

סקרים

בתנאים אופטימליים

ולמעשה - 2

סקרים וזעת קהל

"סקרים ודעת קהל" – סילבוס

- ◆ סקרים בתנאים אופטימליים – המדגם והאוכלוסייה – רמת הדיוק והטעות המקרית
- ◆ אם כל כך טוב למה כל כך רע? – הביקורת על הסקרים כלל והסקרים בישראל בפרט, הצעות חוק לאיסור פרסום תוצאות סקרים לפני בחירות
- ◆ הסקרים כמשפיעים על דעת הקהל – רק שלילה?
- ◆ הבסיס המדעי לעריכת הסקרים – תורת הדגימה, דגימה מקרית פשוטה, דגימת אשכולות (משקי בית), דגימת שכבות (אזורים גיאוגרפיים) – חלוקה לשכבות לאחר הסקר והשקלול בסקרים
- ◆ הסקרים המעשיים - סקרים טלפוניים, סקרים פנים אל פנים, וקבוצות מיקוד
- ◆ הסקרים כמקרה פרטי של מחקרים סטטיסטיים – המחקר הגדול ביותר התולדות האנושות
- ◆ "הסקרים האחרים" – סקרי שביעות רצון במסחר ובתעשייה, סקרים לקביעת מדיניות מסחרית
- ◆ הטיות בסקרים, חוסר תשובה, אורך השאלון, ניסוח השאלות, ומיקום השאלה בשאלון.
- ◆ בעיות מבניות – "רצייה חברתית", "לשאל את חסר המידע"
- ◆ סקרים ככלי לעיצוב אסטרטגית בחירות וקביעת מדיניות ציבורית
- ◆ סקרים טלפונים במערכת בחירות ומדגם קלפיות
- ◆ השפעת מבנה השאלה על המשיב: "השאלה הפתוחה", "השאלה הסגורה" ו"השאלה החצי פתוחה"
- ◆ השפעת מיקום השאלה על התשובות בשאלונים טלפוניים וויזואליים
- ◆ מניפולציה של הסקרים והתקשורת - על הסקרים כברומטרים מוטעים לדעת קהל
- ◆ מניפולציה של הסקרים- המשך- פירסום ופירסומת בעזרת סקרים
- ◆ **רשימה ביבליוגרפית:**
- ◆ פוקס ק., בר-לב ש. "אמת וסקר" הוצאת הקיבוץ המאוחד ואוניברסיטת חיפה, 1998
- ◆ Cochran, W.G., "Sampling techniques", 3rd Edition, Wiley, NY, 1977
- ◆ Waksberg, J. "Sampling Methods for Random Digit Dialing", Journal of the American Statistical Association, 1978, 40-46
- ◆ Utts, J. "Seeing through statistics", 3rd Edition, Thomson Brooks/Cole, CA, 2005
- ◆ Crossen, C. "Tainted Truth – The manipulation of fact in America"

אמירות הסקרים התקפים

מושגי יסוד-סיכום

◆ **פרמטר** - הערך באוכלוסייה (20%), אינו ידוע לעורך המחקר (וכנראה לעומם לא יהיה ידוע, מלבד בסימולציה)

■ - ערך המשתנה נחקר עליו יש להסיק מהמדגם **אחוז**
ה"מיוחדים באוכלוסייה"

◆ **האומדן: תוצאת המדגם** האומדת את ערך **הפרמטר**

◆ **טווח הטעות המקרית (א')**: טווח **מסביב לערך הפרמטר** עליו נוכל לומר (ברמת בטחון נתונה, למשל 95%) שתוצאת המדגם **הבודד** תימצא (לבדיקה דרוש **ערך הפרמטר**)

◆ **טווח הטעות המקרית (ב')** "רווח בר סמך": טווח **מסביב תוצאת המדגם** עליו נוכל לומר (ברמת בטחון נתונה, למשל 95%) שהוא "מכסה" (כולל) את **ערך הפרמטר**

