

# פיתוח שאלון אישיות ממוחשב בשיטת הבחירה הכפויה (FC)

כנס אפי 2023, מעלה החמישה

Joel Rapp  
Tzur Karelitz  
Dvir Kleper  
Itamar Cohen



# שאלון אישיות חדש

✓ המרכז הארצי לבחינות ולהערכה (מאל"ו) מפתח בימים אלו שאלון אישיות ממוחשב ואדפטיבי

✓ השאלון מעריך את חמשת ממדי האישיות על פי מודל ה- "Big five":

מוחצנות

נעימות

מצפוניות

יציבות רגשית

פתיחות מחשבתית

✓ השאלון, נבנה במקביל בעברית ובערבית מנתונים שנאספו בישראל ומותאם תרבותית לשני המגזרים



# שאלון בשיטת הבחירה הכפויה

השאלון בנוי בשיטת הבחירה הכפויה (Forced Choice) – בה מוצג לנבדק בכל פעם זוג היגדים המתארים התנהגות או הלך רוח מסוימים ועליו לבחור את ההיגד שלדעתו מתאר אותו טוב יותר.

זוג ההיגדים יכול להיות מאותו ממד אישיות או מממדים שונים. הוא יכול לתאר התנהגויות המקובלות כחיוביות או כשליליות.



# דוגמא לזוג היגדים מאותו מימד אישיותי

- אני מקפיד לסדר את שולחן העבודה שלי בסוף יום העבודה.
- אני אוהב להיות בסביבה מסודרת.



# דוגמא לזוג היגדים "חיוביים"

○ בעבודה, אני נוהג לקחת אחריות.

○ אני אדם אופטימי מטבעי.

אם הנבדק חש ששני ההיגדים מתארים אותו היטב, הוא מחוייב להכריע ביניהם, ולבחור את ההיגד שמתאר אותו טוב יותר.



# דוגמא של זוג היגדים "שליליים"

- אני נוהג לבקר את עצמי בכל דבר ועניין.
- אני לא אוהב להיפגש עם חברים.

במקרה כזה, אף אם הנבדק חושב ששני ההיגדים אינם מתארים אותו, הוא יאלץ לבחור את ההיגד הפחות דחוי בעיניו.



# השאלון הוא אדפטיבי

- ✓ השאלון הוא אדפטיבי.
- ✓ הבניה, ההעברה והציינון של השאלון מבוססים על תורת התגובה לפריט (IRT).
- ✓ השימוש ב-IRT מאפשר לדגום בכל צעד את זוג ההיגדים שמפיק את מרב האינפורמציה על הנבדק בזוג הממדים שמהם נדגמו היגדים.

✓ הבסיס ליישום הבחירה הכפויה במסגרת שאלון אדפטיבי המבוסס על IRT נוסח על ידי Stark, et.al. (2012 ; 2005) במודל המכונה MUPP (Multi-Unidimensional Pairwise Preference).

✓ בבסיס המודל ההנחה כי כאשר נבחן מחוייב לבחור בין שני היגדים, ( $t$ - $s$ ) הליך ההעדפה מוגדר ככזה שבו הנבחן מקבל את ההיגד שנבחר ודוחה את ההיגד האחר.

✓ מרחב ההסתברות של הבחירה יכול להיות אפוא רק: קבלת  $s$  ודחיית  $t$  או דחיית  $s$  וקבלת  $t$ . לא קיימת אפשרות שבה מקבלים את שניהם או דוחים את שניהם.



# הנוסחה היסודית במודל ה-MUPP

✓ ההסתברות לבחור את היגד  $s$  על פני היגד  $t$ , המתוארת בספרות כ-  $P_{s>t}$  מנוסחת באופן הבא:

$$P_{(s>t)_i}(\theta_{d_s}, \theta_{d_t}) = \frac{P_{st}\{1, 0\}}{P_{st}\{1, 0\} + P_{st}\{0, 1\}} \approx \frac{P_s\{1\}P_t\{0\}}{P_s\{1\}P_t\{0\} + P_s\{0\}P_t\{1\}}$$

✓ כלומר שמתייחסים לשני התהליכים כבלתי תלויים ולכן מכפילים את ההסתברויות.

