

סטודנטים יקרים

אנו גאים להציג בפניכם חוברת זו בקורס סטטיסטיקה ב' 30112, המהווה חלק קטן ממערך גדול של חומר עזר לסטודנטים באתר OpenBook.

באתר קיימים הסברים מוקלטים בווידאו עם שלל אמצעי המחשה שמטרתם להנגיש את החומר ולהפוך את חווית הלמידה למהנה ומעניינת.

סימונים:

✓ קיים פתרון מוקלט באתר - בלחיצה על הסימן תועבר לדף הרלוונטי באתר.

מצאתם טעות? נשמח שתשלחו לנו הודעה לכתובת המייל info@OpenBook.co.il

openbook
המרכז לקידום אקדמי

אנו מאחלים לכם הנאה בלמידה,
התעשרות בידע ובתובנות וכמובן הרבה הצלחה!

המרכז לקידום אקדמי OpenBook.

תוכן עניינים

16	התפלגות נורמלית ✓
16	ציון תקן ✓
17	הכרות עם התפלגות נורמלית: ✓
17	שאלה 1
18	שאלה 2
18	שאלה 3
19	שאלה 4
19	שאלה 5
20	שאלה 6
20	שאלה 7
20	שאלה 8
21	שאלה 9
21	שאלה 10
22	התפלגות בינומית ✓
23	ניסוי בינומי ✓
24	הסתברות בינומית ✓
24	מודל בינומי של ניסויים ✓
24	שאלה 1
24	שאלה 2
25	שאלה 3
25	שאלה 4
26	התפלגות הדגימה ומשפט הגבול המרכזי ✓
26	מושגים בסיסיים ✓
26	תצפית/ערך
26	מדדים מספריים עבור משתנים כמותיים
26	אוכלוסייה
26	מדגם
27	שאלה 1
28	התפלגות הדגימה של הממוצע ✓
29	שאלה 2
30	דגימה מתוך התפלגות נורמלית ✓
30	שאלה 3
32	דגימה מתוך התפלגות כללית – משפט הגבול המרכזי ✓

32	משפט הגבול המרכזי.....
32	הקירוב ההתפלגות הבינומית להתפלגות הנורמלית.....
32	תיקון רציפות.....
32	בחירת גודל מדגם.....
33	שאלה 4.....
33	שאלה 5.....
33	שאלה 6.....
34	שאלה 7.....
34	שאלה 8.....
34	שאלה 9.....
34	שאלה 10.....
35	שאלה 11.....
35	שאלה 12.....
35	שאלה 13.....
35	שאלה 14.....
36	שאלה 15.....
36	שאלה 16.....
36	שאלה 17.....
36	שאלה 18.....
37	שאלה 19.....
37	שאלה 20.....
37	שאלה 21.....
38	שאלה 22.....
38	שאלה 23.....
38	שאלה 24.....
38	שאלה 25.....
39	שאלה 26.....
39	שאלה 27.....
39	שאלה 28.....
39	שאלה 29.....
40	שאלה 30.....
40	שאלה 31.....
41	שאלה 32.....
41	שאלה 33.....
42	שאלה 34.....

openbook
המרכז לקידום אקדמי

42	שאלה 35
42	שאלה 36
43	שאלה 37
43	שאלה 38
43	שאלה 39
43	שאלה 40
44	שאלה 41
44	שאלה 42
44	שאלה 43
45	אמידה ✓
46	אמידה נקודתית ✓
46	פרמטר וסטטיסטי
47	דוגמא 1
47	דוגמא 2
47	דוגמא 3
47	דוגמא 4
47	המשך דוגמא 4
49	סוגי אמידה
49	המשך דוגמא 4
49	תרגיל 1
50	אמידה מרווחית ✓
51	תכונות רווח סמך
53	מודל (1) רווח בר סמך לתוחלת כשהשונות באוכלוסייה σ^2 ידועה ✓
53	שאלה 1
54	שאלה 2
54	שאלה 3
54	שאלה 4
54	שאלה 5
55	שאלה 6
55	שאלה 7
56	שאלה 8
56	שאלה 9
57	שאלה 10
58	מודל (2) רווח בר סמך לתוחלת כשהשונות באוכלוסייה σ^2 לא ידועה ✓
59	התפלגות t ✓

60	שאלה 1
60	שאלה 2
60	שאלה 3
60	שאלה 4
61	שאלה 5
62	מודל (3) רווח בר סמך להפרש תוחלות $\mu_1 - \mu_2$ ✓
63	שאלה 1
63	שאלה 2
63	שאלה 3
64	שאלה 4
64	שאלה 5
64	שאלה 6
66	מודל (4) – רווח בר סמך לפרופורציה ✓
67	שאלה 1
67	שאלה 2
67	שאלה 3
67	שאלה 4
68	שאלה 5
68	שאלה 6
68	שאלה 7
68	שאלה 8
69	שאלה 9
69	שאלה 10
69	שאלה 11
69	שאלה 12
70	שאלה 13
70	שאלה 14
70	שאלה 15
72	מודל (5) – רווח סמך להפרש פרופורציות $p_1 - p_2$ ✓
72	שאלה 1
72	שאלה 2
73	שאלה 3
73	שאלה 4
73	שאלה 5
73	שאלה 6

openbook
המרכז לקידום אקדמי

74	שאלה 7
75	מודל (6) - רווח סמך לשונות האוכלוסייה ✓
75	שאלה 1
75	שאלה 2
76	שאלה 3
76	שאלה 4
76	שאלה 5
77	שאלה 6
77	שאלה 7
79	מבחן סטטיסטי
80	הרציונל של ההסקה הסטטיסטית:
81	מהן הטעויות האפשריות? ✓
82	טעות מסוג ראשון ✓
82	טעות מסוג שני ✓
82	סוגי טעויות ✓
83	ההסתברות לטעות ✓
83	עוצמת המבחן ✓
85	value-P מובהקות התוצאה ✓
85	חישוב v.P בסוגים שונים של השערות
86	מודל 1 – בדיקת השערות על תוחלת אחת כאשר השונות באוכלוסייה ידועה ✓
86	שאלה 1
87	שאלה 2
87	שאלה 3
87	שאלה 4
88	שאלה 5
88	שאלה 6
88	שאלה 7
89	שאלה 8
89	שאלה 9
89	שאלה 10
90	שאלה 11
90	שאלה 12
90	שאלה 13
91	שאלה 14
91	שאלה 15

91	שאלה 16
92	שאלה 17
92	שאלה 1
92	שאלה 19
93	שאלה 20
93	שאלה 21
93	שאלה 1
94	שאלה 22
97	מודל 2: בדיקת השערות של ממוצע – סטיית התקן אינה ידועה ✓
98	שאלה 1
98	שאלה 2
98	שאלה 3
99	שאלה 4
100	מודל 6: בדיקת השערות על פרופורציה ✓
100	פרופורציה
100	בדיקת השערות על פרופורציה ✓
100	שאלה 1
101	שאלה 2
101	שאלה 3
101	שאלה 4
102	שאלה 5
102	שאלה 6
102	שאלה 7
102	שאלה 8
103	שאלה 9
103	שאלה 1
103	שאלה 1
105	הסקה סטטיסטית לפי שני מדגמים
106	מודל 3+4: בדיקת השערות על הפרש תוחלות במדגמים ב"ת
106	מדגמים ב"ת ✓
106	מודל 3 שונות ידועות - בדיקת השערות על הפרש תוחלת האוכלוסייה כאשר השונות ידועה ✓
107	שאלה 1
107	שאלה 1
107	שאלה 3
108	מודל 4 – בדיקת השערות: הפרש תוחלות כאשר השונות אינן ידועות ✓

110	שאלה 1
110	שאלה 1
110	שאלה 3
111	שאלה 4
111	שאלה 5
112	שאלה 6
112	שאלה 7
113	שאלה 8
113	שאלה 9
114	שאלה 1
115	מודל 5: בדיקת השערות על הפרש תוחלות/ממוצעים לפי שני במדגמים מזווגים ✓
115	מדגמים תלויים/מזווגים
115	בדיקת השערות על הפרש תוחלות מדגמים מזווגים
116	שאלה 1
116	שאלה 2
116	שאלה 3
117	שאלה 4
117	שאלה 5
117	שאלה 6
119	מודל 7: בדיקת השערות על הפרש פרופורציות ✓
119	שאלה 1
119	שאלה 2
120	שאלה 3
120	שאלה 4
121	שאלה 1
121	שאלה 1
122	מודל 8: בדיקת השערות על שונות אחת ✓
122	שאלה 1
122	שאלה 2
122	שאלה 3
122	שאלה 4
124	מודל 9: הסקה על שתי שונויות של אוכלוסיות ✓
124	מבחן F לשוויון שונויות
124	התפלגות F עבור השערה דו צדדית
125	שאלה 1

125	שאלה 2
125	שאלה 3
126	שאלה 1
126	שאלה 1
128	תרגילים מסכמים
128	שאלה 1
128	שאלה 2
129	שאלה 3
129	שאלה 4
129	שאלה 5
130	שאלה 6
131	שאלה 7
131	שאלה 8
131	שאלה 9
132	שאלה 10
132	שאלה 11
133	שאלה 12
133	שאלה 13
134	שאלה 14
134	שאלה 15
134	שאלה 16
134	שאלה 17
135	שאלה 18
135	שאלה 19
136	שאלה 20
136	שאלה 21
137	שאלה
137	שאלה
138	שאלה
138	שאלה
138	שאלה
139	שאלה
139	שאלה
140	שאלה
140	שאלה

openbook
המרכז לקידום אקדמי

145	מבחנים א- פרמטריים
145	מדגם אחד – מבחן הבינום או מבחן טיב התאמה
145	מבחן הבינום ✓
145	חישוב
146	הכרעה:
147	מבחן חי בריבוע לטיב התאמה Chi-square for goodness of fit (2χ) ✓
147	אופן החישוב
148	שאלה 1
149	מבחנים א – פרמטריים לשני מדגמים תלויים
149	מבחן הסימן ✓
150	מבחן וילקוקסון למדגמים מזווגים תלויים ✓
151	מבחן מקנמר ✓
152	מבחן וילקוקסון לשני מדגמים בלתי תלויים ✓
153	שאלה 1
154	מבחן פישר ✓
156	שאלות מבחנים א-פרמטריים
156	שאלה
156	שאלה
156	שאלה
156	שאלה
156	שאלה
157	שאלה
157	שאלה
157	שאלה
157	שאלה
157	שאלה
157	שאלה
158	שאלה
158	שאלה
158	שאלה
159	שאלה
159	שאלה
160	שאלה
160	שאלה
161	שאלה
161	שאלה

161	שאלה
162	שאלה
162	שאלה
163	שאלה
163	שאלה
164	שאלה
164	שאלה
164	שאלה
165	שאלה
165	שאלה
166	שאלה
166	שאלה
167	שאלה
167	שאלה
167	שאלה
169	שאלה
169	שאלה
169	שאלה
170	שאלה
170	שאלה
171	שאלה
171	שאלה
171	שאלה
172	שאלה
172	שאלה
172	שאלה
173	שאלה
173	שאלה
173	שאלה
174	שאלה
174	שאלה
174	שאלה
175	שאלה
175	שאלה
176	שאלה

openbook
המרכז לקידום אקדמי

177	שאלה
177	שאלה
177	שאלה
178	שאלה
178	שאלה
178	שאלה
179	שאלה
179	שאלה
179	שאלה
180	שאלה
181	שאלה
181	שאלה
181	שאלה
181	שאלה
181	שאלה
182	שאלה
182	שאלה
182	שאלה
183	שאלה
183	שאלה
183	שאלה
184	שאלה
184	שאלה
184	שאלה
184	שאלה
185	שאלה
185	שאלה
185	שאלה
185	שאלה
186	שאלה
186	שאלה
186	שאלה
187	שאלה
187	שאלה
188	שאלה

openbook
המרכז לקידום אקדמי

188	שאלה
189	שאלה
189	שאלה
189	שאלה
189	שאלה
189	שאלה
190	שאלה
190	שאלה
191	שאלה
191	שאלה
191	שאלה
191	שאלה
191	שאלה
192	שאלה
192	שאלה
192	שאלה
193	שאלה
193	שאלה
193	שאלה
194	שאלה
194	שאלה
198	טענות / ממ"ח 01
198	שאלה 1
198	שאלה 2
198	שאלה 3
198	שאלה 4
198	שאלה 5
199	שאלה 6
199	שאלה 7
199	שאלה 8
199	שאלה 9
200	שאלה 10
200	שאלה 11
200	שאלה 12
200	שאלה 13

200	שאלה 14
201	שאלה 15
201	שאלה 16
201	שאלה 17
201	שאלה 18
201	שאלה 19
202	שאלה 20
202	שאלה 21
202	שאלה 22
203	שאלה 23
203	שאלה 24
203	שאלה 25
204	שאלה 26
204	שאלה 27
205	שאלה 28
205	שאלה 29
205	שאלה 30
205	שאלה 31
206	שאלה 32
206	שאלה 33
206	שאלה 34
206	שאלה 35
207	שאלה 36
211	SPSS
212	סיכום המצבים האפשריים
213	מבחן T לתוחלת אחת Test-One Sample T
213	בדיקת השערות על תוחלת אחת
213	שאלה 1
214	מבחן T להפרש תוחלות במדגמים מזווגים Test-Samples T-Paired
214	שאלה 2
215	מבחן T להפרש תוחלות במדגמים ב"ת Test-Independent Samples T
215	שאלה 3
216	שאלה 4
217	שאלה 5
218	שאלה 6

218	שאלה 7
219	שאלה 8
220	שאלה 9
221	שאלה 10
222	שאלה 11
223	שאלה 12
224	שאלה 13
225	שאלה 14
226	שאלה 15
227	שאלה 16
228	שאלה 17
229	שאלה 18
230	שאלה 19
231	שאלה 20
232	שאלה 21
233	שאלה 22
234	שאלה 23
235	שאלה 24
236	שאלה 25
237	שאלה 26
237	שאלה 27

openbook
המרכז לקידום אקדמי

✓ התפלגות נורמלית

התפלגות נורמלית מוגדרת על ידי שני פרמטרים:

✓ ממוצע התצפיות, שמשמש כנקודת האמצע של העקומה

✓ והשונות שמגדירה את רוחב הפעמון.

בסטטיסטיקה ממוצע מוגדר כנקודת האמצע של קבוצת מספרים, ושונות מוגדרת כמדד של הפיזור של קבוצת מספרים

התפלגות נורמלית היא התפלגות פעמונית וסימטרית שבקו האמצע נמצאים מד"י המרכז והם:

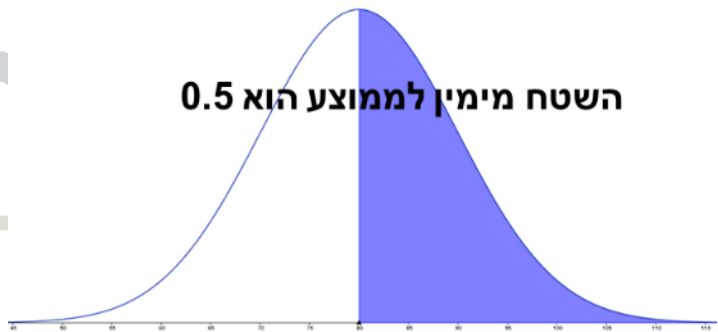
$$\bar{x} = Md = Mo = MR$$

השטח מתחת לעקומה שווה ל 1 (100%):

השטח משמאל לממוצע הוא 0.5



השטח מימין לממוצע הוא 0.5



✓ ציון תקן

ציון תקן, הוא בעצם ציון יחסי, ציון תקן מודד את המרחק של הציון הנידון מהממוצע של קבוצתו,

ומתרגם מרחק זה ליחידת פיזור של קבוצה זו, כך שאפשר להשוות בעזרתו התפלגויות שונות-ממוצע או שונות-פיזור.

ציון תקן של x מסומן z_x ומוגדר כמרחק של x מהממוצע $\bar{x}$, כאשר מרחק זה נמדד ביחידות של סטיית תקן S_x

:

$$z_x = \frac{x - \bar{x}}{s_x}$$

ניתן להציג כל ערך של המשתנה X המתפלג נורמאלי במונחים של מספר סטיות התקן שהוא מרוחק מן הממוצע - שם אחר: ציון סטנדרטי, והתפלגותו נקראת התפלגות נורמאלית סטנדרטית עם ממוצע שווה ל-0 וסטיות תקן שווה ל-1, $S_z = 1$, $\bar{z} = 0$.

את השטח משמאל לערך Z מסוים נסמן כך: $\phi(z)$

מה אנחנו צריכים לזכור?

$$P(Z < \text{Number}) = \phi(\text{Number})$$

$$P(Z > \text{Number}) = 1 - \phi(\text{Number})$$

$$\phi(-\text{Number}) = 1 - \phi(\text{Number})$$

$$\phi(a < Z < b) = \phi(b) - \phi(a)$$

הכרות עם התפלגות נורמלית:

(1) אורית ושי לומדים באותה כיתה ונבחנו באותו מבחן במתמטיקה. הציון של אורית בבחינה היה 90 והציון של שי 80. מי הצליח יותר?

(2) אורית בכיתה יא' 9 קיבלה במבחן 90. שי בכיתה יא' 5 קיבל במבחן 80. מי הצליח יותר?

ניסוי 1

התפלגות נורמלית נוצרת מקבוצת תצפיות אקראיות ובלתי תלויות, למשל כשכדור נופל במורד לוח סיכות, מרבית הכדורים יפלו באזור המרכז, ואחוז קטן יפול בצדדים. כשנספור כמה כדורים נפלו בכל חור ונציג אותם בגרף (היסטוגרמה) נקבל צורה גלית דמוית פעמון.

ניסוי 2

נבחר מספר מינמלי(הנמוך ביותר) של שורות 5- ובהסתברות של 0.5.

פירוש הדבר הוא שלכדורים הנופלים יהיו רק חמישה מקומות שאליהם יוכלו ליפול וההסתברות שיפלו במרכז היא הגבוהה ביותר.



שאלה 1

תלמיד קיבל ציון $x=70$ בחשבון וציון $y=75$ בתנ"ך.

המידע על התפלגויות שני הציונים הללו בכיתתו הוא כדלהלן:

בחשבון - $\bar{x} = 65$, $s_x = 3$

בתנ"ך- $\bar{y} = 70$, $s_y = 4$

באיזה מקצוע קיבל ציון טוב יותר באופן יחסי לכיתתו?



שאלה 2

נחשב לפי הטבלה להתפלגות נורמלית

$$\phi(1.2) =$$

$$\phi(0.3) =$$

$$\phi(2.87) =$$

$$\phi(-0.82) =$$

$$\phi(???) = 0.500$$

$$\phi(???) = 0.9750$$

$$\phi(???) = 0.1003$$

$$\phi(???) = 0.9904$$



שאלה 3

1. ידוע כי התפלגות הטמפרטורה הנמדדת מדי יום בת"א בשעה עשר בבוקר היא נורמלית עם ממוצע 20 וסטיית תקן 4 (כל הנתונים במעלות)

א. מהו אחוז הימים שבהם הטמפרטורה עולה על 22?

ב. מהו אחוז הימים בהם הטמפרטורה היא פחות מ- 21?

ג. מהו אחוז הימים שבהם הטמפרטורה נמוכה מ- 17?

ד. מהו אחוז הימים שבהם הטמפרטורה גבוהה מ- 16?

ה. מהו אחוז הימים שבהם הטמפרטורה היא בין 18 ל- 23?

ו. חמסין מוגדר כטמפרטורה שגבוהה מ-30. מהו אחוז הימים בהם יש חמסין בת"א?

ז. מהו אחוז הימים שבהם הטמפרטורה עולה על 15?

ח. מהי הטמפרטורה ש- 67% מהימים יותר קרים ממנה?

ט. מהי הטמפרטורה ש- 33% מהימים יותר חמים ממנה?

י. מהי הטמפרטורה ש- 64% מהימים יותר חמים ממנה?

יא. מהי הטמפרטורה ש- 50% מהימים יותר קרים ממנה?

יב. מהי הטמפרטורה ש- 25% מהימים יותר חמים ממנה?

יג. מהי הטמפרטורה ש- 55% מהימים פחות קרים ממנה? (פחות קרים = יותר חמים)

יד. מהי הטמפרטורה ש- 80% מהימים יותר חמים ממנה?

טו. מהי הטמפרטורה ש- 10% מהימים יותר חמים ממנה?



שאלה 4

מילוי קרטון חלב של 1 ליטר נעשה בעזרת מכונה אוטומטית. ניתן לכוון את המכונה כדי שתיתן כמות ממוצעת נתונה של חלב. כמות המילוי מתפלגת בקירוב נורמלית עם סטיית התקן של 0.01 ליטר.

א. מהי פרופורציית הקרטונים שבהם פחות מ- 0.97 ליטר חלב, אם המכונה כוונה למילוי ממוצע של 1 ליטר?

ב. מהי פרופורציית הקרטונים שבהם יותר מ- 1 ליטר חלב, אם המכונה כוונה למילוי ממוצע של 1.02 ליטר?

ג. לאיזה מילוי ממוצע חייב היצרן לכוון את המכונה, אם על פי התקן, אחוז הקרטונים שבהם פחות מ- 1 ליטר חלב אינו יכול לעלות על 1.5%?

ד. המכונה כוונה לממוצע התקני, שמצאת בסעיף ג', מהו אחוז הקרטונים שכמות החלב בהם אינה סוטה מהממוצע ביותר מסטיית תקן אחת?



שאלה 5

התפלגות הציונים בקורסים בכלכלה היא בקירוב נורמלית.

לבחינה במבוא לכלכלה במועד א' ניגשו 500 סטודנטים, הציון הממוצע בבחינה היה 70 וסטיית התקן 12.

