

תוכן העניינים

מ ב ו א

1. מהי סטטיסטיקה? 3
- 1.1 סטטיסטיקה תיאורית..... 3
- 1.2 חסקה סטטיסטית..... 3
2. אוכלוסיית המחקר והמשתנים..... 5
- 2.1 אוכלוסיית מחקר..... 5
- 2.2 אוכלוסייה סטטיסטית..... 6
- 2.3 משתנים..... 7
- 2.4 מיון משתנים..... 7
- 2.4.1 מיון לפי מהות המשתנה..... 7
- 2.4.2 מיון לפי רמת המדידה..... 9
- 2.4.3 השוואה בין שיטות המיון..... 12
- 2.4.4 מיון לפי כיוון התיאור..... 12
- 2.5 כלליות המחקר הסטטיסטי..... 13
- מושגים לחזרה..... 15
- תרגילים - פרק 2..... 16

ס ט י ס ט י ק ה ת י א ו ר י ת

3. לוחות סטטיסטיים..... 17
- 3.1 לוח חד-מימדי..... 17
- 3.2 לוח דו-מימדי..... 21
- 3.3 לוח תלת-מימדי..... 22
- מושגים לחזרה..... 24
- תרגילים - פרק 3..... 25
4. הצגה גרפית..... 27
- 4.1 דיאגרמת המעגל..... 27
- 4.2 דיאגרמת מקלות..... 29

30	4.3 היסטוגרמה
36	4.4 מצולע שכיחויות
37	4.5 צורת ההתפלגות
41	מושגים לחזרה
42	תרגילים - פרק 4
47	5. ערכים מרכזיים
47	5.1 השכיח
47	5.1.1 השכיח למשתנה איכותי וכמותי-בדיד
47	5.1.2 השכיח למשתנה כמותי-רציף
50	5.1.3 תכונות השכיח
51	5.2 החציון
51	5.2.1 החציון בסדרת ערכים בודדים
53	5.2.2 החציון בטבלת שכיחויות
57	5.2.3 תכונות החציון
58	5.3 הממוצע האריתמטי
59	5.3.1 הממוצע בסדרת ערכים בודדים
60	5.3.2 הסימן Σ
62	5.3.3 הממוצע בטבלת שכיחויות
65	5.3.4 תכונות הממוצע
68	5.4 השוואה בין הערכים המרכזיים
68	5.4.1 לפי סוג המשתנה
69	5.4.2 לפי צורת ההתפלגות
72	5.4.3 לפי פונקצית ההפסד בניבוי
77	מושגים לחזרה
78	תרגילים - פרק 5
83	6. מדדי פיזור
83	6.1 התחום
84	6.1.1 חישוב התחום
84	6.1.2 תכונות התחום
84	6.2 התחום הבין רביעוני
86	6.2.1 הרביעונים למשתנה בדיד
87	6.2.2 הרביעונים למשתנה רציף
90	6.2.3 תכונות התחום הבין רביעוני
90	6.3 ערכי חלוקה

94	6.4 ממוצע הסטיות המוחלטות
94	6.4.1 ממוצע הסטיות המוחלטות בסדרת ערכים בודדים
96	6.4.2 ממוצע הסטיות המוחלטות בטבלת שכיחויות
98	6.5 חשונות וסטיית התקן
99	6.5.1 חשונות וסטיית התקן בסדרת ערכים בודדים
102	6.5.2 חשונות וסטיית התקן בטבלת שכיחויות
104	6.5.3 דרך נוספת לחישוב חשונות וסטיית התקן
108	6.5.4 סטיית התקן למשתנה רציף
110	6.5.5 תכונות חשונות וסטיית התקן
113	6.6 השוואת סטיות תקן
116	מושגים לחזרה
117	תרגילים - פרק 6
121	7. ציוני תקן
122	7.1 התכונות של ציוני תקן
127	מושגים לחזרה
128	תרגילים - פרק 7
129	8. מדדי אסימטריה וגבנוניות
129	8.1 מדדי אסימטריה
130	8.1.1 מדד אסימטריה לפי הרביעונים
133	8.1.2 מדדי פירסון לאסימטריה
135	8.1.3 מדד אסימטריה לפי המומנט השלישי של ציוני תקן
137	8.2 גבנוניות
139	מושגים לחזרה
140	תרגילים - פרק 8
143	9. התפלגות נורמלית
153	תרגילים - פרק 9
154	סיכום לסטטיסטיקה תיאורית
156	פלט מחשב לסטטיסטיקה תיאורית

