



מרכז ארצי לבחינות ולהערכה (ע"ר)
NATIONAL INSTITUTE FOR TESTING & EVALUATION
المركز القطري للامتحانات والتقييم
מיסודן של האוניברסיטאות בישראל

מידת הדיוק של תיקון לקיצוץ תחום
עבור בררה חד-משתנית ועבור בררה רב-משתנית
בהעדר מידע על משקלות משתני הבררה
ועל אוכלוסיית המועמדים

דוח מחקר

RR
17-03

פברואר 2017

תמר קנת-כהן | דביר קלפר

דוח מחקר

RR-17-03

ISBN:978-965-502-206-3

מידת הדיוק של תיקון לקיצוץ תחום
עבור בררה חד-משתנית ועבור בררה רב-משתנית
בהעדר מידע על משקלות משתני הבררה
ועל אוכלוסיית המועמדים

תמר קנת-כהן
דביר קלפר

פברואר 2017

עמוד

תוכן עניינים

1.	תקציר	2
2.	מבוא	3
3.	חלק א' – שיטה	6
4.	חלק א' – תוצאות	10
5.	חלק ב' – שיטה	14
6.	חלק ב' – תוצאות	14
7.	סיכום ודיון	18
8.	מקורות	21
9.	נספחים	22

עמוד

לוחות

1:	המידע שנדרש בתיקון לקיצוץ תחום של המתאם בין הקריטריון, y , לחזאי, x_1	4
2:	הנסיבות שבהן נעשתה השוואה בין שני תיקונים לקיצוץ תחום	5
3:	פרמטרים של נבחני מכפ"ל בערבית ושל נבחני מכפ"ל בשפות אחרות בקרב המועמדים	7
4:	שכיחות נבחני מכפ"ל בערבית ונבחני מכפ"ל בשפות אחרות בקרב המועמדים	7
5:	פרמטרים של שתי הקבוצות בהדמיית אוכלוסיות המועמדים	8
6:	השגיאה (באחוזים) באמידת המתאם בין x_1 ו- y ובין x_2 ו- y ע"י תיקון לקיצוץ תחום בשתי שיטות, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים	11
7:	השוואה בין שתי שיטות לתיקון לקיצוץ תחום – ממצאים ממחקר קודם ונוכחי	19
8:	השגיאה (באחוזים) באמידת המתאם בין x_1 ו- y ע"י תיקון חד-משתני לקיצוץ תחום, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)	24
9:	השגיאה (באחוזים) באמידת המתאם בין x_1 ו- y ע"י תיקון רב-משתני לקיצוץ תחום, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)	25
10:	השגיאה (באחוזים) באמידת המתאם בין x_2 ו- y ע"י תיקון חד-משתני לקיצוץ תחום, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)	26
11:	השגיאה (באחוזים) באמידת המתאם בין x_2 ו- y ע"י תיקון רב-משתני לקיצוץ תחום, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)	27

עמוד

תרשימים

1:	השגיאה (באחוזים) באמידת המתאם בין x_1 ו- y ובין x_2 ו- y ע"י תיקון לקיצוץ תחום בשתי שיטות, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים	12
2:	השגיאה (באחוזים) באמידת המתאם בין x_1 ו- y ובין x_2 ו- y ע"י תיקון לקיצוץ תחום בשתי שיטות, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)	15

1. תקציר

במחקרי תוקף ניבוי שנערכים במאל"ו משתמשים כיום בתיקון לקיצוץ תחום שמניח שהבררה מתבססת על משתנה יחיד – ציון סכם – שהוא צירוף של ציון מכפ"ל וממוצע הבגרות במשקלות שווים (תיקון חד-משתני). לאחרונה נאספו עדויות לכך שהנחה זו על תהליך הבררה אינה מתקיימת במלואה תמיד. לאור זאת נבדקת האפשרות לעבור לשימוש בנוסחה כללית לתיקון, שמניחה שהבררה מתבססת על צירוף כלשהו של ציון מכפ"ל וממוצע הבגרות (תיקון רב-משתני).

מחקר קודם, שהשווה, באמצעות הדמיה, בין שתי שיטות התיקון במצב שבו הפרמטרים של אוכלוסיית המועמדים (שנחוצים לתיקון) ידועים, והדבר היחידי שלא ידוע הוא משקלות משתני הבררה, מצא יתרון ברור לתיקון הרב-משתני. אלא שבנסיבות שמאפיינות מחקרי תוקף רבים לא תמיד מצוי בידי החוקר מידע מדויק על הפרמטרים של אוכלוסיית המועמדים ומתעוררת השאלה אם מסקנות המחקר הקודם תקפות גם בנסיבות אלו. המחקר הנוכחי, על שני חלקיו, בודק שאלה זו. הוא משווה, באמצעות הדמיה, בין שתי שיטות התיקון (תיקון חד-משתני ותיקון רב-משתני), במצב שבו הפרמטרים של אוכלוסיית המועמדים אינם ידועים, מה שמחייב שימוש בערכי פרמטרים שמתקבלים מאוכלוסיות מועמדים דומות. בחלקו הראשון של המחקר משקלות משתני הבררה ידועים, ובחלקו השני – לא. באופן זה מספקים שני חלקי המחקר השלמה של התמונה שהתקבלה במחקר הקודם, על ידי בדיקת מידת הדיוק של שתי שיטות התיקון כפונקציה של שני גורמים: האחד, מידת השוני בין אוכלוסיית המועמדים ה"אמתית" לבין האוכלוסייה שבפרמטרים שלה משתמשים בתיקון, והשני, מידת השוני בין ההנחה שמניח התיקון החד-משתני בנוגע למשקלות משתני הבררה לבין המשקלות ששימשו בפועל. מחקר זה, כקודמו, מתמקד במקרה הפרטי של שני חזאים ששימשו בבררה וקריטריון יחיד.

בחלק הראשון, שבו משקלות משתני הבררה היו ידועים והפרמטרים של המועמדים לא, נמצא שעבור אחד החזאים, התיקון הרב-משתני היה עדיף, ועבור החזאי השני, התיקון החד-משתני היה עדיף. ניתוח מפורט יותר של התוצאות ושל השלכותיהן המעשיות הוביל למסקנה שגם בנסיבות אלה יש יתרון מסוים לתיקון הרב-משתני. בחלק השני, שבו גם משקלות משתני הבררה וגם הפרמטרים של המועמדים לא היו ידועים, מצב שמאפיין את ההקשר שבו נערכים מחקרי תוקף הניבוי במאל"ו, נמצאה עדיפות לתיקון הרב-משתני.

הדוח כולל דיון במגבלות ההכללה של הממצאים ומציג כיוונים למחקר המשך.

2. מבוא

סוגיה רווחת בהקשר של תיקוף כלים למיון מועמדים (ללימודים, למשל), היא תיקון לקיצוץ תחום של המתאם בין חזאי (למשל, ממוצע הבגרות x_1 , או ציון מכפ"ל x_2) לבין קריטריון (למשל, ציון שנה א' y). מתאם זה, שמחושב באופן טיפוסי במדגם הלומדים, כיוון שרק לגביהם יש מידע על y , מספק אומדן חסר למתאם ההיפותטי באוכלוסיית המועמדים, בשל קיצוץ בתחום הציונים הן בחזאי והן בקריטריון. תיקון סטטיסטי, המכונה "תיקון לקיצוץ תחום", מיועד לתקן הטיה זו.

במחקר קודם (קנת-כהן, קלפר וטורול, 2016) הושוּו באמצעות הדמיה שתי שיטות לתיקון לקיצוץ תחום. השיטה הראשונה היא "תיקון עבור בררה חד-משתנית" (Gulliksen, 1987, Chap. 12), שמניחה שהבררה התבססה על משתנה יחיד כלשהו (ציון קבלה, שהוא צירוף מסוים וידוע של x_1 ו- x_2 , ו/או משתנים אחרים). השיטה השנייה היא "תיקון עבור בררה רב-משתנית" (Gulliksen, 1987, Chap. 13), שמניחה שהבררה התבססה על משתנים כלשהם (x_1, x_2 ו/או משתנים אחרים), והיא אדישה לאופן שבו הם מצורפים (למשל, מה המשקלות שלהם בציון הקבלה).

ההשוואה בין שתי שיטות התיקון, שיכנונו להלן "תיקון חד-משתני" ו"תיקון רב-משתני", נעשתה תחת מספר תנאים, ובפרט תחת שיטות שונות לבררה. כל שיטות הבררה שנבדקו התבססו על x_1 ועל x_2 , והן נבדלו זו מזו במשקל שניתן בהן ל- x_1 ול- x_2 . התיקון החד-משתני הניח שציון הקבלה התבסס על צירוף **מסוים** (וידוע) של x_1 ו- x_2 (באופן ספציפי, במשקלות שווים), ואילו התיקון הרב-משתני הניח שהבררה התבססה על צירוף **כלשהו** (לא בהכרח ידוע) של x_1 ו- x_2 . במחקר נמצא שהתיקון הרב-משתני סיפק אומדן מצוין של המתאמים תחת מגוון של שיטות בררה. לעומתו, התיקון החד-משתני תפקד טוב בעיקר כאשר ההנחה שעמדה בבסיסו בדבר שיטת הבררה הייתה נכונה. אמנם תחת תנאים מסוימים (כמו יחס בררה גבוה) התיקון החד-משתני תפקד באופן סביר גם כאשר ההנחה שעמדה בבסיסו בדבר תהליך הבררה הייתה שגויה, אך התיקון הכללי, הרב-משתני, הפגין יתרון על פני התיקון החד-משתני גם בתנאים כאלה.

בשתי שיטות התיקון דרוש מידע מאוכלוסיית המועמדים (נוסחאות התיקון מוצגות בנספח 1). מידע זה מוצג בלוח 1, עבור כל שיטת תיקון, במצב שבו אנו רוצים לאמוד את המתאם בין הקריטריון, y , לבין חזאי, x_1 . בתיקון החד-משתני אנו מניחים שהבררה התבססה על משתנה יחיד כלשהו, z . בתיקון הרב-משתני אנו מניחים שהבררה התבססה על צירוף כלשהו

של שני משתנים, x_1 ו- x_2 . הפרמטרים שנדרשים מהמועמדים מיוצגים ע"י אות גדולה (S) והסטטיסטים שנדרשים מהלומדים ע"י אות קטנה (s).

