מבוא לסטטיסטיקה והסתברות א

פרק 78 - מדדי קשר- מדד הקשר אתא

תוכן העניינים

1. מדד הקשר אתא 1

מדד הקשר אתא:

רקע:

מדד הקשר אתא הינו מדד קשר אסימטרי. כלומר, יש שני מדדי קשר מסוג זה:

$\eta_{y|x}$ - אתא Y לפי X. אפשרי לחישוב כאשר Y מסולם רווחים או מנה ו- X מכל סולם.

$\eta_{x|y}$ - אתא X לפי Y. אפשרי לחישוב כאשר X מסולם רווחים או מנה ו- Y מכל סולם.

מדד הקשר לא בודק כיוון קשר אלא רק את העוצמה , לכן מקבל ערכים בין 0 ל-1.

שלבים בחישוב המדד: $\eta_{y|x}$ (ניבוי Y בהינתן X):

שלב א: נחשב את הממוצע הכללי של Y שיסומן $\bar{\bar{Y}}$.
נחשב לכל ערך של X את הממוצע של Y.

שלב ב: חישוב השונות הכללית של Y, כלומר השונות של ההתפלגות השולית
של Y: $S_y^2 = \frac{\sum_y y^2 \cdot f(y)}{n} - \bar{\bar{y}}^2$.

שלב ג: חישוב שונות הניבויים של Y, כלומר השונות של הממוצעים הקבוצתיים
של Y: $S_{\bar{y}}^2 = \frac{\sum_y \bar{y}_x^2 \cdot f(x)}{n} - \bar{\bar{y}}^2$.

שלב ד: נחשב: $\eta^2_{y|x} = \frac{S^2_{\bar{y}}}{S^2_y}$.

שלב ה: נוציא שורש: $\eta_{y|x} = \sqrt{\eta^2_{y|x}}$.

דוגמה (פתרון בהקלטה):

להלן טבלת שכיחות משותפת עבור קבוצת סטודנטים, כאשר:

Y / X	1	2	3
גבר	5	5	5
אישה	10	5	0

X - מגדר.

Y - מספר הקורסים שנרשם הסטודנט הסמסטר.

א. איזה מדד קשר של אתא ניתן לחשב כאן?

ב. חשבו את המדד האפשרי.

© כל הזכויות שמורות לברק קנדל

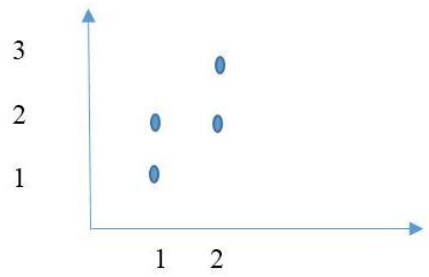
שאלות:

1) להלן נתונים על משפחות בישראל. בדקו לגבי כל משפחה באיזה אזור היא גרה בארץ - Y, וכמו כן כמה מקלטי טלוויזיה יש להם בבית - X.
להלן התוצאות שהתקבלו:

אזור/ מספר טלוויזיות	מרכז	צפון	דרום
0	5	15	0
1	5	5	20
2	5	0	0

א. איזה מדד קשר של אתא ניתן לחשב?
ב. חשבו את מדד הקשר של אתא שניתן לחישוב.

2) להלן דיאגרמת פיזור של 4 תצפיות:



א. בנו טבלת שכיחות משותפת לנתונים.
ב. חשבו את $\eta_{y/x}$ ואת $\eta_{x/y}$.

© כל הזכויות שמורות לברק קנדל

תשובות סופיות:

1) א. $\eta_{x/y}$. ב. 0.585.

2) א. להלן טבלה: ב. $\eta_{x/y}=\eta_{y/x}=0.707$.

$x\backslash y$	**1**	**2**	**3**	$f(x)$
1	1	1	0	2
2	0	1	1	2
$f(y)$	1	2	1	$n=4$

© כל הזכויות שמורות לברק קנדל