

הוגנות הליך הקבלה להשכלה הגבוהה כלפי מועמדים מבקשי התאמות (בשפה העברית)

דביר קלפר

אליוט טורוול, תמר קנת-כהן, כרמל אורן

המרכז הארצי לבחינות ולהערכה

כנס האגודה הישראלית לפסיכומטריקה,

פברואר 2013

נושאים להצגה

- ☐ מטרת המחקר
- ☐ אוכלוסיית המחקר ומשתני המחקר
- ☐ עיבוד הנתונים והמודלים
 - השוואת מדגמים מזווגים באמצעות ציוני נטייה
(Propensity Score Matching)
 - המודל של קלירי ותוקף דיפרנציאלי
- ☐ תוצאות

מטרת המחקר

- ❑ לחקור את הוגנות הליך הקבלה להשכלה הגבוהה בישראל ביחס למועמדים המבקשים התאמות במבחן הכניסה הפסיכומטרי לאוניברסיטאות (מכפ"ל) כדי לפצות על לקויות שונות
- ❑ מחקר זה מתעניין לא רק באלו שקיבלו התאמות, אלא גם באלה שבקשתם סורבה, אם מסיבה טכנית (מסיבות של אי-הגשת כל המסמכים הדרושים בזמן ו/או בקשה שהוגשה מאוחר מדי) ואם על בסיס מקצועי

אוכלוסיית המחקר

- האוכלוסייה הכללית שממנה נגזרו מדגמי המחקר כוללת 124,501 רשומות של סטודנטים מכל האוניברסיטאות בישראל אשר
 - נבחנו במכפ"ל בשפה העברית החל ממועד יולי 2000 ועד תחילת לימודם
 - למדו בשנה א' במחזורים תשס"ב (2002/3) עד תשס"ט (2008/9)
 - דיווחו על המגדר ועל תאריך הלידה שלהם
 - נתונים ציוניהם בקריטריון ובכל החזאים
- סה"כ התקבלו 2,036 יחידות עיבוד, כאשר יחידת עיבוד מוגדרת כחוג לימוד במוסד ובמחזור

אוכלוסיית המחקר

□ האוכלוסייה הכללית חולקה לשש קבוצות זרות: רגילים, בעלי לקות למידה, בעלי בעיות קשב וריכוז, בעלי מגבלות פיזיות, חסרי נתונים ומסורבים מקצועית

		קבוצה	שכיחות	אחוזים
התייחסות	לא ביקשו התאמות	רגילים	120,503	96.79
	ביקשו התאמות וקיבלו	לקות למידה	958	0.77
מיון	ביקשו התאמות וקיבלו	קשב וריכוז	187	0.15
		מגבלות פיזיות	1,096	0.88
	ביקשו התאמות ולא קיבלו	"חסרי נתונים"	299	0.24
דיון		מסורבים מקצועית	1,458	1.17

משתני המחקר

❑ קריטריון: ציון שנה א'

❑ חזאים:

❖ ציון הבגרות

❖ הציון הפסיכומטרי

▪ חשיבה מילולית

▪ חשיבה כמותית

▪ אנגלית

❖ ציון הסכם

סטטיסטיקה תיאורית

קבוצה						מדד	משתנה
מסורבים מקצועית	"חסרי נתונים"	מגבלות פיזיות	קשב ורכוז	לקות למידה	רגילים		
81.8	81.8	80.0	80.8	80.7	82.2	ממוצע	ציון שנה א'
9.8	9.3	10.7	10.6	9.8	9.6	סטיית תקן	
96.1	94.3	98.6	94.3	94.2	99.0	ממוצע	בגרות
8.6	8.9	8.3	7.3	7.9	8.8	סטיית תקן	
585.9	547.9	632.2	613.7	593.2	630.5	ממוצע	מכפ"ל
87.6	97.3	87.0	83.0	83.8	85.0	סטיית תקן	
54.6	51.4	58.8	55.2	54.1	59.1	ממוצע	סכמ
9.4	10.2	9.4	8.5	8.7	9.4	סטיית תקן	
114.4	107.0	123.6	120.1	116.8	122.4	ממוצע	מילולי
16.7	18.3	16.1	14.8	15.7	16.4	סטיית תקן	
115.5	109.4	122.9	119.5	116.6	122.6	ממוצע	כמותי
17.7	19.7	17.6	18.1	20.4	17.0	סטיית תקן	
115.3	109.1	122.9	120.7	114.7	124.6	ממוצע	אנגלית
20.9	22.9	20.6	20.4	21.0	19.0	סטיית תקן	