לוח 1: המידע שנדרש בתיקון לקיצוץ תחום של המתאם בין הקריטריון, y , לחזאי, x_1

תיקון רב-משתני	תיקון חד-משתני	
שונויות משתני הבררה ($S_{x_1x_1}, S_{x_2x_2}$) והשונויות המשותפת שלהם ($S_{x_1x_2}$)	שונויות משתנה הבררה (S_{zz})	פרמטרים מהמועמדים
שונויות ($s_{x_1x_1}, s_{x_2x_2}, s_{yy}$) ושונויות משותפות ($s_{x_1x_2}, s_{yx_1}, s_{yx_2}$) של המשתנים המעורבים	שונויות ($s_{zz}, s_{x_1x_1}, s_{yy}$) ושונויות משותפות ($s_{x_1z}, s_{yz}, s_{yx_1}$) של המשתנים המעורבים ¹	סטטיסטים מהלומדים

לרוב, הסטטיסטים ממדגם הלומדים זמינים לנו. זמינותם של הפרמטרים מאוכלוסיית המועמדים היא עניין מורכב יותר בדרך-כלל. המחקר הקודם (בחלקו ההדמייתי) נעשה בתנאים שבהם הפרמטרים של אוכלוסיית המועמדים, שנחוצים לתיקון, ידועים. מסקנת המחקר הייתה, כאמור, שבתנאים אלה התיקון הרב-משתני עדיף. אלא שבמחקרי תוקף רבים לא מצוי בידי החוקר מידע מדויק על הפרמטרים של אוכלוסיית המועמדים שממנה התקבל מדגם הלומדים של המחקר, והתיקון לקיצוץ תחום נאלץ להתבסס על קירוב לפרמטרים הללו. בנסיבות כאלה נשאלות השאלות: איזה תיקון רגיש יותר לאי ודאות בנוגע לפרמטרים של אוכלוסיית המועמדים, ומה חומרת הפגיעה הפוטנציאלית ממנה בהשוואה לפגיעה הפוטנציאלית מאי ודאות בנוגע למשקלות משתני הבררה. יש סיבה לצפות לכך שההשלכות של אי ודאות בנוגע לפרמטרים של אוכלוסיית המועמדים על מידת הדיוק של שתי השיטות יהיו דיפרנציאליות: מצד אחד, התיקון הרב-משתני משתמש בשלושה פרמטרים של אוכלוסיית המועמדים, בעוד שהתיקון החד-משתני נזקק רק לפרמטר אחד (ראו לוח 1), מה שעשוי לעשותו לרגיש פחות לאי ודאות בנוגע לפרמטרים הללו. מצד שני, כפי שיפורט בפרק השיטה, בתיקון החד-משתני יש לחשב את משתנה הבררה (ציון הסכם). חישוב זה מחייב תקנון מוקדם של מרכיביו, כשהתקנון מתבסס על הפרמטרים של המרכיבים בקרב המועמדים. לפיכך, בתיקון זה יש מקור פוטנציאלי נוסף לרגישות לחוסר ודאות בפרמטרים אלו, מקור שאינו רלוונטי לתיקון הרב-משתני. חלק א' של המחקר הנוכחי בודק את ההשלכות של שימוש באומדנים לא מדויקים של הפרמטרים של אוכלוסיית המועמדים במצב שבו משקלות משתני הבררה בחישוב ציון הקבלה ידועים. חלק ב' של המחקר בודק בו-זמנית את ההשלכות של אי ידיעת הפרמטרים של אוכלוסיית המועמדים ושל אי ידיעת

¹ מסיבות היסטוריות מקובל להציג את הפתרונות של התיקון החד-משתני במונחי סטיות תקן ומתאמים. כאשר משתמשים במונחים אלה, הפתרון עבור המתאם בין y ל- x_1 מחייב ידיעת 4 סטטיסטים בלבד מתוך הלומדים: סטיית התקן של משתנה הבררה והמתאמים בין שלושת המשתנים המעורבים.

המשקלות של משתני הבררה, במטרה לספק את התמונה הכללית שממנה נוכל להסיק מהו התיקון המועדף.

לוח 2 מסכם את מה שנאמר לעיל בהציגו את הנסיבות שבהן הושוו שתי שיטות התיקון במחקר הקודם ובמחקר הנוכחי:

לוח 2: הנסיבות שבהן נעשתה השוואה בין שני תיקונים לקיצוץ תחום

פרמטרים של המועמדים			
לא ידועים	ידועים		
מחקר נוכחי חלק א'	מחקר קודם	ידועים	משקלות משתני הבררה
מחקר נוכחי חלק ב'	מחקר קודם	לא ידועים	

בשני חלקי המחקר הנוכחי, התבסס התיקון החד-משתני על ההנחה שהבררה נעשתה על פי ציון קבלה שמחושב כצירוף של x_1 ו- x_2 במשקלות שווים, ואילו התיקון הרב-משתני התבסס על הנחה חלשה יותר, שהבררה נעשתה על פי צירוף כלשהו של x_1 ו- x_2 . תהליך הבררה שנעשה בפועל השתנה בין שני החלקים של המחקר. בחלק הראשון של המחקר התבססה הבררה על צירוף של x_1 ו- x_2 במשקלות שווים, כלומר הנחת התיקון החד-משתני הייתה נכונה. בחלק זה, אם כן, התקיימו תנאים אופטימליים עבור התיקון החד-משתני, מבחינת נכונות ההנחה על תהליך הבררה, וההשוואה בין שתי שיטות התיקון התמקדה בהשלכות של שימוש בקירובים לפרמטרים של אוכלוסיית המועמדים. חלק זה אפשר לבדוק איזה תיקון רגיש יותר לחוסר דיוק בפרמטרים של אוכלוסיית המועמדים במצב שבו משקלות משתני הבררה ידועים. החלק השני של המחקר היווה הכללה של החלק הראשון בכך שהשווה בין שתי שיטות התיקון ברמות שונות של חוסר דיוק בפרמטרים של אוכלוסיית המועמדים (כמו החלק הראשון), אך הפעם גם תחת שיטות שונות של בררה, כאשר ההבדלים בין השיטות התבטאו במשקלות שניתנו ל- x_1 ול- x_2 בתהליך הבררה. חלק זה אפשר לבדוק בו-זמנית את ההשלכות של שימוש בפרמטרים לא מדויקים של אוכלוסיית המועמדים, ושל אי ודאות בנוגע לאופיו של תהליך הבררה².

המחקר הנוכחי הוא מחקר הדמייתי. הגדרנו בו "אוכלוסיית מועמדים היסטורית" שהיא אוכלוסיית מועמדים שבה ידועים לנו הפרמטרים הנחוצים לתיקון. במקביל הוגדרו "אוכלוסיות מועמדים אמיתיות" שהפרמטרים שלהן נבדלים, בדרגות משתנות של הבדל, מאלה של אוכלוסיית המועמדים ההיסטורית (ואינם ידועים בשלב התיקון לקיצוץ תחום). בכל אוכלוסיית

² נדגיש שאי הוודאות נוגעת למשקלות משתני הבררה ולא להרכב המשתנים, כלומר: ההנחה החלשה יותר שמניח התיקון הרב-משתני על תהליך הבררה מתקיימת.

מועמדים אמיתית נעשה תהליך בררה שבעקבותיו התקבל מדגם לומדים. במדגם הלומדים חושבו מתאמים בין כל אחד מן החזאים, x_1 ו- x_2 , לבין הקריטריון, y . מתאמים אלו תוקנו לקיצוץ תחום בשתי שיטות התיקון – תיקון חד-משתני ותיקון רב-משתני – על פי הפרמטרים של אוכלוסיית המועמדים ההיסטורית. מידת הדיוק של שתי שיטות התיקון בתנאים אלה נבדקה על ידי הפער בין המתאם הרלוונטי באוכלוסיית המועמדים האמיתית (שידוע לנו כיוון שמדובר במחקר הדמייתי) לבין המתאם המתקון לקיצוץ תחום.

ההבדל בין הפרמטרים של אוכלוסיית המועמדים ההיסטורית לאלה של אוכלוסיות המועמדים האמיתיות תופעל על ידי כך שאוכלוסיית המועמדים ההיסטורית הוגדרה כמורכבת משתי קבוצות מועמדים בפרופורציות נתונות. אוכלוסיות המועמדים האמיתיות הוגדרו כמורכבות מאותן שתי קבוצות, אך בפרופורציות משתנות. הפרמטרים של כל אחת משתי הקבוצות היו קבועים באוכלוסיות המועמדים השונות – ההיסטורית והאמיתיות. מאחר שהפרמטרים הנחוצים לתיקון הם פרמטרים מעבר לשתי הקבוצות, אזי על ידי יצירת הבדל בפרופורציות של שתי הקבוצות נוצר הבדל בין הפרמטרים הנכונים (הפרמטרים של אוכלוסיית המועמדים האמיתית) לבין הפרמטרים שעל פיהם מתקנים (הפרמטרים של אוכלוסיית המועמדים ההיסטורית).

ההתמקדות של המחקר הנוכחי בנסיבות שבהן הפרמטרים של המועמדים אינם ידועים נובעת מהרצון לבדוק את תפקוד שתי שיטות התיקון בהקשר שבו נערכים מחקרי תוקף הניבוי במאל"ו. נכון להיום נתוני המועמדים שעומדים לרשותנו הם עבור המועמדים למחזורים תשנ"ב (1991/92) ותשנ"ג (1992/93), והפרמטרים של אוכלוסיית מועמדים זו משמשים אותנו בתיקון לקיצוץ תחום. התיקון שבו משתמשים כיום הוא התיקון החד-משתני, תחת ההנחה שהבררה התבססה על ציון סכמ, שהוא צירוף של ממוצע הבגרות וציון מכפ"ל במשקלות שווים, ואנו בודקים את האפשרות לעבור לשימוש בתיקון רב-משתני. לאור זאת, נבחרו מאפייני ההדמיות שנועדו לבדוק את תפקודן של שתי שיטות התיקון כך שידמו את ההקשר שתואר לעיל.

3. חלק א' – שיטה

3.1 הדמיה של אוכלוסיית המועמדים ההיסטורית

3.1.1 נתונים אמפיריים לצורך בחירת המאפיינים של אוכלוסיית מועמדים היסטורית

כאמור, אוכלוסיית המועמדים ההיסטורית הוגדרה כמורכבת משתי קבוצות מועמדים בפרופורציות נתונות. הפרמטרים של שתי הקבוצות והפרופורציות שלהן באוכלוסייה נקבעו על פי הנתונים של שתי קבוצות שקיימות במציאות: נבחנו מכפ"ל בערבית ונבחנו מכפ"ל בשפות אחרות (כפי שיובהר להלן, הקבוצה השנייה מורכבת בעיקר מנבחני עברית).

