**מדדי קישור**

**חיפוש הקשר בין שני משתנים**

נעשה שימוש במדדי קשר רק כאשר לפחות אחד מהמשתנים הינם בעלי סולם שמי.

תזכורת: סולם שמי – מילולי (זהות) , סולם סדר – דירוג , סולמות רווחים ומנה – כמותי (מספרי)

ישנם 5 מדדי קשר.

* מדד "למדה" - 
* מדד "קרמר" – rc
* מדד "ספירמן" – rs
* מדד קשר לינארי "פירסון" – r
* מדד קשר לא לינארי "אתה" - 

**מדד למדה:**

ניתן לקבל ערך שנע בין 0-1, כאשר 1- קשר חזק , 0- קשר חלש. ( )

הנוסחאות האפשריות ומשמעותם:

Ly = סך כל ההפסד (הטעויות) בניבוי Y בלי ה X.

Lx = סך כל ההפסד (הטעויות) בניבוי X בלי ה Y.

Ly/x = סך כל ההפסד (הטעויות) בניבוי Y עם ה X.

Lx/y = סך כל ההפסד (הטעויות) בניבוי X עם ה Y.

חישוב נתונים מטבלה נעשה ב5 שלבים פעם אחת ל **X** ופעם אחת ל **Y** :

* תמיד נוודא קודם שמדובר בסולם שמי
* אם לא סימנו עבורינו את x וy אנו נבחר כרצוננו.
1. נחפש את השכיח הרלוונטי.
2. חישוב **Ly** על ידי החסרת השכיח מהסך הכולל של הנתונים.
3. נעשה שימוש ב X/Y שלא עליהם נעשה הנוסחה על ידי החסרת שכיחות מהשורה שלהם בטבלה (כלומר ניקח כל פרמטר ופרמטר ונקח את השכיח ונחסיר מהסך הכולל של השורה).
4. חישוב נסכם את כל מרווח הטעויות שמצאנו ב 3. (כל מרווח הטעויות של כל הפרמטרים השונים)
5. נציב בנוסחה.

דוגמה 1:

מצא את עוצמת הקשר בין רמת ההשכלה לרמת ההכנסה על סמך הנתונים הבאים:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| סה"כ | גבוהה | בינונית | נמוכה | YהכנסהX/ השכלה |
| 154 | 5 | 40 | 100 | יסודי |
| 225 | 20 | 160 | 45 | תיכונית |
| 130 | 25 | 100 | 5 | אקדמאית |
| 500 איש | 50 | 300 | 150 | סה"כ |

 :

1. תיכונית – 225
2. (500-225) 275 - Lx
3. בשימוש ב Y:
* נמוכה :50 ( 150-100 )
* בינונית :140 (300-160)
* גבוהה :25 (50-25)
1. Lx/y - 215 (140+50+25)

תוצאה = 0.21 : עוצמת קשר נמוכה.

 :

1. בינונית – 300
2. (500-300) 200 - Ly
3. בשימוש ב Y:
* יסודית :45 ( 145-100 )
* תיכונית :65 (225-160)
* אקדמאית :30 (130-100)
1. Ly/x - 140 (45+65+30)

תוצאה = 0.3 : עוצמת קשר נמוכה.

דוגמה 2:

נתונה הטבלה הבאה:

האם לאור התוצאות ניתן לומר שיש קשר בין מין הבוחר לבין המפלגה שהצביע עבורה?

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| סה"כ | אחרות | דתיות | מרץ | ליכוד | עבודה |  Y מפלגה X מין |
| 300 | 20 | 40 | 40 | 160 | 40 | זכר |
| 200 | 5 | 10 | 35 | 40 | 110 | נקבה |
| 500 איש | 25 | 50 | 75 | 200 | 150 | סה"כ |

 :

1. תיכונית – 300
2. (500-300) 200 - Lx
3. בשימוש ב Y:
* עבודה :40 ( 150-110 )
* ליכוד : 40 (200-160)
* מרץ :35 (75-40)
* דתיות :10 (50-40)
* אחרות :5 (25-20)
1. Lx/y - 130 (40+40+35+5+10)

תוצאה = 0.35 : עוצמת קשר נמוכה.