א. אם ציון "מספיק" מוגדר כציון בתחום 55-64, מהו מספר הסטודנטים במועד זה, שקיבלו "מספיק" בבחינה במבוא לכלכלה?

ב. 2% מהסטודנטים שציוניהם הטובים ביותר יקבלו ציון "מצוין" מהו הציון המינימלי לקבלת "מצוין" בבחינה במבוא לכלכלה?

ג. לבחינה במקרו כלכלה במועד א' ניגשו 400 סטודנטים.

50% מהסטודנטים שנבחנו בבחינה קיבלו ציון גבוה מ 74 ו- 2.5% מהסטודנטים קיבלו ציון גבוה מ 93.6. מה היה הציון הממוצע ומה הייתה סטיית התקן בבחינה במקרו כלכלה?

ד. לבחינה בתורת המחירים במועד א' ניגשו 100 סטודנטים, הציון הממוצע היה 68 וסטיית תקן 8. חשב:

(1) את הציון הממוצע של כל הנבחנים בבחינות בכלכלה.

(2) את סטיית התקן של הציונים בבחינות בכלכלה.



שאלה 6

במפעל לייצור קפה נמס קיימת מכונת מילוי אוטומטית.

הכמות בפועל שהמכונה מכניסה מתפלגת נורמלית עם סטיית תקן של 5 גרם לקופסא.

ידוע כי רק 2% מהקופסאות מכילות פחות מ- 500 גרם קפה.

א. מהו ממוצע כמות הקפה בקופסאות?

ב. מה אחוז הקופסאות המכילות יותר מ- 515 גרם?



שאלה 7

ידוע כי התפלגות הציונים בבחינת גמר בקורס מסוים באוניברסיטה הפתוחה היא נורמלית.

א. בסמסטר א בשנה מסוימת קיבלו 50% מהסטודנטים ציון הגבוה מ- 72.5 ו- 6% קיבלו ציון גבוה מ- 91.16. מה היה הציון הממוצע ומה הייתה סטית התקן של הציונים בסמסטר זה?

ב. בסמסטר א ידוע כי התלמידים שציונם מעל 96 קיבלו ציון "מעולה". כמו כן ידוע כי באותו סמסטר 18 תלמידים קיבלו ציון מעולה. מצא מה היה מספר הנבחנים באותו סמסטר.

ג. בסמסטר ב היה הציון הממוצע 70 וסטית התקן 10. סטודנטים שציונם בין 76-85 קיבלו את הציון טוב. מהו אחוז הסטודנטים בסמסטר זה שקיבלו "טוב" בבחינת הגמר?

ד. מהו הציון בסמסטר ב השקול (טוב באותה מידה) לציון 67 בסמסטר א?



שאלה 8

חברת "האבוב" מתכוונת לשווק סוג חדש של צמיגים רדיאליים. מבדיקות החברה התברר שממוצע אורך החיים של צמיגים אלו הוא 38,000 ק"מ עם סטיית תקן 3,000 ק"מ.

בהנחה שאורך חיי הצמיג מתפלג נורמלית ובדיקות החברה נכונות מצא:

א. מהו אחוז הצמיגים שניתן לצפות שיחזיקו מעמד יותר מ- 38,000 ק"מ?

ב. מהו אחוז הצמיגים שניתן לצפות שיחזיקו מעמד לכל היותר 35,000 ק"מ?

ג. חברת "האבוב" יצאה במסע פרסום והודיע שתפצה בצמיג חדש, כל בעל מכונת שהצמיג שקנה החזיק מעמד פחות מ- 35,000 ק"מ. החברה קבעה את הערך K כך שלא תצטרך לפצות יותר מאשר 10% מרוכשי הצמיגים. מהו הערך K שנקבע?

ד. הסבר, ללא חישוב מחדש, כיצד ולאיזה כוון (אם בכלל) היו משתנות תשובותיך לסעיפים א, ב, ו- ג לו סטיית התקן היתה 5,000 ק"מ?



שאלה 9

משקלם של עכברי מעבדה לבנים מתפלג נורמלית.

ידוע כי מתחת למשקל 171.15 גרם נמצאים 99.2% מהעכברים ומעל למשקל 152.55 נמצאים 12.1% מהעכברים.

א. מהם הממוצע והשונות של משקל עכברי המעבדה ?

ב. מהו אחוז העכברים שמשקלם בין 105 גרם לבין 150 גרם ?

ג. מהו משקל עכבר שרק 10% מעכברי המעבדה שוקלים יותר ממנו ?

ד. מהו טווח הבינרבעוני של משקל עכברי המעבדה?



שאלה 10

משקלו של מחשב מחברת (Note Book) מתפלג נורמלית עם ממוצע 2000 גרם.

משקלם של 95% ממחשבים אלה נמצא בין 1706 גרם לבין 2294 גרם.

א. מצא את סטיית התקן של המשקל של מחשב מחברת.

ב. מהו המשקל שרק 10% מהמחשבים הנ"ל שוקלים פחות ממנו?

openbook
המרכז לקידום אקדמי

✓ התפלגות בינומית

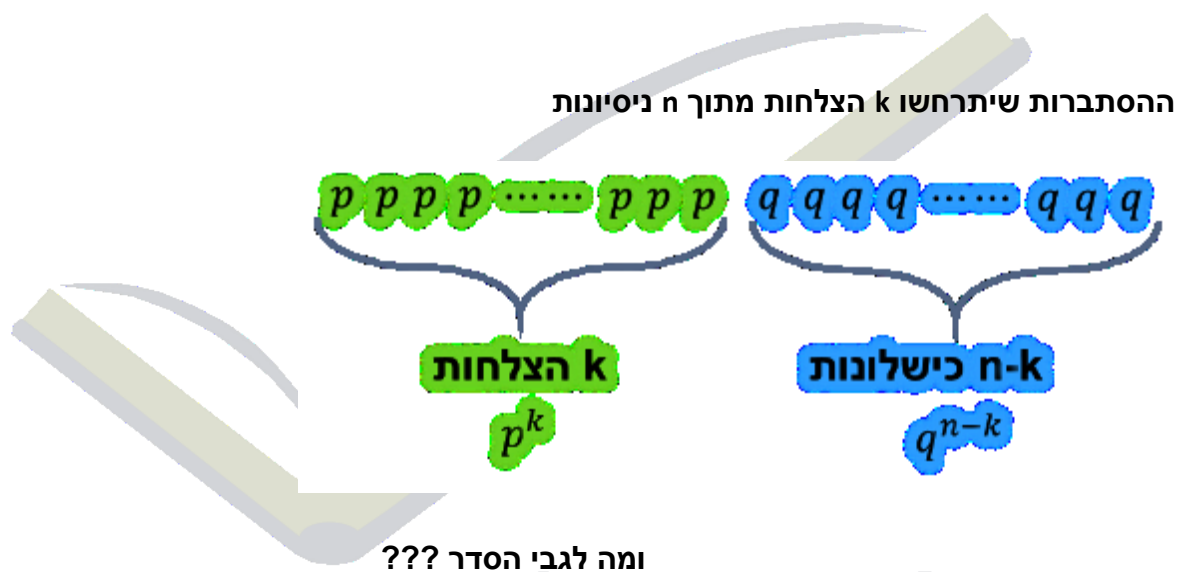
מבצעים n ניסויים בלתי תלויים זה בזה,

כאשר התוצאה של ניסוי כלשהו אינה תלויה בתוצאה של ניסיונות האחרים.

לכל ניסוי בודד יש שתי תוצאות אפשריות בלבד: הצלחה או כישלון.

ההסתברות להצלחה בכל ניסוי קבועה ושווה ל- p

אנו נחשב את ההסתברות שיתרחשו k הצלחות מתוך n ניסיונות.



מספר האפשרויות לסידור k הצלחות מתוך n ניסיונות זה המקדם הבינומי !

המרכז לקידום אקדמי

$\binom{n}{k} p^k q^{n-k}$

$$P(k \text{ הצלחות } n \text{ להצלחה הסתברות עם חזרות } p) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \binom{n}{k} p^k (1-p)^{n-k}$$

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

נוסחה זו נקראת **נוסחת ההסתברות הבינומית**.

כדי להשתמש בנוסחה לפתרון בעיה, מציבים את נתוני הבעיה המתאימים במקום p , n , k ו- $\binom{n}{k}$

סימון: $\binom{n}{k}$ - מספר האפשרויות לסידור k הצלחות ב- n ניסיונות, או, מספר הקומבינציות של k מתוך n .

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

ההסתברות ל- k הצלחות מתוך n ניסויים
 מס' האפשרויות לבחירת המקום של k ההצלחות
 ההסתברות ל- k הצלחות
 ההסתברות ל- $n-k$ כישלונות

ניסוי בינומי

ניסוי ובו 2 תוצאות אפשריות: 0 או 1, הצלחה או כישלון.

מקובל לסמן את ההסתברות להצלחה באות p .

ואת ההסתברות המשלימה ב- $q=1-p$

ניסוי שמתקיימים בו שלושה תנאים:

I. הניסוי מורכב ממספר מסוים של חזרות על אותו ניסוי (כגון 10 הטלות מטבע, או בחירת 5 אנשים).

II. לכל חזרה על הניסוי יש שתי תוצאות אפשריות (כגון 'ציור' - 'מספר')

III. הניסויים החוזרים אינם תלויים זה בזה: מידע אודות תוצאה של ניסוי אחד אינו משנה את ההסתברות

של התוצאה בניסוי החוזר (למשל, בניסוי של הטלת מטבע, ההסתברות (ציור) P קבועה לכל אחד מהניסויים החוזרים)

המרכז לקידום אקדמי

נניח שבוחרים n תלמידים עם החזרה, מכיתה שבה החלק היחסי של הבנים הוא p . מה ההסתברות ל- k בנים (k הצלחות)?

נוכל להציג נקודה טיפוסית במאורע "הצלחות ב- n חזרות" בצורה סכמתית, כך:

$$\underbrace{S S S \dots S}_k \quad \underbrace{F F F \dots F}_{(n-k)}$$

k הצלחות
 $(n-k)$ כישלונות

כל נקודה במאורע המבוקש היא סידור של n הסמלים, k פעמים S ו- $(n-k)$ פעמים F .

ההסתברות של כל נקודה במאורע היא:



$$P\left(\begin{matrix} n \text{ חזרות} \\ k \text{ הצלחות} \end{matrix} \right) = \left(\begin{matrix} n \\ k \end{matrix}\right) p^k (1-p)^{n-k} = \left(\begin{matrix} n \\ k \end{matrix}\right) p^k (1-p)^{n-k}$$

נוסחה זו נקראת **נוסחת ההסתברות הבינומית**.

כדי להשתמש בנוסחה לפתרון בעיה, מציבים את נתוני הבעיה המתאימים במקום k, n, p ו- $\binom{n}{k}$.

סימון: $\binom{n}{k}$. מספר האפשרויות לסידור k הצלחות ב- n ניסיונות, או, מספר הקומבינציות של k מתוך n .

הסתברות בינומית

את ההסתברות ל- k הצלחות בניסוי בינומי, ובקיצור ההסתברות בינומית, אנו מחשבים אם כן בעזרת הנוסחה -

$$P(k \text{ הצלחות}) = \binom{n}{k} p^k (1-p)^{n-k}$$

n - מספר החזרות הבלתי תלויות בניסוי בינומי, ו- p היא ההסתברות להצלחה בניסוי ברנולי.

מודל בינומי של ניסויים

מבצעים סדרה של n ניסויי ברנולי ב"ת, בעלי ההסתברות להצלחה p .

ההסתברות לקבל בדיוק k הצלחות ב- n ניסויים היא $P(k)$.

התוחלת של התפלגות בינומית: $E(X) = np$

השונות של התפלגות בינומית: $V(X) = npq$



שאלה 1

מה ההסתברות שב- 8 הטלות קובייה תתקבל הספרה 5 בדיוק פעמים?



שאלה 2

ידוע שההסתברות שצלף יפגע במטרה בירייה בודדת גדולה פי 3 מההסתברות שהוא יחטיא בירייה בודדת.

א. מה ההסתברות שהצלף יפגע בירייה בודדת.

ב. פגיעה אחת – מפילה את המטרה. צלף יורה למטרה 5 יריות, מה ההסתברות שהמטרה תיפול



שאלה 3

ידוע שאם בוחרים באקראי 3 מכוניות מכלל המכוניות בארץ מסוימת ,
ההסתברות שבכל אחת מ- 3 המכוניות יש מזגן היא 0.064 .
בוחרים באקראי 6 מכוניות מהי ההסתברות שלפחות בשתי מכוניות יש מזגן ?



שאלה 4

הסיכוי לפגוע במטרה במטווח הוא 0.6.
א. מה הסיכוי לפגוע במטרה 4 פעמים מתוך 5 ניסיונות?
ב. מה תוחלת הפגיעות במטרה?

openbook
המרכז לקידום אקדמי

✓ התפלגות הדגימה ומשפט הגבול המרכזי

✓ מושגים בסיסיים

תצפית/ערך

✓ תוצאת המדידה של המשתנה, הערכים שהמשתנה מקבל.

✓ תצפית של המשתנה X מסומנת ב- X_i כאשר האינדקס i משמש לזיהוי התצפית.

✓ למשל: X_i - מספר ילדים במשפחה ה- i , $X=7=3$ - במשפחה מספר 7 יש 3 ילדים.

מדדים מספריים עבור משתנים כמותיים

✓ n - מספר התצפיות

✓ $X_1, X_2, \dots, X_n$ תצפיות לפי סדר הגעתן

✓ $\sum_{i=1}^n X_i$ - סכום התצפיות

✓ $\sum_{i=1}^n X_i^2$ - סכום ריבועי התצפיות

אוכלוסייה

אוסף של פריטים שמעוניינים לחקור/לבדוק. קבוצת כל הפרטים שעונים על קריטריון מסוים

למשל: כל תלמידי התיכון בארץ, נשים מעל גיל 35, משפחות שבהן לפחות 5 ילדים וכו'.

מדגם

תת-קבוצה (קבוצה חלקית) של האוכלוסייה שנבחרה, ועל-פי הנתונים שלה יבוצע מחקר סטטיסטי במטרה להסיק מסקנות על האוכלוסייה

מדגם בגודל n של המשתנה X מסומן ע"י: $X_1, X_2, X_3, \dots, X_n$.

אוכלוסייה ומדגם הם מושגים יחסיים אותה קבוצה יכולה להיות אוכלוסייה או מדגם.

סטטיסטי – תכונה של המדגם. כל אחד ממדדי המרכז והפיזור מהווה סטטיסטי.

פרמטר – תכונה של האוכלוסייה. פרמטר הוא גודל קבוע המאפיין את האוכלוסייה, לעומת סטטיסטי שהוא משתנה מקרי, שמקבל בכל מדגם תוצאה שונה.

פרמטר- כל האובייקטים שרלוונטיים לתכונה מסוימת	מדגם- תת קבוצה של האוכלוסייה $\hat{\theta}$ סטטיסטי - מדד שלקוח מתוך מדגם.
θ פרמטר-מדד שלקוח מתוך האוכלוסייה (המאפיין)	
ממוצע אוכלוסייה = תוחלת $\mu = E(x)$	ממוצע מדגם $\bar{X} = \frac{\sum x_i}{n}$
שונות באוכלוסייה = שונות של מ"מ $V(X_i) = \sigma^2$	שונות במדגם $\hat{S}^2 = \sum \frac{(x_i - \bar{x})^2}{n-1}$
סטיית תקן באוכלוסייה σ	סטיית תקן במדגם $\hat{S}$
פרופורציה באוכלוסייה P	פרופורציה במדגם $\hat{P}$
ערך הפרמטר של האוכלוסייה הוא קבוע ובדיד לא ידוע. הערך לא תלוי במדגם שאנו דוגמים ואינו משתנה במדגם למדגם.	הערך של הסטטיסטי תלוי במדגם שנבחר ומשתנה ממדגם למדגם מכיוון שהסטטיסטים משתנים ממדגם למדגם. יש להם התפלגות. להתפלגות של הסטטיסטי קוראים התפלגות דגימה.

התפלגות הדגימה של סטטיסטי מסוים היא פונקציית ההסתברות או ההתפלגות שלו. התפלגות זו תלויה בצורתו המתמטית של הסטטיסטי, בגודל המדגם ובתכונות האוכלוסייה ממנה נדגם.

על מדגם אפשר להגדיר סטטיסטים שונים: ממוצע, חציון, שונות וכו'.
כל אחד מהם הוא סטטיסטי שיש לו התפלגות דגימה משלו התלויה בהגדרתו התמטית של הסטטיסטי, בגודל המדגם ובהתפלגות האוכלוסייה (או המ"מ) ממנה נדגם.

בתהליך ההסקה הסטטיסטית אנו מעוניינים ללמוד מתוך הסטטיסטים המחושבים במדגם על הפרמטרים של האוכלוסייה או של המשתנה המקרי.

התפלגות הדגימה היא כלי עזר שמאפשר ללמוד ממדדי המדגם (הסטטיסטים) על תכונות באוכלוסייה (הפרמטרים).



שאלה 1

בישוב חופית 5000 משפחות שמהן ל- 500 משפחות 4 רכבים, ל-1500 משפחות 3 רכבים, ל-1250 משפחות 2 רכבים, ל-1000 משפחות רכב אחד ול-750 משפחות אין רכב כלל.

א. מצא את התפלגות מספר הרכבים של תושבי חופית

ב. משפחה מחופית נבחרה באופן מקרי.

נסמן ב- X את המשתנה המקרי "מספר הרכבים שיש למשפחה שנבחרה." מהי פונקציית ההסתברות של X ?

ג. מצא ממוצע ושונות מספר הרכבים באוכלוסייה.

ד. מצא תוחלת ושונות המשתנה המקרי X .

✓ התפלגות הדגימה של הממוצע

כאמור, ממוצע המדגם הוא משתנה מקרי בעל פונקציית הסתברות ולכן יש לו התפלגות.

$$E(\bar{x}) = E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = E\left(\frac{1}{n}X_1 + \frac{1}{n}X_2 + \dots + \frac{1}{n}X_n\right)$$

נשתמש בשתי תכונות התוחלת:

$$1. \text{ תוחלת של סכום שווה לסכום התוחלות: } E(X + Y) = E(X) + E(Y)$$

$$2. E(aX) = aE(X)$$

$$\text{ונקבל: } E(\bar{x}) = \frac{1}{n}E(X_1) + \frac{1}{n}E(X_2) + \dots + \frac{1}{n}E(X_n)$$

לכל M "מ" X_i אותה פונקציית הסתברות כמו למשתנה X עליו נערכות התצפיות, בפרט $E(X_i) = E(X)$ ולכן

נקבל:

$$E(\bar{x}) = \frac{1}{n}E(X) + \frac{1}{n}E(X) + \dots + \frac{1}{n}E(X) = E(X)$$

$$E(\bar{X}) = E(X)$$

מדד הפיזור המקובל ביותר הוא השונות, נבדוק מהי שונות התפלגות הדגימה של הממוצע:

$$\sigma_{\bar{X}}^2 = V(\bar{X}) = V\left[\frac{1}{n}(X_1 + X_2 + \dots + X_n)\right] = \frac{1}{n^2}V(X_1 + X_2 + \dots + X_n)$$

השתמשנו בתכונת השונות: $V(aX) = a^2V(X)$

במדגם מקרי המ"מ $X_1, X_2, \dots, X_n$ בלתי תלויים, והשונות של סכום מ"מ ב"ת שווה לסכום השונויות שלהם.

$$V(\bar{X}) = \frac{1}{n^2}[V(X_1) + V(X_2) + \dots + V(X_n)]$$

לכל מ"מ X_i אותה פונקציית הסתברות כמו למשתנה X , כלומר $V(X_i) = V(X)$ ולכן נקבל:

$$V(\bar{X}) = \frac{1}{n^2}[V(X) + V(X) + \dots + V(X)] = \frac{V(X)}{n}$$

בעזרת הסימון של סטיות תקן σ_X^2 , $V(X) = \sigma_X^2$, $V(\bar{X}) = \sigma_{\bar{X}}^2$

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n}$$

תוחלת הממוצע שווה לתוחלת המ"מ ממנו דוגמים. $\mu_{\bar{x}} = \mu_x$

סטיית התקן של ממוצע המדגם שווה לסטיית התקן של מ"מ ממנו דוגמים, מחולקת בשורש הריבועי של גודל

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} \text{ המדגמים.}$$



שאלה 2

בכד שלושה כדורים הממוספרים כך: 2,4,6.

א. יהי X – מספר הכדור המוצא מהכד באופן מקרי.

חשב את פונקציית ההסתברות של X , את התוחלת והשונות של X .

ב. דוגמים 2 כדורים באופן מקרי מהכד (עם החזרה) ומחשבים את ממוצע ערך הכדורים. מצא את פונקציית התפלגות הדגימה של הממוצע. חשב את תוחלת הממוצע ואת שונות הממוצע.

ג. דוגמים 2 כדורים באופן מקרי מהכד (עם החזרה) ורושמים את הערך המקסימלי. מצא את פונקציית התפלגות הדגימה של הערך המקסימאלי.

openbook
המרכז לקידום אקדמי

דגימה מתוך התפלגות נורמלית

אנו דנים במ"מ $\bar{x}$ שהוא ממוצע המדגם מתוך מ"מ X (או אוכלוסייה), בעל תוחלת μ וסטיית תקן σ .

$$E(\bar{x}) = \mu \text{ ו- } V(\bar{x}) = \frac{\sigma^2}{n} \Rightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

משפט:

בדגימת מדגם שגודלו n מתוך מ"מ נורמלי X בעל תוחלת μ ושונות σ^2 ,

יהיה ממוצע המדגם $\bar{x}$ גם הוא מ"מ נורמלי, בעל תוחלת μ ושונות $\frac{\sigma^2}{n}$.