✓ ההסתברויות  $P_s\{1\}$  ו-  $P_t\{0\}$ , הן פונקציה של "ערכי" הנבדק בכל אחת מהתכונות, כלומר  $\theta_{d_s}$  ו-  $\theta_{d_t}$ .

# באיזה מודל IRT להשתמש לצורך חישוב ההסתברויות?

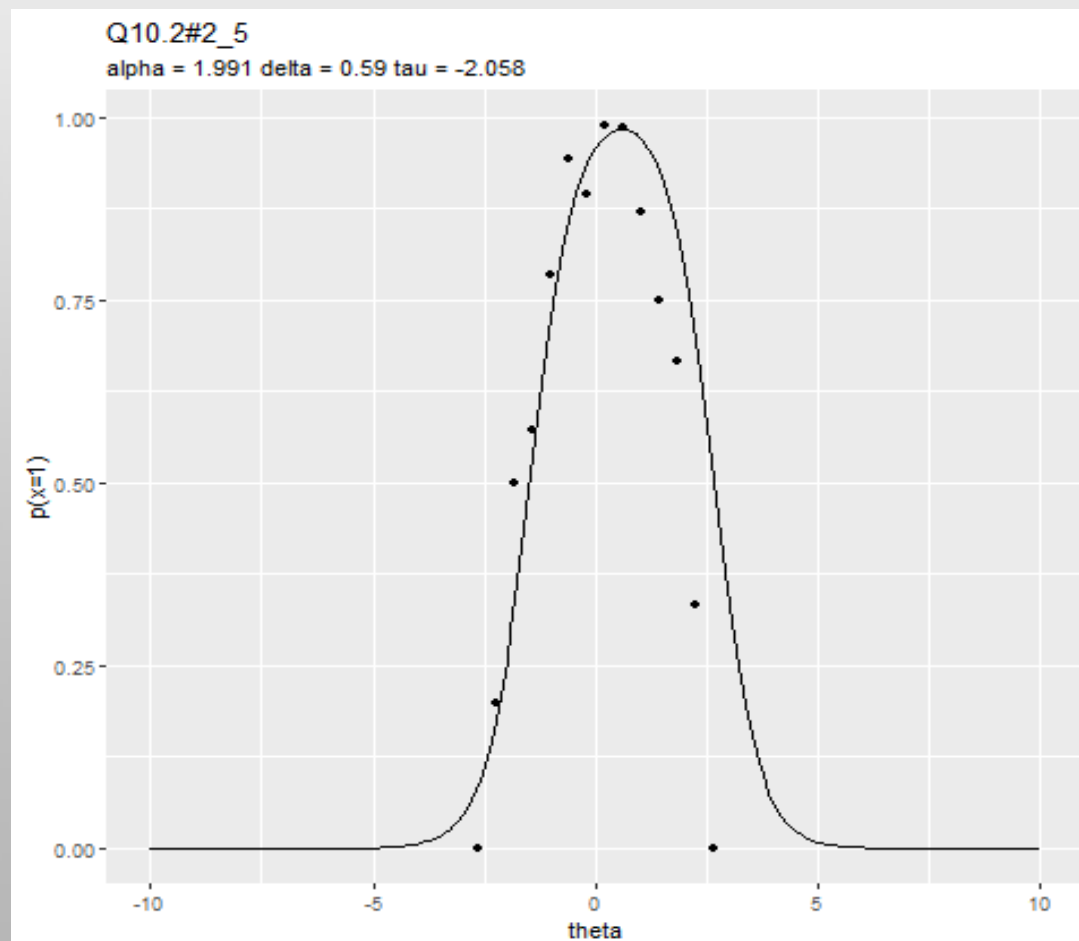
✓ נשקלו שני מודלים של IRT לשימוש ב-MUPP. הראשון הוא זה שהשתמשו בו Stark et.al, והאחר מוכר ונמצא בשימוש רב במבחני הישגים/ידע אדפטיביים. ההבדל ביניהם מתמקד בהנחה הנוגעת לאופי הקשר המתקיים בין ההסתברות לקבל היגד מסוים לבין "הציון" של הנבדק בתכונה.