163	קשר סטטיסטי בין משתנים
163	10.1 קשר סטטיסטי בין שני משתנים
165	10.2 קשר לינארי בין שני משתנים
168	10.3 מקדם המתאם בין המשתנים X ו-Y
173	10.4 התחזית
173	10.4.1 קו הרגרסיה לחיזוי Y לפי X
178	10.4.2 קו הרגרסיה לחיזוי X לפי Y
180	10.5 הקשר בין מקדם המתאם וקווי הרגרסיה
181	10.6 הגדרת מקדם המתאם לפי פונקצית הפסד בניבוי
183	10.6.1 חישוב r^2 לפי פונקצית החפסד
185	10.7 הגדרת מקדם המתאם לפי יחס שונויות
190	מושגים לחזרה
191	תרגילי - פרק 10

ה ס ת ב ר ו ת

193	11. ניסוי מקרי
196	11.1 פעולות בין מאורעות
197	11.2 הגדרת מודל הסתברותי
200	11.3 חוקים בחישוב הסתברויות
207	מושגים לחזרה
208	תרגילים - פרק 11
211	12. הסתברות מותנה
212	12.1 הסתברות של חיתוך מאורעות
213	12.2 הסתברות של חיתוך מאורעות בלתי תלויים
217	12.3 נוסחת בייס להסתברות מותנה
222	מושגים לחזרה
223	תרגילים - פרק 12
225	13. נוסחאות בקומבינטוריקה
228	13.1 דוגמאות לחישוב הסתברויות
229	13.2 ניסוי היפרגיאומטרי

232	13.3 מודל של מכפלה
234	13.4 ניסוי בינומי
236	מושגים לחזרה
237	תרגילים - פרק 13
239	14. משתנה מקרי ופונקצית הסתברות של משתנה מקרי
239	14.1 משתנה מקרי
240	14.2 פונקצית הסתברות
241	14.3 תיאור פונקצית הסתברות
246	14.3.1 תכונות התוחלת
246	14.3.2 תכונות השונות וסטיית התקן
246	14.4 פונקצית הסתברות בינומית
248	14.5 פונקצית הסתברות היפרגיאומטרית
251	14.6 פונקצית הסתברות נורמלית
253	14.7 קירוב נורמלי לבינום
256	מושגים לחזרה
257	תרגילים - פרק 14

ה ס ק ה ס ט י ס ט י ת

259	15. מבוא לחסקה סטטיסטית
260	15.1 שיטות דגימה
265	15.2 התפלגות הדגימה של הממוצע
268	15.3 משפט הגבול המרכזי לממוצע
275	מושגים לחזרה
276	תרגילים - פרק 15
277	16. אמידה
277	16.1 אומד נקודתי
280	16.2 אומד על ידי תחום
280	16.3 אומד נקודתי ל- μ
281	16.4 אומד על ידי תחום ל- μ
283	16.5 רווח בר-סמך ל- μ כאשר σ ידועה

285	16.5.1 תכונות רווח - הסמך
286	16.6 מציאת גודל המדגם
291	מושגים לחזרה
292	תרגילים - פרק 16
295	17. בדיקת השערות על הפרמטר μ כאשר σ נתונה
295	17.1 ניתוח בעיה של בדיקת השערות על μ
296	17.2 תהליך המבחן
301	17.3 דרך נוספת לבדיקת השערות
304	17.4 חישוב β
306	17.4.1 תכונות β
307	17 4.2 קביעת גודל המדגם
312	17.5 מבחן דו-כיווני
315	17.6 הקשר בין רווח בר-סמך ובדיקת השערות
317	מושגים לחזרה
318	תרגילים - פרק 17
321	18. הסקה על פרופורציה באוכלוסייה
323	18.1 משפט הגבול המרכזי לפרופורציה
325	18.2 אמידת הפרמטר P
326	18.2.1 אומד נקודתי
326	18.2.2 רווח בר-סמך לפרמטר P
328	18.2.3 גודל המדגם
332	18.3 בדיקת השערות על הפרמטר P
336	18.4 גודל המדגם עבור α ו- β נתונות
338	מושגים לחזרה
339	תרגילים - פרק 18
341	19. הסקה על הפרמטר μ כאשר σ לא נתונה
341	19.1 אמידת חשונות
342	19.2 התפלגות t
346	19.3 רווח בר-סמך ל- μ כאשר σ לא נתונה
349	19.4 בדיקת השערות על μ כאשר σ לא נתונה