◆ בד"כ **טווח הטעות המקרית** נתון בצורה של "שיעור הטעות לא יעלה על $X\%$ " או "שיעור הטעות הוא פלוס מינוס $X\%$ "

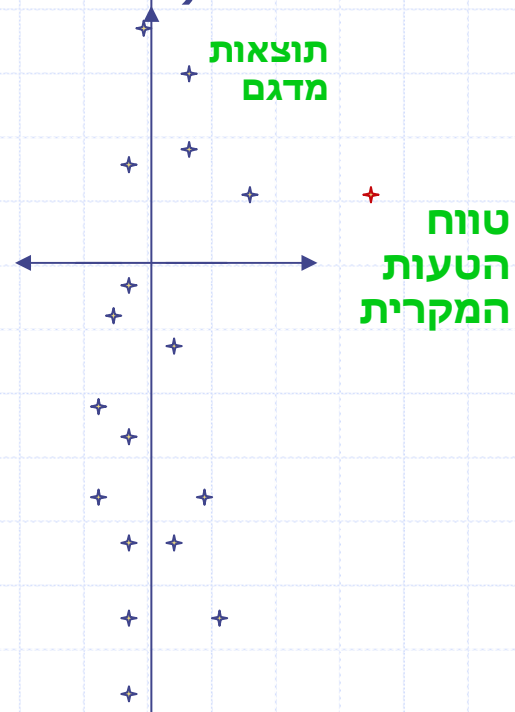
◆ **רמת בטחון (של 95%)**: לו לקחנו מדגמים רבים, אזי:

■ ב 95% מהפעמים (בערך) ה**טווח מסביב תוצאת המדגם**, יכיל את **ערך הפרמטר** ו

■ ב 5% מהפעמים ה**טווח מסביב תוצאת המדגם** לא יכיל את **ערך הפרמטר**

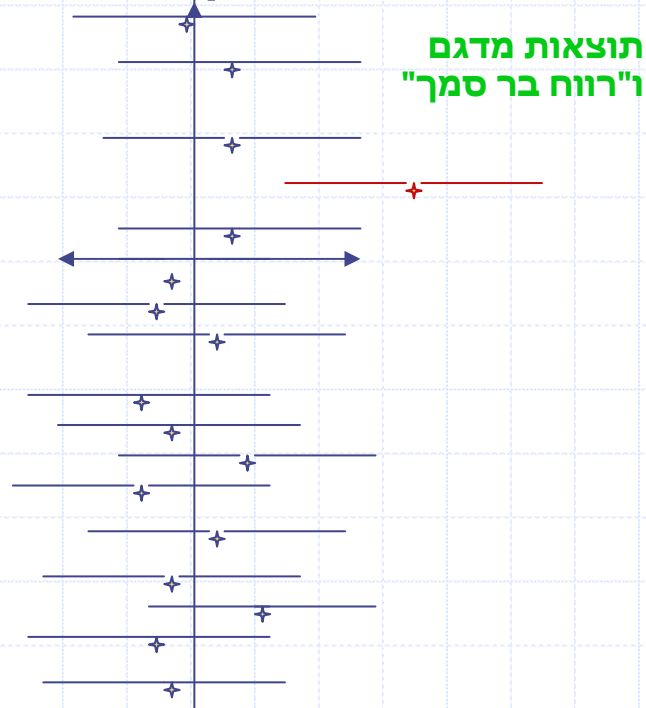
תוצאות מדגמים,
טווח טעות מקרית, "בסיכוי של 95% תוצאת המדגם.."

הפרמטר (20% - הידוע בסימולציה)



תוצאות מדגמים,
טעות מקרית מסביב לתוצאת מדגם
"רווח בר סמך" (אורך רווח משתנה), "בסיכוי של 95% מכסה."