כלומר: אם $X \sim N(\mu, \sigma^2)$ אזי Z_X שהוא $\frac{X-\mu}{\sigma}$ מתפלג אף הוא נורמלית עם ממוצע 0 וסטיית תקן 1, כלומר:

$$Z_X = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

באופן דומה, אם $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ אזי:

$$Z_{\bar{X}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

שאלה 3

נדגים את התפלגות הדגימה של הממוצע בעזרת מדגם מאוכלוסיית משפחות מישוב חופית:

בישוב חופית 5000 משפחות שמהן ל- 500 משפחות 4 רכבים, ל-1500 משפחות 3 רכבים, ל-1250 משפחות 2 רכבים, ל-1000 משפחות רכב אחד ול-750 משפחות אין רכב כלל.

התפלגות מספר הרכבים של תושבי חופית היא:

מספר רכבים	מספר משפחות
0	750
1	1000
2	1250
3	1500
4	500
סה"כ	5000

ה. שתי משפחות מחופית נבחרו באופן מקרי ועם החזרה.

נסמן ב- X_1 את מספר הרכבים שיש למשפחה הראשונה, וב- X_2 את מספר הרכבים שיש למשפחה השנייה שנבחרה.

חשב את התפלגות ממוצע המדגם (פונקציית הסתברות של המשתנה מקרי $\bar{x}$)

ו. בהמשך לסעיף ה' מצא תוחלת ושונות התפלגות הדגימה של ממוצע המדגם.



דגימה מתוך התפלגות כללית – משפט הגבול המרכזי

אם נתונה התפלגות כלשהי של מ"מ X שממנה דוגמים מדגם בגודל n , ומתעניינים במ"מ $\bar{x}$, מתברר שכאשר n גדול במידה מספקת התפלגות הדגימה של $\bar{x}$ קרובה בצורתה להתפלגות הנורמלית.

משפט הגבול המרכזי

יהיה X מ"מ כלשהו בעל תוחלת μ וסטיית תקן σ , ויהיו $x_1, x_2, \dots, x_n$ מ"מ ב"ת שלכל אחד מהם התפלגות כמו זו של x , אזי התפלגות של הממוצע $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$ שואפת להתפלג נורמלית עם תוחלת $x_1, x_2, \dots, x_n$ וסטיית תקן $\frac{\sigma}{\sqrt{n}}$ כאשר n שואף לאינסוף. $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

לפי משפט הגבול המרכזי תהיה התפלגות הממוצע X , עבור n מספיק גדול ($n \geq 30$), קרובה להתפלגות נורמלית אזי $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

הקירוב ההתפלגות הבינומית להתפלגות הנורמלית

כאשר $x \sim B(n, p)$, $E(x) = np$, $V(x) = npq$ לפי משפט הגבול המרכזי תהיה התפלגות הממוצע X , עבור n מספיק גדול, קרובה להתפלגות נורמלית עם תוחלת np ושונות npq , ואז $k \sim N(np, npq)$ יש לבדוק שמתקיימים 2 התנאים הבאים: $n \cdot p \geq 10$, $n \cdot q \geq 10$

תיקון רציפות

(עבור $n < 100$) מכיוון שההתפלגות הנורמלית היא התפלגות של מ"מ רציף ואילו הבינומית היא של מ"מ בדיד, נהפוך את המ"מ הבינומי הבדיד למ"מ רציף ע"י תיקון רציפות. לנקודה בודדת a מפצלים לשני נקודות: $a \pm 0.5$

$$\begin{aligned} P(x \leq a) &\Rightarrow P(x < a + 0.5) \\ P(x \geq a) &\Rightarrow P(x > a - 0.5) \\ P(x = a) &\Rightarrow P(a - 0.5 < x < a + 0.5) \end{aligned}$$

בחירת גודל מדגם

ממוצע המדגם הוא אומד של התוחלת ואנו רוצים להבטיח בהסתברות גבוהה שהסטייה שלו מהתוחלת תהיה קטנה. השאלה היא – מהו גודל המדגם שיבטיח זאת?

מהו גודל המדגם שיש לבחור על מנת שבהסתברות של $1 - \alpha$ לפחות המרחק בין ממוצע המדגם $\bar{x}$ והתוחלת μ לא יעלה על ε ?

עלינו לחשב את גודל המדגם n שעבורו מתקיים: $P(|\bar{x} - \mu| \leq \varepsilon) \geq 1 - \alpha$

$$n \geq \left(\frac{\sigma \cdot Z_{1-\frac{\alpha}{2}}}{\varepsilon} \right)^2$$

גודל המדגם שיש לבחור הוא:



שאלה 4

מספר הביצים שצורכת משפחת ישראלי בחודש הוא משתנה מקרי בעל תוחלת של 30 ביצים וסטית תקן של 5 ביצים.

א. משפחת ישראלי רשמה כמה ביצים צרכה כל חודש במשך שלוש שנים. מה ההסתברות, בקירוב, שהיא צרכה במשך 3 שנים יותר מ-1128 ביצים?

ב. רוצים לבדוק האם תוחלת מספר הביצים שצורכת משפחת ישראלי בחודש היא אכן 30. מה גודל המדגם המינימלי שיש לבדוק כך שההסתברות, שממוצע המדגם יסטה מתוחלתו בפחות מ-1, תהיה לפחות 0.95?



שאלה 5

משקל ביצה מתפלג נורמלית עם תוחלת 60 גרם וסטיית תקן 10 גרם. ביום מסוים נאספו 100 ביצים.

א. ביצה שמשקלה מתחת ל-55 גרם היא ביצה מספר 3. מהי בקירוב ההסתברות שנוכל למלא תבנית אחת לפחות של ביצים מספר 3 (בתבנית מקום ל-30 ביצים).

ב. ביצה שמשקלה מעל ל-70 גרם היא ביצה מספר 1. מהי בקירוב ההסתברות שמבין 100 הביצים שנאספו לא ימצאו מספיק ביצים מספר 1 על מנת למלא תבנית המכילה 18 ביצים?



שאלה 6

למסיבה הוזמנו 450 איש. ההסתברות שאורח שהוזמן יגיע למסיבה היא 0.9 והחלטות המוזמנים השונים אינן תלויות זו בזו.

א. מהי בקירוב ההסתברות שמספר האורחים שיגיעו למסיבה לא יעלה על 420?

ב. מהי בקירוב ההסתברות שלכל היותר 15 מוזמנים לא יגיעו למסיבה?

ג. מהו המספר המינימלי של מנות שיש להכין כדי שלכל אורח תהיה מנה בהסתברות של 0.95 לפחות?



שאלה 7

תהליך ביקורת איכות של ברגים הוא כדלקמן: בודקים 150 ברגים שנוצרו במשך היום ונבחרו באופן מקרי מתוך כל תהליך הייצור של אלפי ברגים. אם במדגם יש לכל היותר 18 ברגים פגומים, הייצור היומי מתקבל, אחרת הוא נדחה. חשב את ההסתברויות הבאות:

(א) הייצור היומי יתקבל, אם ידוע כי בתהליך הייצור 20% פגומים.

(ב) הייצור היומי ידחה, אם ידוע כי בתהליך הייצור 10% פגומים.



שאלה 8

למפגש חזרה בקורס "מבוא לסטטיסטיקה א" הוזמנו 200 סטודנטים. ההסתברות שסטודנט יופיע למפגש היא 0.6, והחלטות הסטודנטים הן בלתי תלויות זו בזו.

א. כמה כסאות יש להכין בכיתה, כדי שבהסתברות של 0.9, לכל סטודנט שיגיע יהיה מקום ישיבה? (העזר במשפט הגבול המרכזי).

ב. רכז מרכז הלימוד, החליט לשבץ את המפגש באודיטוריום המכיל לכל היותר 140 איש (כאשר יש לו כיתה המכילה 99 איש). מה ההסתברות שאכן יזדקקו לאודיטוריום (הכיתה לא תספיק) והוא יספיק לכל הסטודנטים שיגיעו? (העזר במשפט הגבול המרכזי).

ג. מה בקרוב ההסתברות שיגיעו בדיוק 115 סטודנטים?



שאלה 9

לפי טענת חברת האוטובוסים, זמן ההמתנה לאוטובוס בקו "85" מתפלג נורמלית עם תוחלת 10 דקות וסטיית תקן 4 דקות.

א. מה ההסתברות שמשך זמן ההמתנה הממוצע של מדגם מקרי של 36 נוסעים יהיה גבוה מ 12 דקות?

ב. מהו גודל המדגם המינימלי שיש לדגום על מנת להבטיח בהסתברות 0.95 לפחות שמשך זמן ההמתנה הממוצע של הנוסעים יסטה מתוחלת זמן ההמתנה בלא יותר מדקה אחת?



שאלה 10

התקן עבור בלוקים המשמשים לבניה קובע כי תוחלת משקלם צריכה להיות 2 ק"ג עם סטיית תקן של 200 גרם. התפלגות משקל הבלוקים היא נורמלית.

א. טנדר קל הוא בעל יכולת נשיאה של 1810 ק"ג. מה ההסתברות שהטנדר יוכל לשאת 900 בלוקים העומדים בתקן?

ב. מכון התקנים רוצה לבדוק את תוחלת המשקל של בלוקים המיוצרים במפעל מסוים.

מה גודל המדגם שעליו לבדוק אם רוצים שבהסתברות 0.95 לפחות, ממוצע המדגם יסטה מהתוחלת בלא יותר מ- 50 גרם? (הנח שסטיית התקן היא לפי התקן)



שאלה 11

מפעל "תדיר" לייצור סוללות מייצר סוללות שאורך החיים שלהן מתפלג נורמלית עם תוחלת 120 שעות. נתון כי ההסתברות שממוצע אורך החיים של 36 סוללות שנבחרו באופן מקרי יהיה מעל 126 היא 0.0228.

א. מצא את סטיית התקן של אורך חיי סוללה מתוצרת "תדיר".

ב. מהי ההסתברות שממוצע אורך החיים של 81 סוללות מתוצרת "תדיר" שנבחרו באופן מקרי יהיה לכל היותר 118 שעות?

ג. מהו גודל המדגם המינימלי שיבטיח שממוצע אורך החיים במדגם יסטה מתוחלתו בלא יותר מ- 3 שעות בהסתברות 0.98 לפחות?



שאלה 12

אורך החיים של סוללה מתפלג נורמלית עם ממוצע 16 שעות וסטיית תקן 4 שעות.

בביקורת איכות נבדקו 25 סוללות באופן הבא: נבחרה סוללה באופן מקרי והשתמשו בה עד שהתרוקנה, לאחר שהתרוקנה נבחרה סוללה נוספת באופן מקרי וכך הלאה. לזמן שחלף עד שהתרוקנה הסוללה ה- 25 נקרא זמן השימוש.

א. מהם התוחלת וסטיית התקן של זמן השימוש?

ב. מה ההסתברות שזמן השימוש יעלה על 450 שעות?

ג. מה ההסתברות שזמן השימוש יהיה בתחום שבין 360 לבין 420 שעות?

ד. מה ההסתברות שזמן השימוש לא יסטה מתוחלתו ביותר מ- 10 שעות?



שאלה 13

על-פי נתוני משרד הבריאות הסיכוי שאדם המקבל חיסון נגד שפעת יפתח נוגדנים נגד המחלה הוא 0.6. מדגם מקרי של 30 אנשים קיבלו את החיסון. חשב:

א. מהי בקירוב ההסתברות שלכל היותר 20 אנשים יפתחו נוגדנים?

ב. מהי בקירוב ההסתברות שבדיוק 8 אנשים לא יפתחו נוגדנים?

ג. מהי בקירוב ההסתברות שמספר האנשים שיפתחו נוגדנים יהיה לפחות 15 ולא יותר מ- 22?



שאלה 14

הזמנת לחתונה 400 אורחים. ההסתברות שאורח שהוזמן יגיע לחתונה היא 0.8,

והחלטות המוזמנים בלתי תלויות זו בזו.

א. מה ההסתברות שבחתונה ישתתפו לכל היותר 335 אורחים ?

ב. מה ההסתברות שבחתונה ישתתפו לפחות 310 אורחים ?

ג. כמה מנות יש להכין, כדי שבהסתברות 0.98 לכל אורח תהיה מנה ?



שאלה 15

מטילים מטבע הוגן 50 פעמים.

א. מהי בקירוב ההסתברות שמספר ה"עצים" יהיה שונה מתוחלת מספר ה"עצים" ביותר מ- 10 ?

ב. הועלה חשד שהמטבע אינו הוגן והוחלט לדחות את ההנחה שהמטבע הוגן אם מספר ה"עצים"

שיתקבלו ב- 50 הטלות יהיה גדול מ- 25 או קטן מ- 21.

מה ההסתברות שנדחה את ההנחה שהמטבע הוגן למרות שהיא נכונה ?



שאלה 16

מתוך אוכלוסייה שבה 30% מעשנים, נבחר מדגם מקרי של 40 אנשים.

א. מהי בקירוב ההסתברות שיתקבלו במדגם לכל היותר 10 מעשנים ?

ב. מהי בקירוב ההסתברות שבמדגם יהיו בדיוק 10 מעשנים ?

ג. מהי בקירוב ההסתברות שבמדגם יהיו יותר מ- 15 מעשנים ?

ד. מהו גודל המדגם המינימלי כדי שבהסתברות 0.95 לפחות,

הסטייה בין פרופורציה המעשנים במדגם לבין הפרופורציה באוכלוסייה, לא תעלה על 2% ?



שאלה 17

בתיאטרון מסוים יש 500 מקומות ישיבה.

ידוע מניסיון העבר כי 80% מבין המזמינים כרטיסים אכן מגיעים לצפות בהצגה.

להצגה מסוימת הוזמנו 550 כרטיסים.

מה בקירוב ההסתברות שיהיו באולם יותר מ- 35 מקומות פנויים?



שאלה 18

ההנחה היא שההסתברות ללידת בן זהה להסתברות ללידת בת. נדגמו באופן מקרי 120 לידות

א. מהי בקירוב ההסתברות שמספר הבנים יהיה שונה מתוחלת מספר הבנים ביותר מ- 10

ב. חוקר חושד שההסתברות ללידת בן אינה זהה להסתברות ללידת בת.
הוא החליט לדחות את ההנחה שההסתברות זהה, אם מספר הבנים שיתקבלו ב- 120 לידות יהיה גדול מ- 60 או קטן מ- 46. מהי בקרוב ההסתברות לכך שידחה את ההנחה?



שאלה 19

מכונה למילוי בקבוקים בתרופה מסוימת, מכוונת למלא בממוצע 200 מ"ג לבקבוק עם סטיית תקן של 8 מ"ג.

א. הועלה חשד כי המכונה התקלקלה והיא ממלאה כמות נמוכה מהנדרש.

נבדק מדגם מקרי של 100 בקבוקים, ובמדגם נתקבל ממוצע של 197.8 מ"ג.

1. מהי מובהקות התוצאה, כלומר, מהי רמת המובהקות המינימלית עבורה יוחלט שהחשד מוצדק?

2. האם החשד מוצדק ברמת מובהקות 0.05? נמק.

ב. בהמשך לסעיף א(2), מהי עצמת המבחן אם אכן החשד מוצדק והמכונה ממלאה בפועל כמות של 197 מ"ג?

ג. מה גודל המדגם המינימלי שעבורו ההסתברות לדחייה מוטעית של השערת האפס לא תעלה על 0.01

וההסתברות לקבלה מוטעית של השערת האפס, אם אכן החשד מוצדק והמכונה ממלאה בפועל כמות של 197 מ"ג,

לא תעלה על 0.02?



שאלה 20

משקל תושבי הארץ מתפלג נורמלית עם ממוצע 70 ק"ג וסטיית תקן 6 ק"ג.

בבנין רב קומות הותקנה מעלית המיועדת לשאת 9 אנשים.

העומס אותו מסוגלת המעלית לשאת הוא 684 ק"ג

מה ההסתברות שהמעלית תחזיק מעמד בעומס אם נכנסו למעלית 9 אנשים?



שאלה 21

חנות לתכשיטים מוכרת בין שאר מוצריה מדליון מיוחד שעוצב ע"י אמן ישראלי ידוע.

מניסיון ידוע ש - 20% מהלקוחות הנכנסים לחנות רוכשים מדליון זה.

במשך חודש בקרו בחנות 625 לקוחות.

מהי בקרוב ההסתברות שנרכשו בין 120 ל - 140 מדליונים באותו חודש?



שאלה 22

חברת התעופה "רקיע" מעוניינת לדעת מהו משקלו הממוצע של המטען האישי של כל נוסע. סטטיסטיקאי החברה הציע לבדוק מדגם מקרי של n מנוסעים ולחשב את המשקל הממוצע של מטעניהם. הערכת החברה שהתפלגות המשקל היא נורמלית עם סטיית תקן של 5 ק"ג. מה צריך להיות n כך שבהסתברות 0.95, לא יסטה ממוצע המדגם מממוצע האוכלוסייה ביותר מ- 0.49 ק"ג?



שאלה 23

ממוצע הציונים במבחן במתמטיקה באוכלוסייה מסוימת הוא 72, ושונות הציונים 49. נלקח מדגם מקרי של 64 תצפיות מאוכלוסייה זו. א. מהן התוחלת, השונות וסטיית התקן של ממוצע הציונים במדגם הציונים במדגם? ב. מה ההסתברות שממוצע הציונים במדגם ינוע בין 70 ל- 80? ג. מהו הציון במבחן ש- 60% ממוצעי המדגמים בגודל 64 יהיו נמוכים ממנו?



שאלה 24

אחוז בעלי כלי הרכב בקרב הסטודנטים הוא 25%. מבין 1000 סטודנטים שנבחרו באופן מקרי. מצא: א. מה ההסתברות שיהיו לפחות 220 בעלי רכב? ב. מה ההסתברות שהשכיחות היחסית של בעלי רכב בקרב המדגם תהיה בין 0.24 לבין 0.27? ג. כמה סטודנטים יש לבחור, כך שבסיכוי 0.95 לפחות, השכיחות היחסית של בעלי הרכב במדגם תהיה לפחות 0.23? (השתמש במשפט הגבול המרכזי)



שאלה 25

נתונה אוכלוסייה המתפלגת נורמלית עם תוחלת $\mu = 10$ ושונות $\sigma^2 = 9$. א. מה ההסתברות שתצפית שנדגמה באופן מקרי מתוך האוכלוסייה תהיה בעלת ערך הגדול מ- 16? ב. מה ההסתברות שמדגם מקרי בגודל n יהיה בעל ממוצע קטן מ- 10? ג. מה אחוז המדגמים בגודל 100 מתוך האוכלוסייה בעלי ממוצע בתחום $9.85 \leq \bar{X} \leq 10.45$? ד. מה ההסתברות שממוצע מדגם מקרי בגודל 400 מתוך האוכלוסייה יהיה קטן מ- 9.85?



שאלה 26

במיכל דלק של מכונית מסוימת יש 40 ליטר דלק.

מספר הליטרים שבעל המכונית צריך ליום מתפלג נורמלית עם ממוצע 5 ליטרים וסטיית תקן 2 ליטר.

מה ההסתברות שכמות הדלק תספיק ל- 9 ימי נסיעה?



שאלה 27

משקל של ארגז פירות בק"ג שלמים (X) הוא בעל פונקציית ההסתברות הבאה:

X	2	4	5	8
$P(X)$	0.125	0.25	0.25	0.375

מושבניק מהדרום מעוניין להביא יום יום 750 ארגזי פירות לשוק.

א. האם משאית של 4275 ק"ג תספיק כדי להבטיח שההסתברות לכך שהוא יוכל להוביל מטען זה לשוק תהיה לפחות 0.9?

ב. מה גודל המשאית (המשקל המקסימלי המותר) שעליו לשכור כדי להבטיח את הדרישה בסעיף א' בהסתברות של 0.95 לפחות?



שאלה 28

באוניברסיטה גדולה מסוימת נמצא כי 54% הן סטודנטיות. למחקר מסוים נבחרו באופן מקרי 35 נחקרים.

א. מהי בקירוב ההסתברות שבמחקר ישתתפו לכל היותר 20 סטודנטיות?

ב. מהי בקירוב ההסתברות שבמחקר ישתתפו 15 סטודנטים בדיוק?

ג. מהי בקירוב ההסתברות שבמחקר מהוות הסטודנטיות רוב?

ד. מהי בקירוב ההסתברות שהשכיחות היחסית של הסטודנטיות במחקר

לא תסטה משכיחותן היחסית באוכלוסייה כולה ביותר מ-0.05?

ה. במחקר אחר נדגמו באופן מקרי 250 נחקרים. מהי בקירוב ההסתברות שהשכיחות היחסית של

הסטודנטיות במחקר הנוסף לא תסטה משכיחותן היחסית באוניברסיטה כולה ביותר מ-0.03?



שאלה 29

במפעל מסוים מייצרים צלחות. 10% מהצלחות נפסלות בגלל פגמים.

א. דגמו 20 צלחות. מצאו חסם להסתברות כי פרופורציית הצלחות הפסולות במדגם תסטה מ-10% בפחות מ-7%.

ב. נלקח מדגם של 400 צלחות.

מה ההסתברות, בקירוב, שלכל היותר 43 צלחות במדגם תהיינה פסולות?

ג. במחלקת ביקורת האיכות טוענים כי אחוז הצלחות הפסולות במפעל הוא לכל היותר 9%.

לבדיקת טענתם נלקח מדגם של 400 צלחות, מתוכן 30 נמצאו פסולות.

האם תקבלו את טענת מחלקת בקורת האיכות על פי רווח סמך ברמת סמך של 95%? נמקו!