הפרמטרים של שתי הקבוצות: הממוצע וסטיית התקן של ממוצע הבגרות ושל ציון מכפ"ל והמתאם בין שני החזאים בתוך כל קבוצה חושבו על פי נתוני המועמדים למחזורי תשנ"ב ותשנ"ג. ערכים אלה חושבו בחוגים (חוג במוסד ובמחזור) שבהם היו לפחות 5 תצפיות מכל קבוצה³, ומוצעו מעבר לחוגים תוך שקלול במספר התצפיות בקבוצה בחוג. הממוצע וסטיית התקן של הקריטריון והמתאם בין כל אחד מהחזאים ובין הקריטריון נלקחו ממחקר קודם (קנת-כהן, טורוול ואורן, 2011).

הערכים שהתקבלו מוצגים בלוח 3.

לוח 3: פרמטרים של נבחני מכפ"ל בערבית ושל נבחני מכפ"ל בשפות אחרות בקרב המועמדים*

	ממוצעים			סטיות תקן			מתאמים		
	בגרות	מכפ"ל	ציון שנה א'	בגרות	מכפ"ל	ציון שנה א'	בגרות ומכפ"ל	בגרות וציון שנה א'	מכפ"ל וציון שנה א'
נבחני ערבית	91.2	483	72.2	9.2	64	10.6	0.40	0.33	0.32
נבחני שפות אחרות	90.9	595	82.5	8.5	66	8.5	0.41	0.48	0.41

* חישוב הממוצע וסטיית התקן של ממוצע הבגרות ושל ציון מכפ"ל והמתאם בין שני החזאים התבסס על 14,855 נבחני ערבית ו-130,771 נבחני שפות אחרות (על פי נתוני המועמדים למחזורי תשנ"ב ותשנ"ג). החישובים שבהם היה מעורב ציון א' התבססו על 8,009 נבחני ערבית ו-30,561 נבחני ערבית (קנת-כהן ואחרים, 2011). כדאי לציין, שבניגוד לכל יתר הסטטיסטים שמוצגים בלוח, הממוצע וסטיית התקן של הקריטריון משקפים נתוני **לומדים**. עובדה זו אינה קריטית, כיוון שכל מטרתם של החישובים שנעשו בשלב זה היא למצוא פרמטרים סבירים להדמיה.

הפרופורציות של שתי הקבוצות באוכלוסייה: מספר התצפיות מכל קבוצה חושב בהתבסס על תצפית אחד לאדם במוסד ובמחזור על פי נתוני המועמדים למחזורי תשנ"ב ותשנ"ג⁴. הערכים שהתקבלו מוצגים בלוח 4. נציין ש-97.5% מקבוצת נבחני שפות אחרות הם נבחני ערבית.

לוח 4: שכיחות נבחני מכפ"ל בערבית ונבחני מכפ"ל בשפות אחרות בקרב המועמדים

	מספר תצפיות	פרופורציה
נבחני ערבית	6,173	8.7%
נבחני שפות אחרות	64,522	91.3%

³ בדיקה מקבילה נעשתה עבור כל קבוצה גם בחוגים (חוג במוסד ובמחזור) שבהם היו לפחות 5 תצפיות מהקבוצה (כלומר, ללא התניה שגם מהקבוצה השנייה יהיו לפחות 5 תצפיות). הערכים שהתקבלו היו דומים.
⁴ בדיקה מקבילה נעשתה בהתבסס על מופעים חוזרים לאדם במוסד ובמחזור. הערכים שהתקבלו היו דומים.

3.1.2 הערכים שנבחרו כמאפייני אוכלוסיית המועמדים ההיסטורית

אוכלוסיית המועמדים ההיסטורית הוגדרה, אם כן, כמורכבת משתי קבוצות, בהתאמה לקבוצת נבחני מכפ"ל בערבית (קבוצה 1) ולקבוצת נבחני מכפ"ל בשפות אחרות (קבוצה 2). שלושת המשתנים המעורבים – שני "חזאים" ו"קריטריון" – נבנו ברוח הפרמטרים של ממוצע הבגרות (x_1), ציון מכפ"ל (x_2), וציון שנה א' (y). בהתאם לנתונים האמפיריים שהוצגו לעיל הוחלט שהפרמטרים של שתי הקבוצות בהדמיית אוכלוסיות המועמדים (ההיסטורית והאמיתית) יהיו לפי הנתונים שמוצגים בלוח 5 שלהלן.

לוח 5: פרמטרים של שתי הקבוצות בהדמיית אוכלוסיות המועמדים

	ממוצעים			סטיות תקן			מתאמים		
	x_1	x_2	y	x_1	x_2	y	x_2 - x_1	y - x_1	y - x_2
קבוצה 1	90	500	70	9	65	11	0.40	0.40	0.40
קבוצה 2	90	600	80	9	65	11	0.40	0.40	0.40

כמו כן הוחלט שפרופורציית קבוצה 2 באוכלוסיית המועמדים ההיסטורית תהיה 0.9.

3.1.3 הדמיה של אוכלוסיית המועמדים ההיסטורית על פי הערכים שנבחרו

חישוב ממוצעים, סטיות תקן ומתאמים של x_2 , x_1 ו- y : נדגמו 1,000,000 שלשות של משתנים – שני "חזאים" (x_2 ו- x_1) ו"קריטריון" (y) – מהתפלגות רב-נורמלית לפי הפרמטרים של קבוצה 1 ו-9,000,000 שלשות של משתנים מהתפלגות רב-נורמלית לפי הפרמטרים של קבוצה 2. שתי הקבוצות אוחדו, ובקבוצה המאוחדת חושבו תשעה פרמטרים – 3 ממוצעים (של x_2 , x_1 ו- y), 3 סטיות תקן (של x_2 , x_1 ו- y) ו-3 מתאמים (בין x_2 ו- x_1 , בין x_1 ו- y ובין x_2 ו- y).⁵ סטיות התקן של x_1 ושל x_2 והמתאם בין x_2 ו- x_1 שימשו לחישוב שונויות x_2 ו- x_1 והשונויות המשותפת שלהם, שהן הפרמטרים שנדרשים מהמועמדים לצורך התיקון הרב-משתני. סטיות התקן הללו (וכן הממוצעים של x_2 ו- x_1) שימשו גם לתקנון x_2 ו- x_1 לצורך חישוב ציון הסכם (להלן).

חישוב ציון סכם (ציון קבלה) ושונותו: נדגמו 10,000,000 שלשות של משתנים – x_2 , x_1 ו- y – מהתפלגות רב-נורמלית לפי הפרמטרים של הקבוצה המאוחדת. באוכלוסייה זו הוגדר ציון סכם כסכום של x_2 ו- x_1 במשקלות שווים (אחרי שתוקנו לפי הפרמטרים שלהם בקבוצה

⁵ כפי שמפורט להלן, מן הבחינה המעשית, לא כל תשעת הפרמטרים של אוכלוסיית המועמדים ההיסטורית נחוצים לנו. אוכלוסייה זו שימשה אותנו רק כספקית של פרמטרים לצורך התיקונים לקיצוץ תחום, ולפיכך כל הפרמטרים שמעריבים את y אינם נחוצים. בנינו אותה באופן שתואר כדי להדגיש את המקבילות שלה לאוכלוסיות המועמדים האמיתיות.

המאוחדת), וחושבה שונותו. שונות ציון הסכם היא הפרמטר שנדרש מהמועמדים לצורך התיקון החד-משתני.

3.2 הדמיה של אוכלוסיות המועמדים האמתיות

עבור כל ערך של p בין 0 ל-1 (בקפיצות של 0.01) נעשו שני החישובים הבאים: חישוב ממוצעים, סטיות תקן ומתאמים של x_1, x_2 ו- y : נדגמו $100,000 \cdot (1-p)$ שלשות של משתנים x_1, x_2 ו- y – מהתפלגות רב-נורמלית לפי הפרמטרים של קבוצה 1 ו- $100,000 \cdot p$ שלשות לפי הפרמטרים של קבוצה 2. שתי הקבוצות אוחדו, ובקבוצה המאוחדת חושבו תשעת הפרמטרים – 3 ממוצעים, 3 סטיות תקן ו-3 מתאמים, כפי שתואר לעיל לגבי אוכלוסיית המועמדים ההיסטורית.

חישוב ציון סכם "נכון": נדגמו $100,000$ שלשות של משתנים x_1, x_2 ו- y – מהתפלגות רב-נורמלית לפי הפרמטרים של הקבוצה המאוחדת – אלה אוכלוסיות המועמדים האמתיות. בכל אוכלוסיית מועמדים חושב ציון הסכם ה"נכון", כסכום של x_1 ו- x_2 במשקלות שווים, אחרי שתוקנו לפי הפרמטרים שלהם באוכלוסייה זו. זהו ציון הסכם שעל-פיו נעשה תהליך הבררה.

3.3 הגדרת מדגמי הלומדים

עבור כל אוכלוסיית מועמדים אמיתית, 50% הגבוהים בסכם הנכון היוו את מדגם הלומדים. בכל מדגם לומדים חושב ציון הסכם ה"שגוי", כסכום של x_1 ו- x_2 במשקלות שווים, אחרי שתוקנו לפי הפרמטרים שלהם באוכלוסיית המועמדים ההיסטורית. זהו המשתנה שהתיקון החד-משתני מתייחס אליו כאל משתנה הבררה (למרות שבפועל מוינו המועמדים לפי ציון הסכם הנכון).

3.4 חישוב מתאמים מתוקנים לקיצוץ תחום

חישוב מתאמים נצפים: בכל מדגם לומדים חושבו 2 מתאמים – בין x_1 ו- y , ובין x_2 ו- y . חישוב סטטיסטיים לצורך התיקונים: בכל מדגם לומדים חושבו הסטטיסטים הנחוצים לכל שיטת תיקון (ראו לוח 1). תיקון המתאמים לקיצוץ תחום: המתאמים תוקנו לקיצוץ תחום בשתי השיטות – תיקון חד-משתני ותיקון רב-משתני. כאמור, הפרמטרים מרמת המועמדים שבהם משתמשים נלקחו תמיד מאוכלוסיית המועמדים ההיסטורית (נוסחאות התיקון מוצגות בנספח 1).