1. התגובות להיגדים מתנהגות על פי מודל "פרוש" (GGUM)

2. התגובות להיגדים מתנהגות על פי מודל מונוטוני (דומיננטי)

# המודל הפרוש (unfolding)

ה- GGUM (Generalized Graded Unfolding Model)



# המודל הפרוש (unfolding)

✓ "המודל הפרוש" של רוברטס ושותפיו, מתאים בעיקר לפריטי "מסכים – לא-מסכים" או לפריטים שהתגובה עליהם ניתנת על סולם ליקרט. בד"כ בשאלוני דיווח עצמי, אישיות ו/או עמדות.

✓ לאופיין ההסתברות להסכים עם הפריט יש צורה פעמונית, כלומר, שישנו ערך "אידיאלי" על סולם התכונה שבו ההסתברות להסכים עם ההיגד היא מקסימלית (קרובה ל-1).

✓ כעקרון, ככל שאדם ניחן יותר בתכונה הנמדדת, כך ההסתברות להסכים עם ההיגד עולה. זאת, עד לערך מסוים של התכונה, שממנו, ההסתברות יורדת ככל שאדם ניחן יותר בתכונה.

✓ לפי כך, ישנם שני איזורי אי-הסכמה: "אי הסכמה מלמטה" ו-"אי הסכמה מלמעלה".

# דוגמא להיגד המתאים למודל פרוש

"אני מסדר את הדוא"ל שלי במחשב פעם בשבוע"

ישנם שני אזורי אי הסכמה: ✓

- אי הסכמה מלמטה (אני לא/פחות מסודר ולכן ההיגד לא מתאר אותי)
- אי הסכמה מלמעלה (אני מסודר הרבה יותר ממה שמתואר בהיגד ולכן התיאור הזה לא מתאים לי כי אני מסדר את הדוא"ל שלי כל יום)



# מודל מונוטוני / דומיננטי

✓ מודל IRT מוכר ופופולרי המשמש במבחני הישגים/ידע/  
מיומנויות (לא דיווח עצמי). יש בו שני פרמטרים: פרמטר קושי  
(b), ופרמטר אבחנה (a).

✓ מתאים לפריטים שבהם אפשר להגדיר מהי התשובה הנכונה.

✓ ככל שאדם בעל ידע/יכולת/מיומנויות גבוהה יותר, ההסתברות  
להצליח להשיב נכון על פריט עולה.


✓ יכול להתאים גם להיגדים של שאלוני אישיות. כמו למשל בהיגד  
"אני אוהב להיפגש עם חברים"

✓ הקשר בין מידת התכונה בה ניחן האדם לבין ההסתברות  
להסכים עם ההיגד הוא מונוטוני. במקרה זה, ככל שאדם

חברותי יותר הוא ייטה להסכים עם ההיגד. לא צפויה ירידה **NITE**  
בהסכמה גם אצל אנשים מאוד מאוד חברותיים.

# מה מתאים יותר להיגדי אישיות ועמדות?

✓ בספרות ניטש ויכוח בין תומכי מודל פרוש לבין תומכי מודל מונוטוני לתיאור היגדי אישיות ועמדות.

✓ Roberts, Starks, Drasgow  Brown, Maydeu-Olivares (TIRT models), Morillo



# מה מתאים יותר להיגדי אישיות ועמדות?

- ✓ תומכי הגישה הפרושה טוענים שההסכמה עם היגדים המודדים אישיות ועמדות מתאימים יותר מתנהגת על פי ערך אידאלי בתכונה, שממנה והלאה יש פחות הסכמה. **זאת, במיוחד בהיגדים ממוקמים במרכז ההתפלגות.** לטענתם, להניח שהם מתנהגים באופן מונוטוני זה להתעלם ממידע חשוב שאפשר באמצעותו לשפר את אמידת התכונה. הם בנו גם מודלים של מבחנים אדפטיביים העושים שימוש במודל.
- ✓ תומכי המונוטוניות טוענים שישנן תכונות (כמו פתולוגיה) שמתיישבים יותר עם מודל מונוטוני, שהמודלים פשוטים יותר וחסכניים יותר, ושסולם הציונים שנוצר מהיגדים מונוטוניים הוא מהימן יותר ובעל תכונות פסיכומטריות טובות יותר.
- ✓ ממילא קשה לפתח פריטים במרכז הסקלה.