שאלה 30

בפארק שעשועים מופעל משחק מזל.

פונקציית ההסתברות של הרווח X בשקלים בהפעלה אחת של המשחק היא:

X	10	0	-20
$P(X)$	0.7	0.2	0.1

א. חשב את התוחלת והשונות של X .

ב. המשחק מופעל 300 פעם. מהי בקירוב ההסתברות שהרווח הממוצע יהיה בין 3 ל-7 ש"ח?

ג. כמה הפעלות דרושות אם מעוניינים שהרווח הממוצע לא יסטה מתוחלתו ביותר מ-2 ש"ח בהסתברות של 0.99 לפחות?



שאלה 31

במכון "העין הטובה" מבצעים ניתוחים להסרת משקפיים, שיעור ההצלחה הוא 84%.

בחודש מרץ נקבעו ניתוחים ל-120 אנשים.

א. מהו בקירוב ההסתברות שבחודש מרץ יהיו במכון "העין הטובה" 105 ניתוחים מוצלחים?

ב. מהי בקירוב ההסתברות שבחודש מרץ יהיו במכון "העין הטובה" לפחות 105 ניתוחים מוצלחים?

ג. מהי בקירוב ההסתברות ששיעור הניתוחים המוצלחים בחודש מרץ יהיה בין 78% ל-82%?

ד. בחודש אפריל נקבעו ניתוחים ל-87 אנשים.

מה בקירוב ההסתברות שבחודש אפריל יהיו במכון "העין הטובה" יותר מ-68 ניתוחים מוצלחים?

ה. כמה ניתוחים יש לקבוע לחודש מאי, אם מעוניינים שבהסתברות 0.98 לפחות, הסטייה בין השכיחות היחסית של ניתוחים מוצלחים במאי לבין הפרופורציה האמיתית שלהם לא תעלה על 0.015? ענה תוך שימוש במשפט הגבול המרכזי ללא תיקון רציפות.



שאלה 32

מספר המקומות במטוס ג'מבו הוא 400. מנתונים סטטיסטיים של חברת תעופה מסוימת ידוע שההסתברות שנוסע הרשום לטיסה אכן יגיע אליה היא 0.9. לאור זאת מאפשרת חברת התעופה לסוכני הנסיעות לרשום לטיסה מעל ל-400 נוסעים (פעולה המכונה Over Booking).

א. לטיסה מסוימת נרשמו 450 נוסעים.

מהי בקירוב ההסתברות שהמטוס יספיק לכל הנוסעים שאכן יגיעו?

ב. אם נרשמו לטיסה 425 נוסעים. מה בקירוב ההסתברות שיהיו במטוס יותר מ-20 מקומות פנויים?



שאלה 33

במבחן מסויים ניתן לקבל אחד משלושת הציונים הבאים: 100, 80, 60.

30% מהנבחנים מקבלים 60, 50% מקבלים 80 והיתר מקבלים 100.

א. נגדיר X – הציון במבחן, רשום את פונקציית ההסתברות של X .

ומצא את התוחלת, השונות וסטית התקן של X .

ב. מה ההסתברות שהציון של תלמיד הניגש למבחן יהיה מעל 60?

ג. יהי $\bar{X}$ ממוצע מדגם מקרי של 2 תלמידים מאוכלוסיית הנבחנים במבחן זה.

מצא את פונקציית ההסתברות של $\bar{X}$.

ד. חשב את התוחלת והשונות של ממוצע המדגם בשתי דרכים:

(1) בחישוב ישיר מתוך פונקציית ההסתברות שמצאת בסעיף ג'

(2) תוך שימוש בתכונות התפלגות הדגימה של הממוצע.

ה. מה ההסתברות שהציון הממוצע של 2 תלמידים יהיה לפחות 80?

ו. מהן התוחלת והשונות של ממוצע מדגם של 4 תלמידים מאוכלוסיית הנבחנים במבחן זה?

ז. מה ההסתברות שהציון הממוצע של 36 תלמידים הנבחנים במבחן זה יהיה לפחות 80?



שאלה 34

במבחן מסויים ניתן לקבל אחד משלושת הציונים הבאים: 100, 80, 60

30% מהנבחנים מקבלים 60, 50% מקבלים 80 והיתר מקבלים 100.

א. נגדיר X – הציון במבחן, רשום את פונקציית ההסתברות של X .

ומצא את התוחלת, השונות וסטית התקן של X .

ב. מה ההסתברות שהציון של תלמיד הניגש למבחן יהיה לכל היותר 80?

ג. יהי $\bar{X}$ ממוצע מדגם מקרי של 2 תלמידים מאוכלוסיית הנבחנים במבחן זה.

מצא את פונקציית ההסתברות של $\bar{X}$.

ד. חשב את התוחלת והשונות של ממוצע המדגם בשתי דרכים:

(1) בחישוב ישיר מתוך פונקציית ההסתברות שמצאת בסעיף ג'

(2) תוך שימוש בתכונות התפלגות הדגימה של הממוצע.

ה. מה ההסתברות שהציון הממוצע של 2 תלמידים יהיה לכל היותר 80?

ו. מהן התוחלת והשונות של ממוצע מדגם של 4 תלמידים מאוכלוסיית הנבחנים במבחן זה?

ז. מה ההסתברות שהציון הממוצע של 36 תלמידים הנבחנים במבחן זה יהיה לכל היותר 80?



שאלה 35

מסובבים רולטה הוגנת שעליה רשומות הספרות 0-4

נסמן ב- X_i את תוצאת הרולטה בסיבוב ה- i .

א. חשב את $E(X)$ ואת $V(X)$.

ב. מסובבים את הרולטה 1,000 פעם.

מהי בקירוב ההסתברות שממוצע הספרות לא יסטה מתוחלתו ביותר מ- 0.1?

ג. כמה פעמים יש לסובב את הרולטה כך שבהסתברות 0.9 לפחות ממוצע הספרות המתקבלות יהיה בין 1.9

ל- 2.1? העזר במשפט הגבול המרכזי.



שאלה 36

המשטרה נוהגת לערוך מדי פעם ביקורת תקינות רכב בכבישים.

מנתוני משרד התחבורה נוסעים על הכבישים בארץ 20% מכוניות שאינן תקינות

א. מהי בקרוב ההסתברות שמתוך 64 מכוניות שנבדקות ביום אחד יותר מ- 15 תימצאנה בלתי תקינות?

ב. מה בקרוב ההסתברות שמתוך 64 מכוניות שנבדקות ביום אחד תימצאנה בדיוק 14 בלתי תקינות?

ג. כמה מכוניות יש לבדוק אם מעוניינים שבהסתברות 0.9 לפחות, הסטייה בין

השכיחות היחסית של המכונות שאינן תקינות במדגם לבין הפרופורציה שלהן באוכלוסיה לא תעלה על 0.04?



שאלה 37

מילוי קרטון חלב נעשה בעזרת מכונה אוטומטית.

כמות המילוי מתפלגת בקירוב נורמלית עם ממוצע 1 ליטר וסטיית תקן של 0.1 ליטר.

מה ההסתברות ש- 16 קרטוני חלב שנבחרו באורח מקרי יהיו בעלי כמות חלב ממוצעת הפחותה מ 0.95 ליטר?



שאלה 38

בקופסה נמצאים 5 כרטיסים ממוספרים בספרות 0,2,3,4,6.

מוציאים מהקופסה באופן מקרי ועם החזרה 400 כרטיסים.

א. מה ההסתברות שסכום המספרים שהופיעו על הכרטיסים שהוצאו גדול מ- 1250?

ב. מה ההסתברות שהכרטיס הממוספר 6 הופיע פחות מ- 100 פעמים?

ג. כמה כרטיסים יש להוציא עם החזרה מהקופסה, כך שבהסתברות 0.9 לפחות ממוצע המספרים שהופיעו על הכרטיסים יהיה בין 2.8 ל- 3.2? העזר במשפט הגבול המרכזי.



שאלה 39

אחוז הנשים שסיימו לימודיהן במדעי המחשב באוניברסיטה הפתוחה ומוצאות עבודה במקצוען, הינו 40%.

מה ההסתברות שמבין 250 נשים, בוגרות מדעי המחשב, שנבחרו באופן מקרי, לפחות 115 ימצאו עבודה במקצוען?



שאלה 40

חוקר שווקים המתמחה בשוק מוצרי החלב מעריך כי 40% מהקונים מעדיפים לרכוש חלב בקרטונים

על פני שקיות הניילון המקובלות. לבדיקת הערכתו בחר מדגם מקרי של 140 קונים.

בהנחה שהערכת החוקר נכונה,

מהי בקירוב ההסתברות שבמדגם יותר מ- 60 נשאלים יעדיפו חלק בקרטונים?



שאלה 41

בקניון הותקנו מדרגות נעות המיועדות לשאת בשעות העומס קיבולת מקסימלית של 120 אנשים. המדרגות תוכננו לשאת עומס מירבי של 8900 ק"ג. תוחלת המשקל של אוכלוסיית המשתמשים במדרגות הנעות היא 72 ק"ג עם סטיית תקן 8 ק"ג. מהי בקירוב ההסתברות שהמדרגות יתקלקלו בגלל עומס יתר, כאשר הן עמוסות ב- 120 אנשים בשעות העומס?



שאלה 42

לפי טענת משרד הבריאות, זמן ההמתנה הממוצע לבדיקת MRI הוא 10 ימים וסטיית תקן 4 ימים. א. מה ההסתברות שמשך זמן ההמתנה הממוצע של מדגם מקרי של 36 חולים יהיה גבוה מ- 12 ימים? ב. מה ההסתברות שמשך זמן ההמתנה הממוצע של מדגם מקרי של 36 חולים יסטה מתוחלת זמן ההמתנה בלא יותר מיום אחד? ג. מהו גודל המדגם המינימלי שיש לדגום על מנת להבטיח בהסתברות 0.95 לפחות שמשך זמן ההמתנה הממוצע של החולים יסטה מתוחלת זמן ההמתנה בלא יותר מיום אחד?



שאלה 43

התפלגות גובה החיילים המתגייסים לצה"ל היא נורמלית עם תוחלת 170 ס"מ וסטיית תקן 10 ס"מ. מאמן כדורסל מעוניין להקים נבחרת כדורסל מגדוד חיילים המונה 150 חיילים. המאמן מעוניין בחיילים שגובהם 180 ס"מ לפחות. מהי בקירוב ההסתברות שהמאמן ימצא בגדוד יותר מ- 16 חיילים לנבחרת?

✓ אמידה

ביחידה זו נלמד לאמוד מדדים (פרמטרים) של אוכלוסייה ומשתנים מקריים בעזרת מדדים (סטטיסטיים) של המדגם

ונבטא את דיוק האומדן הזה ע"י רווח שנקרא "רווח סמך".

נדון בבעיות של הערכת גודלו של פרמטר מסוים באוכלוסייה נתונה, על סמך מדגם מקרי מתוך האוכלוסייה הערכה כזו נקראת **אמידה**.

אמידה נכנסת לשימוש כאשר לא ידוע הפרמטר באוכלוסייה ואנו צריכים לתת הערכה באמצעות אחד הסטטיסטיים.

האמידה נעשית בשני שלבים:

✓ **אמידה נקודתית** – כאשר אנו מעריכים את הפרמטר באמצעות ערך מספרי אחד ויחיד. מציאת סטטיסטי שיהווה אומד נקודתי לפרמטר הלא ידוע של האוכלוסייה.

✓ **אמידה מרווחית** – הערכה באמצעות טווח של ערכים. מציאת רווח מקרי שנקבע ע"י האומד הנקודתי ומכיל בתוכו בהסתברות גבוהה את הפרמטר הנאמד. הרווח קרוי רווח סמך.

openbook
המרכז לקידום אקדמי

פרמטר וסטטיסטי

פרמטר – תכונה של האוכלוסייה.

פרמטר הוא גודל קבוע המאפיין את האוכלוסייה, לעומת סטטיסטי שהוא משתנה מקרי, שמקבל בכל מדגם תוצאה שונה.

סטטיסטי – תכונה של המדגם.

כל אחד ממדדי המרכז והפיזור מהווה סטטיסטי.

אוכלוסייה – כל האובייקטים שרלוונטיים לתכונה מסוימת	מדגם – תת קבוצה של האוכלוסייה
θ פרמטר – מדד שלקוח מתוך אוכלוסייה (מאפיין של האוכלוסייה)	$\hat{\theta}$ סטטיסטי – מדד שלקוח מתוך המדגם. המדד המחושב מהמדגם.
$\mu = E(X)$ – ממוצע אוכלוסייה = תוחלת	$\bar{X}$ – ממוצע המדגם $\bar{X} = \frac{\sum x}{n}$
$\sigma^2 = V(X)$ – שונות באוכלוסייה = שונות של מ"מ	$\hat{\sigma}^2$ – שונות במדגם $\hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$
σ סטיית תקן באוכלוסייה	$\hat{\sigma}$ – סטיית תקן במדגם
P – פרופורציית בעלי התכונה באוכלוסייה (שכיחות יחסית/אחוז)	$\hat{P}$ – פרופורציית בעלי התכונה במדגם.
ערך הפרמטר של האוכלוסייה הוא קבוע ובד"כ לא ידוע. הערך לא תלוי במדגם שאנו דוגמים ואינו משתנה ממדגם למדגם.	הערך של הסטטיסטי תלוי במדגם שנבחר ומשתנה ממדגם למדגם. מכיוון שהסטטיסטי משתנים ממדגם למדגם, יש להם התפלגות. להתפלגות של הסטטיסטי קוראים התפלגות הדגימה.

כיצד אומדים פרמטר?

אומדים פרמטר בעזרת סטטיסטי מתאים.

סטטיסטי הוא ביטוי מתימטי מוגדר (פונקציה) של תצפיות המדגם:

$$X_1, X_2, \dots, X_n$$

באופן כללי נהוג לסמן את הפרמטר אותו יש לאמוד ב- θ (תטא),

את האומדן (הסטטיסטי) המשמש לאמידת הפרמטר ב- $\hat{\theta}$

הסימן $\hat{\theta}$ מעל הסטטיסטי הוא ציון מקובל לאומדן

אנו אומרים במקרה זה ש- $\hat{\theta}$ הוא אומדן נקודתי ל- θ

ואת הערך שקיבל האומד $\hat{\theta}$ במדגם מסוים ב- θ

ההבחנה בין אומד לאומדן

אומד estimator - השם של האומד: ממוצע, סטיית תקן, פרופורציה.

סטטיסטי המשמש לאמידת פרמטר נקרא אומד.

אומדן estimate - הערך המספרי של האומד.

הערך הספציפי שמקבל האומד במדגם שנבחר.



דוגמא 1

לפי מדגם מקרי של 100 נשים יש לאמוד את הגיל הממוצע בו נישאות נשים בישובים העירוניים בישראל.



דוגמא 2

מכון התקנים מעוניין לאמוד את תוחלת אורך החיים של נורה מתוצרת מפעל מסוים. המכון יעשה זאת ע"י בדיקת מדגם מקרי של נורות מתוך אלו המיוצרות במפעל הנידון



דוגמא 3

ממוצע האוכלוסיה או תוחלת המשתנה המקרי הוא הפרמטר μ ($\theta = \mu$)



דוגמא 4

משתנה מקרי X הוא תוצאת הטלת קובייה תקינה.

תוחלתו $\mu = E(X) = 3.5$.

נניח ש- μ לא ידוע לנו ואנחנו רוצים לאמוד אותו



המשך דוגמא 4

מוצע לאמוד את μ , תוחלת תוצאת הקובייה, בעזרת מדגם בן שתי הטלות X_1, X_2 . ומוצעים שלושה אומדים:

$$1. \hat{\theta}_1 = \frac{1}{2}(X_1 + X_2)$$

$$2. \hat{\theta}_2 = \frac{1}{3}(2X_1 + X_2)$$

$$3. \hat{\theta}_3 = \frac{1}{3}(X_1 + X_2 + 3)$$

בדוק לגבי כל אחד מהאומדים אם הוא חסר הטיה, נזכיר כי מהגדרת המדגם נובע שגם ל- X_1 , וגם ל- X_2

אותה התפלגות כמו ל- X ובפרט $E(X_1) = E(X_2) = 3.5$



openbook

המרכז לקידום אקדמי

סוגי אמידה

(1) אמידת התוחלת באוכלוסייה – אם $E(\bar{x}) = \mu$ אזי ממוצע המדגם $\bar{x}$ הוא חסר הטיה μ

(2) אמידת השונות באוכלוסייה –

$$\hat{S}^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} = \frac{\sum x_i^2}{n} - \bar{x}^2 \quad \text{כך: השונות במדגם מחושבת כ:}$$

$\hat{S}_x^2$ השונות במדגם – הוא אומד מוטה לשונות באוכלוסייה כיוון ש- $E(S_x^2) \neq \sigma_x^2$

$$\hat{S}_x^2 - \text{הוא אומד נקודתי חסר הטיה (הטוב ביותר) לשונות באוכלוסייה ומחושב כך: } \hat{S}^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1}$$

כאשר השונות באוכלוסייה לא ידועה, נשתמש במקומה ב- $\hat{S}_x^2$ כאשר $\hat{S}_x^2 = \frac{n}{n-1} \cdot S_x^2$

(3) פרופורציה באוכלוסייה p , האומד הנקודתי החסר הטיה לפרופורציה באוכלוסייה הוא $\hat{p}$ פרופורציה

במדגם.

המשך דוגמא 4

בסדרה מסוימת של 5 הטלות (מדגם מסוים) קיבלנו את התוצאות: 3,6,1,1,4

$$\bar{x} = \frac{1}{5} (3 + 6 + 1 + 1 + 4) = 3 \quad \text{שלפיהן אמדנו את התוחלת:}$$

נשתמש עתה באותן הטלות לאמוד את השונות σ^2 בעזרת האומד $\hat{S}^2$.

תרגיל 1

במדגם מקרי של 20 תלמידים נמדדו גובהיהם (בס"מ):

183,162,160,156,172,177,158,166,169,201,191,170,168,172,168,185,171,

179,165,174.

אמוד על פי מדגם זה את תוחלת ושונות הגובה של התלמידים.

אמידה מרווחית

אמד נקודתי הינו "ניחוש" לערך הפרמטר הלא ידוע.

הפרדוקס הוא שהערך שהוא נותן כמעט תמיד שגוי.

כמעט בשום מקרה לא יהיה הערך של האומדן שווה בדיוק לערך הלא ידוע של הפרמטר.

לכן, נרצה למצוא רווח "סביר" עבור הפרמטר באופן שנוכל להיות כמעט בטוחים שהפרמטר אמנם כלול ברווח שיתקבל.

לתחום של ערכי הפרמטר $[L, U]$ שעבורם סביר לקבל את התוצאה שהתקבלה במדגם קוראים רווח סמך עבור הפרמטר הנאמד.

מידת הביטחון נקראת רמת הסמך: $(1 - \alpha) \cdot 100$ אחוז עבור θ .

$$P(L < \theta < U) = 1 - \alpha$$

אמידה באמצעות רב"ס (רווח בר סמך)

רווח בר סמך הוא טווח של ערכים שמעריכים את הפרמטר באוכלוסייה.

הטווח נבנה באמצעות האומד הנקודתי חסר ההטיה, שלו מוסיפים ומחסירים גודל מסוים.

הגודל שמוסיפים ומחסירים הוא שרירותי והוא תלוי בעד כמה אנחנו רוצים להיות בטוחים שהפרמטר כלול בתוך הטווח.

רווח סמך – תחום מספרים סימטרי המבוסס על נתוני המדגם, שניתן לומר שיש סיכוי של $1 - \alpha$ שהפרמטר הנאמד נמצא בתוכו.