3.5 חישוב השגיאה של המתאמים המתוקנים

עבור כל אחד מארבעת המתאמים (מתאם בין x_1 ו- y , ומתאם בין x_2 ו- y , כאשר כל אחד מהם מתוקן בשתי שיטות – תיקון חד-משתני ותיקון רב-משתני) שחושבו בכל מדגם, חושבה

שגיאת האמידה לפי הנוסחה: $\frac{\text{מתאם אמת} - \text{מתאם מתוקן}}{\text{מתאם אמת}} * 100 = \text{שגיאה}$,

כאשר המתאם האמתי הוא המתאם באוכלוסיית המועמדים האמתית.

על חלק זה חזרנו 100 פעמים עבור כל אוכלוסיית מועמדים אמתית (כלומר: עבור כל ערך של p), וזאת על מנת לקבל תוצאות יציבות. התוצאות המוצגות הינן ממוצעי השגיאות על פני 100 החזרות.

4. חלק א' – תוצאות

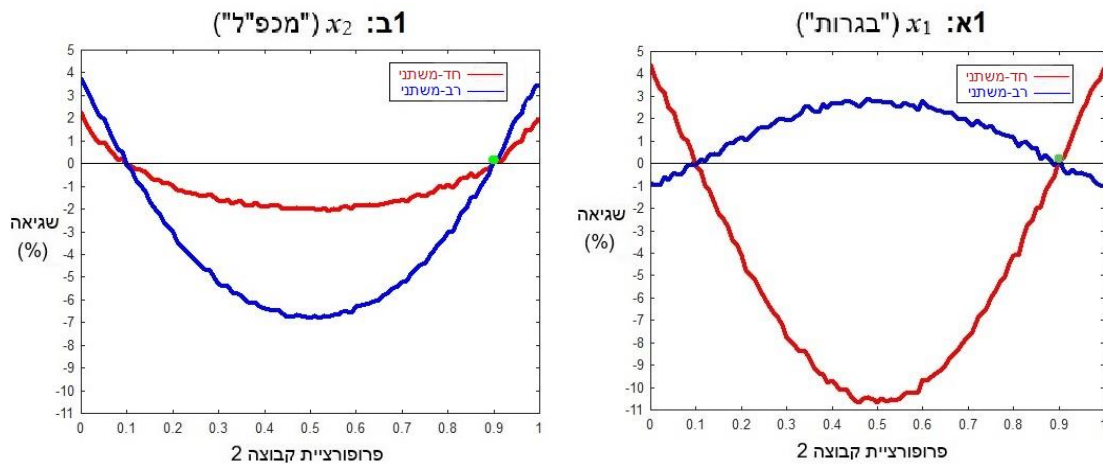
ממוצעי השגיאות של ארבעת המתאמים המתוקנים באוכלוסיות מועמדים אמתיות (11 אוכלוסיות נבחרות מתוך 101 האוכלוסיות האמתיות שהגדרנו), שנבדלות זו מזו בפרופורציית קבוצה 2, מוצגים בלוח 6. ערכים אלו מוצגים גם בתרשים 1 עבור כל האוכלוסיות האמתיות שהגדרנו.

לוח 6: השגיאה (באחוזים) * באמידת המתאם בין x_1 ו- y ובין x_2 ו- y ע"י תיקון לקיצוץ תחום בשתי שיטות, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים

x_2 ("מכפ"ל")		x_1 ("בגרות")		פרופורציית קבוצה 2
רב-משתני	חד-משתני	רב-משתני	חד-משתני	
3.76	2.23	-0.88	4.41	0.00
-0.04	-0.07	-0.10	-0.08	0.10
-3.01	-0.95	1.17	-4.09	0.20
-5.29	-1.62	1.92	-7.78	0.30
-6.38	-1.92	2.72	-9.69	0.40
-6.80	-2.00	2.74	-10.65	0.50
-6.30	-1.80	2.83	-9.64	0.60
-5.20	-1.59	1.89	-7.73	0.70
-3.02	-0.92	1.17	-4.08	0.80
0.18	0.17	0.21	0.22	0.90
3.49	1.99	-1.04	4.26	1.00

* ערך חיובי פירושו שהתיקון מספק אומדן יתר וערך שלילי פירושו שהתיקון מספק אומדן חסר של המתאם האמתי.

תרשים 1: השגיאה (באחוזים) באמידת המתאם בין x_1 ו- y ובין x_2 ו- y ע"י תיקון לקיצוץ תחום בשתי שיטות, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים



כפי שנאמר קודם, פרופורציית קבוצה 2 (שמדמה את קבוצת נבחני מכפ"ל בשפות שאינן ערבית) באוכלוסיית המועמדים ההיסטורית היא 0.9. בהתאם לכך, ניתן לראות בלוח 6 ובתרשים 1 שכאשר פרופורציית קבוצה 2 באוכלוסיית המועמדים האמתית גם היא 0.9 (או 0.1)⁶, השגיאה בתיקון המתאמים בשתי השיטות היא אפסית (כזכור, בחלק זה הנחת התיקון החד-משתני בנוגע לתהליך הבררה היא נכונה). נקודה זו מסומנת בתרשים בנקודה ירוקה.

מה קורה כאשר אוכלוסיית המועמדים האמתית הולכת ומתרחקת מאוכלוסיית המועמדים ההיסטורית (עד לנקודת הסימטריה)?

1. השגיאה באמידת המתאם הולכת וגדלה, אם כי באופן כללי ממדיה אינם גדולים מאוד: אומדן של פרמטר נחשב למצוין אם ההטיה היא בטווח של $\pm 5\%$ ולסביר אם ההטיה היא בטווח של $\pm 10\%$ (Hoogland & Boomsma, 1998). כמעט בכל המצבים, אם כן, האמידה היא סבירה, ולרוב אף מצוינת.

2. כשמשווים בין שתי שיטות התיקון, אזי מעבר לשני החזאים, השגיאה (אם נתעלם מכיוונה) גדולה יותר תחת התיקון החד-משתני, אם כי ההבדל בין השיטות אינו גדול: הטעות בתיקון הרב-משתני אינה עולה על 7%, בעוד שבתיקון החד-משתני היא עשויה להגיע ל-11%. תוחלת הערך המוחלט של הטעויות היא 2.73% בתיקון הרב-משתני ו-3.54% בתיקון החד-משתני.

3. הממצא הבולט הוא קיומה של אינטראקציה בין השיטה ובין החזאי: ביחס ל- x_1 ("בגרות") התיקון החד-משתני מספק אומדן חסר של המתאם בעוד שהתיקון הרב-

⁶ הסימטריה בגודל השגיאה סביב פרופורציה של 0.5 נובעת מן העובדה שהפרמטרים הרלוונטיים (סטיות התקן של המשתנים והמתאמים בין המשתנים) **בתוך** שתי הקבוצות הם זהים (ראו לוח 5).

משתני מספק אומדן יתר קל. ביחס ל- x_2 ("מכפ"ל") שני התיקונים מספקים אומדן חסר, כאשר השגיאה של התיקון החד-משתני היא קטנה יותר (נזכיר שההבדל בין החזאים הוא שב- x_1 ממוצעי שתי הקבוצות הם זהים ואילו ב- x_2 הממוצע של קבוצה 2 גבוה מהממוצע של קבוצה 1). מבחינת x_1 , אם כן, התיקון הרב-משתני עדיף, ואילו מבחינת x_2 התיקון החד-משתני עדיף. המסקנה החשובה שניתן לגזור מממצאים אלה היא שאין יתרון עקבי של שיטת תיקון אחת על פני האחרת כאשר נאלצים להשתמש בקירובים לפרמטרים של המועמדים.

4. חשוב להזכיר שבדיקת התפקוד של שתי השיטות נעשתה תחת הדמיה של ההקשר שבו נערכים מחקרי התוקף של מאל"ו. כדי להעריך את יכולת ההכללה של הממצאים בדקנו תנאים נוספים שבהם הכנסנו הבדלים בין קבוצה 1 לקבוצה 2 במתאמים ובסטטיסטיים נוספים (למשל, במקום שוויון בין הקבוצות בממוצע של x_1 , הגדרנו מצב שבו הממוצע של קבוצה 1 – 95 – גבוה מזה של קבוצה 2 – 85). תמונת הממצאים הייתה שונה מאוד בין התנאים, כך שלא ניתן להכליל מהממצאים שתוארו לעיל על תפקוד שתי שיטות התיקון. בנוסף, יש להביא בחשבון שהגישה שנקטנו בה כדי לדמות שינויים הולכים וגדלים בין אוכלוסיית המועמדים ההיסטורית לאוכלוסיית המועמדים האמתית – שתי קבוצות עם מאפיינים קבועים שהפרופורציה שלהן באוכלוסיית המועמדים משתנה – מהווה אילוסטרציה לשינוי אפשרי, ואינה מייצגת בהכרח את השינוי בפועל (שאינו ידוע לנו, כמובן). לפיכך, ממצאי המחקר ומסקנותיו הם מקומיים ומתייחסים להדמיה המסוימת שעשינו.

5. אשר למסקנה המעשית, בכפוף לסימני השאלה ולהסתייגויות שהוצגו לעיל, נראה שבדומה למה שמצאנו במקרה שבו הפרמטרים של המועמדים ידועים (קנת-כהן ואחרים, 2016), גם כאשר הפרמטרים הללו אינם ידועים, אין לתיקון החד-משתני יתרון כלשהו על פני התיקון הרב-משתני, אפילו במצב שבו התנאים לתפקודו הם אופטימליים (כלומר, גם כאשר ההנחה שעומדת בבסיסו בנוגע לתהליך הבררה היא נכונה). למעשה, בהקשר של נתוני מאל"ו, השגיאה הגדולה ביותר באמידה (ראו תרשים 1) היא שגיאה **שלילית** באמידת תוקף x_1 ("בגרות"), שקורית כאשר משתמשים בתיקון החד-משתני – שגיאה שעשויה להיחשב, מן הסתם, בעייתית מן הבחינה הפוליטית. בנוסף, ניתן להיווכח שהמחיר של שימוש בתיקון הרב-משתני כשאינו אופטימלי קטן מהמחיר של שימוש בתיקון החד-משתני כשאינו אופטימלי (התוספת המקסימלית לטעות היא 5% כשמשתמשים בתיקון הרב-משתני במקום בתיקון החד-משתני עבור x_2 , לעומת תוספת של 8% כשמשתמשים בתיקון החד-משתני במקום בתיקון הרב-משתני עבור x_1).