355	מושגים לחזרה
356	תרגילים - פרק 19
359	20. הסקה על שונות האוכלוסיה σ^2
360	20.1 התפלגות הדגימה של הסטטיסטי $\hat{\sigma}^2$
360	20.2 התפלגות χ^2
362	20.2.1 הקשר בין התפלגות נורמלית להתפלגות χ^2
363	20.3 בדיקת השערות על σ^2
365	20.4 רווח בר-סמך ל- σ^2
367	מושגים לחזרה
368	תרגילים - פרק 20
369	21. הסקה סטטיסטית לפי שני מדגמים
371	21.1 השוואת ממוצעים לפי שני מדגמים בלתי תלויים
373	21.2 בדיקת השערות על $(\mu_1 - \mu_2)$ כאשר σ_1 ו- σ_2 ידועות
377	21.3 רווח בר-סמך ל- $(\mu_1 - \mu_2)$ כאשר σ_1 ו- σ_2 ידועות
379	21.4 הקשר בין רווח בר-סמך לבדיקת השערות
	21.5 בדיקת השערות על $(\mu_1 - \mu_2)$ כאשר σ_1 ו- σ_2
381	לא ידועות אך שוות
387	21.6 רווח בר-סמך ל- $(\mu_1 - \mu_2)$ כאשר σ_1 ו- σ_2 לא ידועות אך שוות
388	21.7 הסקה על $(\mu_1 - \mu_2)$ כאשר σ_1 ו- σ_2 לא ידועות ואינן שוות
392	ניתוח פלט מחשב
399	21.8 הסקה על הפרש ממוצעים לפי שני מדגמים מזווגים
401	21.9 בדיקת השערות על μ_D
405	מושגים לחזרה
406	תרגילים - פרק 21
409	22. השוואת פרופורציות לפי שני מדגמים בלתי תלויים
410	22.1 בדיקת השערות על $(P_1 - P_2)$
414	22.2 רווח בר-סמך ל- $(P_1 - P_2)$
416	מושגים לחזרה
417	תרגילים - פרק 22
419	23. הסקה על שונות של שתי אוכלוסיות

419	23.1	הסקה על יחס השונות σ_1^2/σ_2^2
421	23.2	התפלגות F
424	23.3	בדיקת השערות על σ_1^2/σ_2^2
426	23.4	רווח בר-סמך ל- σ_1^2/σ_2^2
428		מושגים לחזרה
429		תרגילים פרק 23
430		טבלת סיכום להסקה סטטיסטית
435		נספח - טבלאות סטטיסטיות:
437		טבלה מס' 1 - התפלגות נורמלית
438		טבלה מס' 2 - התפלגות t
439		טבלה מס' 3 - התפלגות X^2
440		טבלה מס' 4 - התפלגות F
444		טבלה מס' 5 - מספרים מקריים
445		ביבליוגרפיה

הקדמה

סטטיסטיקה ל"לא סטטיסטיקאים".... לכל אלה שלא בחרו להיות סטטיסטיקאים, אך נדרשים ללמוד סטטיסטיקה במסגרת לימודיהם העיקריים בתחומים שונים כמו: רפואת, סוציולוגיה, כלכלה, עבודה סוציאלית, חינוך, פסיכולוגיה ועוד ועוד, מיועד ספר זה.