## סקר לאיסוף נתונים על ההיגדים

- ✓ כל ההיגדים בכל אחת מהתכונות/מימדים הועברו למדגם של כ- 500 ישראלים בעלי פרופיל דמוגרפי מתאים (כולל דוברי עברית וערבית).
- ✓ לכל נבדק הוצגו ההיגדים בזה אחר זה (לא בזוגות) והוא התבקש להשיב בכנות לגבי כל היגד באיזו מידה ההיגד מתאר אותו בסולם 1-6. בהמשך, הפכנו את התגובות לדיכוטומיות (1 → 4-6; 0 → 1-3).
- ✓ אמדנו את הפרמטרים של כל ההיגדים בשני המודלים.
- ✓ בנינו שאלון אדפטיבי המבוסס על פרמטרים אלו והסתברויות התואמות ובחננו את התוצאות.
- ✓ דימינו הרצה של השאלונים על אלפי נבחנים סימולטיביים, שהציונים שלהם נגזרו מהתפלגויות ידועות מראש. בדקנו את הקשר בין אומדני הציונים שהתקבלו לבין הציונים האמיתיים.

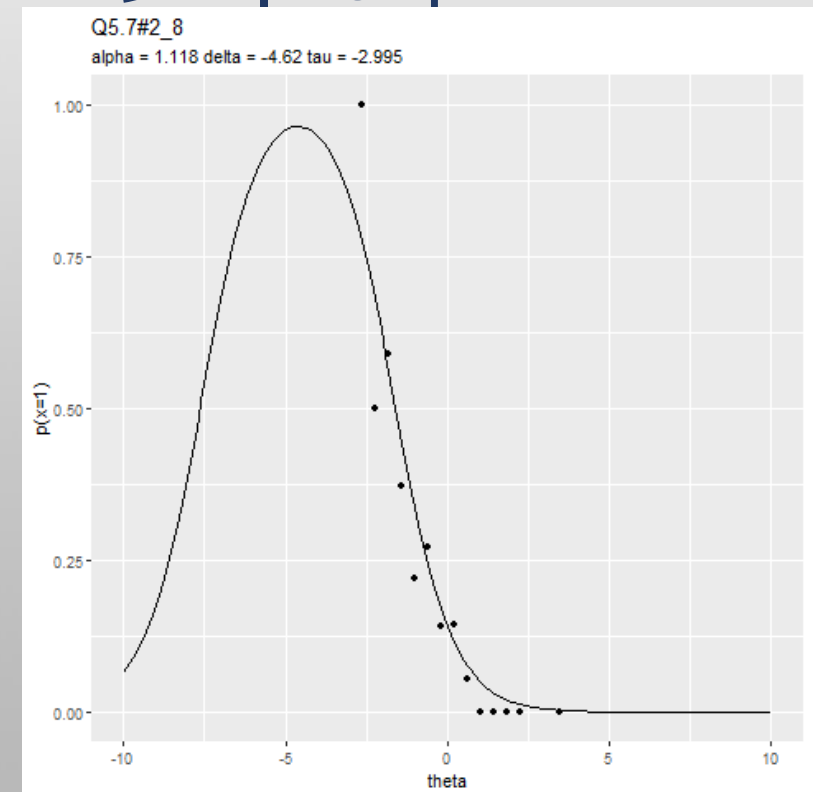
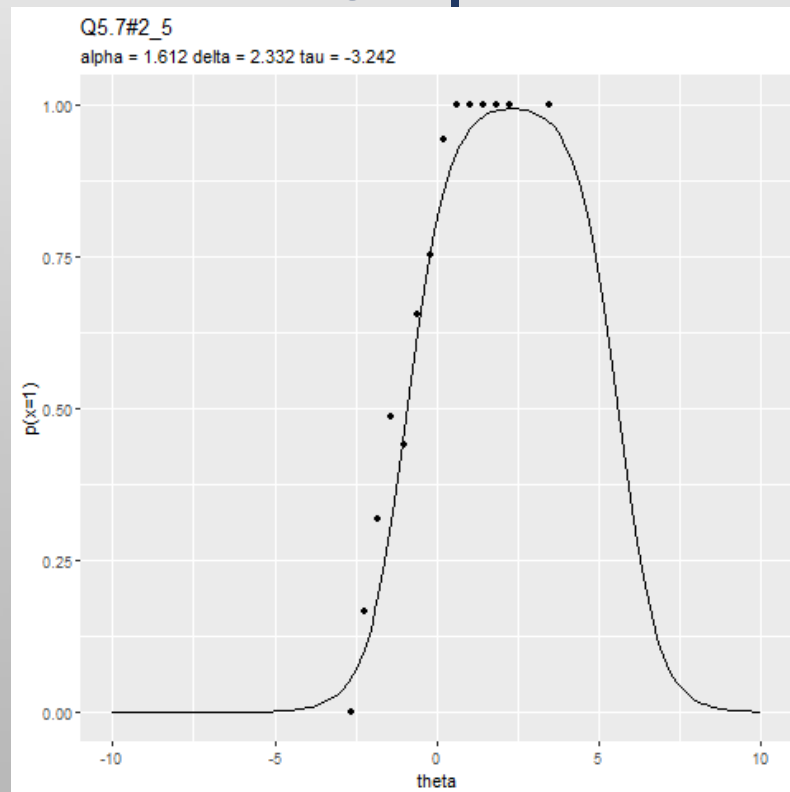
# תוצאות:

## בשני המודלים:

- ✓ בסך הכל, הייתה התאמה טובה בין הנתונים האמפיריים של התגובות שנאספו לעקומות התאורטיות.
- ✓ בסמילוציות של הרצת המבדק, הצלחנו לשחזר היטב את רמות היכולת של הנבחנים הסימולטיביים (דהיינו, התקבל מתאם גבוה בין תטות אמיתיות לתטות נאמדות וטעויות האמידה היו בסך הכל קטנות).

# עבור המודל הפרוש

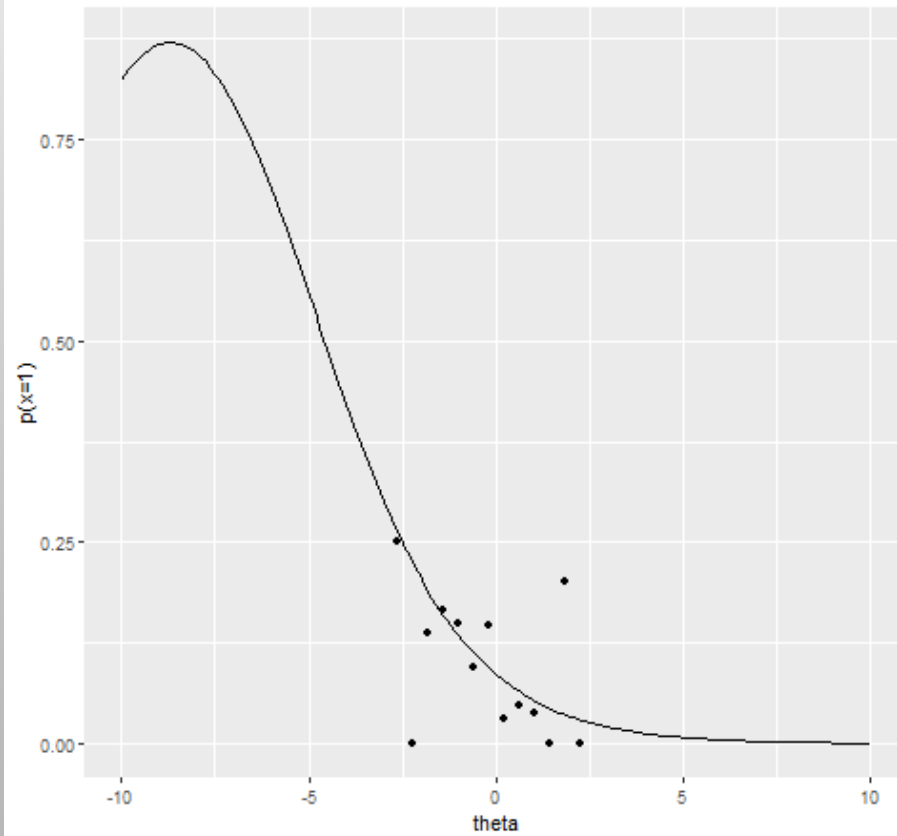
✓ ניתוח ויזואלי של התוצאות הצביע על כי ברוב המקרים (כ-95%), הנתונים התאימו היטב לעקומה הפעמונית. ואולם, הנתונים התמקדו רק באחד הצדדים של הפעמון. בפרטים החיוביים- בעליה, ובפרטים השליליים בירידה. כך שאין לדעת מה צפוי להיות במדרון השני.