בנסיבות אלה, אין בממצאים של חלק זה מכדי לשנות את מסקנת המחקר הקודם (קנת-כהן ואחרים, 2016), שהיא שיש מקום לשקול מעבר לתיקון הרב-משתני.

5. חלק ב' – שיטה

חלק זה מהווה הכללה של חלק א' וככזה, הוא נבדל ממנו בדבר אחד: ציון הסכם ה"נכון", זה שעל-פיו נעשתה הבררה, לא הוגדר תמיד כסכום של x_1 ו- x_2 במשקלות שווים, כפי שנעשה בחלק א'. במקום זה, הגדרנו בחלק זה פרמטר w , שמבטא את המשקל של x_2 בתהליך הבררה שנעשה באוכלוסיות המועמדים האמתיות: $w=0$ פירושו שהבררה בפועל של הלומדים נעשתה אך ורק לפי x_1 , ו- $w=1$ פירושו שהבררה נעשתה אך ורק לפי x_2 . בחלק א', אם כן, נבדק המקרה שבו $w=0.5$. חשוב להדגיש שרק ציון הסכם ה"נכון" ותהליך הבררה שנעשה על פיו משתנים לפי הערך של w . חישוב הסכם ה"שגוי" כסכום של x_1 ו- x_2 במשקלות שווים (אחרי שתוקננו לפי הפרמטרים שלהם באוכלוסיית המועמדים ההיסטורית) לא השתנה.

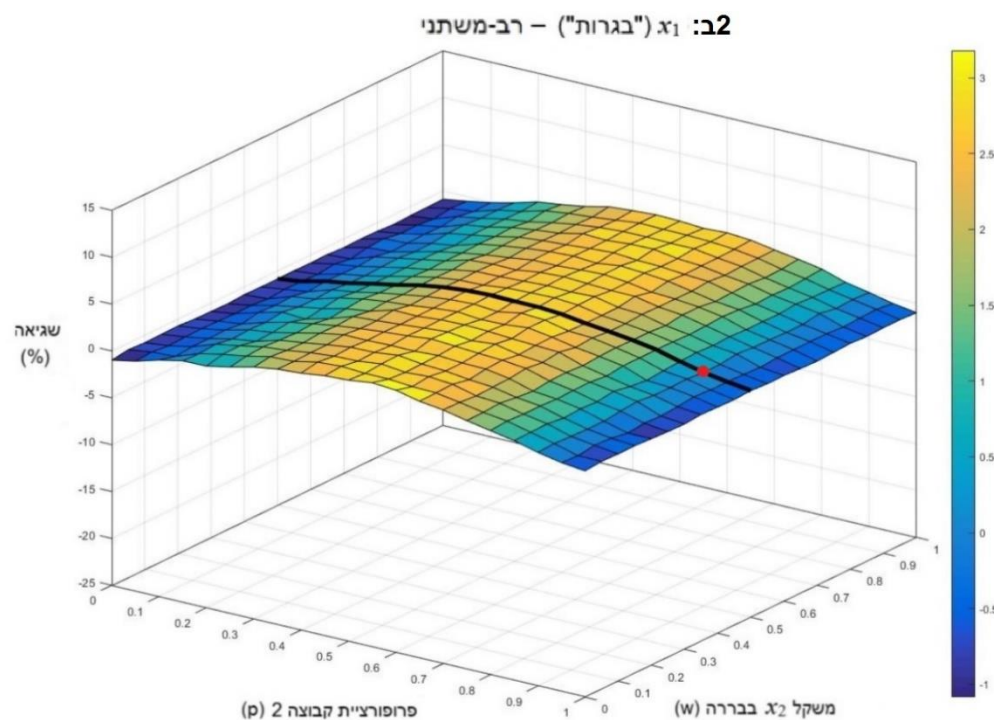
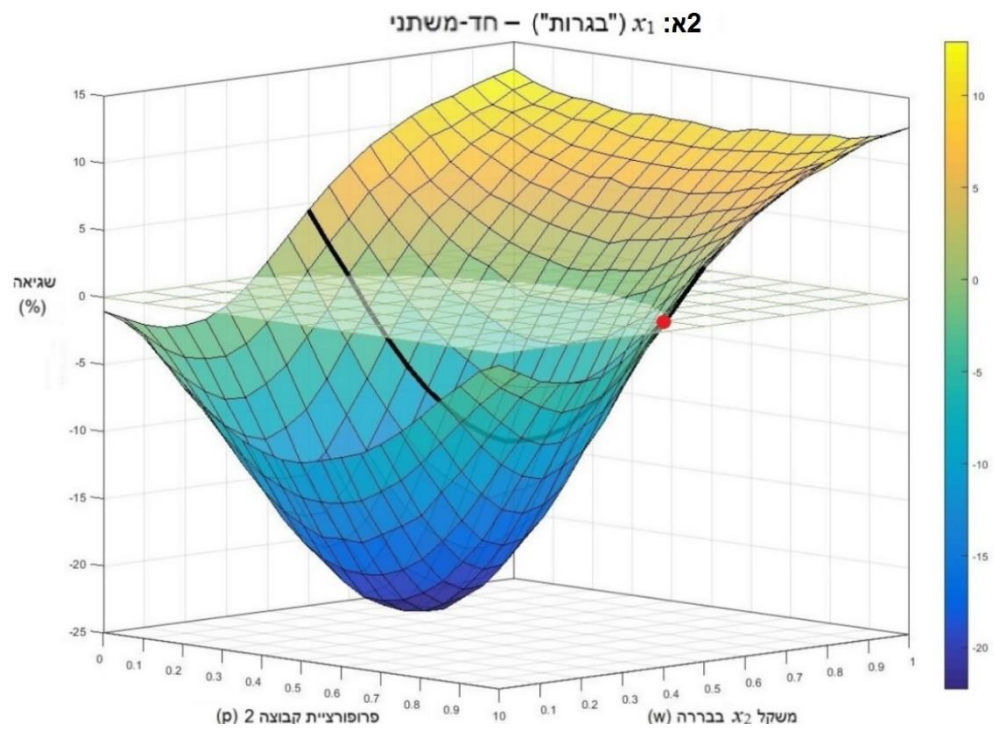
לסיכום, הגדרת אוכלוסיות המועמדים – הן ההיסטורית והן האמתיות – זהה בשני החלקים. מה ששונה הוא שחישוב ציון הסכם ה"נכון" באוכלוסיות המועמדים האמתיות והבררה שנעשתה על פיו השתנו כפונקציה של w . כל יתר העיבודים (חישוב מתאמים מתוקנים לקיצוץ תחום והשגיאה שלהם) נעשו כפי שתואר בחלק א'.

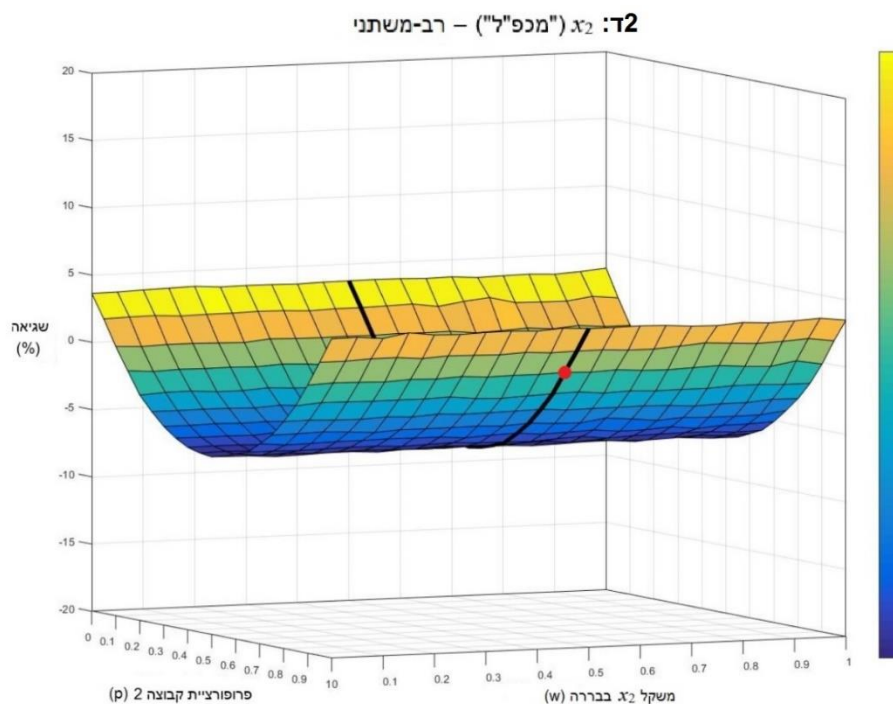
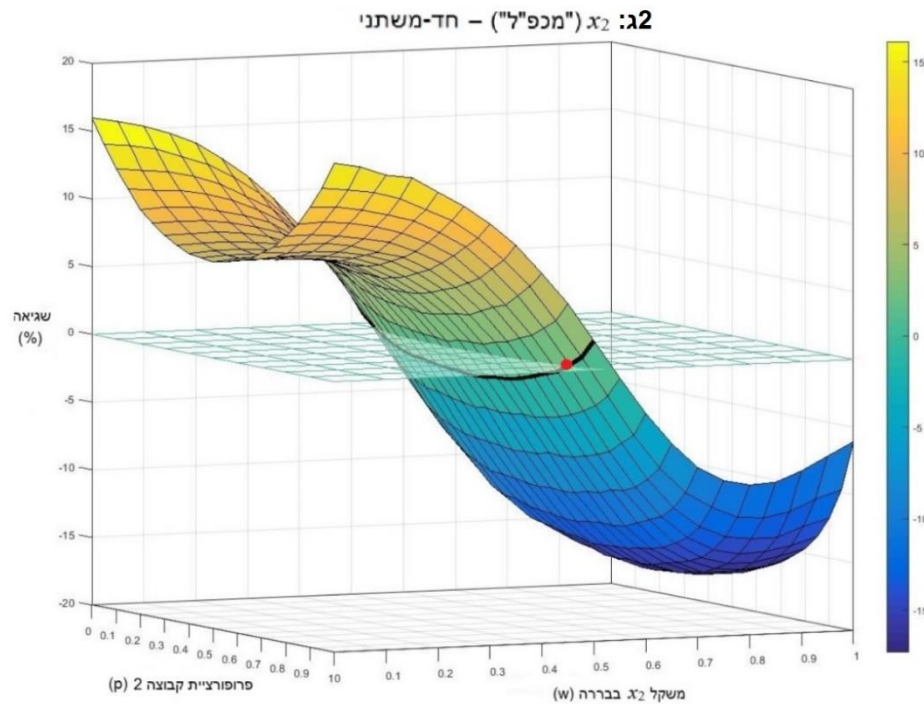
שני הפרמטרים, p ו- w , שונו ביחד על פני הריבוע $[0,1] \times [0,1]$ בקפיצות של 0.05 (ולא של 0.01, כפי שנעשה בחלק א'). בסך הכול היו, אם כן, 21 ערכים אפשריים הן עבור p והן עבור w וביחד 441 צירופים. בדומה למה שנעשה בחלק א', חזרנו 100 פעמים על חלק ב' עבור כל צירוף של ערכים של p ו- w , על מנת לקבל תוצאות יציבות.