עיבוד הנתונים: השוואת ציון נטייה

- ❑ כל אחת מחמש קבוצות המוקד הושוותה בנפרד לקבוצת הרגילים המהווה את קבוצת ההתייחסות
- ❑ לכל אדם מקבוצת המוקד הותאם אדם מקבוצת ההתייחסות שלומד איתו באותה יחידת העיבוד בשיטת ה – Propensity Score Matching (ראה טבלה בהמשך)
- ❑ הזיווג נעשה על-פי מגדר וגיל שעבורם מספר הנתונים החסרים היה קטן יחסית

ציון נטייה

□ במחקר תצפיתי הקצאת נבדקים לקבוצת ההתייחסות ולמוקד איננה מקרית ולכן יכולה להוביל לטעות בפרשנות התוצאות

□ מטרת ציון הנטייה לספק מספר אחד (סקלאר) במקום מספר משתני רקע שבעזרתו יהיה ניתן להתאים בין הקבוצות ולדמות הקצאה מקרית

באופן לא פורמלי: ציון נטייה הוא הסיכוי של נבדק לקבל טיפול (במקרה שלנו טיפול=מתן/סירוב להתאמות מיוחדות) בהינתן רק משתני הרקע שלו

באופן פורמלי: ציון נטייה למשתתף i מוגדר כהסתברות המותנית לקבלת טיפול מסויים ($T_i = 1$) לעומת אי קבלתו ($T_i = 0$) בהינתן וקטור משתני רקע $\vec{x}_i$

$$e(\vec{x}_i) = P(T_i = 1 | \vec{X}_i = \vec{x}_i)$$

כאשר ההנחה היא שבהינתן ה- X -ים, ה- T -ים בלתי תלויים

$$P(T_1 = t_1, \dots, T_N = t_N | \vec{X}_1 = \vec{x}_1, \dots, \vec{X}_N = \vec{x}_N) = \prod_{i=1}^N e(\vec{x}_i)^{t_i} \cdot (1 - e(\vec{x}_i))^{1-t_i}$$

ציון נטייה - המשך

מהגדרה זו נובע כי T ו- X בלתי תלויים בהינתן $e(x)$; ולכן, אנשים בקבוצת המוקד וההתייחסות בעלי אותו ציון (או ציון קרוב) יטו להיות בעלי התפלגות דומה במשתני הרקע

- לצורך הערכת ציון הנטייה אנו משתמשים ברגרסיה לוגיסטית
- על ציון הנטייה אנו מפעילים אלגוריתמים של השוואה (מיון אקראי של הנבדקים בכל קבוצה, ביצוע מכפלה קרטזית מותנית בין הקבוצות, חישוב המרחק האוקלידי בין הזוגות ובחירת הזוג בעל המרחק המינימלי)

