מובהקות 0.05?

2) מחקר מעוניין לבדוק האם אכילת שוקולד משפיעה על רמת הסרטונין בדם. נלקחו 102 אנשים ברמת בריאות דומה. הם חולקו באקראי לשתי קבוצות: 52 קיבלו שוקולד ו-50 לא. בדקו את רמת הסרוטונין בדם בשתי הקבוצות: בקבוצת השוקולד ל-18 אנשים הייתה רמת סרטונין הנחשבת לגבוהה. בקבוצה שלא קבלה שוקולד ל-11 הייתה רמת סרטונין הנחשבת לגבוהה.

א. מהי רמת המובהקות המינימלית לדחיית השערת האפס?

ב. מה המסקנה ברמת מובהקות של 5%?

תשובות סופיות

1) נדחה H_0.

2) א. 1 ב. לא נדחה H_0.

© כל הזכויות שמורות לברק קנדל