6. חלק ב' – תוצאות

ממוצעי השגיאות של ארבעת המתאמים המתוקנים עבור 11×11 צירופים נבחרים של p (פרופורציית קבוצה 2 באוכלוסיית המועמדים האמתית) ו- w (המשקל של x_2 בתהליך הבררה שנעשה בפועל) מוצגים בלוחות 8 עד 11 בנספח 2. ערכים אלו מוצגים בתרשים 2 עבור כל הצירופים של p ו- w שהגדרנו.

תרשים 2: השגיאה (באחוזים) באמידת המתאם בין x_1 ו- y ובין x_2 ו- y ע"י תיקון לקיצוץ תחום בשתי שיטות, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)





בכל התרשימים, מראה הקו השחור העבה את התוצאות במצב שבו $w=0.5$, כלומר במצב שבו הבררה התבססה על סכום של x_1 ו- x_2 במשקלות שווים. ניתן לראות שהעקומות שנוצרות זהות לעקומות המקבילות בתרשים 1. כפי שנאמר קודם, חלק א' הוא מקרה פרטי של חלק ב', כפי שממחישים התרשימים⁷. לפיכך, כפי שהראינו בחלק א', במצב שבו $w=0.5$

⁷ בחרנו שלא להסתפק בחלק ב', הכללי יותר, מסיבות דידקטיות – חלק זה מחייב הצגה מורכבת (תרשימים תלת-ממדיים) של התוצאות, מה שעשוי להכביד על ההבנה.

וכן $p=0.9$ (כלומר שפרופורציית קבוצה 2 באוכלוסיית המועמדים האמתית זהה לפרופורציה באוכלוסיית המועמדים ההיסטורית), השגיאה בתיקון המתאמים בשתי השיטות היא אפסית, וזאת כיוון שגם הנחת התיקון החד-משתני בנוגע לתהליך הבררה נכונה וגם הפרמטרים של המועמדים שבהם משתמשים בתיקון נכונים. נקודה זו מסומנת בתרשים 2 בנקודה אדומה (מקבילה לנקודה הירוקה בתרשים 1).

מה קורה כאשר w שונה מ-0.5 וכשאוכלוסיית המועמדים האמתית הולכת ומתרחקת מאוכלוסיית המועמדים ההיסטורית (עד לנקודת הסימטריה)⁸?

1. השפעת w בתיקון הרב-משתני (תרשימים 2ב' ו-2ד'): כצפוי, גודל השגיאה בתיקון הרב-משתני, אינו תלוי ב- w : גובה הפונקציה הוא קבוע כשנעים על ציר ה- w .

2. השפעת w בתיקון החד-משתני (תרשימים 2א' ו-2ג'): המשטח האופקי השקוף למחצה מייצג את קו הגובה 0, כלומר, מצב של העדר שגיאה: בשני התרשימים ערכו של w משפיע על גודל הטעות (האפקט אינו מונוטוני על פני כל הטווח של w). כך, למשל, אם אוכלוסיית המועמדים האמתית דומה לאוכלוסיית המועמדים ההיסטורית ($p=0.9$), אזי אם ההנחה בדבר תהליך הבררה היא נכונה ($w=0.5$), התיקון יהיה מדויק (הנקודה האדומה בתרשים), אך אם משקל x_2 בבררה גבוה מ-0.5, אזי, במקרה של x_1 (תרשים 2א'), במצב החדש התיקון החד-משתני עושה טעות חיובית, כלומר: אומד את התוקף של x_1 כגבוה מדי, ולהיפך אם משקל x_2 בבררה נמוך מ-0.5. תמונה הפוכה מתקבלת במקרה של x_2 (תרשים 2ג').

כדאי לציין שבמקרה של x_1 , במצב שבו $p < 0.9$ ו- $w > 0.5$ הטעויות משני המקורות לחוסר דיוק של התיקון הן בכיוונים מנוגדים, ועשויות לקזז זו את זו. מצב כזה עשוי לאפיין מוסד שבו פרופורציית קבוצה 1 ("נבחני ערבית") בקרב המועמדים גדולה מ-0.1 ומשקל x_2 ("מכפ"ל") בבררה גבוה מ-0.5. במקרה של x_2 שתי הטעויות האלה מצטברות והמצב שתואר גורר טעות שלילית באמידת התוקף של x_2 .

3. השפעת w בתיקון החד-משתני אינה קבועה ברמות שונות של p . כלומר, קיימת אינטראקציה (אורדינלית), וזאת הן במקרה של x_1 (תרשים 2א') והן במקרה של x_2 (תרשים 2ג'). האינטראקציה חזקה יותר במקרה של x_1 . כך, למשל, בתרשים 2א', כש- $p=0.9$ הטעויות ב- w -ים השונים נעות בין -8% ל-+12%; ואילו כש- $p=0.5$ הטעויות ב- w -ים השונים נעות בין -22% ל-+11%.

⁸ מכאן ואילך נתייחס רק לשינוי של p בטווח שבין 0.5 ל-1. כפי שניתן לראות בתרשים 2 ובטבלאות 8-11, התנהגות הפונקציה בטווח שבין 0 ל-0.5 היא זהה. כלומר, הפונקציה סימטרית סביב הערך $p=0.5$. העובדה שהסימטריה אינה מושלמת מקורה בטעות דגימה.

4. השפעת ρ בתיקון הרב-משתני (תרשימים 2ב' ו-2ד'): במקרה של x_1 , השפעת ρ קטנה מאוד (קמירות קלה בתרשים 2ב'): הטעויות נעות בין 1%-ל-3%+. במקרה של x_2 , השפעת ρ גדולה יותר (קעירות לא מבטלת בתרשים 2ד'): הטעויות נעות בין 7%-ל-4%+.

5. השפעת ρ בתיקון החד-משתני (תרשימים 2א' ו-2ג'): בשני המקרים ערכו של ρ משפיע על גודל הטעות, וכאמור, בשניהם יש אינטראקציה (אורדינלית) בין ρ ל- w , כאשר במקרה של x_1 האינטראקציה חזקה יותר. כך, למשל, בתרשים 2א' כש- $w=1.0$ אין ל- ρ השפעה על גודל הטעות; כש- $w=0.5$ הטעויות ב- ρ -ים השונים נעות בין 11%-ל-4%+; וכש- $w=0.3$ הטעויות ב- ρ -ים השונים נעות בין 22%-ל-2%-.

מעבר לחקירת ההשפעה של שני הגורמים על מידת הדיוק של שני סוגי התיקונים ביחס לשני המשתנים, השאלה העיקרית שמעסיקה אותנו בהקשר הנוכחי היא: איזו שיטה יותר מדויקת כאשר גם w וגם ρ אינם ידועים. ניתן להיווכח שבמצב זה יש יתרון ברור לשימוש בתיקון הרב-משתני כיוון שתפקודו יציב יותר, כלומר: קרוב יותר בכל הנסיבות לתפקודו במצב של ידע מושלם. הוא מושפע מ- ρ בלבד, וההשפעה הזו אינה דרמטית – היא מינימלית ביחס ל- x_1 וקצת יותר גדולה ביחס ל- x_2 (ראו תרשימים 2ב' ו-2ד'). שגיאת האמידה בתיקון הרב-משתני לא עברה את ה-3% במקרה של x_1 , ואת ה-7% במקרה של x_2 . התיקון החד-משתני, לעומת זאת, מושפע לא רק מ- ρ אלא גם מ- w , והשפעה זו היא דרמטית הן במקרה של x_1 והן במקרה של x_2 : ניתן לראות בתרשימים של התיקון החד-משתני (2א' ו-2ג') שהשינוי בצבעים הרבה יותר דרמטי על פני הטווח של w מאשר על פני הטווח של ρ . שגיאת האמידה בתיקון החד-משתני עלולה להגיע ל-22% במקרה של x_1 , ול-17% במקרה של x_2 . במילים אחרות, אי הוודאות בנוגע ל- w היא קריטית יותר מאי הוודאות בנוגע ל- ρ (בנסיבות שבדקנו), ולפיכך התיקון הרב-משתני, שמתמודד היטב עם אי ודאות כזו, עדיף.

7. סיכום ודין

מחקר קודם (קנת-כהן ואחרים, 2016), שהתמקד במקרה הפרטי של שני חזאים ששימשו בבררה (x_1 ו- x_2) וקריטריון יחיד (y), השווה בין שתי שיטות לתיקון לקיצוץ תחום – תיקון חד-משתני ותיקון רב-משתני – במצב שבו הפרמטרים של אוכלוסיית המועמדים, שנחוצים לתיקון, ידועים. המחקר מצא שהתיקון הרב-משתני דייק באמידת המתאמים תחת מגוון של שיטות בררה. לעומתו, התיקון החד-משתני תפקד טוב בעיקר כאשר ההנחה שעמדה בבסיסו בדבר שיטת הבררה הייתה נכונה.