לרבים מבין "הלא סטטיסטיקאים" יש רתיעה מהמקצוע סטטיסטיקה בגלל הקשר למספרים ונוסחאות, ומוכרים המשפטים כמו: "סטטיסטיקה מקצוע יבש".... "לעולם לא אדע סטטיסטיקה".... "אינני מסוגל לעבוד עם נוסחאות".... "לא יודע מתימטיקה ופחד מכל דבר הקשור למספרים".... וכי'.

ובכן, אפשר ללמוד סטטיסטיקה גם ללא דמעות ופחד - זהו הקו המנחה בספר זה.

בכל מחקר מבוסס על נתונים דרוש ידע בסטטיסטיקה כדי לארגן, לתאר ולחסיק מסקנות מחנותונים. כל מי שעוסק במחקר אמפירי חייב בידע בסיסי בשיטות סטטיסטיות, גם אם במחקר שותף סטטיסטיקאי האחראי לעיבודים. הסטטיסטיים. לחוקר ה"לא סטטיסטיקאי" חייבת להיות שפה משותפת עם הסטטיסטיקאי של המחקר, כדי שיוכל לשאול את השאלות המתאימות וכמובן להבין את התוצאות המחקריות. גם מי שאינו עוסק במחקר אמפירי, אך קורא פרסומים מדעיים, חייב בידע בסטטיסטיקה כדי להבין את משמעות הממצאים. על כן, בתחומים רבים תלמידי תואר ראשון חייבים בקורס בסיסי בסטטיסטיקה, ותלמידי תואר שני חייבים בקורס מתקדם בשיטות סטטיסטיות.

הספר מבוסס על רשימות של קורסים בסטטיסטיקה שלימדתי במשך למעלה מ-20 שנה באוניברסיטה העברית בירושלים ובאוניברסיטת תל-אביב, לתלמידי תואר ראשון ותואר שני במדעי החברה, בביה"ס לעבודה סוציאלית ובביה"ס לרפואת.

הדגש בספר הוא על הצד היישומי של הסטטיסטיקה. החסברים אינטואיטיביים ומבוססים על התגיון העומד מאחורי השיטות הסטטיסטיות. אין הוכחות מתימטיות מורכבות ואין דגש על חרקע התיאורטי-מתימטי, שהוא מעניינם של הסטטיסטיקאים. בספר משולבים פלטי מחשב של תוכניות סטנדרטיות מ-SPSS ו-SPSS-X. נלמד לקרוא ולפרש את התוצאות בפלטים אלה.

לספר שני חלקים:

חלק א' - מתאים לקורס בסיסי במבוא לסטטיסטיקה הניתן במסגרת מדעי החברה והרפואה.

חלק ב' - עוסק בשיטות סטטיסטיות מתקדמות, ומיועד לתלמידי מוסמך במדעי החברה והרפואה.

בסיום כל פרק רשימת מושגים לחזרה ותרגילים. לתרגילים חישוביים מובאים הפתרונות אחרי כל תרגיל.

במהדורה השנייה של הספר נוסף פרק העוסק בנושא: קשר לינארי בין שני משתנים (פרק מסי 10). הנושא הרחב של קשר סטטיסטי בין משתנים יופיע בחלק ב' של הספר הנמצא בהכנה.

במהדורה השלישית של הספר תוכנסו מספר תיקונים.

תודתי נתונה לדייר אלונה רביב על הערותיה המועילות אשר יושמו במהדורה זו. כן תודתי לגבי דרורה ועדיה שהאירה את עיני בענייני לשון.

תודה מיוחדת לבתי שרון על המסירות והדיוק בהדפסת כתב היד, לבעלי צבי ובני עמרי שעודדוני בכתיבת ספר זה.

רונית איזנבך

מ ב ו א

1. מהי סטטיסטיקה ?

המדע סטטיסטיקה עוסק בשיטות לתיאור וניתוח נתונים אמפיריים, ושיטות להסקת מסקנות מנתונים אלה.

בכל מחקר המבוסס על נתונים אמפיריים דרוש ידע בסטטיסטיקה כדי לארגן את הנתונים, לנתח ולחסיק מהם מסקנות. על כן הסטטיסטיקה משמשת כמקצוע עזר בתחומים רבים כמו: רפואה, כלכלה, סוציולוגיה, פסיכולוגיה, חינוך, ביולוגיה, הנדסת ועוד.