הפרמטר (20% - הלא ידוע)



סיכום תוצאות הסימולציות

מדגמים של 4800	מדגמים של 1200	
1.13%	2.26%	שיעור הטעות (בסיכוי של 95%)
18.87%-21.13%	17.74%-22.26%	הטווח (מסביב ל 20%) – לטעות מקרית בסיכוי של 95%
185	188	מספר המדגמים (מתוך 200) בטווח
92.5%	94%	אחוז המדגמים (מתוך 200) בטווח
100%	100%	אחוז המדגמים (מתוך 200) בטווח 17% - 23%

תוצאות הסקרים בתקשורת

לעתים קרובות בליווי "שיעור הטעות המירבית"

- כללית, אין דבר כזה

- שיעור הטעות – רק:

- בסיכוי מסויים (כמו 95%)

- ומחושב תחת הנחה שהמדגם נלקח תחת התנאים של הדגימה ההסתברותית

- **כיצד יודעים אם התנאים מתקיימים ומה שיעור הסטייה באמת?**

- בדרך כלל לא יודעים, האוכלוסייה נשארת "נעלמת"

- יוצאים מן הכלל בודדים: תוצאות האוכלוסייה מתפרסמות:

- מדד המחירים לצרכן

- תוצאות הבחירות

"תוצאות סקר" מול "תוצאות אמת"

אך השוואה בין ערכי התחזיות ל"תוצאות האמת" אינה מאורע שכיח. ובאין השוואה עם "תוצאות האמת", לא מתעוררת גם השאלה של אובדן אמינות. אלא שבסקרים בהם התוצאות "האמיתיות" לא נודעות לעולם, "ההפסד החברתי" מעריכת סקרים לא תקפים אינו קטן יותר, אלא להיפך, הוא עשוי להיות גדול פי כמה מאובדן האמינות. ההחלטות בשטחים השונים המתבססות על תוצאות סקרים שאינם תקפים, עלולות להיות שגויות וליצור נזק כלכלי, חברתי וגם אחר. כאשר אין השוואה ל"תוצאות האמת" אין גם הפקת לקחים, ואין גם "שכר ועונש" לסוקרים לפי טיב מלאכתם.

חוסר היכולת לשפוט את "טיב המלאכה"

חוסר היכולת לשפוט את טיב מלאכתו של עוסק במקצוע כלשהו לפי התוצאות, הוא מצב מיוחד. קשה למצוא מקצועות רבים כאלה. האם יכול מישהו לדמיין עוסק ברפואה, או בעיתונאות, או במכונאות או בכל מקצוע אחר, שטיב עבודתו נישאר עלום? אך בעריכת סקרים זהו אכן המצב. לא ניתן לבטל את ההשערה שלתפוצתם של הסקרים הלא תקפים יש מיתאם עם חוסר היכולת להעריך את מלאכתם של עורכי הסקרים על פי טיב תוצאותיהם.

אמינות הסקרים – סימולציה מעשית

◆ **האוכלוסייה** - אינדונזיה – נניח 100,000,000 (מאה מיליון) -
המאופיינת חד משמעית על ידי מספר ת"ז (הנע בין 1 ל
(100,000,000

◆ **מדגם (מדגמים)** של 1,200 אנשים (ואחר כך של 4,800)

◆ **ה"תכונה"** מאופיינת על ידי הספרה האחרונה של ת"ז : אם
הספרה היא 0 או 5 אז יש התכונה, אחרת אין

■ או אם כל המספר מתחלק ב 5 יש התכונה)

◆ אחוז ה"מיוחדים באוכלוסייה" – 20%

◆ **נבדוק** הסטייה בין התוצאה המדגם לערך ה"אמיתי"
באוכלוסייה – כיצד "שיעור הטעות לא יעלה על 3%" (או
פחות, בהתאם לגודל המדגם ולשיעור ה"מיוחדים באוכלוסייה")

◆ **נבדוק** מה פרוש האמרה ברמת בטחון של 95% על ידי :

■ לקיחת 200 מדגמים של 1,200 כל אחד (או של 4,800)

■ בדיקת אחוז המדגמים הנמצאים בטווח הטעות (95% ?)