המחקר הנוכחי התמקד באותו מקרה פרטי, והשווה בין שיטות התיקון במצב שבו הפרמטרים של אוכלוסיית המועמדים אינם ידועים, מה שמחייב שימוש בערכי פרמטרים שמתקבלים מאוכלוסיות מועמדים דומות. המחקר הנוכחי, בדומה לקודמו, היה מחקר הדמייתי, וככזה הוא אפשר לבדוק את מידת הדיוק של המתאמים שתוקנו לקיצוץ תחום. מידת הדיוק של שתי שיטות התיקון נבדקה כפונקציה של שני גורמים: האחד, מידת השוני בין אוכלוסיית המועמדים ה"אמתית" לבין האוכלוסייה שבפרמטרים שלה השתמשנו לצורך התיקון, והשני, מידת השוני בין ההנחה שהניח התיקון החד-משתני בנוגע למשקלות משתני הבררה לבין המשקלות ששימשו בפועל. ממצאי המחקר הנוכחי מסוכמים להלן בלוח 7, לצד ממצאי המחקר הקודם, על פי המבנה שבו הוצג לוח 2.

לוח 7: השוואה בין שתי שיטות לתיקון לקיצוץ תחום – ממצאים ממחקר קודם ונוכחי

הפרמטרים של המועמדים			
לא ידועים	ידועים		
אין עדיפות עקבית לאף שיטה (מחקר נוכחי חלק א')	שתי השיטות מתפקדות היטב (מחקר קודם)	ידועים	משקלות משתני הבררה
תיקון רב-משתני עדיף (מחקר נוכחי חלק ב')	תיקון רב-משתני עדיף (מחקר קודם)	לא ידועים	

במחקר הנוכחי ("הפרמטרים של המועמדים לא ידועים") נמצא שכאשר ההנחה שעמדה בבסיס התיקון החד-משתני בדבר תהליך הבררה הייתה נכונה ("משקלות משתני הבררה ידועים"), אזי במקרה של x_1 (שמדמה את ממוצע הבגרות) התיקון הרב-משתני נמצא עדיף, ובמקרה של x_2 (שמדמה את ציון מכפ"ל) התיקון החד-משתני היה עדיף. כאשר ההנחה שעמדה בבסיס התיקון החד-משתני בדבר תהליך הבררה לא הייתה נכונה ("משקלות משתני הבררה לא ידועים"), התיקון הרב-משתני נמצא עדיף הן במקרה של x_1 והן במקרה של x_2 .

ניתן לאפיין את מצב העולם שבו נערכים מחקרי תוקף הניבוי במאל"ו כמצב שבו תהליך הבררה לא ידוע: במקרים רבים משתמשים בציון קבלה שבו המשקלות של ממוצע הבגרות וציון מכפ"ל הם שווים, אך במקרים רבים אחרים המשקלות אינם שווים. לעתים השימוש בשני החזאים אינו מבוסס על צירופם במשקלות ידועים וקבועים כלשהם (למשל, כאשר משתמשים בציון הגבוה מבין השניים). לבסוף, בשנים האחרונות התרחבה האפשרות להתקבל לתחומי לימוד שונים גם על סמך ממוצע הבגרות בלבד או ציון מכפ"ל בלבד. מחקרי

התוקף הללו נערכים, אם כן, בתנאים שבהם w לא ידוע, וכפי שמצאנו, במצב זה יש עדיפות לתיקון הרב-משתני. כדאי להוסיף שמשקלות מרכיבי הבררה משתנים בין תחומי לימוד, מוסדות ומחזורים, כך שגם אם היינו יודעים את המשקלות המדויקים שבהם משתמשים בכל צירוף של מוסד, תחום לימודים ומחזור, לא היה זה מעשי להפעיל תיקון חד-משתני דיפרנציאלי לכל צירוף כזה. התיקון הרב-משתני מתאים לסיטואציה כזו, וממצאי המחקר הנוכחי מלמדים שהוא מתמודד עם האתגר היטב, גם כאשר נתוני המועמדים אינם ידועים באופן מדויק.

במצב שבו משקלות משתני הבררה ידועים אין, לכאורה, יתרון לאף שיטה, אך עיון בתוצאות המפורטות מלמד שגם בנסיבות אלה מסתמן יתרון קל לתיקון הרב-משתני: המחיר של שימוש בתיקון הרב-משתני במקום בתיקון החד-משתני במקרה של x_2 נמוך מהמחיר של שימוש בתיקון החד-משתני במקום בתיקון הרב-משתני במקרה של x_1 . בנוסף, כאשר משתמשים בתיקון הרב-משתני במקום בתיקון החד-משתני המחיר הוא הגדלת אומדן החסר⁹ של תוקף x_2 ("מכפ"ל") – כיוון שמן הבחינה הפוליטית הוא קביל, בעוד שכאשר משתמשים בתיקון החד-משתני במקום בתיקון הרב-משתני המחיר הוא יצירת אומדן חסר ניכר של תוקף x_1 ("בגרות") לעומת אומדן יתר קל אם משתמשים בתיקון המתאים יותר – כיוון שמן הבחינה הפוליטית הוא פחות נוח.

לצד המסקנה העיקרית של המחקר, קרי: שהתמודדות התיקון הרב-משתני עם המידע המוגבל על נתוני המועמדים היא סבירה (שגיאה של פחות מ-10% באמידת המתאמים), עולה מסקנה חשובה נוספת, בדבר חשיבותו של מידע מדויק על אוכלוסיית המועמדים האמתית: כפי שהמחשנו בהדמיות שערכנו, מידע עדכני ושוטף על נתוני המועמדים (ולו רק ברמת הפרמטרים הנחוצים לתיקון) יוכל להעלות מאוד את איכות התיקון.

לבסוף, יש מקום להזכיר ולחדד את מגבלות המחקר. ראשית, בהדמיות שעשינו נבחרו מאפיינים – הממוצעים וסטיות התקן של המשתנים, המתאמים בין המשתנים וכן הפרופורציות של שתי הקבוצות שהרכיבו את אוכלוסיות המועמדים – שמשקפים בקירוב את ההקשר שבו נערכים מחקרי התוקף של מאל"ו. כאשר שינינו חלק מן הפרמטרים הרלוונטיים, התקבלו ממצאים שונים מאוד. לפיכך, זהו אינו מחקר תאורטי כללי ולא ניתן להכליל ממנו על תפקוד שתי השיטות מעבר לכל מצבי העולם האפשריים. המסקנות שגזרנו מן הממצאים ישימות, לפיכך, להקשר המוגבל שנבדק במחקר זה.

יתרה מזו, במחקר זה בחרנו לתפעל את ההבדל בין הפרמטרים של אוכלוסיית המועמדים ההיסטורית, ששימשה אותנו בתיקון, לאלו של אוכלוסיות המועמדים האמתיות, על ידי שינוי

⁹ בהנחה שפרופורציית קבוצה 2 נמצאת בין הערכים הקיצוניים (0.9 ו-0.1) ולא מחוצה להם.

בפרופורציות של שתי קבוצות שנבדלו זו מזו בפרמטרים אחדים. זוהי, כמובן, רק המחשה אפשרית אחת מני רבות לאפיון דיפרנציאלי של אוכלוסיות מועמדים. לפיכך, גם בהקשר המוגבל של מחקרי התוקף של מאל"ו, מציגות ההדמיות המחשה של שינוי אפשרי אחד באוכלוסיית המועמדים, ואין לראות בהן מיצוי של כל טווח השינויים שעשויים היו להתרחש על פני השנים באוכלוסייה זו.

לבסוף, חשוב להזכיר שגם התיקון הרב-משתני מניח הנחה על תהליך הבררה: הוא מניח שהבררה התבססה על צירוף כלשהו של ממוצע הבגרות ושל ציון מכפ"ל, ושאינן משתנים נוספים שמעורבים בבררה. הנחה זו אמנם חלשה יותר מן ההנחה שמניח התיקון החד-משתני, אך לא ניתן להתעלם ממנה. נתונים ראשוניים מסקר שנערך לאחרונה בקרב ראשי יחידות המיון האוניברסיטאיות מלמד שהחלטות הקבלה מתבססות לעתים על משתנים נוספים, כמו ראיונות אישיים, מבחני קבלה חוגיים, ציונים במקצועות ספציפיים בתעודת הבגרות, ועוד. לפיכך, במידה שבה הנחת התיקון הרב-משתני בנוגע לתהליך הבררה אינה מדויקת, ניתן לצפות לפגיעה באיכות האמידה שהוא מספק. בימים אלה אנו מנסים להעריך את היקף החריגה מההנחה שעושה התיקון הרב-משתני בנוגע לתהליך הבררה. במחקר הבא, השלישי בסדרת המחקרים על התיקון הרב-משתני, נביא תוצאות מפורטות יותר מן הסקר שהוזכר לעיל. בנוסף יכלול המחקר יישום של המקרה הכללי של התיקון הרב-משתני, שבו יש יותר משני משתנים שעברו בררה ישירה ויותר ממשתנה אחד שעבר בררה עקיפה.

8. מקורות

קנת-כהן, ת', טורוול א' ואורן, כ' (2011). **הוגנות מערכת המיון לאוניברסיטאות לפי שפת היבחנות במכפ"ל (ערבית או עברית) ומגדר** (דוח מס' 376). ירושלים: מרכז ארצי לבחינות ולהערכה.

קנת-כהן, ת', קלפר ד' וטורוול, א' (2016). **תיקון לקיצוץ תחום עבור בררה חד-משתנית (במקרה התלת-משתני) ועבור בררה רב-משתנית: השוואה בנתונים הדמייתיים ובנתונים אמפיריים** (דוח מס' 16-07). ירושלים: מרכז ארצי לבחינות ולהערכה.

Gulliksen, H. (1987). *Theory of mental tests*. Hillsdale, NJ: Erlbaum. (Original work published 1950)

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, 26, 329–367.

נספח 1 – נוסחאות התיקונים לקיצוץ תחום¹⁰

תיקון לקיצוץ תחום עבור בררה חד-משתנית (במקרה התלת-משתני)

להלן תוצג נוסחת התיקון עבור המתאם באוכלוסייה בין זוג משתנים, x ו- y (R_{xy}), כאשר תהליך הבררה התבסס על משתנה שלישי, z .

