

תהליך דגימה במפקד אוכלוסין 2022

1. מבוא

מפקד אוכלוסין מעצם מהותו הוא ספירה של אנשים, אך תוצאת הספירה אינה מטרה בפני עצמה אלא אמצעי להשגת מטרות אחרות המתרחבות ומשתנות על פני זמן. יתר על כן, כמעט שלא היו מפקדים בעולם שבהם היה אפשר להסתפק בספירה בלבד. המפקדים בישראל מאז 1961 זהים במטרות זה לזה כמו גם למפקדים מודרניים בעולם: ספירת האוכלוסייה ואפיונה, דמוגרפיה וחברתית-כלכלית, בקבוצות אוכלוסייה קטנות ובאזורים גאוגרפיים קטנים עד רמת פירוט של אזור סטטיסטי. ייחודו של המפקד לעומת פעולות סטטיסטיות אחרות הוא ביכולתו לספק מידע מפורט ההכרחי לקביעת מדיניות אפקטיבית בכל תחומי החיים ליחידות גאוגרפיות קטנות.

תהליך הדגימה תוכנן כדי להשיג את המטרות הרבות המצוינות לעיל בדגש על כיסוי כל האוכלוסיות (רגילות או מיוחדות) המתגוררות באופן קבוע במדינת ישראל ביחידות הגאוגרפיות הקטנות ביותר.

2. תהליך דגימה 2.1 מסגרות דגימה

שלוש מסגרות דגימה זרות (כיסוי האוכלוסייה שונה בכל מסגרת) שימשו לדגימה עבור מפקד אוכלוסין 2022:

- מרשם אוכלוסין משופר המעודכן לסוף אוקטובר 2021, שכולל 9,245,000 פרטים ללא אוכלוסייה שנמצאת במוסדות (כ-285,000 בדוים בפזורה וביישובי קבע בדרום וכ-8,960,000 פרטים אחרים);
- רשימת דיירי מוסדות ממפקד המוסדות האחרון ל-31.12.2020 (כ-110,000 דיירים שמשותפים בפקידה במפקד האוכלוסין);
- מרשם "דרכונים" ¹.

2.2 דגימה ממרשם אוכלוסין משופר

תהליך הדגימה כלל את השלבים האלה: קביעת גודל המדגם לאזור סטטיסטי, בניית משפחות מינהליות גרעיניות, קביעת משקל רצוי למשפחות האלה, מיון, חלוקה לשכבות דגימה, הקצאה לשכבות, ודגימה.

¹ יהודים ללא אזרחות ישראלית שנכנסו לישראל ברישיון סטודנט לרוב כדי ללמוד בשיבות, ובני משפחה הנלווים אליהם.

• קביעת סך גודל המדגם לאזור סטטיסטי

במרשם האוכלוסין המשופר נוצרו 3,476 אזורים סטטיסטיים לדגימה לאחר מחיקת אזורים שבהם אין אוכלוסייה או יש אוכלוסייה קטנה מאוד, והוספת 19 אזורים (פוליגונים) עבור אוכלוסייה בפזורה בדווית בדרום ואזורים פיקטיביים המיועדים לדגימה בלבד עבור פרטים שהאזור שלהם נמחק.

גודל המדגם נקבע לכ-7% מאוכלוסיית מדינת ישראל. ככל שאזור הדגימה קטן יותר הוגדל אחוז הדגימה בו כדי לאפשר דיוק דומה של גודל המדגם בכל אזור. באזורים שבהם עד 65 נפשות נדגמו נציגים של כל המשפחות הגרעיניות, ובאלה שבהם 66 עד 130 נפשות נדגמו 65 נציגים של משפחות גרעיניות. גודל המדגם עולה בצורת פרבולה הפוכה באזורים שבהם 131 עד כ-10,000 תושבים. באזורים גדולים חריגים שבהם מעל 10,000 תושבים גודל המדגם עולה באופן ליניארי מקדקוד הפרבולה במקדם 0.019. ביישובים ערביים ללא אזורים סטטיסטיים המוגדרים היטב במסגרת ובאזורים פיקטיביים שנוצרו (איחוד של אזורים קטנים ופרטים ללא עיגון) נדגמו מעל 6.5% מהפרטים.

• קביעת משפחות גרעיניות לדגימה

נקבעו משפחות גרעיניות (של הורים וילדים) שכל חבריהן רשומים באותו אזור, על פי יחסי קרבת משפחה, למשל, גם אם סבא וסבתא מתגוררים באותה כתובת של ילדיהם ונכדיהם, נקבעה להם משפחה גרעינית נפרדת. נוסף לכך, גרושים ללא ילדים ורווקים יהודים מעל גיל 22 אשר רשומים יחד עם המשפחה שלהם מהווים גם הם משפחה גרעינית נפרדת. בסך הכל נוצרו כ-3,570,000 משפחות גרעיניות.