✓ $\bar{x}$ ממוצע המדגם הוא אומד נקודתי לממוצע באוכלוסייה (תוחלת μ)

✓ הערך $\bar{x}$ שיקבל $\bar{X}$ במדגם מסוים יהיה בד"כ שונה מהערך של הפרמטר μ

⇓

שגיאת האמידה $(\bar{x} - \mu)$ תהיה שונה מאפס

כדי לאמוד את ממוצע האוכלוסייה אנו בונים מרווח בטחון = רווח סמך סביב $\bar{x}$ וסביבו נבנה את טווח הערכים

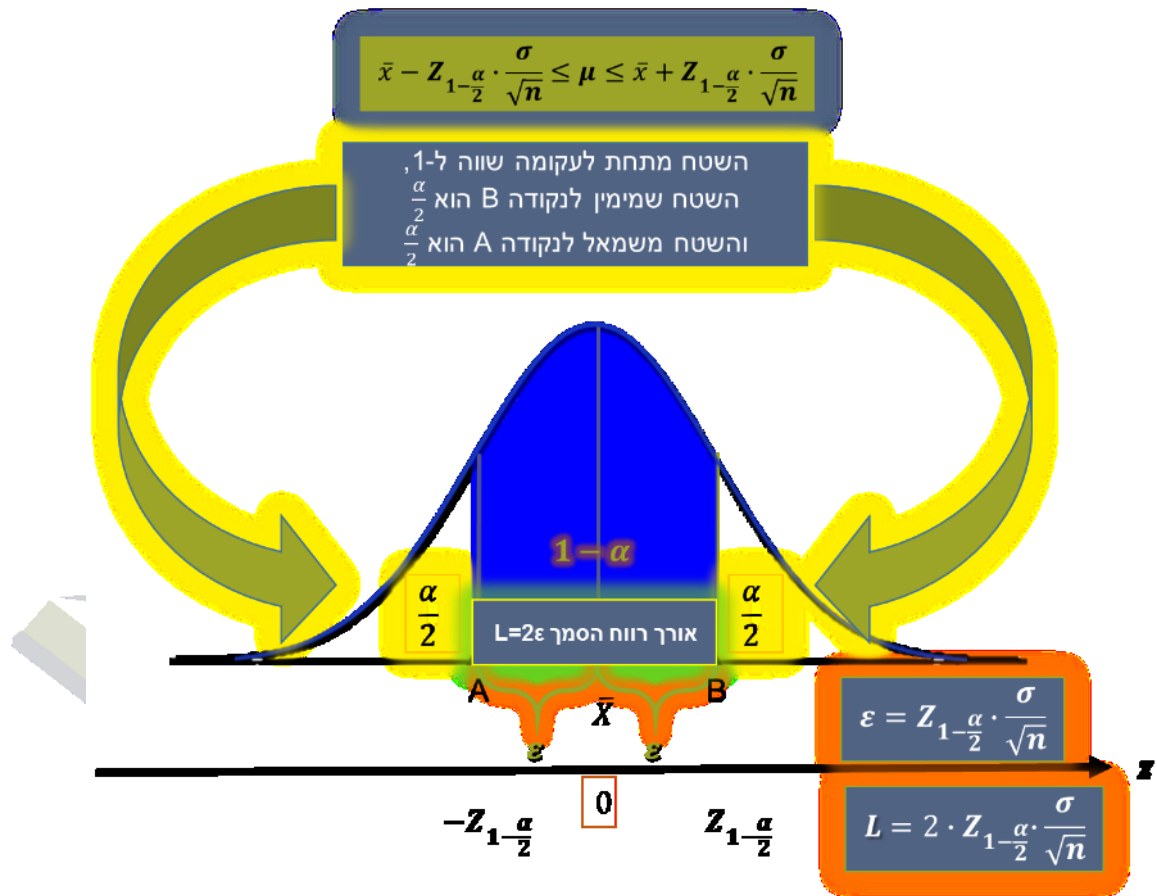
$1 - \alpha$: רווח הסמך = מרווח ביטחון = השטח

רווח הסמך מגדיר את הסיכוי/ההסתברות ש- μ הפרמטר אכן כלול בתוך רווח הסמך (A, B) .

$$P(A < \mu < B) = 1 - \alpha$$

A – גבול תחתון של רווח הסמך

B – גבול עליון של רווח הסמך



תכונות רווח סמך

✓ הרווח סמך סימטרי סביב ממוצע המדגם $\bar{X}$.

✓ אורך הרווח, $L = 2 \cdot Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$, הוא קבוע ותלוי בשלושה גורמים:

(1) ברמת הסמך, $1 - \alpha$: ככל שרמת הסמך גבוהה יותר אורך הרווח גדל.

(2) בסטיית התקן (σ) של האוכלוסייה ממנה הוצא המדגם: ככל שסטיית התקן גדולה יותר אורך הרווח גדול יותר.

(3) בגודל המדגם (n) ככל שמגדילים את המדגם אורך הרווח קטן.

אם המדגם יגדל פי k אזי אורך הרווח יקטן פי $\sqrt{k}$.

אם נרצה להקטין את אורך הרווח פי k אזי נגדיל את המדגם פי k^2

✓ הסטייה בין האומד לפרמטר, $|\bar{X} - \mu|$, נקראת **שגיאת האמידה** (או טעות הדגימה) ושווה למחצית אורך הרווח סמך.

גודל המדגם שיבטיח ברמת סמך $1 - \alpha$, שהסטייה בין האומד $\bar{X}$ לפרמטר μ , לא תעלה על גודל מסוים

$$n \geq \left(\frac{Z_{1-\frac{\alpha}{2}} \cdot \sigma}{\varepsilon} \right)^2 \quad (\text{כלומר אורך הרווח לא יעלה על } 2\varepsilon) \text{ הוא:}$$

רווח בר סמך לתוחלת יש 2 אפשרויות:

מקרה ראשון: השונות σ^2 ידועה:

1. אוכלוסייה נתונה נורמלית כאשר יש n כלשהו. $X \sim N(\mu = ?, \sigma^2)$

2. אוכלוסייה מהתפלגות כלשהי ו $n \geq 30$

בשני המקרים
$$\bar{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

מקרה שני: השונות σ^2 לא ידועה:

1. התפלגות כלשהי (יתכן גם נורמלית), $n \geq 30$ אז
$$\bar{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{s}}{\sqrt{n}}$$

2. אוכלוסייה נורמלית אבל $n < 30$ אז:
$$\bar{x} \pm t_{\frac{\alpha}{2}}^{n-1} \cdot \frac{\hat{s}}{\sqrt{n}}$$

openbook
המרכז לקידום אקדמי

מודל (1) רווח בר סמך לתוחלת כשהשונויות באוכלוסייה σ^2 ידועה

לצורך אמידת התוחלת נחשב (או אם נתון נשתמש) את האומד הנקודתי חסר ההטיה $\bar{x}$ וסביבו נבנה את טווח הערכים, כאשר ידוע ש- $\bar{x}$ מתפלג נורמלית $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$.

לצורך בניית הרב"ס יש להגדיר את **רמת הסמך = רמת הביטחון** שמסומנת ב-

$1 - \alpha$ **רווח הסמך = השטח** = מגדירה את הסיכוי/ההסתברות ש- μ הפרמטר אכן כלול בתוך רווח הסמך.

אורך/טווח הרב"ס $L = 2 \cdot Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ – גבול עליון – גבול תחתון

$$2\varepsilon = L \Rightarrow \varepsilon = \frac{L}{2} = Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

כאשר ε היא שגיאת האמידה המקסימלית

$$\bar{x} - Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

$$n \geq \left(\frac{Z_{1-\frac{\alpha}{2}} \cdot \sigma}{\varepsilon} \right)^2 \quad \text{חישוב } n \text{ מינימלי:}$$

רווח סמך לאמידת μ בעזרת האומד $\bar{x}$ נקבע בעזרת ההתפלגות של $\bar{X}$:

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

לכל $z > 0$ מתקיים:

$$P(-z < Z_{\bar{X}} < z) = \Phi(z) - \Phi(-z) = 2\Phi(z) - 1$$



שאלה 1

אורך החיים (בשעות) של נורות מתוצרת מפעל 'נורלי' מתפלג נורמלית עם סטיית תקן $\sigma = 22$.

לאמידת התוחלת של אורך חיי הנורות נבדקו 16 נורות מקריות ונמצא שאורך החיים הממוצע שלהן הוא $\bar{x} = 863$.

(1) מצא רווח סמך ברמה של 90% לתוחלת μ .

(2) נניח שבמדגם של 100 נורות התקבל $\bar{x} = 868$. מהו עתה רווח-סמך ברמה של 90% עבור μ ?

(3) מהו גודל המדגם, n , שיבטיח ברמה של 90% שהערך האמיתי של μ אינו שונה מממוצע המדגם ביותר משעה אחת?



שאלה 2

במכונה אוטומטית למשקאות חמים, כמות המשקה הנמזגת לכוס (בגרמים) מפולגת בקירוב לפי התפלגות נורמלית בעלת סטיית תקן $\sigma = 20$.

א. מהו רווח סמך ברמת סמך של 95% לתוחלת כמות הנוזל הנמזגת, אם במדגם של 16 כוסות התקבל ממוצע של 280 גרם לכוס?

ב. איזה גודל מדגם יבטיח שברמת סמך זו לא ייסטה ממוצע המדגם מן התוחלת ביותר מאשר 2 גרם.



שאלה 3

מפעל צמיגים מייצר צמיגים כך שהלחץ המקסימלי שלהם מתפלג נורמלית עם סטיית תקן 0.7 אטמוספרות. נלקח מדגם בן 50 צמיגים ונמצא כי ממוצע הלחץ המקסימלי שלהם היה 2.8 אטמוספרות.

א. מצא רווח סמך ברמת סמך 95% לתוחלת הלחץ המקסימלי של הצמיגים במפעל.

ב. כיצד תשתנה תשובתך לסעיף א' אם מספר הצמיגים במדגם היה 100?

ג. מהו גודל המדגם שיבטיח ברמת סמך של 95% שהערך האמיתי של התוחלת אינו שונה מממוצע המדגם ביותר מ 0.1 אטמוספרות?



שאלה 4

חוקר התבקש לאמוד את ממוצע שעות העבודה בשבוע של נשים עובדות. ידוע כי סטיית התקן היא 4 שעות. במדגם מקרי של 225 נשים נמצא אומדן נקודתי 36.4 שעות.

א. מצא רווח סמך לממוצע שעות העבודה בשבוע של נשים עובדות, ברמת סמך 0.98.

ב. מהו גודל המדגם שעל החוקר היה לבחור אילו רצה לקבל רווח סמך, באותה רמת סמך, שאורכו מחצית מהאורך שהתקבל בסעיף א'?

ג. אם החוקר יבנה את רווח הסמך לפי גודל המדגם שמצאת בסעיף ב'. האם יוכל לטעון בביטחון רב יותר שרווח זה יכלול את μ ? נמק.



שאלה 5

ידוע שכמות הקפה הנמזגת לספל באוטומט למכירת קפה מתפלגת נורמלי עם סטיית תקן של 10 סמ"ק. בבדיקה תקופתית שנערכה לאוטומטים נבדקו 16 אוטומטים ונמצא כי ממוצע כמות הקפה לספל היא 138.5 סמ"ק.

א. ברמת סמך של 97%, מצא מהו רווח הסמך לתוחלת כמות הקפה הנמזגת לספל באוטומט למכירת קפה?

ב. מהו גודל המדגם שיבטיח שגיאת אמידה שלא תעלה על 2 סמ"ק ברמת סמך של 95%?

ג. מהו גודל המדגם שיבטיח שגיאת אמידה שלא תעלה על 2 סמ"ק ברמת סמך של 99% ?



שאלה 6

להלן נתונים שנאספו על קבוצת 10 נערים העומדים לפני גיוס ואשר השתתפו בתוכנית להפחתת משקל "אין דבר העומד בפני הרצון להתגייס":

משקל בתחילת התוכנית (ק"ג) X	משקל בסיום התוכנית (ק"ג) Y
90	76
84	78
120	122
90	80
108	104
106	108
80	76
92	89
98	96
110	104

לצורך חישוביך הנך ראשי להשתמש בחישובים להלן:

$$\sum_{i=1}^{10} (y_i - \bar{y})^2 = 2304.1, \quad \sum_{i=1}^{10} (x_i - \bar{x})^2 = 1475.6, \quad \sum_{i=1}^{10} y_i = 933, \quad \sum_{i=1}^{10} x_i = 978$$

בהנחה שהתפלגות המשקל היא נורמלית, ענה על השאלות הבאות:

- מהו הרווח סמך לתוחלת המשקל של הנערים לפני גיוס בתחילת התוכנית ברמת סמך 99%?
- מהו הרווח סמך לסטיית תקן המשקל של הנערים לפי גיוס בסיום התוכנית ברמת סמך 99%?
- האם התוכנית להפחתת המשקל עוזרת? בדוק ברמת מובהקות 0.05 ונמק.



שאלה 7

להלן נתונים שנאספו על קבוצת 10 נערים העומדים לפני גיוס ואשר השתתפו בתוכנית להפחתת משקל "אין דבר העומד בפני הרצון להתגייס":

משקל בתחילת התוכנית (ק"ג) X	משקל בסיום התוכנית (ק"ג) Y
90	76

78	84
122	120
80	90
104	108
108	106
76	80
89	92
96	98
104	110

לצורך חישוביך הנך ראשי להשתמש בחישובים להלן:

$$\sum_{i=1}^{10} (y_i - \bar{y})^2 = 2304.1, \quad \sum_{i=1}^{10} (x_i - \bar{x})^2 = 1475.6, \quad \sum_{i=1}^{10} y_i = 933, \quad \sum_{i=1}^{10} x_i = 978$$

בהנחה שהתפלגות המשקל היא נורמלית, ענה על השאלות הבאות:

א. מהו הרווח סמך לתוחלת הפרש המשקל של הנערים לפני גיוס בין המשקל בתחילת התוכנית לבין משקלם בסיום התוכנית ברמת סמך 99%?

ב. מהו הרווח סמך לסטיית תקן המשקל של הנערים לפי גיוס בסיום התוכנית ברמת סמך 99%?



שאלה 8

ידוע שכמות הקפה הנמזגת לספל באוטומט למכירת קפה מתפלגת נורמלית עם סטיית תקן של 10 סמ"ק.

לאמידת כמות הקפה הממוצעת:

א. מהו גודל המדגם שיבטיח שגיאת אמידה שלא תעלה על 2 סמ"ק ברמת סמך של 95%?

ב. מהו גודל המדגם שיבטיח שגיאת אמידה שלא תעלה על 2 סמ"ק ברמת סמך של 99%?



שאלה 9

אגודת הצרכנים מאשימה את חברת ה"תבלין" בהפחתת המשקל הממוצע של התבלינים בצנצנת, האמור להיות 10 גרם.

מניסיון קודם ידוע, כי התפלגות משקל התבלינים בצנצנות מתפלג נורמלית עם סטיית תקן 0.8 גרם.

על מנת לבדוק את טענת אגודת הצרכנים, נלקח מדגם מקרי של 100 צנצנות, ונמצא כי המשקל הממוצע שלהן 9.85 גרם.

- א. מהי רמת המובהקות המינימלית שעבורה ניתן להחליט כי אגודת הצרכנים צודקת בטענתה? נמק.
- ב. מה תהיה מסקנתך עבור רמות המובהקות 0.05 ו- 0.01? נמק על סמך תוצאות סעיף א בלבד וללא חישוב מחדש.
- ג. בנה רווח סמך לתוחלת המשקל של התבלינים בצנצנות ברמת סמך של 95%.
- ד. מה צריך להיות גודל המדגם שיש לדגום על מנת שבאותה רמת סמך שבסעיף א' נקבל רווח סמך ששגיאת האמידה שלו לא תעלה על 0.05 גרם?



שאלה 10

ידוע כי סטיית התקן של משקל בננות מזן מסויים היא 4 גרם.
 על מנת לאמוד את המשקל הממוצע של בננות
 מהו גודל המדגם המינימלי שיבטיח שממוצע משקל הבננות במדגם יסטה מתוחלת המשקל בלא יותר מ- 0.5 גרם
 בהסתברות של 0.95 לפחות?

openbook
 המרכז לקידום אקדמי

מודל (2) רוח בר סמך לתוחלת כשהשונות באוכלוסייה σ^2 לא ידועה

אם השונות באוכלוסייה לא ידועה, נשתמש במקומה באומדן נקודתי חסר הטיה לשונות באוכלוסייה:

$$\hat{s}_x^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n-1} = \frac{n}{n-1} \cdot s_x^2$$

סטיית התקן אינה ידועה עבור כל האוכלוסייה, אלא מחושבת עבור המדגם הספציפי.

אם השונות באוכלוסייה לא ידועה, נשתמש במקומה באומדן נקודתי חסר הטיה לשונות באוכלוסייה $\hat{s}_x^2$.
סטיית התקן אינה ידועה עבור כל האוכלוסייה, אלא מחושבת עבור המדגם הספציפי.

נשים לב ברוח סמך לתוחלת כאשר השונות לא ידועה, האם המדגם גדול דיו.

נשים לב:

1. כאשר המדגם גדול דיו ($n \geq 30$) וההתפלגות כלשהי (יתכן גם נורמלית), אז $\bar{x} \pm Z_{1-\frac{\alpha}{2}} \cdot \frac{\hat{s}}{\sqrt{n}}$

2. כאשר המדגם לא גדול דיו ($n < 30$) והאוכלוסייה נורמלית אז:

$$\bar{x} \pm t_{\frac{\alpha}{2}}^{n-1} \cdot \frac{\hat{s}}{\sqrt{n}}$$

זוהי התפלגות t עם n-1 דרגות חופש.

מכיוון שהשונות לא ידועה, לא ניתן להשתמש בהתפלגות Z ולכן נשתמש בהתפלגות t.

התפלגות t – התפלגות סימטרית סביב 0, כמו התפלגות Z הן התפלגויות נורמליות, אבל ההבדל היחיד בין Z ל-t הוא בשונות:

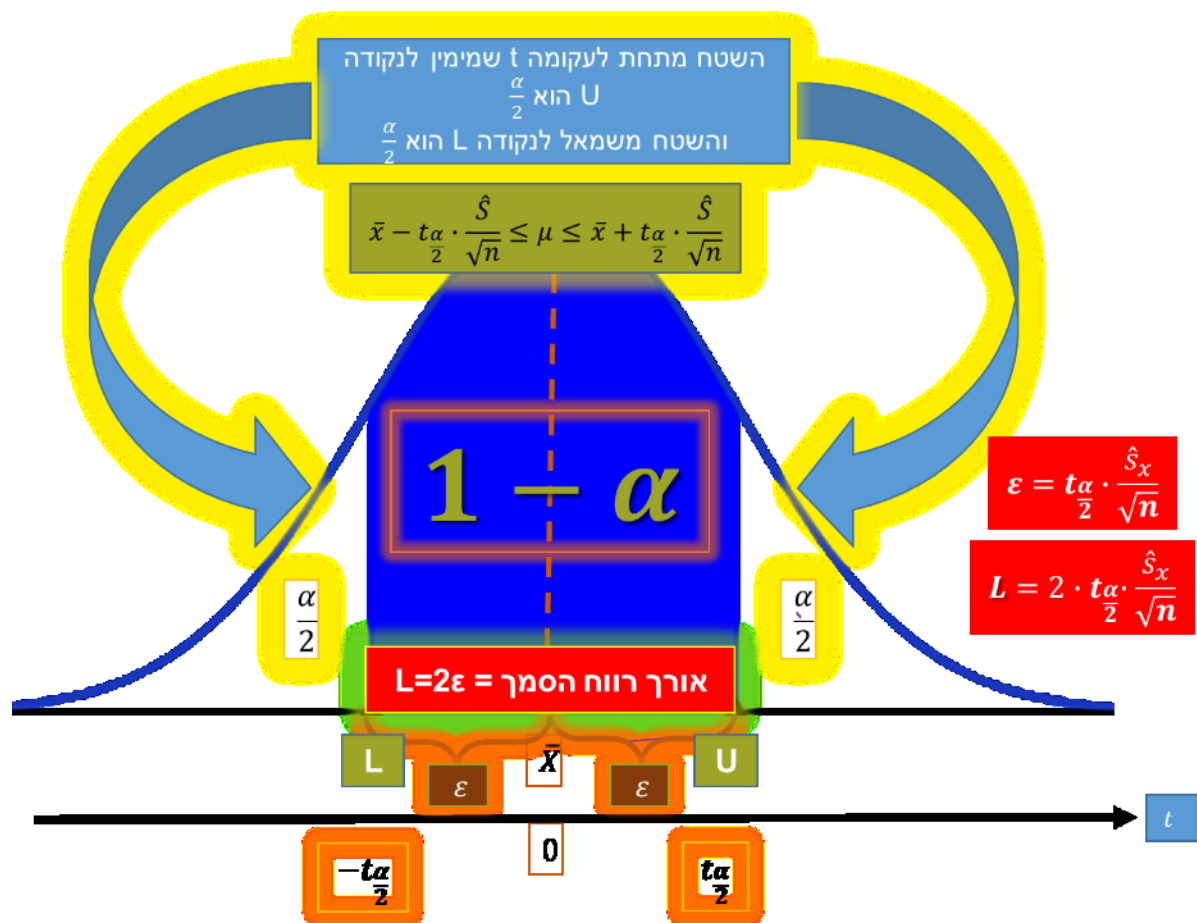
בהתפלגות Z השונות תמיד שווה ל-1.

בהתפלגות t השונות גדולה יותר ומשתנה בהתאם לגודל המדגם $V=n-1$ דרגות חופש

v- שורה בטבלה, $(1-\alpha)$ - עמודה בטבלה ונחשב:

$$\underbrace{\bar{x} - t_{\alpha}^{(n-1)} \cdot \frac{\hat{s}_x}{\sqrt{n}}}_{\text{Lower Limit}} \leq \mu_x \leq \underbrace{\bar{x} + t_{\alpha}^{(n-1)} \cdot \frac{\hat{s}_x}{\sqrt{n}}}_{\text{Upper Limit}}$$

נזכור ש: $\hat{s}^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} = \left(\frac{n}{n-1}\right) \cdot S^2$



✓ התפלגות t

היא התפלגות פעמונית סימטרית סביב 0. הפיזור גדול יותר מאשר בהתפלגות נורמלית סטנדרטית וצורת הפעמון תלויה במספר דרגות החופש (V).

התפלגות פעמונית סימטרית סביב 0, הפיזור גדול יותר מאשר בהתפלגות נורמלית סטנדרטית וצורת הפעמון תלויה במספר דרגות החופש (v).

כמו התפלגות Z הן התפלגויות נורמליות, אבל ההבדל היחיד בין Z ל- t הוא בשונות:

✓ **בהתפלגות Z השונות תמיד שווה ל-1.**

✓ **בהתפלגות t השונות גדולה יותר ומשתנה בהתאם לגודל המדגם $V=n-1$ דרגות חופש**

$V=n-1$ שורה בטבלה

$-\frac{\alpha}{2}$ - עמודה בטבלה ונחשב:

$$\bar{x} - t_{\frac{\alpha}{2}}^{n-1} \cdot \frac{\hat{S}_x}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2}}^{n-1} \cdot \frac{\hat{S}_x}{\sqrt{n}}$$



שאלה 1

במדגם בגודל $n=13$ מאוכלוסיה נורמלית

נמצא ממוצע $\bar{x} = 114$

ואומדן לסטיית התקן באוכלוסיה $\hat{s} = 12$.

מצא ל- μ רווח-סמך ברמה של 95%.