אמינות הסקרים בתנאים אופטימליים-תזכורת

י"עד כמה קרובים (או רחוקים) עשויים להיות האומדנים המתבססים על תוצאות הסקר מהפרמטרים של האוכלוסייה? או, במילים אחרות, מהי מידת האמינות של התוצאות המתקבלות מסקר?

...דוגמה היפוטטית. **נניח** שברצוננו לאמוד את אחוז בעלי האופניים מבין תושבי אינדונזיה, או לאמוד את אחוז תושבי אינדונזיה שגובהם מעל 170 ס"מ, או כל תכונה אחרת באוכלוסיית אינדונזיה, שהתשובה אליה מסווגת לאחת משתי אפשרויות (למשל בעל או חסר אופניים).

נניח שמצויה בידינו רשימה מדויקת של כל תושבי אינדונזיה,

נניח שמתוך הרשימה נדגמים **בכלים הסתברותיים**, 1,200 תושבים,

נניח שכל הנשאלים עונים על שאלת הסקר,

נניח שכל הנשאלים עונים במדויק על שאלת הסקר.

אזי היינו יכולים לטעון **ברמת בטחון של 95% כי :**

הסטייה בין האומדן שהיה מתקבל ממדגם קטן זה לבין הפרופורציה האמיתית באוכלוסייה של בעלי האופניים באוכלוסיית אינדונזיה העצומה **לא תעלה על 3%.**

אמינות הסקרים – סימולציה מעשית

❖ כיצד בודקים בסימולציה את נכונות האמרה

”שיעור הטעות לא יעלה על 3%” (או פחות,

בהתאם לגודל המדגם ובהתאם לשיעור

ה”מיוחדים באוכלוסייה”)?

❖ כיצד בודקים בסימולציה את נכונות את האמרה

”ברמת בטחון של 95%”?

הסקרים בתנאים אופטימליים

הרחבות:

- ◆ מה קורה כאשר לשאלה במקום 2 אפשרויות (כן, לא) – יש יותר אפשרויות (כמו בבחירת מועמדים, או שביעות רצון ב 5 קטגוריות)?
- ◆ מה קורה כאשר יש יותר משאלה אחת בסקר?
- ◆ התשובה הדיכוטומית לשאלה הספציפית כפתרון ביניים
- ◆ ציטוט מ"הארץ": "הסקר נערך בקרב....בסיכוי של 95% הטעות המקרית בכל שאלה בנפרד לא עולה על...."

תכנוני דגימה "משוכללים" יותר

- דגימת שכבות (Stratified)
- חלוקה לשכבות לאחר דגימה (Post-stratification)
- חלוקה לשכבות לאחר הדגימה דו-שלבית
(Two-fold Post-stratification)–
- בשלב ראשון דגימת שכבות רגילה (שכבות דמוגרפיות
(ערבים, עולים, ותיקים, מצביעים בפעם ראשונה)
- בשלב השני Post (למשל לפי הצבעה בבחירות קודמות)
- סקרים בריאות בארה"ב *Health Interview Survey*
- דגימת מכסות (Quota sampling)

דגימת מכסות - (קירוב לדגימת שכבות)

- אם אין מסגרות דגימה (רשימות, בסיסי נתונים) לכל שיכבה (כמו למשל, ותיקים, עולים, ערבים, מצביעים פעם ראשונה) ניתן להגיע לגודל מדגם רצוי בכל שכבה על ידי המשך דגימה עד שמתקבלת ה"מכסה"
- דורש מאמץ ניכר בשלבים הסופיים של עבודת השדה
- בדרך כלל ניתן "חופש" גדול יותר בבחירת הנישאלים
- דגימת מכסות - דגימת שכבות עם דגימה לא אקראית בכל שכבה
- מחקרים שהשוו תוצאות של דגימת שכבות ("אמיתית"), (הסתברותית) עם דגימת מכסות מראות:
 - הטייה במאפיינים כמו השכלה, הכנסה, תעסוקה
 - התאמה טובה בנושאים של דעות ועמדות

הנחות בסיסיות

1. האוכלוסייה מוגדרת היטב, ואנחנו דוגמים מתוכה

2. בחלוקה לשכבות לאחר הדגימה, כאשר נידגם פרט, מתקבלים הערכים שלו ללא טעות.