לכל משפחה גרעינית נקבע משקל רצוי המבטא את פרופורציית הסתברויות הדגימה של המשפחות השונות באותו אזור. משמעות הדבר היא שהמשקל שונה מאדם לאדם, לפי גודל משפחתו. נציין כי חשוב לקבוע משקל זהה לפרטים במשקי בית לצורכי סטטיסטיקה של משקי בית ואין הצדקה לתגבר משקי בית גדולים. נציין כי המשקל היחסי שנבחר למשפחה גרעינית דומה לנפש סטנדרטית, והוא מוגדל לצעירים בני 25–35 (רישומם במרשם אינו מדויק) ומוקטן למבוגרים מעל גיל 65 ומעלה (רישומם נכון). הגיל הקובע במשפחה הוא של האם או של ההורה במשפחה חד-הורית.

המשקל היחסי למשפחה גרעינית חושב לפי הטבלה הזאת:

גיל קובע >25	גיל >25 קובע >35	גיל >35 קובע >65	גיל >65 קובע	גודל משפחה גרעינית לדגימה
1	1.1	1	0.9	1
2.2	2.2	2	1.8	2
2.915	2.915	2.65	*2.65	3
3.52	3.52	3.2	3.2	4
4.125	4.125	3.75	3.75	5
4.25	4.25	4.25	4.25	6
4.75	4.75	4.75	4.75	7
5.2	5.2	5.2	5.2	8+

* משפחות גרעיניות החסרות בקובץ המינהלי מסומנות באפור.

הסתברות הדגימה הראשונית הרצויה למשפחה היא סך גודל המדגם באזור כפול המשקל היחסי של המשפחה חלקי סך המשקלות היחסיים של כל המשפחות באזור.

לצורך קביעת השכבות, מיינו את המשפחות הגרעיניות באזור (מיון "נחש כפול") לפי:

1. גודל משפחה גרעינית מינהלית;
2. בתוכו, משפחה חד-הורית בנפרד ממשפחות שבהן שני הורים;
3. בתוכו, לפי גיל קובע בשיטת "נחש".

המיון נעשה כך שמשפחות סמוכות דומות אחת לשנייה לפחות בשתי תכונות קובעות וברוב רובם של המקרים בכל 3 התכונות הנ"ל.

• חלוקה לשכבות דגימה

החלוקה לשכבות נעשית תוך שילוב של רעיונות מדגימה שיטתית עם דגימה רנדומלית בשכבות בכל אזור סטטיסטי בנפרד. לכל משפחה גרעינית באזור מוצמד משקל קובע של המשפחה חלקי סך כל המשקלות באזור (הסתברות דגימה ראשונית). האזור מחולק למספר מקסימלי של שכבות באופן שהמשפחות לפי המיון מהוות ביחד שכבת דגימה אחת והסתברויות הדגימה הראשוניות הרצויות שלהן מסתכמות ב-2 בקירוב.

• דגימה

מכל שכבת דגימה נדגם ממ"פ (מדגם מקרי פשוט) ללא החזרה של 2 משפחות גרעיניות, כלומר, בהסתברות של שתיים חלקי מספר משפחות גרעיניות בשכבה (גודל השכבה). מכל משפחה (שנדגמה) נדגם נציג אחד (ממ"פ בתוך המשפחה) בהסתברות

של 1 חלקי גודל משפחה. הסתברות דגימה הסופית של פרט הוא מכפלה של שתי ההסתברויות.

בסופו של תהליך זה, נדגמו בסך הכל 675,353 פרטים ממרשם האוכלוסין ללא דיירי מוסדות ו"דרכונים".

בלוח 1 מובאת התפלגות המדגם לפי גודל משפחה גרעינית (קטום ב-6), מין וקבוצת גיל.

לוח 1: התפלגות המדגם, לפי גודל משפחה גרעינית, מין וקבוצת גיל

גודל משפחה גרעינית	קבוצת גיל	מספר נדגמים גברים	מספר נדגמות	מספר נדגמים – סך הכל
1	19-0	2,017	1,979	3,996
1	29-20	33,342	27,206	60,548
1	39-30	22,847	11,059	33,906
1	59-40	22,851	13,813	36,664
1	+60	15,466	31,691	47,157
2	19-0	7,387	7,030	14,417
2	29-20	2,880	5,925	8,805
2	39-30	2,505	5,121	7,626
2	59-40	7,205	16,887	24,092
2	+60	35,737	32,043	67,780
3	19-0	15,511	14,897	30,408
3	29-20	4,417	6,251	10,668
3	39-30	4,693	6,080	10,773
3	59-40	8,229	13,016	21,245
3	+60	3,219	1,638	4,857
4	19-0	24,590	23,183	47,773
4	29-20	4,319	5,339	9,658
4	39-30	7,291	8,640	15,931
4	59-40	11,611	11,989	23,600
4	+60	1,224	485	1,709
5	19-0	27,267	25,475	52,742
5	29-20	3,447	3,578	7,025
5	39-30	5,001	6,753	11,754
5	59-40	11,589	10,196	21,785
5	+60	543	158	701
6	19-0	33,042	32,704	65,746
6	29-20	3,323	3,313	6,636
6	39-30	4,455	6,180	10,635
6	59-40	9,030	7,349	16,379
6	+60	274	63	337
סך הכל		335,312	340,041	675,353