שאלה 2

משקל של 7 שלטים של מכשיר מסויים הוא (בגרמים):

100,98,104,102,98,102,96

בהנחה שמשקל השלטים מתפלג נורמלית,

מצא רווח סמך ברמה של 98% לתוחלת המשקל μ .



שאלה 3

בשתי שכונות נמדדו רמות הרעש [דציבלים] (להלן X) ביום העצמאות במספר מוקדים:

מוקד	1	2	3	4	5	6
שכונה א	100	80	115	95	112	
שכונה ב	110	117	102	109	98	103

א. חשב רווח סמך לתוחלת X בשכונה א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת X בשכונה ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. האם דרושה הנחה כלשהי לצורך החישובים? אם כן - רשום והסבר אותה. אם לא - נמק מדוע.



שאלה 4

חברה העוסקת במחקרי שוק אמדה את ממוצע מספר הקניות החודשי של גבינת קוטג' על סמך מדגם מקרי של 16 צרכנים וקבלה את הרווח סמך: (11.8035, 18.1965). כמו כן, האומדן חסר הטיה לשונות, שחושב מהמדגם היה 36. בהנחה שמספר הקניות מתפלג נורמלית, מהי רמת הסמך לפיה חושב רווח הסמך.



שאלה 5

בשתי ערים גדולות נמדד שיעור החלקיקים (ליחידת נפח) של מזהם (X) שמקורו בכלי רכב. להלן התוצאות:

עיר א	5	5	5.2	4.9	4.7	5.1	5.5
עיר ב	4.8	4.6	5	4.8	4.7	5.2	

א. חשב רווח סמך לתוחלת X בעיר א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת X בעיר א ברמת סמך 98%.



מודל (3) רווח בר סמך להפרש תוחלות $\mu_1 - \mu_2$

במקרים רבים מעוניינים להשוות אוכלוסיות שונות מבחינת תוחלתן.

אחת הדרכים לעשות זאת היא לבנות רווח סמך להפרש התוחלת.

סימונים:

μ_1, μ_2 - תוחלת האוכלוסייה הראשונה והשנייה

σ_1^2, σ_2^2 - שונות האוכלוסייה הראשונה והשנייה

n_1, n_2 - מספר התצפיות מהמדגם הלקוח מהאוכלוסייה הראשונה והשנייה.

$\bar{x}_1, \bar{x}_2$ - ממוצעי המדגמים הלקוחות מהאוכלוסייה הראשונה והשנייה.

$\hat{s}_1^2, \hat{s}_2^2$ - שונות מדגמית של המדגם הלקוח מהאוכלוסייה הראשונה והשנייה.

נבדיל בין שני מקרים:

1. שונות ידועות:

א. $\begin{cases} X_1 \sim N(\mu_1 = ?, \sigma_1^2) \\ X_2 \sim N(\mu_2 = ?, \sigma_2^2) \end{cases}$ אוכלוסיות נורמליות ב"ת, σ_1^2, σ_2^2 ידועות, n_1, n_2 כלשהם.

ב. $\begin{cases} X_1 \sim G_1(\cdot) \\ X_2 \sim G_2(\cdot) \end{cases}$ אוכלוסיות כלשה ב"ת, σ_1^2, σ_2^2 ידועות, $n_1 \geq 30, n_2 \geq 30$.

ב-2 המקרים: $\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

2. שונות אינן ידועות:

א. $\begin{cases} X_1 \sim G_1(\cdot) \\ X_2 \sim G_2(\cdot) \end{cases}$ אוכלוסיות כלשה σ_1^2, σ_2^2 לא ידועות, $n_1 \geq 30, n_2 \geq 30$.

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{s}_1^2}{n_1} + \frac{\hat{s}_2^2}{n_2}}$$

ב. עבור $n_1 < 30 \cup n_2 < 30$ מניחים התפלגויות נורמליות ושוויון השונות:

$$\mu_1 - \mu_2 = (\bar{x}_1 - \bar{x}_2) \pm t_{\frac{\alpha}{2}}^{n_1+n_2-2} \cdot S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \text{ כאשר האומדן:}$$

שאלה 1

בחוות כבשים הגיעה עונת הגז והעובדים עושים תחרות בין שתי קבוצות: בקבוצה א' העובדים הוותיקים ובקבוצה ב' העובדים החדשים שזה להם הגז הראשון. בנתונים שנאספו נמצא כי קבוצה א' גזמה 120 ק"ג צמר ב-40 ימי עבודה שנדגמו מקרית בעוד שקבוצה ב' גזמה 70 ק"ג צמר ב-35 ימי עבודה שנדגמו מקרית.

ידוע כי סטיית התקן בקרב העובדים הוותיקים היא 1.5 ובקרב העובדים החדשים היא 2.5.

א. מצא רווח סמך להפרש ממוצע כמות הצמר ליממה שנאסף, בין שתי הקבוצות ברמת סמך 98%.

ב. על סמך תוצאות סעיף א' האם ניתן לומר שיש הבדל בין שתי הקבוצות? הסבר

שאלה 2

חברה לאריזת סוכר אורזת סוכר לפי משקל. משקל השקיות מתפלג נורמלית עם סטיית תקן 0.3 ק"ג. בדיקות קבועות נעשות כדי לשמור על אחדות המשקל. נלקחו שני מדגמים בלתי תלויים בחודשים ינואר ויוני והתקבלו התוצאות הבאות בק"ג:

מדגם יוני

מדגם ינואר

$$n_2 = 25$$

$$n_1 = 25$$

$$\bar{x}_2 = 2.15$$

$$\bar{x}_1 = 1.9$$

א. מצא רווח סמך למשקל הממוצע של שקיות הסוכר ביוני ברמת סמך 0.99.

ב. מצא רווח סמך להפרש בין המשקל הממוצע של שקיות הסוכר בשני התאריכים ברמת סמך 0.95. האם ניתן לומר שחל שינוי במשקל הממוצע בשני התאריכים? הסבר.

שאלה 3

כדי להשוות את ההישגים במתמטיקה בשני בתי ספר,

נבדקו בשנה מסוימת ציוני בחינות הבגרות במתמטיקה של 50 תלמידים,

שנבחרו באופן מקרי בבית הספר א ושל 75 תלמידים,

שנבחרו באופן מקרי בבית ספר ב.

במדגם א נמצא ממוצע ציונים 82 ואומדן לסטיית תקן של האוכלוסייה $(s_1^2 = 6)$.

ובמדגם ב נמצא ממוצע ציונים 76 ואומדן לסטיית התקן $(s_2=8)$.

מהו רווח-סמך ברמה של 95% להפרש הממוצעים, $\mu_1 - \mu_2$, בין שני בתי הספר?



שאלה 4

בכדי להשוות את ההישגים בסטטיסטיקה ב-2 מכללות של האוניברסיטה הפתוחה,

נבדקו בסמסטר מסוים ציוני הבחינה בסטטיסטיקה של 80 תלמידים שנבחרו באופן מקרי ממכללה א' ושל 90 תלמידים שנבחרו באופן מקרי ממכללה ב'.

במדגם ממכללה א', נמצא ממוצע ציונים 74 ואומדן חסר הטיה לשונות הציונים באוכלוסייה 36.

במדגם ממכללה ב', נמצא ממוצע ציונים 72 ואומדן חסר הטיה לשונות הציונים באוכלוסייה 64.

א. מצא רווח סמך להפרש ממוצעי הציונים בין 2 המכללות ברמת סמך 98%.

ב. על סמך תוצאות סעיף א' האם ניתן לומר שיש הבדל בין הציונים ב-2 המכללות? הסבר.



שאלה 5

בנק בודק אפשרות של פתיחת סניף חדש באחד משני מקומות.

נבחר מדגם מקרי מכל אחד משני המקומות $n_1 = 10$, $n_2 = 8$ והתקבל כי ההכנסה הממוצעת במקום הראשון היא \$650 ובמקום השני \$600.

בהנחה כי סטיות התקן הן \$40 במקום הראשון ו-\$30 במקום השני.

א. מצא רווח סמך להפרש התוחלות של ההכנסות הממוצעות בשני המקומות ברמת סמך 98%.

ב. לאור התוצאות האם ניתן לומר כי ההכנסה הממוצעת במקום הראשון גבוהה יותר מזו במקום השני? נמק.

שאלה 6

חוקרת מעוניינת להשוות 2 שיטות שונות להורדה במשקל: שיטת "רזה לתמיד" ושיטת "הבטן לגב". לשם כך היא דוגמת, באופן מקרי, 50 אנשים שעברו דיאטה בשיטת "רזה לתמיד" ו-70 אנשים שעשו דיאטה בשיטת "הבטן לגב". 50 האנשים שעברו דיאטה בשיטת "רזה לתמיד" ירדו בממוצע 11 ק"ג ו-70 האנשים שעשו דיאטה בשיטת "הבטן לגב" ירדו בממוצע 13 ק"ג. סטיות התקן של האוכלוסיות זהות ושוות ל-1.5 ק"ג.

א. מצא רווח סמך להפרש ממוצעי הירידה במשקל בשתי השיטות ברמת סמך 95%.

ב. על סמך תוצאות סעיף א' האם ניתן לומר שיש הבדל בין השיטות? הסבר.



openbook
המרכז לקידום אקדמי

מודל (4) – רווח בר סמך לפרופורציה

$\hat{p}$ אומדן חסר הטיה לפרופורציה באוכלוסייה

באמידת פרופורציה אנו מעוניינים לחשב אחוז כלשהו מהאוכלוסייה.

p – הפרמטר האמיתי = הפרופורציה באוכלוסייה

ניתן לומר ש- p הוא ההסתברות להצלחה בהתפלגות בינומית $X \sim B(n, p)$.

לכן אמידת פרופורציה היא ניסוי בינומי בעל הסתברות להצלחה p , כאשר אנו מבצעים n חזרות ב"ת ומס' איברים בעלי התכונה המבוקשת שנקבל היא x , אז:

$$X \sim \text{Bin}(n, p) \Rightarrow E(X) = np, \quad V(X) = npq$$

משפט הגבול המרכזי לפרופורציה

אם מתוך אוכלוסייה בעלת הפרמטר P היינו מוציאים את כל המדגמים האפשריים בגודל n כלשהו, ובכל המדגמים האפשריים מחשבים את הפרופורציה במדגם P ,

אזי סדרת כל הפרופורציות של כל המדגמים האפשריים עבור n מספיק גדול שואפת להתפלגות נורמלית עם:

ממוצע: P

וטעות תקן: $\sigma_p = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

הנוסחה: $\hat{p} \pm Z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$

כאשר לא ידועה האמידה לפרופורציה ניקח את הערך המקסימלי למכפלה ולכן $\hat{p} = \hat{q} = 0.5$

גודל המדגם המינימלי $\hat{p} \cdot \hat{q} \cdot \left(\frac{Z_{1-\alpha/2}}{\varepsilon}\right)^2$, כאשר $n \geq \frac{\hat{p} \cdot \hat{q}}{\varepsilon^2} \cdot Z_{1-\alpha/2}^2$

גודל המדגם המינימלי

(1) אם ידועה האמידה לפרופורציה: $\hat{p} \cdot \hat{q} \cdot \left(\frac{Z_{1-\alpha/2}}{\varepsilon}\right)^2$, $n \geq$

(2) כאשר לא ידועה האמידה לפרופורציה ניקח את הערך המקסימלי למכפלה

ולכן $\hat{p} = \hat{q} = 0.5$

ולכן נקבל:

$$n \geq \left(\frac{Z_{1-\frac{\alpha}{2}}}{\varepsilon} \right)^2 \cdot 0.5 \cdot 0.5$$

$$n \geq \left(\frac{Z_{1-\frac{\alpha}{2}}}{2\varepsilon} \right)^2 \text{ או}$$



שאלה 1

600 מתוך 800 מרואיינים במדגם מקרי של ישראלים סבורים שאין סיכוי לרסן את העלייה במחירי הדלק. חשב רווח סמך של 96% לפרופורציית הישראלים הסבורים כך.



שאלה 2

מועמד לראשות עירייה מסוימת מעוניין לאמוד את פרופורציית התומכים בו, כך שיוכל להיות בטוח בהסתברות של 98% ששגיאת האמידה לא תעלה על 0.03. איזה גודל מדגם יבטיח לו הסתברות כזאת?



שאלה 3

חוקר מבקש לאמוד את פרופורציית הסובלים מיתר לחץ דם בקרב האוכלוסייה הבוגרת. נחקר מדגם מקרי של 400 נבדקים ונמצא כי 60 מתוכם סובלים מיתר לחץ דם.

א. מהו האומד הנקודתי לפרופורציית הסובלים מיתר לחץ דם באוכלוסייה הבוגרת?

ב. אמוד את פרופורציית הסובלים מיתר לחץ דם באוכלוסייה על ידי רווח בר סמך של

(1) 95% (2) 99% מה המסקנה?



שאלה 4

שרות התעסוקה אומד מדי חודש את אחוז המובטלים במדינה.

א. מהו גודל המדגם אשר על שירות התעסוקה לקחת אם דורשים שהאומד לא יסטה מהאחוז האמיתי באוכלוסייה ביותר מאחוז אחד בביטחון של 95%?

ב. חזור על סעיף א' בהנחה שידוע שאין יותר מ-14% מובטלים במדינה.



שאלה 5

במסגרת הבחירות בארה"ב נשאלו 757 אנשים בעד מי יצביעו. 400 מהם ענו שיבחרו במועמד בוש. מערכת העיתון בו פורסמו התוצאות הוסיפה בהערות שוליים: "תיתכן סטייה של 4% + בתוצאות". באיזו רמת סמך השתמש העיתון?



שאלה 6

על-פי נתוני בתי המשפט לתעבורה, פרופורציית המשפטים המסתיימים בשלילת רישיון נהיגה היא 0.6. מתוך 120 משפטי תעבורה שהתקיימו בפני שופט א' הסתיימו 90 בשלילת רישיון.
א. מצא רווח סמך ברמת סמך 99% לפרופורציית השלילות שפוסק שופט א'.
ב. לאורך כמה משפטים יש לבחון את החלטות השופט כדי לקבל רווח סמך ברמת סמך 99% שאורכו קטן מ- 0.05?

שאלה 7

מפלגת "ישראל יפה" מעוניינת לאמוד את אחוז התמיכה בה בציבור.
א. מהו גודל המדגם המינימלי שעל המפלגה לבחור כדי שבסקר שיתקבל היא תדע את פרופורציית התמיכה בה בביטחון של 90%, ועם שגיאת אמידה שלא תעלה על 0.02?
ב. כיצד ניתן להקטין את הוצאות הסקר, אם ידוע שאין סיכוי לתמיכה של יותר מ- 7% מהציבור במפלגה "ישראל יפה".
ג. במדגם שאת גודלו מצאת בסעיף ב', נמצאו 4.5% התומכים במפלגה "ישראל יפה". בנה רווח סמך ברמת סמך 95% עבור פרופורציית התומכים בה באוכלוסייה כולה.
ד. בהסתמך על תוצאות סעיף ג', אם בישראל כ- 5.5 מיליון בעלי זכות בחירה, בנה רווח סמך עבור מספר הקולות שתקבל מפלגת "ישראל יפה" בבחירות הכלליות, ברמת סמך 95%.



שאלה 8

150 אנשים נשאלו כמה פעמים בשבוע הם קונים עיתון יומי.

להלן התפלגות התשובות:

מספר פעמים	0	1	2	3	4	5	6
מספר אנשים	27	8	22	40	24	4	25

א. מצא רווח סמך לממוצע מספר העיתונים שקונה אדם בשבוע ברמת סמך 95%.

ב. מצא רווח סמך לפרופורצית האנשים הקוראים לפחות 5 עיתונים בשבוע, על סמך הנתונים, ברמת סמך 98%.

שאלה 9

חברת שיווק מעוניינת לאמוד את פרופורצית המשתמשים במוצר מסוים. לשם כך שכרה את שרותי חברה למחקרי שוק. החברה בדקה מדגם מקרי של 100 איש חישה ומצאה את רווח הסמך $(0.438, 0.702)$

- מהו אחוז האנשים במדגם המשתמשים במוצר זה?
- מהי רמת הסמך לפיה חושב רווח הסמך?
- לפי נתוני אותו מדגם מה יהיה רווח הסמך לפרופורצית המשתמשים במוצר ברמת סמך 95%?
- בהמשך לסעיף ג', מה צריך להיות גודל המדגם אם מעוניינים ששגיאת האמידה תקטן פי 2?

שאלה 10

100 אנשים הועמדו למשפט על עבירות תנועה, 19 יצאו זכאים ו-81 הורשעו וקיבלו קנס של 750 שקל במוצק עם סטיית תקן (s) 200 שקל.

- מצא רווח סמך ברמת סמך 0.95 לסיכוי לצאת זכאי במשפט על עבירת תנועה.
- מצא רווח סמך ברמת סמך 0.9 לגובה הקנס הממוצע במשפט תעבורה, בהנחה שגובה הקנס מתפלג נורמלית.
- הסטטיסטיקאי של בית המשפט קבע כי ע"ס הנתונים ניתן לומר שהקנס הממוצע על עבירת תנועה הוא בין 708 לבין 792 שקל. מהי רמת הסמך של קביעתו זו?

שאלה 11

ארגון "נשים עובדות" טוען שגברים מועדפים על פני נשים בעלות כישורים דומים, בקבלת עבודה.

לבדיקת הטענה נבחרו מקרית 100 נשים ו-200 גברים בעלי כישורים דומים, שהציעו עצמם לעבודה. מבין הנשים התקבלו לעבודה 70, ומבין הגברים 160.

א. בדוק את טענת הארגון ברמת מובהקות 0.05.

ב. על סמך הנתונים, מצא רווח סמך לפרופורציית הנשים המתקבלות לעבודה, ברמת סמך 99%.

שאלה 12

במפעל מסוים מייצרים צלחות. 10% מהצלחות נפסלות בגלל פגמים.

א. דגמו 20 צלחות. מצאו חסם להסתברות כי פרופורציית הצלחות הפסולות במדגם תסטה מ-10% בפחות מ-7%.

ב. נלקח מדגם של 400 צלחות.

מה ההסתברות, בקירוב, שלכל היותר 43 צלחות במדגם תהיינה פסולות?

ג. במחלקת ביקורת האיכות טוענים כי אחוז הצלחות הפסולות במפעל הוא לכל היותר 9%.

לבדיקת טענתם נלקח מדגם של 400 צלחות, מתוך 30 נמצאו פסולות.

האם תקבלו את טענת מחלקת בקורת האיכות על פי רווח סמך ברמת סמך של 95%? נמקו!



שאלה 13

הנהלת חברת התקשורת "השוק" מעוניינת לאמוד את אחוז הקוראים הפוטנציאלי למגזין כלכלי חדש אותו היא עומדת להפיץ.

מהו גודל המדגם המינימלי שעליה לדגום, על מנת להבטיח שהסטייה בין האומדן לפרמטר לא תעלה על 0.04 ברמת סמך 98%?



שאלה 14

יצרן משקאות קלים מעוניין לבדוק את השפעתו של מבצע פרסום שמטרתו להחדיר משקה חדש לשוק.

שני מדגמים מקריים נבחרו לפני ואחרי מבצע הפרסום ובו נשאלו צרכנים האם הם שמעו על המשקה החדש.

התוצאות הראו שלפני המבצע שמעו על המשקה 450 מתוך 1200 צרכנים ואחריו 700 מתוך 1250.

א. מצא רווח סמך ברמת סמך 95% להפרש הפרופורציות של הצרכנים שמכירים את המשקה לפני מבצע הפרסום ולאחריו.

ב. משרד הפרסום טוען כי המבצע היה אפקטיבי בהעלאת המודעות למשקה החדש.

האם הטענה מוצדקת ברמת מובהקות 0.01? נמק.

ג. מצא רווח סמך ברמת סמך 90% לפרופורציית הצרכנים שמכירים את המשקה אחרי המבצע.



שאלה 15

ארגון "נשים עובדות" טוען שבקבלת לעבודה גברים מועדפים על פני נשים בעלות כישורים דומים.

לבדיקת הטענה נבחרו מקרית 100 נשים ו-200 גברים בעלי כישורים דומים, שהציעו עצמם לעבודה.

מבין הנשים התקבלו לעבודה 70, ומבין הגברים 160.



openbook

המרכז לקידום אקדמי

מודל (5) – רווח סמך להפרש פרופורציות $p_1 - p_2$

הנחות: (1) $\hat{p}_1, \hat{p}_2 \sim N$

(2) תק"ל- תנאי קרוב לנורמלי $n_1\hat{p}_1, n_1\hat{q}_1, n_2\hat{p}_2, n_2\hat{q}_2 \geq 10$

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

$$(\hat{p}_1 - \hat{p}_2) - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

במקרים רבים מעוניינים לאמוד את ההבדלים בין הפרופורציות

למשל: כדי לאמוד את ההבדלים במידת התמיכה במפלגה מסוימת בקרב אזרחים בעלי השכלה שונה,

נבדק מדגם מקרי של 400 משפחות בעלי השכלה נמוכה (עד 12 שנות לימוד) מתוכם 162 הביעו את תמיכתם במפלגה, והשאר אמרו שהם מתנגדים לה.

לעומת זאת התברר שבמדגם מקרי של 200 משפחות בעלי השכלה גבוהה, הביעו 132 את תמיכתם במפלגה, והשאר הביעו התנגדות.

נסמן ב- p_1 וב- p_2 את פרופורציית התומכים בקרב בעלי השכלה הנמוכה ובעלי ההשכלה הגבוהה, בהתאמה,

אנו מעוניינים לאמוד את ההפרש $p_1 - p_2$.