 :

1. הליכוד – 200
2. (500-200) 300 - Ly
3. בשימוש ב Y:
* זכר :140 ( 300-160 )
* נקבה :90 (200-110)
1. Ly/x - 230 (140+90)

תוצאה = 0.233 : עוצמת קשר נמוכה.

**מדד קרמר:**

סימולים :

 = **O** התוצאות בניסוי. (מה שמופיע בטבלה)

 **E** = מה שאנו מצפים. (סה"כ שורה \* סה"כ טור לחלק למספר הנתונים)

= נתון שנציב בנוסחה ונמצא אותו לפי הפורמט הבא:

**∑**

 **L** = המספר הקטן ביותר מבין השורות והטורים.

(אם יש 4 שורות ו2 טורים ניקח את הטורים – כלומר (L=2

לאחר שנגלה :

1. את **E**.
2. את **O**.
3. את .
4. נציב את הנתונים בנוסחה הבאה:

*דוגמה :*

*בסקר שנערך התקבלו התוצאות הבאות:*

*האם לאור התוצאות ניתן לומר שיש קשר בין מין הילד לבין רמת האינטלגנציה שלו?*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *סה"כ* | *140+* | *120-139* | *100-119* | *100-* |  *IQ y | x מין* |
| *40* | *8**E=5.6* | *10**E=12* | *14**E=16.4* | *8**E=6* | *בנים* |
| *60* | *6**E=8.4* | *20**E=18* | *27**E=24.6* | *7**E=9* | *בנות* |
| *100* | *14* | *30* | *41* | *15* | *סה"כ* |

 ****

*****(0.71+0.44+0.35+0.23+0.22+1.02+0.68) =* ***3.97***

נציב את הנתונים בנוסחה:

אין קשר בין מין הילד לרמת האינטלגנציה.

*דוגמה לשאלה מלאה הכוללת את מדד למדה וקרמר:*

*במחקר שנערך התקבלו התוצאות הבאות: חשב בעזרת כל מדדי הקשר המתאימים.*

*האם ניתן לומר שיש קשר בין הקורס שלמד הסטודנט לבין מידת ההנאה מהקורס? (משתנים שמים)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *סה"כ* | *הנאה מרובה* | *הנאה בינונית* | *הנאה מועטה* | *yהנאה x קורס* |
| *13* | *5* | *5* | *3* | *מתמטיקה* |
| *27* | *15* | *6* | *6* | *סטיסטיקה* |
| *60* | *30* | *9* | *21* | *אנגלית* |
| *100* | *50* | *20* | *30* | *סה"כ* |

*למדה:*

 :

1. אנגלית – 60
2. (100-60) 40 - Lx
3. בשימוש ב Y:

הנאה מועטה : 20 (50-30)

הנאה בינונית : 11 (20-9)

הנאה מרובה : 9 (30-21)

1. Lx/y : 40 (20+11+9)

תוצאה = **0** : אין קשר!.

קרמר:

 :

1. הנאה מרובה – 50
2. (100-50) 50 - Ly
3. בשימוש ב Y:

מתמטיקה : 8 (13-5)

סטטיסטיקה : 12 (27-15)

אנגלית : 30 (60-30)

1. Ly/x : 50 (8+12+30)

תוצאה = **0** : אין קשר!.

****

*****(0.20+0.54+0..5+2.21+0.06+0.75++0.34+0.160) =* ***4.76***

***מדד ספירמן:***

*נשתמש במדד ספירמן כאשר אחד מהמשתנים בסולם סדר (דירוג) והשני מסולם סדר או מנה.*

*(לא יכול להיות שמי)*

*המדד תמיד יראה בפורמט זה:*

|  |  |
| --- | --- |
| *מס האיברים*  | *X* |
|  | *Y* |

*שלבי חישוב :*

1. *סידור הדרגות.*

***Rx*** *– x דרגות המשתנה* ***Ry*** *– y דרגות המשתנה*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *10* | *72* | *65* | *72* | *50* | *X* |
| *1* | *4.5.* | *3* | *4.5* | *2* | *RX* |

