

## מהימנות האומדנים

הממצאים המובאים בפרסום זה הם אומדנים עבור כלל האוכלוסייה. אומדנים אלה מבוססים על סקר מדגמי ועלולים להיות חשופים לטעויות משני סוגים:

### 1. טעויות מדגמיות

טעויות מדגמיות נובעות מכך שנחקר רק מדגם ולא כלל האוכלוסייה. המדגם שהסקר הנוכחי מבוסס עליו הוא רק אחד מתוך מספר רב של מדגמים אפשריים, שהיו יכולים להתקבל באותה שיטת דגימה ובאותו גודל. ברור כי אומדנים המבוססים על מדגמים שונים יהיו שונים זה מזה, וכמעט כולם יהיו שונים מהערך שהיה מתקבל אילו נערך מפקד – "הערך המפקדי".

**האומדן**  $X'$  הוא הערך הנאמד לפי המדגם המסוים של סקר זה בעבור הערך המקביל לו  $X$ , שהיה מתקבל אילו נערך מפקד מלא.

**טעות דגימה של אומדן**  $\sigma(X')$  היא מידה ממוצעת של השוני בין כל האומדנים השונים שהיו יכולים להתקבל מכל המדגמים האפשריים באותו גודל ובאותה שיטה, לבין הערך שהיה מתקבל לו נערך מפקד מלא, באותם תנאים של איסוף הנתונים.

כיוון שברור שאומדן מסוים סוטה (כמעט בוודאות) מהערך המפקדי, לא רצוי להתייחס בלעדית לערך הנאמד ממדגם אלא להתייחס גם לרווח שבו הערך המפקדי עשוי להימצא בהסתברות נתונה (ברמת ביטחון מסוימת). ניתן לעשות זאת באמצעות **טעות הדגימה**.

**רווח-סמך לאומדן** הוא רווח שמכיל את הערך המפקדי  $X$  ברמת ביטחון מסוימת, נתונה מראש. האומדן  $X'$  על פי המדגם, והאומדן לטעות הדגימה שלו -  $\sigma'(X')$ , מאפשרים לבנות רווח-סמך עם רמת ביטחון מסוימת, נתונה מראש, כך שהרווח מכיל את הערך המפקדי  $X$  בביטחון שנקבע.

מקובל להציג רווחי-סמך ברמת ביטחון של 95%, לכן רווח הסמך ברמת ביטחון זו מתקבל על-ידי הרווח שבין  $X' - 2\sigma'(X')$  לבין  $X' + 2\sigma'(X')$ . ניתן לומר בביטחון של 95% שהערך המפקדי נמצא ברווח זה.

ניתן לקבוע רמת ביטחון גבוהה יותר או נמוכה יותר, ולחשב את רווח-הסמך כך:

$\alpha$	67%	80%	90%	95%	99.5%
$K(\alpha)$	1.0	1.3	1.7	2.0	2.8

כאשר  $K(\alpha)$  (שהוא מספר טעויות הדגימה לכל כיוון) נקבע בהתאם לרמת הביטחון הדרושה,  $\alpha$ .

רווחי-הסמך הם בדרך כלל סימטריים סביב האומדן, אך בעבור אומדנים המבוססים על מספר קטן של מקרים במדגם (פחות מ-40), רווחי-הסמך אינם סימטריים. במקרים אלה, הן האומדן עצמו והן האומדן לטעות הדגימה שלו חשופים לטעות גבוהה.

בנספח המתודולוגי מובאים לוחות של טעויות דגימה **מקורבות** לאומדנים השונים המוצגים בפרסום בליווי הסבר על אופן השימוש בלוחות, הדגמת בניית רווחי-סמך והשוואת אומדנים מקבוצות זרות.

אומדני טעויות דגימה **מקורבות** נקבעו לפי השפעת **אפקט התכנון** על המדגם בקבוצות שונות של האוכלוסייה. **אפקט התכנון** על המדגם מוגדר על-ידי היחס שבין שונות הדגימה (טעות הדגימה בריבוע) של האומדן על-פי המדגם המורכב, לבין שונות הדגימה של האומדן אילו התקבל ממדגם מקרי פשוט באותו גודל.

יש להעיר, שלעתים נוח להעריך את דיוק האומדן לפי **טעות הדגימה היחסית** המוגדרת כטעות הדגימה של האומדן, חלקי הערך הנאמד.

כדי להזהיר את הקורא משימוש באומדנים החשופים לטעויות גבוהות, הוצגו אומדנים עם טעויות דגימה יחסיות של למעלה מ-20% (מספרים קטנים מ-2,000) בסוגריים עגולים ( ).

## 2. טעויות לא-מדגמיות

טעויות לא-מדגמיות בסקר יכולות לנבוע ממקורות רבים בכל שלבי איסוף הנתונים ועיבודם, והן מצויות גם במפקד מלא.

טעויות לא-מדגמיות עיקריות בסקר הן:

א. **טעויות הנובעות מאי-היענות:** טעויות הנובעות מכך שמשקי בית אינם נחקרים בגלל היעדרות או סירוב. דבר זה עלול לגרום להטיה מסוימת באומדנים, משום שתכונותיהם של האנשים הכלולים במשקי בית אלה יכולות להיות שונות מאלה של אנשים שנחקרו.

ב. **טעויות תשובה:** טעויות הנובעות מאי הבנת השאלות, חוסר רצון או חוסר יכולת להשיב נכון או הצגת שאלה בדרך לא נכונה.

ג. **טעויות בשלבי העיבוד:** טעויות בתהליך עיבוד החומר, כגון: טעויות בסימול ובקלידת השאלונים. חלק מהטעויות מתוקנות באמצעות ביקורות הנערכות על החומר.

ד. חלק ממשק הבית נחקר בשבוע השונה מ"השבוע הקובע" (כפי שהוגדר בסעיף 4 בפרק ב). דבר זה גורם גם הוא להטיה מסוימת באומדנים.

בעוד שאת טעויות הדגימה ניתן לאמוד מתוך נתוני הסקר עצמו, קשה (ואף בלתי אפשרי) להעריך טעויות לא-מדגמיות. בפרסום זה, אין מביאים הערכות כמותיות על הטעויות הלא-מדגמיות. חשוב להדגיש, שבתכנון הסקר ובביצועו נעשה מאמץ להקטין ככל האפשר את מספר הטעויות.