3. הנחה נוספת של המודל:

- כל הנישאל מסכים להשתתף בסקר, או
- אלטרנטיבית, מניחים *Missing at random* MAR (Rubin, 1976) "תצפיות חסרות באקראי" הסיכוי לתצפית שתהיה "חסרה" לא תלויה בערכי הנתונים

תוצאות סקרי כוונות הצבעה לאור התיאוריה

האם $M_i = 1$? סירובים בסקרים טלפוניים

תאריך	משיבים	נישאלים
16.1	530	872
22.1	876	1,168
26.1	864	1,025
	820	1,366
13.11 - מרשימות עבודה	698	646
17.11 - מרשימות ליכוד	535	545
25.11 - מרשימות ליכוד	505	552

האם "חסר באקראי"? סקר I-2002 "הארץ" – Post-stratification תוצאות : הצבעה ב-99

	תוצאות סקר		99
העבודה	25.0%	31.3%	21.7%
הליכוד	32.1%	40.3%	15.1%
ש"ס	5.4%	6.8%	13.9%
מר"צ	4.8%	6.0%	8.1%
ישראל בעליה	1.1%	1.4%	6.0%
שינוי	3.4%	4.3%	5.4%
מפלגת המרכז	0.5%	0.6%	5.9%
מפד"ל	2.2%	2.7%	4.9%
אגודת ישראל	2.3%	2.9%	4.3%
האיחוד הלאומי	1.9%	2.3%	3.5%
ישראל ביתנו	0.6%	0.8%	3.0%
עם אחד	0.5%	0.6%	2.2%
מפלגה אחרת	2.5%		
פתק לבן	0.5%		
מסרב להשיב	5.4%		
לא הצבעתי	11.5%		
			100.0%

כתבה ב"הארץ 29.11.2002

תאריך חלק: 29/11/2002 עיתון: הארץ
 עמוד: 14 מילים: 93
 קר מבחנות: הליכוד זכה ליותר קולות בסקר
 כתב: ורטר ירסי
 לוגו: בחירות 2003
 בותרת על:

טקסט

טקסט

אלו גאקתה דא ובעת תכבה לטור תאקטקרו "הארץ" - "די בליכוד שגילו רבים מהנשאלים.

חירות תשס"ט קוטלונ מרר לאנזמה מסלעב העביו בבחירות תשס"ט קוטלונ מרר לאנזמה מסלעב העביו וזי להבובעבאסלהם שאיגטי מואמיס 1999 את אחוזי להבובעבאסלהם שאיגטי מואמיס 1999 את או פוסלסרפעת הנוכחי תלסדה לטכודס גאפסתיים: הפוסלסרפעת הנוכחי תלסדה לטכודס גאפסתיים: תופלאנשאלטלשבו חאלסורבאלת העבוסולי מעיין תופלאנשאלטלשבו חאלסורבאלת העבוסולי מעיין ולות שהמקרו העביו עליוסוגבה 1999 גרמטיבו לשאלות שהמקרו העביו עליוסוגבה 1999 גרמטיבו לשא שים יטלת המבילסו שלעכצו למפסות אחרות; אנשים יטלת המבילסו שלעכצו למפסות אחרות; או ה בא 2001 אמר לשלסן קבבם רבתו מראשות הממשל לב או שלא בחום לב, כי הצביעו לליכוד.