לסטטיסטיקה שני כיוונים עיקריים:

א. סטטיסטיקה תיאורית.

ב. הסקה סטטיסטית.

1.1 סטטיסטיקה תיאורית

הסטטיסטיקה התיאורית עוסקת בשיטות לארגון, תיאור ותמצות הנתונים שנאספו במחקר הסטטיסטי.

הנושאים העיקריים בתחום זה הם: בניית לוחות סטטיסטיים, הצגות גרפיות של נתונים, סיכום הנתונים על ידי ערכים מרכזיים ומדדי פיזור, וניתוח קשרים סטטיסטיים.

בסטטיסטיקה תיאורית מתרכזים בתיאור אותם נתונים שנאספו במחקר כאוכלוסיית היעד של המחקר, ללא הכללות לקבוצות רחבות יותר. הקבוצה הנחקרת היא אוכלוסיית המחקר, אותה מתארים באמצעות השיטות של הסטטיסטיקה התיאורית.

1.2 הסקה סטטיסטית

הסקה סטטיסטית עוסקת בשיטות להכללה מתוצאות של מדגם מייצג על כל האוכלוסייה.

פונקצית ההסתברות של X לפי הנוסחה:

$$N=8 \quad n=3 \quad R=3 \quad X \sim H(8, 3, 3)$$

$$P(X=x) = \frac{\binom{R}{x} \binom{N-R}{n-x}}{\binom{N}{n}}$$

נציב בנוסחה את הנתונים:

$$P(X=x) = \frac{\binom{3}{x} \binom{5}{3-x}}{\binom{8}{3}}$$

x	0	1	2	3
P(x)	10/56	30/56	15/56	1/56

התוחלת של X עבור פונקצית ההסתברות ההיפרגיאומטרית:

$$EX = n \frac{R}{N} = 3 * \frac{3}{8} = 1.125$$

סטיית התקן של X עבור פונקצית ההסתברות ההיפרגיאומטרית:

$$\sigma_X = \sqrt{n \frac{R}{N} \left(1 - \frac{R}{N}\right) \frac{N-n}{N-1}} = \sqrt{3 * \frac{3}{8} * \frac{5}{8} * \frac{8-3}{8-1}} = 0.708$$

14.6 פונקציית הסתברות נורמלית

פונקציית הסתברות של משתנה מקרי X רציף, $-\infty < X < \infty$, צורתה פעמון סימטרי.

למשתנה המקרי X :

μ תוחלת:

σ סטיית התקן:

$$Z = (X - \mu) / \sigma \quad \text{ציון התקן}$$

$\Phi(Z)$ חשטח המצטבר עד ל- Z :

פונקציית הסתברות נורמלית:

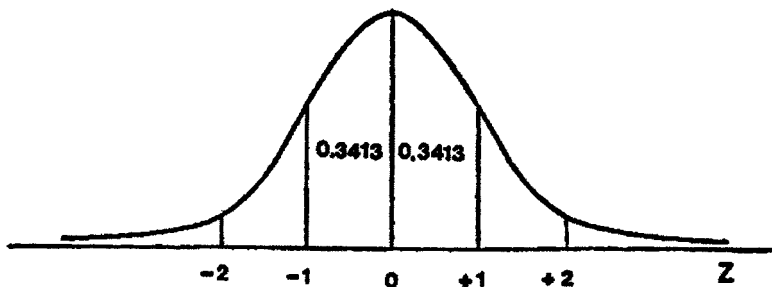
$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(X-\mu)^2}{\sigma^2}}$$

$$e=2.71828... \quad \pi=3.14159...$$

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}Z^2}$$

עבור ציון התקן Z :

$$\mu=0 \quad \sigma=1$$



דנו בתכונות ההתפלגות הנורמלית במסגרת הסטטיסטיקה חתיאורית, כאשר חשטח מציין אחוז המקרים באוכלוסייה הנחקרת.

כפונקציית הסתברות - חשטח מציין הסתברות של המשתנה המקרי X .