# דוגמאות להתאמה גרועה

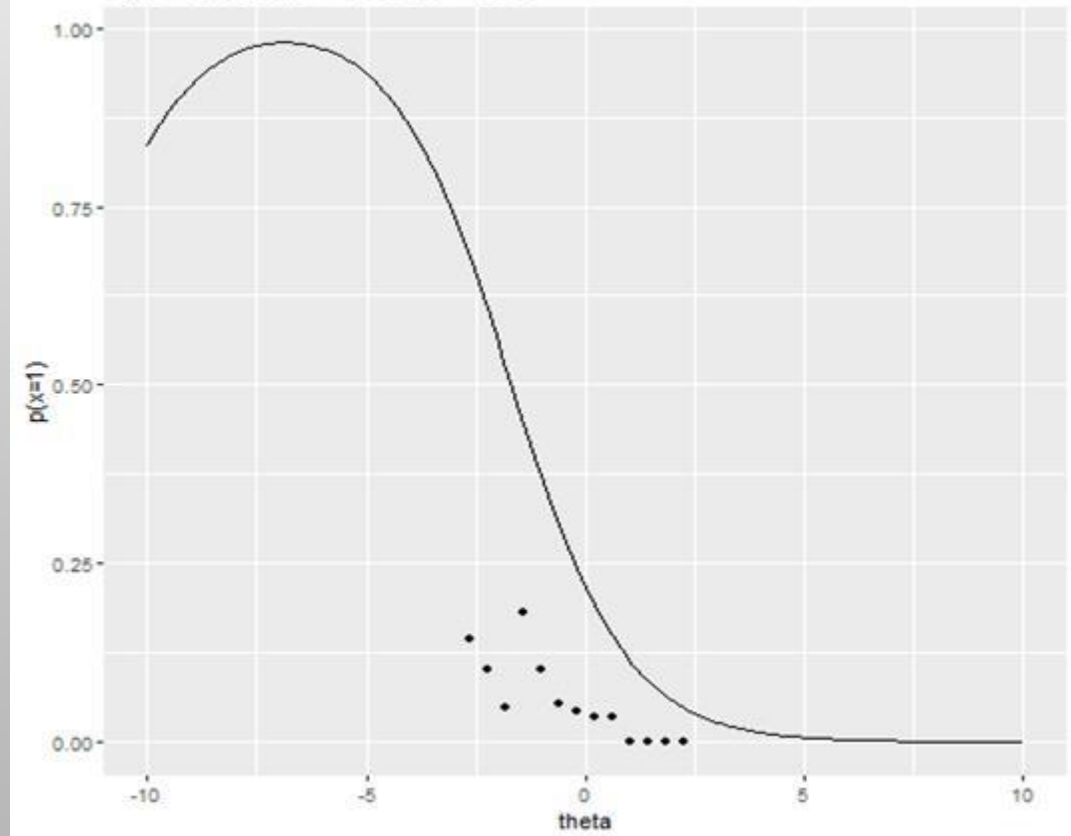
Q15.1#2\_5

alpha = 0.493 delta = -8.693 tau = -3.866

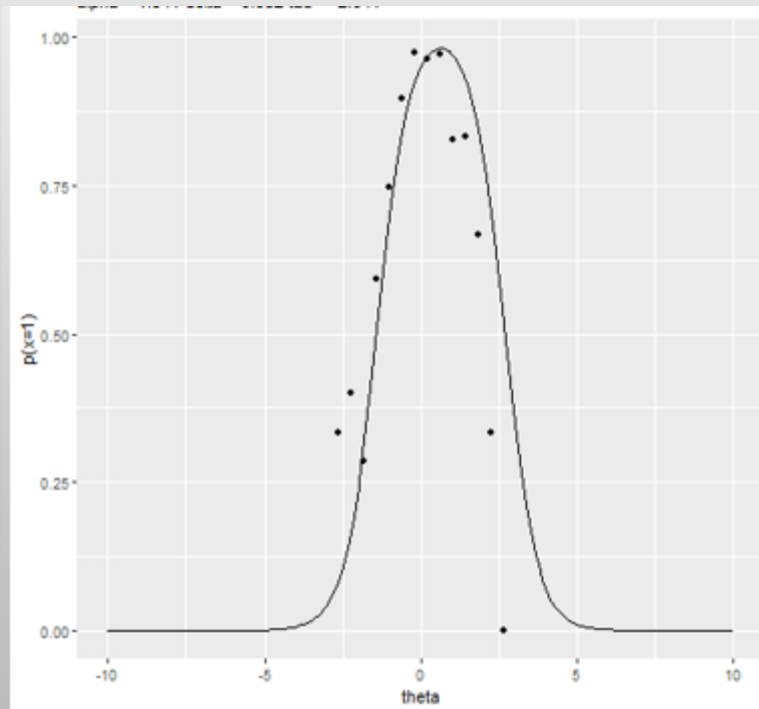


Q9.2#2\_8

alpha = 0.755 delta = -6.865 tau = -5.155



# דוגמא לפריט עם התאמה טובה למודל הפרוש



## עוד קשיים במודל הפרוש

✓ היו יחסית יותר מקרים של אי התאמה במודל הפרוש מאשר במודל הדו פרמטרי.

✓ מאחר שהנתונים הסתדרו רק על צד אחד, עלה החשש שהפרמטר שמציין את מיקום ההיגד על הסקלה (דלתה) והפרמטר שמציין את רוחב הפעמון הם שרירותיים (טאו).