השוואה לפני ואחרי הזיווג

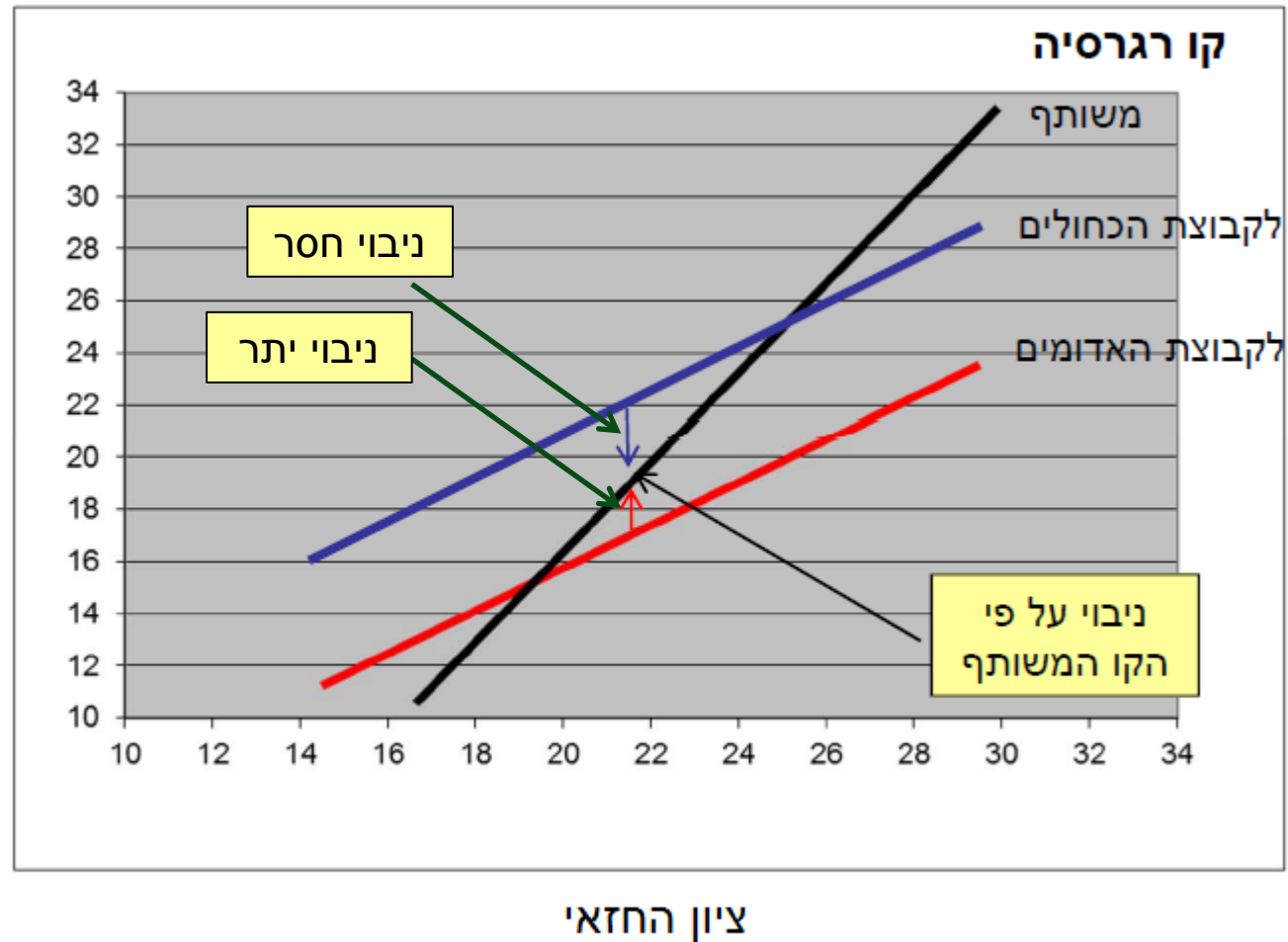
לאחר הדיווג לפי ציוני נטייה					לפני הדיווג לפי ציוני נטייה					קבוצות
גיל		מין		אנשים	גיל		מין		אנשים	
ממוצע	סטיית תקן	אחוז גברים	אחוז נשים		ממוצע	סטיית תקן	אחוז גברים	אחוז נשים		
22.51	2.09	49	51	913	21.62	2.53	40	60	59,283	רגילים
22.58	2.07	49	51	913	22.58	2.07	49	51	913	לקות למידה
22.86	2.16	61	39	182	21.64	2.59	40	60	23,571	רגילים
22.91	2.32	61	39	182	22.91	2.32	61	39	182	קשב וריכוז
21.96	2.07	53	47	1,046	21.51	2.51	41	59	64,930	רגילים
22.02	2.32	53	47	1,046	22.02	2.32	53	47	1,046	מגבלות פיזיות
22.03	2.08	43	57	1,398	21.56	2.49	41	59	70,615	רגילים
22.06	2.05	43	57	1,398	22.06	2.05	43	57	1,398	"חסרי נתונים"
22.86	2.16	61	39	182	21.64	2.59	40	60	23,571	רגילים
22.91	2.32	61	39	182	22.91	2.32	61	39	182	מסורבים מקצועית

הטיה בברירה: המודל של קלירי

- יש לקשר את בדיקת ההוגנות של כלי המיון עם הקריטריון שאותו הוא מיועד לנבא – הצלחה בלימודים במוסדות להשכלה גבוהה
- מודל זה מגדיר הטיה בברירה כמצב שבו שימוש בקן רגרסיה משותף לניבוי הקריטריון על ידי החזאי גורר ניבוי יתר או חסר של הקריטריון
- ביישום המודל מחושבת משוואת רגרסיה לניבוי הקריטריון מעבר לשתי הקבוצות. עבור כל קבוצת מוקד מחושבת הסטייה הממוצעת בין ציון הקריטריון המנובא על פי משוואת הרגרסיה המשותפת לבין הציון בפועל (ציון מנובא פחות ציון בפועל)