תוספת בדו"ח שנשלח ל"הארץ"

.....ניתוח התוצאות בסקר הנוכחי נערך תחת ההנחה שתשובות הנישאלים הן אמינות (דהינו, אוכלוסיות מסוימות נוטות יותר מאחרות להשתתף בסקר). אם לעומת זאת, ההסבר השני הוא ההסבר הנכון, (דהינו, יש בעיה באמינות הנתונים) הדבר מוסיף רמה של הטיה בשיעור לא ידוע בתוצאות הסקר (וגם של סקרים אחרים המתפרסמים חדשות לבקרים). הטיה זו מתווספת לטעות המקרית.

שינויים בשאלונים בסקרים הבאים ל"הארץ"

- מעבר מ"שאלה פתוחה" לשאלה סגורה" בדבר ההצבעה ב – 1999
- קביעת סדר הקראת המפלגות כך שהמפלגות "הנשכחות מוזכרות ראשונות
- שלוש המפלגות הראשונות המוזכרות היו :
 - ש"ס
 - איחוד לאומי
 - מפד"ל
- שיעור האנשים שאמרו בסקרים שהצביעו לליכוד ב – 1999 ירד בהמשך מ – 40% ל – 30% (שיעור ההצבעה הכללי ב – 1999 היה, כאמור 15%)
- ולמקורות אחרים של הטייה...

סקר של 1000 איש – "לוקח מכולם"



1 מ-איחוד לאומי, וגם 1 מ-מר"צ,

2 מ-ש"ס וגם 3 מ-שינוי, 1 מ-ליכוד וגם 2 מ-ערבים

האוכלוסייה לא מוגדרת – ראה "ביום הבחירות" ו"בתוך הקלפי"



ולצורת הדגימה בפועל...

The New York Times | CBS NEWS Poll

How the Poll Was Taken

The latest New York Times/CBS News Poll is based on telephone interviews conducted Friday through Sunday with 1,138 adults throughout the United States. Of these, 856 said they were registered to vote.

The sample of telephone exchanges called was randomly selected by a computer from a complete list of more than 36,000 active residential exchanges across the country.

Within each exchange, random digits were added to form a complete telephone number, thus permitting access to both listed and unlisted numbers. Within each household, one adult was designated by a random procedure to be the respondent for the survey.

The results have been weighted to take account of household size and number of telephone lines into the residence and to

adjust for variations in the sample relating to geographic region, race, sex, age, and education.

In theory, in 19 cases out of 20 the results based on such samples will differ by no more than three percentage points in either direction from what would have been obtained by seeking out all American adults.

For smaller subgroups the potential sampling error is larger. For example, it is plus or minus five percentage points for either Democrats or Republicans or independents taken as a separate group.

In addition to sampling error, the practical difficulties of conducting any survey of public opinion may introduce other sources of error into the poll. Variations in question wording or the order of questions, for instance, can lead to somewhat different results.

"אין להניח שרבים מאנשי המקצוע בישראל יחלקו על כך של רמת התכנון של הסקרים (כגון זה דלעיל) שונה מהותית מזו של רוב הסקרים הטלפוניים הנערכים ע"י מכונים בישראל".

סיכום מקורות הטייה עד כה:

- (1) חוסר תשובה דיפרנציאלי (אין MAR)
 - (2) אוכלוסיית היעד (להבדיל מהאוכלוסייה הוירטואלית של "לו הבחירות היו מתקיימות היום") לא מוגדרת היטב בעת עריכת הסקר
 - (3) נתונים חסרים על גודלי תת-אוכלוסיות בחלוקה לשכבות לאחר הדגימה
 - (4) ביצוע לא אופטימלי של הסקר בשטח
 - (5) בעיות באמינות נתונים, הן ב Y – והן ב X
- אבל, בביצוע סקרים לאורך זמן, רמת האמינות של המגמות (דלתה) עשויה להיות גדולה יותר**

תחזיות והתוצאות בפועל....