✓ במקרים שבהם נראה היה כי הנתונים התאימו לפעמון (עליה וירידה משני הצדדים), לא נמצאה התאמה תכנית. (כמו בדוגמא הקודמת)

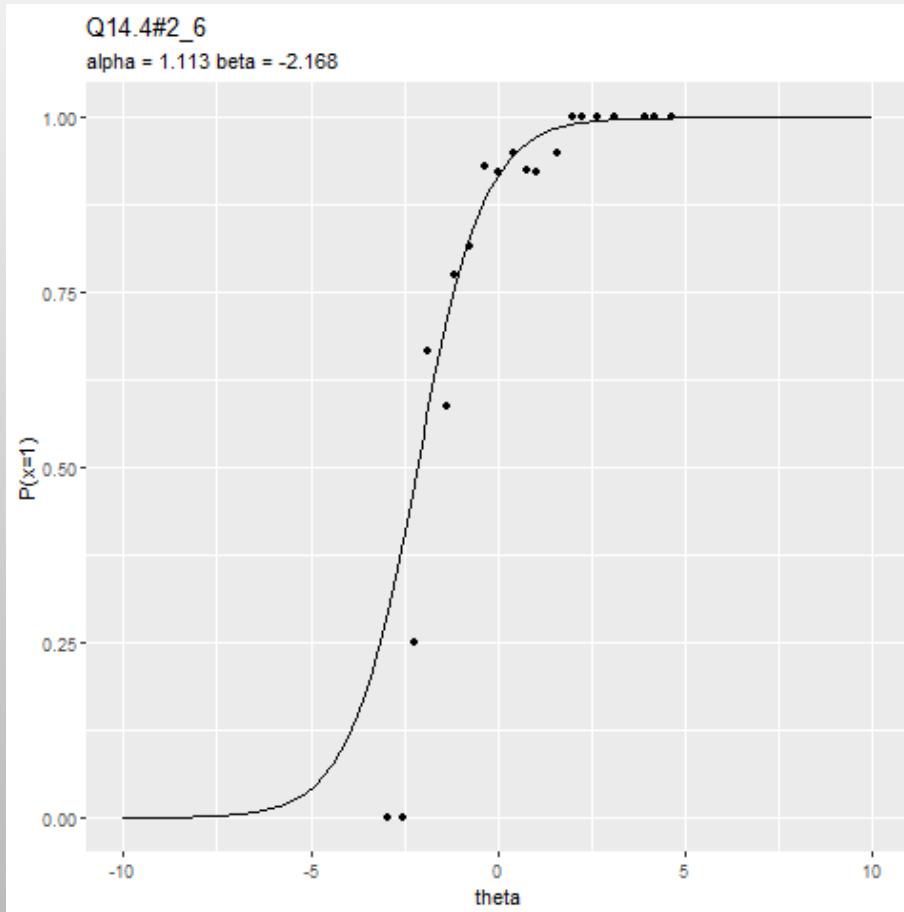
✓ המהלך האדפטיבי של השאלון לא סיפק אותנו. בחלק מהשאלונים, היו שינויים גדולים ("קפיצות") בלתי מוסברים באומדנים של ציוני התכונות, בשלבים מתקדמים של השאלון. עלה החשש שהדבר נובע מכך שישנם שני אזורים בתכונה שבהם צפויה תגובה "אפס".