נתונים:

- מתאמים בין המשתנים (r_{xy}, r_{xz}, r_{yz}) במדגם המתקבלים
- סטיית התקן של משתנה הבררה במדגם המתקבלים (s_z)
- סטיית התקן של משתנה הבררה באוכלוסייה (S_z)

אזי עבור $w_z = \left(\frac{s_z^2}{S_z^2} - 1 \right)$, המתאם המתוקן הוא:

$$R_{xy} = \frac{r_{xy} + w_z r_{xz} r_{yz}}{\sqrt{(1 + w_z r_{xz}^2)(1 + w_z r_{yz}^2)}}$$

תיקון לקיצוץ תחום עבור בררה רב-משתנית (במקרה הפרטי של שני חזאים ששימשו בבררה וקריטריון יחיד)

להלן תוצג נוסחת התיקון עבור המתאם באוכלוסייה בין זוג משתנים, x_1 ו- y (R_{yx_1}), כאשר תהליך הבררה התבסס על צירוף כלשהו של x_1 ו- x_2 .

נתונים:

- השונות ($s_{x_1x_1}, s_{x_2x_2}, s_{yy}$) והשונות המשותפות ($s_{x_1x_2}, s_{yx_1}, s_{yx_2}$) של המשתנים במדגם המתקבלים
- השונות ($S_{x_1x_1}, S_{x_2x_2}$) והשונות המשותפת ($S_{x_1x_2}$) של משתני הבררה באוכלוסיית המועמדים

אזי:

השונות המשותפת בין y ל- x_1 באוכלוסיית המועמדים:

$$S_{yx_1} = \frac{1}{s_{x_1x_1}s_{x_2x_2} - s_{x_1x_2}^2} \cdot (s_{x_1x_1}s_{yx_1}s_{x_2x_2} + s_{x_1x_2}s_{yx_2}s_{x_1x_1} - s_{x_1x_2}s_{yx_1}s_{x_1x_2} - s_{x_1x_1}s_{yx_2}s_{x_1x_2})$$

¹⁰ הסבר נוסף, כולל פירוט ההנחות של כל שיטת תיקון, אפשר למצוא אצל קנת-כהן ואחרים, 2016.

השונויות של y באוכלוסיית המועמדים:

$$\begin{aligned}
 S_{yy} = & s_{yy} + \frac{1}{(s_{x_1x_1}s_{x_2x_2} - s_{x_1x_2}^2)^2} \\
 & \cdot [S_{x_1x_1}(s_{yx_1}^2s_{x_2x_2} + s_{yx_2}^2s_{x_1x_2} - 2s_{yx_1}s_{yx_2}s_{x_1x_2}s_{x_2x_2}) \\
 & + S_{x_2x_2}(s_{yx_2}^2s_{x_1x_1} + s_{yx_1}^2s_{x_1x_2} - 2s_{yx_1}s_{yx_2}s_{x_1x_2}s_{x_1x_1}) \\
 & + 2S_{x_1x_2}(s_{x_1x_2}^2s_{yx_1}s_{yx_2} + s_{x_1x_1}s_{x_2x_2}s_{yx_1}s_{yx_2} - s_{yx_1}^2s_{x_1x_2}s_{x_2x_2} - s_{yx_2}^2s_{x_1x_2}s_{x_1x_1})] \\
 & - \frac{1}{s_{x_1x_1}s_{x_2x_2} - s_{x_1x_2}^2} \cdot (s_{yx_1}^2s_{x_2x_2} - 2s_{yx_1}s_{yx_2}s_{x_1x_2} + s_{yx_2}^2s_{x_1x_1})
 \end{aligned}$$

והמתאם המתוקן הוא:

$$R_{yx_1} = \frac{S_{yx_1}}{\sqrt{S_{yy}S_{x_1x_1}}}$$

נספח 2 – תוצאות חלק ב' עבור צירופים נבחרים של w - p

לוח 8: השגיאה (באחוזים) באמידת המתאם בין x_1 ו- y ע"י תיקון חד-משתני לקיצוץ תחום, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)

משקלו של x_2 בבררה (w)											פרופורציית קבוצה 2 (p)
1.00	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00	
12.94	12.11	11.12	9.54	7.41	4.46	1.04	-1.85	-2.88	-2.45	-1.04	0.00
11.72	10.53	8.60	6.29	3.75	0.08	-4.15	-7.24	-7.99	-5.96	-2.50	0.10
11.38	9.31	7.07	4.27	0.67	-4.25	-9.54	-13.61	-13.39	-10.00	-5.80	0.20
10.88	8.54	6.02	2.57	-1.58	-7.62	-14.15	-18.07	-17.46	-13.53	-8.14	0.30
10.90	8.29	5.16	1.60	-3.25	-9.77	-17.15	-21.14	-20.32	-15.69	-9.81	0.40
10.72	8.15	5.15	1.21	-3.73	-10.67	-18.22	-22.20	-21.32	-16.30	-10.25	0.50
10.97	8.38	5.23	1.67	-3.01	-9.88	-16.89	-21.01	-20.33	-15.56	-9.43	0.60
11.02	8.70	6.06	2.69	-1.60	-7.59	-14.30	-18.18	-17.58	-13.73	-8.02	0.70
11.41	9.13	6.85	4.45	0.71	-4.13	-9.64	-13.31	-13.29	-10.02	-5.59	0.80
11.68	10.23	8.52	6.51	3.96	-0.10	-4.28	-7.10	-8.03	-5.96	-2.85	0.90
12.78	12.14	11.11	9.34	7.39	4.45	1.02	-1.81	-2.95	-2.52	-0.88	1.00

לוח 9: השגיאה (באחוזים) באמידת המתאם בין x_1 ו- y ע"י תיקון רב-משתני לקיצוץ תחום, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)

משקלו של x_2 בבררה (w)											פרופורציית קבוצה 2 (p)
1.00	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00	
-0.87	-0.93	-0.89	-0.93	-0.85	-0.86	-0.82	-0.90	-0.74	-0.83	-0.98	0.00
-0.04	0.19	-0.01	-0.16	-0.01	0.07	0.18	0.18	0.01	-0.05	0.22	0.10
1.23	1.10	1.12	1.05	1.03	1.00	1.24	0.67	0.96	1.23	0.77	0.20
1.94	1.93	2.08	1.89	2.02	2.03	2.03	2.15	2.07	1.94	1.86	0.30
2.69	2.61	2.46	2.54	2.48	2.67	2.53	2.50	2.49	2.52	2.41	0.40
2.77	2.79	2.81	2.69	2.72	2.72	2.50	2.89	2.54	2.88	2.86	0.50
2.71	2.72	2.52	2.55	2.72	2.54	2.81	2.78	2.48	2.72	2.98	0.60
2.08	2.03	2.12	1.96	2.02	1.89	1.92	2.01	2.00	1.86	2.07	0.70
1.23	0.94	0.92	1.27	1.06	1.16	1.13	1.11	1.01	1.11	1.05	0.80
-0.08	-0.07	-0.05	0.00	0.19	-0.09	0.02	0.24	-0.15	-0.07	-0.08	0.90
-1.03	-0.93	-0.89	-1.08	-0.90	-0.82	-0.88	-0.97	-0.83	-0.75	-0.86	1.00

לוח 10: השגיאה (באחוזים) באמידת המתאם בין x_2 ו- y ע"י תיקון חד-משתני לקיצוץ תחום, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)

משקלו של x_2 בבררה (w)											פרופורציית קבוצה 2 (p)
1.00	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00	
-5.87	-8.10	-8.84	-7.39	-3.30	2.24	7.34	11.18	13.81	15.20	15.98	0.00
-12.04	-13.14	-13.02	-10.85	-5.93	0.02	5.34	9.15	11.13	12.19	12.58	0.10
-14.77	-15.57	-14.96	-12.29	-7.03	-1.12	4.26	7.41	9.16	9.95	9.90	0.20
-16.16	-16.67	-15.69	-12.74	-7.58	-1.55	3.32	6.53	7.94	8.45	8.36	0.30
-16.71	-17.21	-16.19	-13.02	-7.79	-1.86	2.90	5.77	7.15	7.53	7.36	0.40
-16.91	-17.31	-16.16	-13.21	-7.89	-2.05	2.70	5.69	7.00	7.35	7.09	0.50
-16.68	-17.08	-16.23	-12.95	-7.61	-1.87	2.95	5.78	7.20	7.65	7.58	0.60
-16.02	-16.55	-15.70	-12.86	-7.57	-1.59	3.24	6.47	7.96	8.40	8.45	0.70
-14.65	-15.52	-15.00	-12.03	-7.04	-0.90	4.15	7.56	9.26	9.88	10.03	0.80
-12.05	-13.32	-13.11	-10.80	-5.86	0.05	5.34	9.11	11.20	12.15	12.51	0.90
-6.06	-8.05	-8.94	-7.49	-3.33	2.21	7.32	11.16	13.66	15.07	16.12	1.00

לוח 11: השגיאה (באחוזים) באמידת המתאם בין x_2 ו- y ע"י תיקון רב-משתני לקיצוץ תחום, לפי פרופורציית קבוצה 2 באוכלוסיית המועמדים (p) ומשקלו של x_2 בבררה (w)

משקלו של x_2 בבררה (w)											פרופורציית קבוצה 2 (p)
1.00	0.90	0.80	0.70	0.60	0.50	0.40	0.30	0.20	0.10	0.00	
3.89	3.60	3.65	3.73	3.75	3.75	3.68	3.73	3.83	3.73	3.63	0.00
-0.02	0.14	0.11	-0.08	0.08	0.05	0.04	0.12	-0.03	0.03	0.04	0.10
-3.16	-3.22	-3.14	-3.31	-3.13	-3.20	-2.96	-3.19	-3.22	-3.08	-3.22	0.20
-5.29	-5.33	-5.16	-5.24	-5.20	-5.15	-5.22	-5.13	-5.17	-5.16	-5.20	0.30
-6.33	-6.43	-6.42	-6.36	-6.36	-6.35	-6.37	-6.40	-6.41	-6.39	-6.41	0.40
-6.74	-6.79	-6.68	-6.86	-6.81	-6.81	-6.80	-6.70	-6.71	-6.67	-6.77	0.50
-6.33	-6.28	-6.47	-6.30	-6.21	-6.36	-6.29	-6.38	-6.38	-6.30	-6.20	0.60
-5.07	-5.19	-5.14	-5.30	-5.21	-5.13	-5.32	-5.17	-5.18	-5.24	-5.11	0.70
-3.02	-3.16	-3.25	-3.07	-3.11	-3.00	-3.09	-3.03	-3.08	-3.15	-3.10	0.80
-0.11	-0.11	0.11	-0.04	0.12	0.09	0.03	0.11	0.06	0.01	-0.04	0.90
3.52	3.62	3.66	3.56	3.74	3.70	3.67	3.73	3.71	3.61	3.75	1.00