הטיה בברירה: המודל של קליר

ציוני הקריטריון



תוקף דיפרנציאלי: מקדם המתאם של פירסון

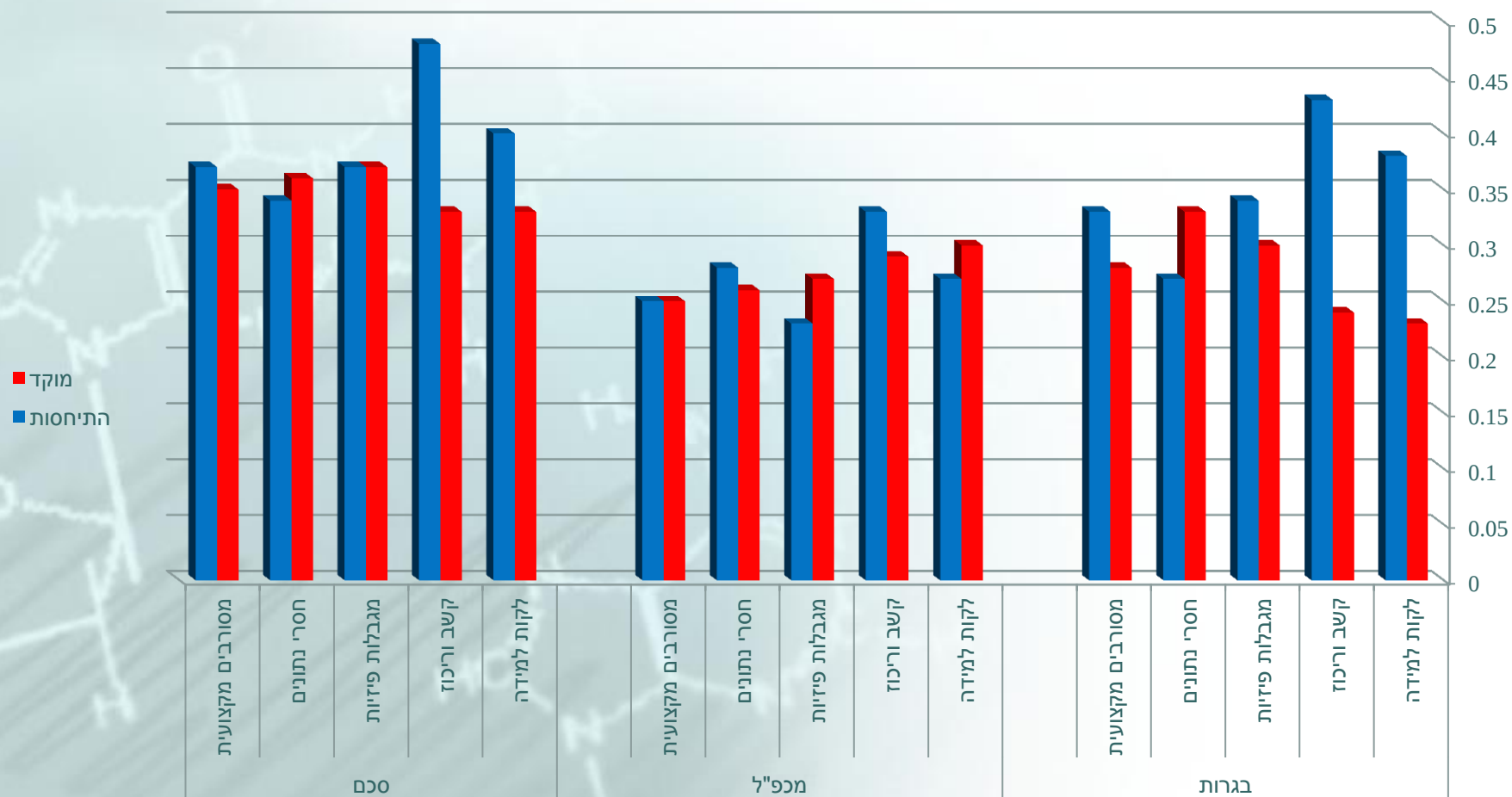
- עבור כל זוג קבוצות המזווגות חושב (בנפרד – הן עבור קבוצת המוקד והן עבור קבוצת ההתייחסות המתאימה) מקדם המתאם של פירסון בין כל אחד מהחזאים והקריטריון
- לא בוצע תיקון לקיצוץ תחום, בעיקר בגלל גדלי קבוצות קטנות במוקד. תוצאת תוקף הניבוי שחושבה משמשת לצרכי השוואה בלבד ולא כדי להעריך את תוקף הניבוי "האמיתי" של מערכת המיון

תוצאות: קלירי

חזאי						קבוצה
אנגלית	כמותי	מילולי	סכמ	מכפ"ל	בגרות	
0.04	0.05	0.04	0.01	0.02	0.03	לקות למידה
0.03	0.02	0.03	0.00	0.01	0.02	קשב וריכוז
0.14	0.14	0.14	0.14	0.14	0.15	מגבלות פיזיות
-0.05	-0.06	-0.08	-0.09	-0.10	-0.03	"חסרי נתונים"
0.02	0.01	0.00	-0.02	-0.02	0.03	מסורבים מקצועית

במונחי סטיות תקן של הקריטריון; פער בין ציון מנובא לציון בפועל – ערך חיובי מבטא הטיה לטובת קבוצת המוקד וערך שלילי הטיה לרעתה

תוצאות: מקדם תוקף - חזאים ראשיים



שאלות?

תודה!