## תוצאות עבור המודל הדו פרמטרי

- ✓ הניתוח הויזואלי הצביע על התאמה טובה של הנתונים למודל המונוטוני, ככל שהציון בתכונה עלה, כך עלתה ההסתברות לקבל את הפריט (והפוך עבור פריטים מנוסחים שלילית). לא נצפו "ירידות" בטווח הגבוה של הסקלה.
- ✓ מעט פריטים היו בעלי התאמה לא טובה למודל ונפסלו (כ-2% בלבד).
- ✓ בסימולציות של המבדק - האדפטיביות פעלה היטב ובהתאם למצופה. לא נצפו קפיצות מוזרות באומדני הציונים לאורך השאלון, והשינויים מצעד לצעד היו הדרגתיים.



# דוגמא להתאמה טובה של הנתונים למודל



# לסיכום

- ✓ נתוני מרבית ההיגדים התאימו למודל מונוטוני, ולצד אחד מונוטוני במודל הפרוש.
- ✓ בהיגדים הבודדים שבהם הנתונים התאימו למודל הפרוש, לא זיהינו תוכן שמתאים לכך.
- ✓ חלק מהפרמטרים במודל הפרוש נראו שרירותיים.
- ✓ היו "קפיצות" מוזרות באמידת התכונות במהלך השאלון האדפטיבי.
- ✓ מודל ה-2PL מוכר מאוד בספרות ובמאל"ו, למאל"ו נסיון רב בעבודה עם מודלים מונוטוניים.
- ✓ התאמנו את ה-MUPP למודל הדו פרמטרי ובנינו שאלון אדפטיבי על פי השיטה של ה-MUPP. השאלון עבד כהלכה בסימולציות והצלחנו לשחזר את הציונים האמיתיים.
- ✓ בהתאם להחלטה זו, יכולנו לקצר את השאלון בכ-10%

# החלטה

- ✓ הוחלט להשתמש במודל המונוטוני הדו-פרמטרי בשאלון האישיות החדש ולהתאים את האלגוריתם בהתאם (בעיקר חישוב נוסחאות ההתסברות/ הנראות).
- ✓ ייערך נסוי כלים עם השאלון ונבדוק את טיבן של התוצאות האמיתיות, ככל שיצטברו כאלו.
- ✓ השאלון צפוי לעבור כתפעולי ביולי 2023

Thank you 

התייחסויות?

תגובות ?

שאלות?



תודה לאנשי מדור מרום, מדור מחקר, מדור ציונים ומדור מערכות  
מידע במאל"ו וכל מי שלקח חלק בפרוייקט זה.