*בכל מקרה כמו זה שיש 2 מספרים זהים ניקח את המקומות בהם הם אמורים לעמוד (מקרה זה 4,5) ונציב את ממוצע הדרגות העוקבות..*

1. *: נבצע החסרה בין שורה עליונה לתחתונה.*

 *(אין משמעות לאיזה שורה מאחר ומעלים בריבוע בשלב 3)*

1. *: מעלים את כל תוצאות ההחסרה בריבוע.*
2. *: מחברים את כל התוצאות.*
3. *מציבים בנוסחה:*
* *ניתן לקבל תוצאות rs = 1 , rs = -1 rs=0*

*-1 : קשר שלילי חזק.*

*0 : אין קשר.*

1. *: קשר חיובי חזק.*
* *במידה ויש שימוש במילים כגון ציונים : מספיק, כמעט טוב ,טוב , טוב מאוד ,*

*נעשה שימוש רגיל למרות שיש מילים כי זה סולם סדר ולא שמי מאחר ויש למילים משמעות מספרית של סדר ודירוג ואין משמעות לשם עצמו.*

*דוגמה 1 :*

*חשב בעזרת מדד ספירמן את הקשר בין הציון בחשבון לבין הציון באינטלגנציה על סמך התוצאות הטובות שהתקבלו אצל 7 תלמידים.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *ציון חשבון* | *85* | *56* | *91* | *70* | *88* | *71* | *68* |
| *ציון אינטלגנציה* | *129* | *109* | *118* | *102* | *127* | *116* | *110* |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Rx | 5 | 1 | 7 | 3 | 6 | 4 | 2 |
| Ry | 7 | 2 | 5 | 1 | 6 | 4 | 3 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| di | -2 | -1 | 2 | 2 | 0 | 0 | -1 |
|  | 4 | 1 | 4 | 4 | 0 | 0 | 1 |

*דוגמה 2 :*

*מנסים לבדוק האם יש קשר בין מחיר שטיחים ומידת איכותם , את איכות השטיחים מדדו בסולם*

 *1-10 כאשר 1 בעל איכות נמוכה ו10 בעל איכות גבוהה.. חשב את מדד הקשר המתאים .*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  *איכות - X*  | *6* | *10* | *4* | *6* | *3* | *8* |
|  *מחיר- Y*  | *3200* | *4000* | *1600* | *2800* | *1000* | *3000* |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Rx | 3.5 | 6 | 2 | 3.5 | 1 | 5 |
| Ry | 5 | 6 | 2 | 3 | 1 | 4 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| di | -1.5 | 0 | 0 | 0.5 | 0 | 1 |
|  | 2.25 | 0 | 0 | 0.25 | 0 | 1 |

***מדד פירסון:***

*נשתמש במדד פירסון כאשר שני המשתנים בסולם מנה.*

*המדד תמיד יראה בפורמט זה:*

|  |  |
| --- | --- |
| *מס האיברים*  | *X* |
|  | *Y* |

*או לחלופין נקבל חלקי נוסחאות (על בסיס המידע החלקי נדרש להשלים את החסר)*

ניתן לקבל ערך שנע בין 1-1-, כאשר 1 קשר חזק , 1- קשר חלש. ( )

*שלבי חישוב:*

*לאחר שוידינו שמדובר בסולם מנה ,*

1. *נמצא את ה COV(x\*y) על ידי שימוש בנתונים : \*==COV(x\*y).*
2. *כדי שנוכל להציב בנוסחה למציאת המקדם הלינארי יש למצוא את הנתונים הבאים:*
3. *ממוצע X. *
4. *סטיית תקן ל X. *
5. *ממוצע Y. *
6. *סטיית תקן לy . *
7. *ממוצע של Y\*X  (הוספת כל הנתונים של X ו Yוחלוקה בN)*
8. *לאחר שמצאנו את כל הנתונים נציב אותם בנוסחה:*