על הסקר הנכון א-פוסטריות

כן, גם כאשר מתבצעים סקרים רבים, בהם הן שיטת הדגימה והן ניתוח הנתונים אינם מתבצעים בהתאם לשיטות המתאימות, עדיין קיימת הסתברות גבוהה לכך שיימצא לפחות סקר אחד שתוצאתו תהיה קרובה ל"תוצאת האמת". אין להסיק מתוצאותיו של סקר אחד מני רבים ש"קלע לתוצאות האמת" כי השיטות הסטטיסטיות בהן השתמשו מבצעיו היו בהכרח מדעיות. במקרה זה לא חל המבחן שבספר דברים שעל פיו נקבע האם הדבר אשר דיבר הנביא הוא דבר אמת או דבר שקר: "...וכי תאמר בלבבך: איכה נדע את הדבר אשר לא דיברו ה' ? אשר דיברו הנביא בשם ה', ולא יהיה הדבר ולא יבוא – הוא הדבר אשר לא דיברו ה'" (דברים, י"ח, כ"א-כ"ב). אך במציאות של "הרבה נביאים והרבה נבואות", גם אם דברם לא מבוסס כדבעי, הרי בהסתברות גבוהה, אחד ה"דברים" יהיה יבוא באופן מקרי. המבחן הוא בשימוש בשיטות המדעיות המבטיחות אפשרות של חזרה על תהליך האמידה בשיעור טעות דגימה שניתן לצפותו מראש. יש לפעול על פי שיטות מדעיות אלה בכל הסקרים, ובמיוחד בסקרים הרבים הנערכים בחיינו בהם, שלא בדומה לבחירות, לא ניתן להשוות את התחזיות עם "תוצאות האמת".

טעות סטטיסטית לעומת טעות פוליטית

בליל הבחירות.

נתמקד כאן בהיבטים של המתודולוגיה ולא דווקא במשמעות הפוליטית של התוצאות. בהינתן תוצאות הבחירות של 1996, משימה זו היא קשה במיוחד. כך למשל, אילו תחזית הטלוויזיה היתה מבשרת שהמועמד הנבחר קיבל 56% מן הקולות ובפועל הוא היה נבחר ב-55%, היו כולם גומרים את ההלל על דיוק התחזית. לעומת זאת, אילו התחזית היתה מצביעה על כך שהמועמד קיבל 49.5% מן הקולות ובפועל היה מקבל 50.5% או אז הסטטיסטיקאים המעורבים בתחזית היו זקוקים לתמיכה נפשית, וכל זאת בשל "ההחמצה" המכרעת בחיזוי המנצח. כזכור, בתחזית של בחירות 1996 בשעה 22:00, המצב היה אכן קרוב להנחה ההיפותטית השנייה.

ברם, יש לקבוע כי מבחינה סטטיסטית, שתי התחזיות ההיפותטיות דלעיל מדויקות באותה מידה. את שיעור הטעות יש לבדוק כפונקציה של תכנון הדגימה ושל המודל לבניית התחזית ולהשוותו עם שיעורי הטעות המתקבלים בשיטות מחקר אחרות. כמו כן חשוב לבדוק את הקשר בין שיעור טעות הדגימה לבין מהימנות הנתונים. בבעיות אלה עוסק מאמר זה.

מסרים כלליים

- (1) בסקרים הטלפוניים יש הטייה מובנית עם מקורות הטייה שונים (שאלון, חוסר תשובה דיפרנציאלי, תשובות לא אמינות). שיעורי טעות עשויים להיות גדולים (לעיתים במידה ניכרת) מה"טעות המקרית".
- (2) התוצאות לגבי המגמות המתקבלות מסקרים עוקבים עשויים להיות יותר אמינים מהתוצאות של הסקרים הבודדים (ראה מכוני דעת קהל כמו באוניברסיטת מישיגן *Ann Arbor*).
- (3) יש צורך במחקר מתודולוגי נוסף לנוכח הגידול הניכר בשיעורי "חוסר תשובה" בשאלונים טלפוניים.
- (4) אבל, אין אלטרנטיבה טובה לסקרים טלפוניים.