זהו מקרה פרטי של אמידת הפרש תוחלות $\mu_1 - \mu_2$ בדומה לנוסחת הפרש התוחלות:

$$(\hat{p}_1 - \hat{p}_2) - Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}} < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + Z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$



שאלה 1

במטרה לבדוק האם שינוי מסוים שהוכנס בתהליך הייצור הביא לשיפור, נבדקו 1500 פריטים שיוצרו לפני השינוי, שביניהם נמצאו 75 פגומים, בעוד שמתוך 2000 פריטים שיוצרו לאחר השינוי נמצאו 80 פגומים.

מצא רווח סמך ברמה של 95% להפרש בין הפרופורציות של הפריטים הפגומים, לפני ואחרי השינוי.



שאלה 2

בסקר דעת קהל לגבי התמיכה בממשלת האחדות, התקבל כי:

300 מתוך 400 בוחרים ימין, תומכים. 200 מתוך 300 בוחרים שמאל, תומכים.

א. בנה רווח סמך ברמת סמך 0.9 להפרש הפרופורציות.

ב. בנה רווח סמך לפרופורציית התמיכה בממשלת האחדות מתוך כלל הנתונים במדגם, ברמת סמך 0.9



שאלה 3

חוקר ערך השוואה של תופעות לוואי בין שתי תרופות שונות לאותה מחלה. במדגם מקרי של 600 נבדקים שקבלו תרופה א' נתגלו תופעות לוואי ל - 120 נבדקים, ואילו במדגם מקרי של 1000 נבדקים שקבלו תרופה ב' נתגלו תופעות לוואי אצל 140 נבדקים. חשב רווח סמך להפרש הפרופורציות ברמת סמך של 0.95, ורמת סמך של 0.98.



שאלה 4

במשאל דעת קהל תמכו במדיניות הממשלה 300 מתוך 450 גברים, ו 180 מתוך 300 נשים.

א. בנה רווח סמך ברמת סמך 90% עבור ההפרש בין פרופורציות התמיכה בממשלה של הגברים ושל הנשים.

ב. בנה רווח סמך לפרופורציית התמיכה בממשלה על סמך המדגם של 750 הנשאים ברמת סמך 95%.



שאלה 5

בסקר דעת קהל על תכנית חדשה בטלוויזיה נמצאו התוצאות הבאות:

	צופים	לא צופים
נשים	38	162
גברים	108	192

א. בדוק את ההשערה שלמין אין כל השפעה על הצפייה בתכנית ברמת מובהקות 0.05 ? נמק.

ב. האם לאור התוצאות ניתן לומר, ברמת מובהקות 5%, שפרופורציית הנשים הצופות בתכנית נמוכה מ - 20% ? נמק.



שאלה 6

ארגון "נשים עובדות" טוען שגברים מועדפים על פני נשים בעלות כישורים דומים, בקבלת עבודה.

לבדיקת הטענה נבחרו מקרית 100 נשים ו- 200 גברים בעלי כישורים דומים, שהציעו עצמם לעבודה.

מבין הנשים התקבלו לעבודה 70, ומבין הגברים 160.

א. בדוק את טענת הארגון ברמת מובהקות 0.05.

ב. על סמך הנתונים, מצא רווח סמך לפרופורציית הנשים המתקבלות לעבודה, ברמת סמך 99%.



שאלה 7

יצרן משקאות קלים מעוניין לבדוק את השפעתו של מבצע פרסום שמטרתו להחדיר משקה חדש לשוק.

שני מדגמים מקריים נבחרו לפני ואחרי מבצע הפרסום ובו נשאלו צרכנים האם הם שמעו על המשקה החדש.

התוצאות הראו שלפני המבצע שמעו על המשקה 450 מתוך 1200 צרכנים ואחריו 700 מתוך 1250.

א. מצא רווח סמך ברמת סמך 95% להפרש הפרופורציות של הצרכנים שמכירים את המשקה לפני מבצע הפרסום ולאחריו.

ב. משרד הפרסום טוען כי המבצע היה אפקטיבי בהעלאת המודעות למשקה החדש.

האם הטענה מוצדקת ברמת מובהקות 0.01? נמק.

ג. מצא רווח סמך ברמת סמך 90% לפרופורציית הצרכנים שמכירים את המשקה אחרי המבצע.



מודל (6) - רווח סמך לשונות האוכלוסייה

במצבים בהם נרצה להסיק ממדגם מסוים על הפיזור הקיים באוכלוסייה.

נוסחת רווח הסמך לפרמטר שונות האוכלוסייה:

$$\frac{\hat{s}^2 \cdot (n-1)}{\chi_{\alpha/2}^2} < \sigma^2 < \frac{\hat{s}^2 \cdot (n-1)}{\chi_{1-\alpha/2}^2}$$

ואם נתונה לנו סטיית התקן S אז הנוסחה תהיה:

$$\frac{nS^2}{\chi_{(\alpha/2)}^2} < \sigma^2 < \frac{nS^2}{\chi_{(1-\alpha/2)}^2}$$

אומד חסר הטיה לשונות האוכלוסייה הוא הסטטיסטי

$$\hat{S}^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

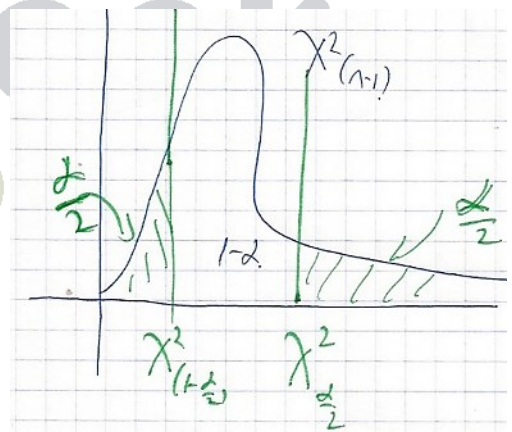
$$\hat{S}^2 = \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n-1}$$

התפלגות חי בריבוע χ^2

התפלגות χ^2 היא התפלגות א-סימטרית חיובית/ימנית, התלויה בפרמטר אחד שהוא ד"ח - צורת התפלגות מושפעת ממספר דרגות החופש: ככל ש-n גדל, הצורה נהיית סימטרית ושטוחה יותר. מקבלת רק ערכים

חיוביים או אפס.

$$X = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma^2} = \frac{(n-1)\hat{S}^2}{\sigma^2} \sim \chi_{(n-1)}^2$$



שאלה 1

במדגם מקרי בגודל $n=20$ מאוכלוסייה נורמלית התקבל ממוצע ו- $\hat{s} = 4.52$.

מצא רווח סמך ברמה של 95% עבור σ^2



שאלה 2

מכונה מייצרת חלקי מתכת גליליים.

במדגם של 9 חלקים נמצאו האורכים הבאים:

25.3 , 24.7 , 25.8 , 26.1 , 25.0 , 25.1 , 24.9 , 25.7 , 25.8

בהנחה שהתפלגות האורך של החלקים המיוצרים היא נורמלית בקירוב,

מצא רווח סמך ברמה של 99% לסטיית התקן של התפלגות זו.



שאלה 3

מפעל המייצר נורות רצה לבדוק את אורך חייהן הממוצע של הנורות ואת סטיית התקן. הוא בדק את אורך

$$\sum_{i=1}^{16} x_i = 12720 ; \sum_{i=1}^{16} (x_i - \bar{x})^2 = 11760$$

חייהן של 16 נורות שנבחרו באופן מקרי מתוצרת המפעל ומצא:

א. אמוד בעזרת רווח סמך את אורך החיים הממוצע של הנורות מתוצרת המפעל, ברמת סמך של 95%

ב. אמוד בעזרת רווח סמך את סטיית התקן של אורך החיים של הנורות, ברמת סמך 95%.



שאלה 4

בשביל לבדוק איזה אורך דיסק צריך בשביל להקליט סינגל, נלקח מדגם מקרי של 9 שירים שיצאו

בחודש האחרון, ונבדק אורך השיר בדקות. להלן התוצאות:

4.5 , 3.0 , 3.5 , 4.0 , 5.0 , 3.6 , 4.8 , 5.2 , 4.2

בהנחה שאורך שיר מתפלג נורמלית, מצא:

א. רווח סמך לתוחלת אורך השיר ברמת סמך 98%.

ב. רווח סמך ברמת סמך 90% עבור סטיית התקן של אורך השיר



שאלה 5

במפעל לייצור ברגים נמדדו קוטריהם של 9 ברגים ב-מ"מ והתקבלו התוצאות הבאות :

10.2 10.2 10.0 10.6 10.1 9.9 9.8 9.9 10.2

ידוע כי קוטר הברגים מתפלג נורמלית.

א. מצא אומדנים חסרי הטיה לתוחלת ושונות קוטר הברגים.

ב. מצא רווח סמך ברמת סמך לתוחלת הקוטר.

ג. מצא רווח סמך ברמת סמך 95% לסטיית התקן של קוטר הברגים.

שאלה 6

במדגם של 30 סטודנטים שנבחרו באופן מקרי ונבחנו ב"מבוא לסטטיסטיקה ב" התקבלו התוצאות הבאות לגבי ציוניהם בבחינה: $\bar{x} = 72$; $s = 7.6$; $n = 30$. 28% מהסטודנטים נכשלו.

בהנחה שהציונים מתפלגים נורמלית אמוד את:

א. ממוצע הציונים ב"מבוא לסטטיסטיקה" ברמת סמך 90%.

ב. סטיית התקן של הציונים ב"מבוא לסטטיסטיקה", ברמת סמך 95%.

ג. פרופורציית הסטודנטים שעברו את המבחן בהצלחה ברמת סמך 98%.

שאלה 7

להלן נתונים שנאספו על קבוצת 10 נערים העומדים לפני גיוס ואשר השתתפו בתוכנית להפחתת משקל "אין דבר העומד בפני הרצון להתגייס":

משקל בתחילת התוכנית (ק"ג) X	משקל בסיום התוכנית (ק"ג) Y
90	76
84	78
120	122
90	80
108	104
106	108
80	76
92	89
98	96
110	104

לצורך חישוביך הנך ראשי להשתמש בחישובים להלן:

$$\sum_{i=1}^{10} (y_i - \bar{y})^2 = 2304.1, \quad \sum_{i=1}^{10} (x_i - \bar{x})^2 = 1475.6, \quad \sum_{i=1}^{10} y_i = 933, \quad \sum_{i=1}^{10} x_i = 978$$

בהנחה שהתפלגות המשקל היא נורמלית, ענה על השאלות הבאות:

א. מהו הרווח סמך לתוחלת המשקל של הנערים לפני גיוס בתחילת התוכנית ברמת סמך 99%?

- ב. מהו הרווח סמך לסטיית תקן המשקל של הנערים לפני גיוס בסיום התכנית ברמת סמך 99%?
- ג. האם התוכנית להפחתת המשקל עוזרת? בדוק ברמת מובהקות 0.05 ונמק.



openbook
המרכז לקידום אקדמי

מבחן סטטיסטי

שלב ראשון – נתונים.

שלב שני – ניסוח השערות:

השערה הינה היגד על פרמטר של האוכלוסייה.

השערה סטטיסטית, מסומנת באות H (מהמילה Hypothesis) היא טיעון לגבי ההתפלגות של האוכלוסייה.

בד"כ השערה היא טענה לגבי ערך מספרי של פרמטר של האוכלוסייה

השערת האפס – H_0 – מתארת בד"כ את המצב הקיים, זו ההשערה שהחוקר מבקש להפריך. אנו מקבלים אותה באופן זמני ובודקים.

השערה אלטרנטיבית – H_1 – השערת החוקר, השערה שאותה מבקש החוקר להוכיח באמצעות הנתונים.

$$\begin{aligned} H_0: \theta &= c \\ H_1: \theta &= c_2 \end{aligned}$$

$$\begin{aligned} H_0: \theta &= c \\ H_1: \theta &\neq c \end{aligned}$$

$$\begin{aligned} H_0: \theta &= c \\ H_1: \theta &< c \end{aligned}$$

$$\begin{aligned} H_0: \theta &= c \\ H_1: \theta &> c \end{aligned}$$

דו צדדי

חד צדדי

הערות

✓ סימן שוויון יופיע תמיד ב H_0 .

✓ ההשערות נקראות חד צדדיות כאשר H_1 טוענת לערכים הנמצאים כולם מצד אחד של הערכים הנקובים ב H_0 . (סימן $>$ או $<$). אזור הדחייה של השערת האפס נמצא כולו בצד אחד של ההתפלגות.

✓ ההשערה נקראת דו צדדית, כאשר H_1 טוענת לערכים הנמצאים משני צדי הערכים הנקובים ב H_0 . (סימן $\neq$). אזור הדחייה של השערת האפס נמצא משני צדי ההתפלגות.

שלב שלישי – הנחות. אנו מניחים הנחות על התפלגות הדגימה באופן מלא כדי שנוכל לומר משהו על ההסתברות לקבל תוצאה מסוימת.

שלב רביעי – שרטוט וקביעת אזור דחיית H_0 . שרטוט ההשערות וקביעת אזור דחייה וקבלה. גודל זה נקבע לפי רמת המובהקות α , ומיקומו בהתפלגות הדגימה תלוי בסוג ההשערה: חד צדדית או דו צדדית

שלב חמישי – קביעת סטטיסטי המבחן. נבחר סטטיסטי שיהווה אומד מתאים לפרמטר הנבדק, ושאותו נחשב מהמדגם.

שלב שישי – הסקת מסקנות. אם הערך שנתקבל נמצא באזור הקבלה, נקבל את H_0 , אחרת נדחה.

תוצאה שנופלת באזור הדחיה תיקרא תוצאה **מובהק**- כלומר שהמדגם לא נלקח מהאוכלוסייה המקורית אלא מהאוכלוסייה האלטרנטיבית.

תוצאה שנופלת באזור הקבלה תיקרא תוצאה **בלתי מובהקת**- כלומר שהמדגם נלקח מהאוכלוסייה המקורית.
אנו דוחים/מקבלים את H_0 בלבד משום שהיא זו שקיבלנו זמנית כנכונה ובדקנו.

הרציונל של ההסקה הסטטיסטית:

אם ברצוננו לבדוק אם שיטת הלימוד של המרכז לקידום אקדמי **OpenBook** משפרת ציוניהם של הסטודנטים בקורס סטטיסטיקה ב, נדגום מספר סטודנטים ונלמד אותם בשיטה החדשה – נבדוק אם הממוצע החדש שלהם רחוק מספיק מהממוצע של כל יתר הסטודנטים שלא למדו בשיטת הלימוד של **OpenBook** כיצד נדע מה זה רחוק מדי?

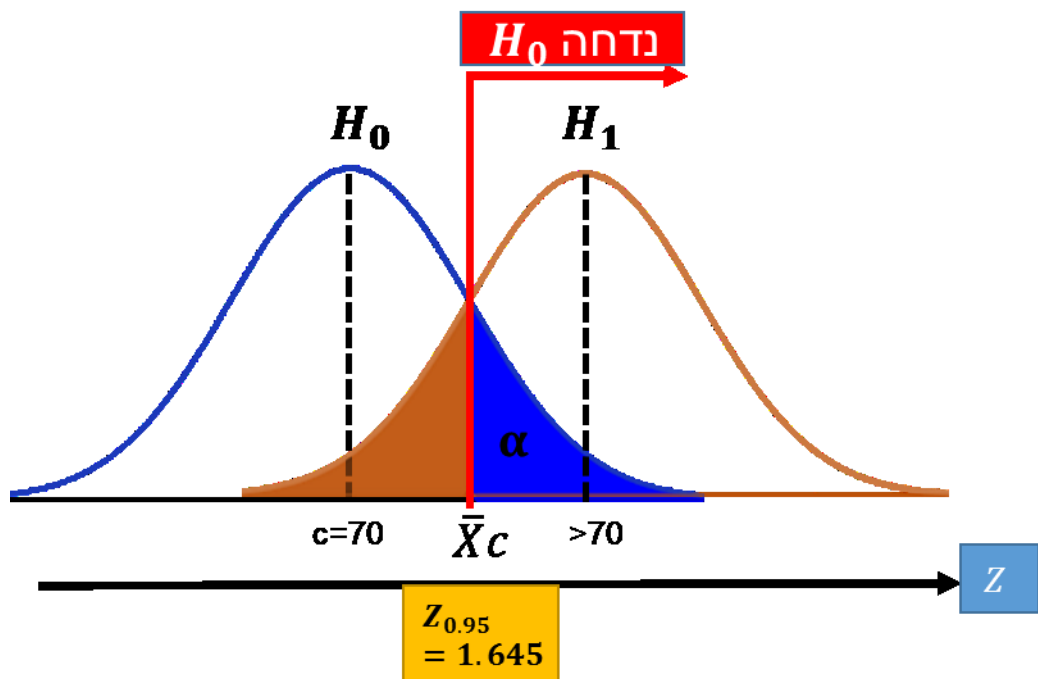
נלקח מדגם של 100 סטודנטים וקיבלתי ממוצע מדגם $\bar{x} = 73$, סטיית התקן היא: $\sigma = 10$ מה יש לנו פה? המרכז לקידום אקדמי **OpenBook** טוען שהוא יכול לשפר את ממוצע האוכלוסייה, מהו המצב הקיים? בדקנו במחשבי האוניברסיטה וקיבלנו $\mu = 70$. ולכן: המצב הקיים אנו רושמים כ: H_0 , טענת החוקר אנו מסומנת ב

$$H_0: \mu = 70 \quad H_1: \mu > 70$$

ולכן מה יש לנו?

המרכז לקידום אקדמי **OpenBook** מנסה להעלות את הציון, ולכן אנו מקציבים לו את רמת המובהקות שאם ייפול בה הממוצע החדש אז אנחנו נקבל את הטענה ששיטת הלימוד של המרכז לקידום אקדמי **OpenBook** מגדילה את הממוצע. השטח הזה בקצה הפעמון – אם הציון יהיה מספיק קיצוני אז החברה צודקת!

הנקודה שמפרידה בין האזור שאנו מקבלים את השערת החוקר לבין דוחים אותה זה אזור הקריטי כי הוא מכריע אם נקבל או נדחה!



מבחן חד צדדי ימני על תוחלת אחת ושונות ידועה ולכן: $Z_{\bar{x}} = \frac{\bar{x} - c}{\sigma / \sqrt{n}}$, $1.645 = \frac{\bar{x} - 70}{10 / \sqrt{100}}$ ולכן $\bar{x} = 71.645$

הממוצע 73 נמצא באזור דחיית H_0 ולכן נקבל את H_1 , ברמת מובהקות 0.05 נקבל את טענת המרכז לקידום אקדמי OpenBook ואכן המרכז לקידום אקדמי OpenBook מעלה את הממוצע האוכלוסייה.

מהן הטעויות האפשריות?

אם דחינו את השערת האפס

α

הטעות האפשרית היא שהשערת האפס נכונה ודחינו אותה בטעות

אם קיבלנו את השערת האפס

β

הטעות האפשרית היא שהשערת האפס לא נכונה וקיבלנו אותה בטעות

✓ טעות מסוג ראשון

הטעות הנגרמת כתוצאה מדחייה לא נכונה של השערת האפס.

דוחים את H_0 כאשר היא נכונה.

ההסתברות לטעות מסוג ראשון, מסומנת ב- α ,

היא רמת המובהקות שקבענו בתחילת המחקר ומוגדרת:

$$\alpha = P(\text{reject } H_0 \mid H_0 \text{ true})$$

α היא החלק של H_0 שנמצא באזור הדחייה.

$1 - \alpha$ היא החלק של H_0 שנמצא באזור הקבלה

✓ טעות מסוג שני

טעות מסוג שני – הטעות הנגרמת כתוצאה מקבלה לא נכונה של השערת האפס.

מקבלים את H_0 כאשר היא אינה נכונה.

ההסתברות לטעות מסוג שני, מסומנת ב- β

$$\beta = P(\text{accept } H_0 \mid H_1 \text{ true})$$
 ומוגדרת ע"י:

β היא החלק של H_1 שנמצא באזור הקבלה של H_0 .

✓ סוגי טעויות

מצב אמיתי			
H1 נכונה	H0 נכונה		
שגיאה מסוג שני	החלטה נכונה	קבל H_0	החלטה
החלטה נכונה	שגיאה מסוג ראשון	דחה H_0	

✓ ההסתברות לטעות

מצב אמיתי			
H1 נכונה	H0 נכונה		
β טעות מסוג שני	$1 - \alpha$	קבל H0	החלטה
$1 - \beta$ עוצמת המבחן	α טעות מסוג ראשון	דחה H0	

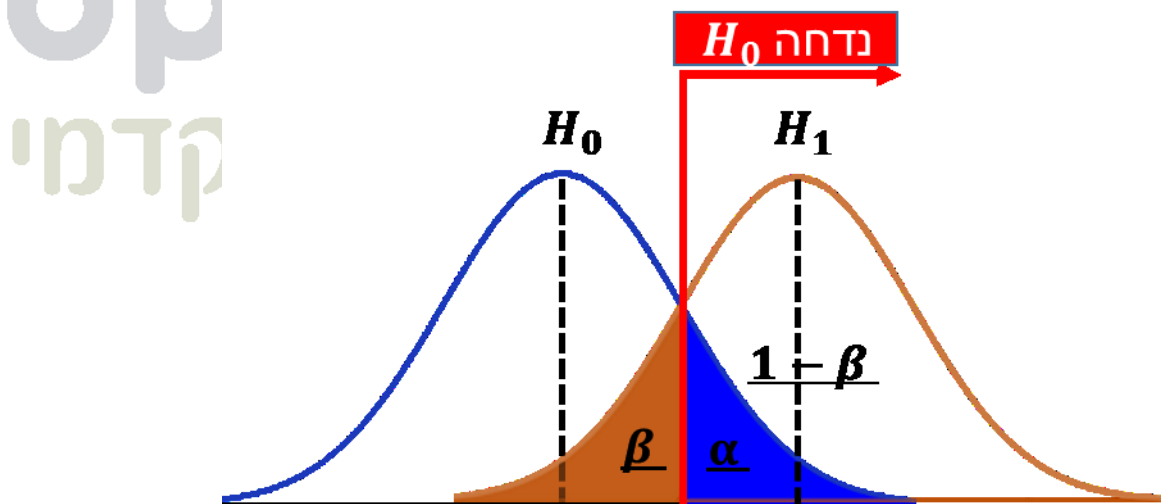
✓ עוצמת המבחן

עוצמת המבחן, מסומנת ב- $1 - \beta$,

הוא החלק של H1 שנמצא באזור הדחייה של H0.