*לאחר מציאת המקדם הלינארי אנו מוצאים את סוג הקשר (שלילי או חיובי)*

*נמשיך כדי למצוא את קו הניבוי ל X ו Y :*

*נמצא אתA B :*

*קודם נמצא את B ולאחר מכן נמצא את הA :*

1. *קו ניבוי : Y () לאחר שימוש בנוסחאות מטה:*

******

1. *קו ניבוי X : () לאחר שימוש בנוסחאות מטה:*

******

1. *מציאת שונות הטעויות :*

*נוסחה למציאת שונות הטעויות:*

**

*במידה וחסר לנו אחד x/y נהיה חייבים להשתמש דווקא בנוסחה השניה עם החלפת הy ב.x*

1. *מציאת שימוש הנתונים :*

* בדף הנוסחאות מופיע כך : *

*במציאת ה X נשתמש ב X במקום Y.*

*דוגמה 1:*

*חוקר החליט לבדוק את הקשר בין X - ציון מתמטיקה לבין Y – ציון בסטטיסטיקה.*

*במדגם 5 תלמידים , התקבלו התוצאות הבאות:*

|  |  |  |
| --- | --- | --- |
| *X\*Y* | *Y - סטטיסטיקה* | *X - מתמטיקה* |
| *30* | *6* | *5* |
| *35* | *7* | *5* |
| *36* | *6* | *6* |
| *48* | *8* | *6* |
| *64* | *8* | *8* |

1. *חשב את עוצמת הקשר הלינארי בין X-.Y*
2. *מהו קו הניבוי בציון הסטטיסטיקה (Y) לפי ציון המתמטיקה(X)?*
3. *אם תלמיד קיבל במתמטיקה 7 מה יהיה הניבוי לציון בסטטיסטיקה?*
4. *אם תלמיד קיבל בסטטיסטיקה 6.5 מה יהיה הניבוי לציון במתמטיקה?*
5. *מהי שונות הניבויים בניבוי הציון בסטטיסטיקה לפי הציון במתמטיקה?*
6. *מהי שונות הטעויות בניבוי הציון בסטטיסטיקה לפי הציון במתמטיקה?*
7. *מהי שונות הטעויות בניבוי הציון במתמטיקה לפי הציון בסטטיסטיקה?*

*לפני הכל נחשב את 5 המדדים הבסיסיים:*

*1.ממוצע X : *

*2.סטיית תקן ל X: *

*3.ממוצע Y: *

*4.סטיית תקן ל Y: *

*. : X\*Y. 5*

*א. לאחר שמצאנו את המדדים הבסיסיים אפשר להתחיל בהצבה בנוסחאות ופתרון.*

 *42.6-6\*7 ->0.6=COV(x\*y).*

*ב. קו ניבוי לY.*

*** 0.5=***

*(**)*

 ג. מציבים את הנתון – 7.



 ד. חישוב דומה רק של הX כולל הצבת נתון:

*** 0.751=***

*(**)*



 ה. חישוב שונות הניבויים לסטטיסטיקה Y:



 ו. שונות הטעויות של Y.

נשתמש בשיטה הבאה:

1-

ז. שונות טעויות בניבוי X:

 מאחר ועוד לא מצאנו את הניבוי של X.

1-

דוגמה 2.

דוגמה בה מוצאים מדדים רק עם חלק מהנתונים:

במחקר לבדיקת הקשר בין X – הציון הממוצע בבגרות לבין Y- הציון במבחן הפסיכומטרי התקבלו התוצאות הבאות במדדים של 50 תלמידים.

1. האם קיים קשר בין X-Y?
2. אם ממוצע הבגרות של שמעון היה 9.5 מה יהיה הניבוי לציון הפסיכומטרי?
3. מהי שונות הטעויות בניבוי ציון מבחן הפסיכומטרי לפי ציון הממוצע בבגרות?

נתונים לנו:

   



נמצא את המדדים הבסיסיים על בסיס הנתונים:

*1.ממוצע X : *

*2.סטיית תקן ל X: *

*3.ממוצע Y: *

*4.סטיית תקן ל Y: *

*. : X\*Y. 5*

*מציאת cov: *

לאחר שמצאנו את הנתונים נציב בנוסחה:

1. הצבה בנוסחה.