דגימת דיירי מוסדות

מסגרת לדגימת מוסדות נבנתה על בסיס מפקד מוסדות (100%) ל-31 בדצמבר 2020, שכלל כ-185,000 פרטים שמתגוררים בהם. לצורך הדגימה, חלק מהמוסדות, כגון מוסדות רווחה ומוסדות קטנים מאוד לא נכללו כלל במסגרת. נוסף לכך, הוצאו מהמסגרת אנשי צוות במוסד, ילדים במוסדות חינוך שהוריהם מכסים בדגימה ממרשם אוכלוסין משופר, ועוד. לאחר כל ההורדות התקבלה מסגרת לדגימה הכוללת כ-110,000 פרטים.

• שכבות דגימה

כדי להגדיר את שכבות דגימה זוהו משפחות במוסדות לפי יחס קרבה הורה-ילד ו/או בני זוג והוגדרו סוגי משפחות במוסד: בודדים רווקים, בודדים שאינם רווקים, זוג, הורה-ילד, ועוד. שכבות דגימה הן חתך של מזהה מוסד וסוג משפחה בו. נציין כי במוסדות יש לכל היותר שני סוגים של משפחות, לכן בכל שכבה יש מספיק יחידות דגימה לצורך הדגימה.

בסך הכל הוגדרו 1,280 שכבות שמתוכן 944 מוגדרות בסמל מוסד בלבד (המוסד מהווה שכבה). שאר 168 המוסדות חולקו כל אחד ל-2 שכבות (סך הכל 336 שכבות).

• דגימה

הקצאת גודל מדגם בשכבות, הוגדרה בעזרת פונקציה קמורה כלפי מעלה, כפי שנעשה עבור אוכלוסייה רגילה, כמובן בפרמטרים שונים. המדגם הופק בשני שלבים: בשלב הראשון דגימת ממ"פ ללא החזרה של משפחות בשכבות לפי ההקצאה שנקבעה באמצעות הפונקציה ובשלב השני דגימת ממ"פ של נציג משפחה. הסתברות הדגימה הסופית היא מכפלה של שתי הסתברויות שהתקבלו בשלב הראשון ובשלב השני.

בשל מרווח הזמן הגדול בין בניית מסגרת הדגימה (31 בדצמבר 2020) עד התחלת המפקד (2 באפריל 2022) נעשו תיקונים הן במסגרת והן במדגם. מהמסגרת הוצאו נפטרים (כ-2,500) וילדים (כ-2,600) שהוריהם מכסים.

גודל המדגם הסופי הוא כ-15,000 דיירי מוסדות, שמהווים כ-14.8% ממסגרת הדגימה המתוקנת, שכללה כ-105,300 דיירי מוסדות. כמובן, הסתברויות דגימה תוקנו בעקבות השינויים האלה.

2.3 דגימת זרים ("דרכוניסטים")

לצורך מפקד האוכלוסין 2022 הוחלט לא לדגום מתוך כלל הזרים הנמצאים בארץ באופן קבוע אלא רק מתוך קבוצת ה"דרכוניסטים". המסגרת המעודכנת של קבוצה זו לסוף שנת 2021 כללה כ-20,800 פרטים. נבנו 6,314 משפחות גרעיניות מינהליות (ללא כתובות) שבהן מבוגר אחד לפחות. 327 משפחות ללא מבוגרים הוסרו מהמסגרת.

- חלוקת המשפחות לשכבות נעשתה לפי גודל המשפחה הגרעינית המינהלית:

שכבה	שכיחות	אחוז
1	1,931	30.58
2	923	14.62
3	1,189	18.83
4	824	13.05
5	510	8.08
6+	937	14.84
סך הכל	6,314	100

• דגימה

בשלב הראשון דגימת ממ"פ ללא החזרה של 24% משפחות בכל שכבה ובשלב השני דגימת ממ"פ של נציג משפחה (בכל משפחה נציג בהסתברות אחת חלקי גודל המשפחה). הסתברות הדגימה הסופית היא מכפלה של שתי ההסתברויות שהתקבלו בשלב הראשון ובשלב השני.

בסך הכל נדגמו כ-1,500 פרטים מקבוצה זו, שהם כ-7% מהמסגרת.