ב. מציאת קו ניבוי לY והצבת הנתון שקיבלנו (9.5).

*** 51=***

*(**)*



1. שונות טעויות בניבוי Y.

(1-

דוגמה 3: במחקר שבדק את הקשר בין X גיל הנישואין לבין Y מספר הילדים התקבלו התוצאות הבאות במדגם של 10 נשים:

1. מהו הממוצע של גיל הנישואין של האשה? (ממוצע X)
2. חשב את עוצמת הקשר הליניארי בין X-Y.

נתונים הפרטים הבאים:

   

בנוסף נתון לנו שניבוי Y לפי X : 

כך ש A = 21 B = 0.75-

חסר לנו פרטים רבים שנצטרך להשלים :

נתחיל עם מה שיש:

*1.ממוצע X : ?*

*2.סטיית תקן ל X: ?*

*3.ממוצע Y: *

*4.סטיית תקן ל Y: *

*. : X\*Y. 5*

1. ננסה למצוא את הX על ידי הצבה של הנתונים שיש לנו בנוסחה למציאת A.

****** נעביר אגפים : x0.75=21-3.

  =24

1. חסר לנו את שונות ה X כדי שנוכל לחשב את **:

נמצא את השונות: **

נציב את הנתונים בנוסחה למציאת B :

**מדד אתה**

מקבל ערכים שנעים בין 0-1 .

מציאת האתה מבוסס על ממוצע והשונות , לא חובה להוציא בריבוע למרות שאתה זה השורש של התוצאה..

מדד זה דומה למדד למדה:

נוסחאות:

דוגמה לשימוש בנוסחה:

בכיתה בת 20 תלמידים נבדקו הציונים באנגלית – X , ובלשון – Y , התקבלו התוצאות הבאות : חשב את מדדי אתה המתאימים.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| סה"כ | 9 | 8 | 7 | Yלשון| Xאנגלית |
| 12 | 3 | 6 | 3 | 8 |
| 8 | 1 | 4 | 3 | 9 |
| 20 | 4 | 10 | 6 | סה"כ |

נחשב את הממוצע בכדי להציב בנוסחה:

1. ממוצע Y: 
2. מציאת טעויות LY: ()



1. עם שימוש ב X :

יש לנו שתי X שורות באנגלית ונחשב כל שורה בנפרד:

 x=8  x=9

1. מציאת Ly/x מציאת ט באמצעות הX:



1. לאחר שמצאנו את כל הפרטים נציב בנוסחה:

*לאחר שמצאנו את Y נציב גם את X:*

1. ממוצע X: 
2. מציאת טעויות LX: ()



1. עם שימוש ב y :

יש לנו שלוש שורות לשון ונחשב כל שורה בנפרד:

y=7   8y= y=9

1. מציאת Lx/y מציאת x באמצעות הy:



1. לאחר שמצאנו את כל הפרטים נציב בנוסחה:

***סיכום מדדים***

|  |  |  |  |
| --- | --- | --- | --- |
| *מדד* | *מתי משתמשים?* | *ערך* | *יראה כך:* |
| *למדה* | *משתנה אחד לפחות מסולם שמי.* | *0-1* | *משולשים בפינה השמאלית למעלה* |
| *קרמר* | *משתנה אחד לפחות מסלום שמי.* | *0-1* | *משולשים בפינה השמאלית למעלה* |
| *ספירמ* | *משתנה אחד מסולם סדר דירוג והשני מסולם סדר או מנה.* | *מינוס 1 - 1* | *שתי שורות שוכבות* |
| *פירסון* | *כאשר שני המשתנים מסולם מנה (מספר)* | *מינוס 1 - 1* | *שתי שורות שוכבות או חלקי נוסחאות פתורות* |
| *אתה* | *כאשר שני המשתנים מסולם מנה או כאשר מתבקשים לחשב סולם אחד מתוך שניים אך הוא מנה* | *0-1* | *שתי שורות שוכבות* |