ומוגדרת ע"י: $1 - \beta = P(\text{rejec } H_0 | H_1 \text{ true})$

עוצמת המבחן היא ההסתברות לעשות את הדבר הנכון, לדחות את השערת האפס כאשר צריך לדחותה.



ה' חוקר	הת' א' /		
קבלה $H_0 = \mu = \mu_0$	דחייה $H_a = \mu \neq \mu_0$		
α	$1 - \alpha$	β התפוצה	H_0
- דיוק גבוה - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות	- דיוק גבוה - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות	- דיוק גבוה - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות	- דיוק גבוה - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות
$1 - \beta$	β	β התפוצה	H_a
- דיוק גבוה - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות	- דיוק גבוה - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות	- דיוק גבוה - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות	- דיוק גבוה - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות - הסתברות אפס של טעות

openbook

המרכז לקידום אקדמי

✓ P-value מובהקות התוצאה

רמת המובהקות הקטנה ביותר (α_{min}) עבורה נדחה את H_0 על סמך תוצאות המדגם.

ההסתברות, בהנחה שהשערת האפס נכונה, שסטטיסטי המבחן שווה לערך שהתקבל במדגם או קיצוני ממנו.

מחשבים את ההסתברות לקבל את התוצאה שהתקבלה או קיצונית ממנה.

להסתברות זו קוראים p-value.

הגדרה 1: ההסתברות לקבל תוצאה מדגמית כמו שקיבלנו או יותר מובהקת (קיצונית לכיוון H_1) תחת ההנחה

- ש H_0 נכונה.

הגדרה 2: רמת המובהקות הקטנה ביותר שעבורה נידחה את השערת האפס. מחושב בהתאם לערכי המדגם

עבור $P.value < \alpha$

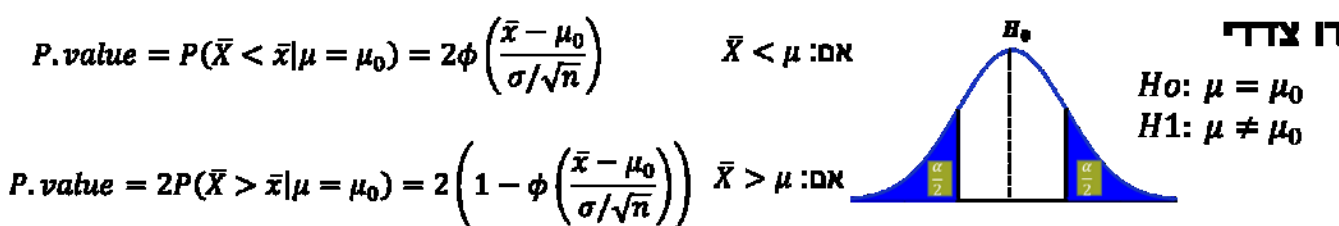
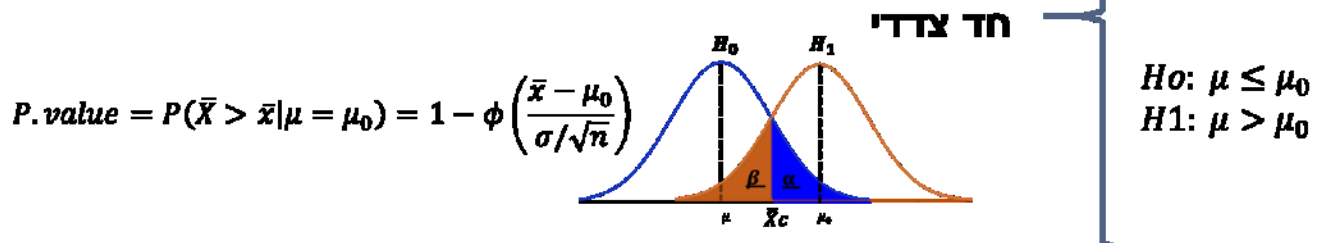
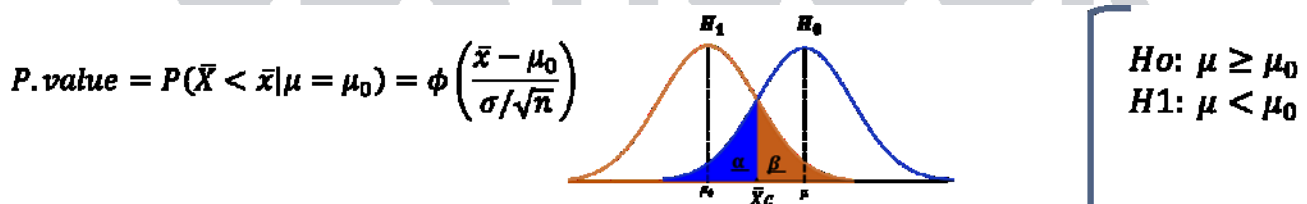
המסקנה: לדחות את H_0

עבור $P.value > \alpha$

המסקנה: לא לדחות את H_0

זוהי שיטה נוספת להחליט האם לדחות או לקבל את H_0

חישוב P.v בסוגים שונים של השערות



מודל 1 – בדיקת השערות על תוחלת אחת כאשר השונות באוכלוסייה ידועה

הנחות:

$x \sim N(\mu, \sigma^2)$ או לפי מ.ג.מ. כאשר $n > 30$.

אזי מתקיים ש: $\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$

לפי משפט הגבול המרכזי ממוצע המדגם מתפלג בקירוב נורמלי עבור גודל מדגם מספיק גדול

$$Z_x = \frac{\bar{x} - c}{\sigma / \sqrt{n}} \quad \text{הסטטיסטי במדגם:}$$

נעסוק ב- 3 סוגים של השערות מורכבות:

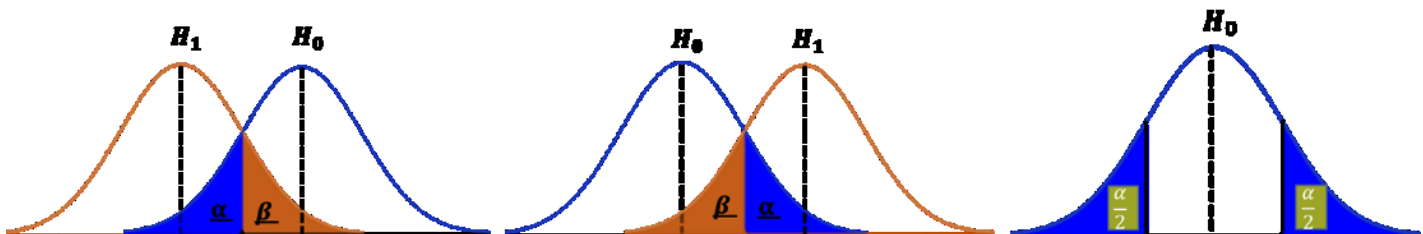
$$\begin{aligned} H_0: \mu &\geq \mu_0 \\ H_1: \mu &< \mu_0 \end{aligned}$$

$$\begin{aligned} H_0: \mu &\leq \mu_0 \\ H_1: \mu &> \mu_0 \end{aligned}$$

$$\begin{aligned} H_0: \mu &= \mu_0 \\ H_1: \mu &\neq \mu_0 \end{aligned}$$

חד צדדי

דו צדדי



שאלה 1

מפעל לצינורות מייצר סוג צינור מסוים שקוטרו הממוצע 50 מ"מ עם סטיית תקן 6 מ"מ. לבדיקת תקינות הייצור נלקח בכל יום מדגם של 64 צינורות מהתוצרת היומית ונמדד קוטרם.

א. נסח את ההשערות וקבע מהו כלל ההכרעה ברמת מובהקות 5%.

ב. ביום מסוים נמצא במדגם קוטר ממוצע של 48.8 ס"מ. האם התוצרת היומית עברה את הביקורת על סמך כלל ההכרעה שקבעת בסעיף א'?

ג. אם ביום כלשהו חלה תקלה בייצור והקוטר הממוצע של הצינורות היה 48 מ"מ בלבד עם סטיית תקן 6 מ"מ, מה ההסתברות שהתקלה לא תתגלה בביקורת היומית? כיצד נקראת הסתברות זו?



שאלה 2

מפעל לייצור צינורות מייצר צינור שקוטרו מתפלג נורמלית עם תוחלת של 50 מ"מ וסטיית תקן של 6 מ"מ. במחלקת ביקורת האיכות דוגמים בכל יום 81 צינורות ומודדים את קוטרם, בכדי לבדוק, בעזרת מבחן סטטיסטי,

האם מכונת הייצור מכוילת כנדרש או שקוטר הצינורות קטן מהדרוש.

א. נסח את ההשערות וקבע מהו כלל ההכרעה לבדיקתו, ברמת מובהקות 0.05.

ב. ביום מסוים נמצא במדגם קוטר של 48.8 מ"מ.

מה הייתה החלטת מחלקת ביקורת האיכות באותו יום לגבי פעולת המכונה?

ג. אם ביום כלשהו מכונת הייצור התקלקלה והיא מייצרת את הצינורות בקוטר שתוחלתו 48 מ"מ בלבד

(סטיית התקן לא משתנה), מה ההסתברות שהתקלה לא תתגלה בביקורת האיכות? כיצד נקראת הסתברות זו?



שאלה 3

ממוצע ציוני הבגרות במתמטיקה בארץ בשנה מסוימת היה 6.5 עם סטיית תקן 1.2. מדגם של 100 בוגרי בי"ס תיכון מסוים השיג באותה שנה ממוצע של 5.7 בבגרות במתמטיקה.

א. האם אפשר לטעון שבאוכלוסיית בית ספר זה נמצאה ירידה שאיננה מקרית בהשוואה לממוצע הארצי ברמת מובהקות של 0.05?

ב. מה תהיה מסקנתך ברמת מובהקות 0.01 ו-0.1 יש להסביר עבור כל רמת מובהקות, האם ניתן להגיע למסקנה ללא בדיקה מחדש.



שאלה 4

מכונה לאריזת קופסאות גבינה מתוכננת לארוז קופסאות במשקל עם תוחלת 200 גרם וסטיית תקן 10 גרם. הועלתה טענה כי הכיול של המכונה נפגם והיא אורזת כעת קופסאות בעלות משקל גבוה יותר ללא שינוי בסטיית התקן. לצורך בדיקת הטענה שקלו 64 קופסאות גבינה.

א. נסח את ההשערות ורשום את המבחן לבדיקתן ברמת מובהקות 0.01.

ב. חשב את עוצמת המבחן המתואר בסעיף א' אם התוחלת האמיתית של משקל קופסת גבינה היא 204 גרם.

ג. כיצד תשתנה עוצמת המבחן, שחישבת בסעיף ב', כתוצאה מהשינויים הבאים: (כל שינוי בנפרד)

1. רמת המובהקות שווה ל-0.05.

2. גודל המדגם הוא 120.

3. סטיית התקן שווה ל-20 גרם.

4. התוחלת האמיתית שווה ל-203 גרם.

בכל מקרה נמק תשובתך או חשב.



שאלה 5

מנתוני משרד הבריאות ידוע כי ממוצע לחץ הדם באוכלוסייה הוא 133 עם סטיית תקן 15. במטרה לבדוק את הטענה כי מנהלים סובלים מלחץ דם גבוה יותר משאר האוכלוסייה, נלקח מדגם מקרי של 64 מנהלים ונמצא שממוצע לחץ הדם שלהם הוא 137.

- מה תהיה מסקנתך ברמת מובהקות של 0.01.
- מהי הטעות האפשרית במסקנתך מסעיף א', אם ידוע שבקרב מנהלים לחץ הדם הממוצע הוא 140? הסבר וחשב.
- חשב גודל מדגם מינימלי כך שיבטיח שאם לחץ הדם בקרב המנהלים הוא אכן 140 נצליח לגלות זאת בהסתברות של 99% לפחות, ואם לחץ הדם אצל מנהלים אינו גבוה מ-133, יוחלט בטעות שהוא גבוה יותר בהסתברות שלא תעלה על 0.02.



שאלה 6

מעוניינים לבדוק אם מאזניים חדשים שהגיעו למעבדה מכילים היטב. לשם כך נלקחה כמות חומר במשקל (ידוע) של 100 גרם, ונשקלה 25 פעמים במאזניים החדשים. לפי פרסומי היצרן של המאזניים, בשקילת 100 גרם יש סטיית תקן של 3.2 גרם וההתפלגות נורמלית.

- במדגם התקבל משקל ממוצע של 101.6 גרם. האם ניתן להסיק שהמאזניים מכילים היטב? נסח את ההשערות ובדוק ברמת מובהקות 0.05.
- האם ניתן להגיע לאותה מסקנה, ללא חישוב נוסף, עבור רמות המובהקות 0.01, 0.1? הסבר.
- במבחן מסעיף א, מהו הסיכוי לגלות שהמאזניים אינם מכילים היטב, אם למעשה יש להם הטיה של 2 גרם כלפי מעלה?



שאלה 7

התפלגות משקל ביצי תרנגולת היא נורמלית עם ממוצע 60 גרם, וסטיית תקן 15 גרם.

חוקר הציע הורמון חדש שאמור להגדיל את המשקל הממוצע ב-5 גרם מבלי לשנות את סטיית התקן. הטענה נבדקה בעזרת מבחן סטטיסטי.

- ההורמון החדש נוסה על 100 תרנגולות. נמצא כי ההסתברות לטעות מסוג שני במבחן היא 0.1, מהי רמת המובהקות של המבחן?
- בהמשך לסעיף א', אם ממוצע משקל הביצים במדגם היה 62 גרם, מה תהיה מסקנת הבדיקה?
- מהו גודל המדגם המינימלי שיבטיח רמת מובהקות של 0.01 ועצמה של 0.95 לפחות?



שאלה 8

- למשקלו של תפוח יש ממוצע 260 גרם וסטיית תקן 38 גרם. אגרונום המתמחה בגידול תפוחים טוען כי שיטת השקיה חדשה שפיתח מעלה את המשקל הממוצע של תפוח ל-275 גרם.
- לבדיקת טענתו נלקח מדגם מקרי של 60 תפוחים מחלקת קרקע שהושקתה בשיטת ההשקיה החדשה. הוחלט ששיטת ההשקיה החדשה תיכנס לשימוש אם משקל התפוחים הממוצע במדגם יהיה מעל 267 גרם.
- א. מהי רמת המובהקות של המבחן המוצע?
- ב. מהי עצמת המבחן המוצע?
- ג. כיצד ישתנו תשובותיך לסעיפים א ו ב אם יוחלט להכניס לשימוש את שיטת ההשקיה החדשה רק אם משקל התפוחים הממוצע במדגם יהיה מעל 273 גרם? נמק. אין צורך לחשב מחדש אך גם חישוב מחדש יתקבל כנימוק.
- ד. מהו גודל המדגם המינימלי שיבטיח שההסתברות להחליט בטעות שהשיטה החדשה יעילה לא תעלה על 0.05 וההסתברות להחליט בטעות שהשיטה החדשה אינה יעילה לא תעלה על 0.01?



שאלה 9

- בשיטת ייצור קיימת מייצר עובד בממוצע 60 מוצרים ליום עם סטיית תקן של 15 מוצרים. מנהל הייצור מציע שיטת תמריצים שתגדיל את התפוקה הממוצעת ב-10%.
- לבדיקת טענתו ניתנו התמריצים ל-100 עובדים שנבחרו באופן מקרי ונמצא כי התפוקה הממוצעת הייתה 64 מוצרים.
- א. נסח את ההשערות ובדוק את הטענה ברמת מובהקות 0.01.
- ב. מהו גודל המדגם שיש לדגום כדי לבנות מבחן סטטיסטי לבדיקת טענתו של מנהל הייצור בהסתברות לטעות בטעות ששיטת התמריצים לא עוזרת לא תעלה על 0.05, וההסתברות להחליט בטעות ששיטת התמריצים עוזרת לא תעלה על 0.06?



שאלה 10

- בשיטת ייצור קיימת מייצר עובד בממוצע 120 מוצרים ליום עם סטיית תקן של 24 מוצרים.
- מנהל המפעל מציע שיטת תמריצים כדי להגביר את התפוקה ב-10%.
- הוחלט לערוך בדיקה מדגמית כדי לבדוק את יעילות השיטה.
- א. מהו גודל המדגם שיש לבדוק אם מעוניינים שההסתברות להחליט בטעות שהשיטה יעילה לא תעלה על 0.05, וההסתברות לגלות שהשיטה אכן מגבירה את התפוקה ב-10% תהיה לפחות 0.97?
- ב. השיטה הופעלה על מדגם מקרי של עובדים בגודל שמצאת בסעיף א', ונמצא כי ממוצע המדגם הוא 126. מהי מסקנת הבדיקה ברמת מובהקות 0.01?
- ג. מהי הטעות האפשרית במסקנה מסעיף ב ומהי הסתברותה?



שאלה 11

קופת חולים בדקה שיטה חדשה לקבלת חולים במרפאה, כדי לקצר את משך ההמתנה לרופא.

במצב הקיים ידוע כי משך ההמתנה לרופא מתפלג נורמלי עם תוחלת 26 דקות וסטיית תקן של 12 דקות.

הוחלט כי השיטה החדשה תופעל לגבי מדגם מקרי של 36 נבדקים.

א. נסח את ההשערות ובנה מבחן לבדיקתן ברמת מובהקות 0.01.

ב. במדגם נמצא כי משך ההמתנה הממוצע הוא 23 דקות. האם השיטה החדשה מקצרת את זמן ההמתנה? נמק.

ג. בהמשך לסעיף ב', מהי מובהקות התוצאה, כלומר מהי רמת המובהקות הקטנה ביותר בה נחליט שהשיטה החדשה אכן מקצרת את זמן ההמתנה, על סמך נתוני המדגם.



שאלה 12

חוקר מעוניין לאמוד את ממוצע לחץ הדם של אוכלוסיית הקשישים.

נבדק מדגם מקרי של 225 קשישים, ובמדגם נמצא ממוצע לחץ דם 148.

ידוע כי סטיית התקן של לחץ הדם באוכלוסייה היא 20.

א. חשב את רווח סמך, לפרמטר μ , ברמת סמך 0.95, ופרש את התוצאה שהתקבלה.

ב. מהי שגיאת האמידה המקסימלית?

ג. אם נגדיל את המדגם פי 4, כיצד יושפע מכך רווח הסמך ל- μ , באותה רמת סמך?



שאלה 13

על פי נתוני משרד הבריאות ממוצע רמת הכולסטרול באוכלוסייה הוא 185 וסטיית התקן 15.

ארגון למניעת מחלות לב טוען שרמת הכולסטרול הממוצעת באוכלוסייה גבוהה יותר.

לבדיקת טענת הארגון נבדקה רמת הכולסטרול של מדגם מקרי של 120 אנשים והתקבלה רמת כולסטרול של 188.

א. האם התוצאות מאששות את טענת הארגון ברמת מובהקות 0.02? נמק.

ב. בהמשך לסעיף א, חשב את עצמת המבחן אם רמת הכולסטרול הממוצעת באוכלוסייה היא 190 ללא שינוי בסטיית התקן.



שאלה 14

נתונות ההשערות:

$$H_0: \mu = 80$$

$$H_1: \mu = 70$$

וכן נתון: $\sigma = 15$

לבדיקת ההשערות נקבע כלל ההכרעה הבא: דחה את H_0 אם $\bar{X} < 75$.

א. מהו גודל המדגם המינימלי הדרוש להשגת עוצמה של 0.99?

ב. עבור גודל המדגם שמצאת בסעיף א, מהי ההסתברות לטעות מסוג ראשון של המבחן שנקבע?

ג. מהו גודל המדגם המינימלי הדרוש להשגת עוצמה של 0.98 והסתברות לטעות מסוג ראשון שאינה עולה על 0.05?



שאלה 15

מדגם בגודל 100 נלקח מתוך אוכלוסייה המתפלגת נורמלית עם $\sigma = 40$.

ברמת מובהקות 0.05 מצא:

א. מהו אזור הדחייה עבור ההשערות: $H_0: \mu = 42$ כנגד $H_1: \mu = 50$

ב. מהו אזור הדחייה עבור ההשערות: $H_0: \mu = 50$ כנגד $H_1: \mu = 42$

ג. מהי החלטתך, בכל אחד מהמקרים אם ממוצע המדגם הוא 45.5

ד. חזור על סעיפים א-ג עבור $\sigma = 20$



שאלה 16

יצרן סיגריות נתבקש לאמוד את תכולת הניקוטין הממוצעת בסיגריות שהוא מייצר. ידוע כי תכולת הניקוטין בסיגריות מתפלגת נורמלית עם סטיית תקן 5.

היצרן דגם באופן מקרי 25 סיגריות ועל סמך תוצאות המדגם בנה רווח סמך לתכולת הניקוטין הממוצעת וקיבל את הרווח $26 < \mu < 30$.

א. באיזו רמת סמך בנה היצרן את הרווח סמך?

ב. היצרן טוען כי תכולת הניקוטין הממוצעת בסיגריות שהוא מייצר היא 25. האם ניתן לאשש את טענתו על סמך הרווח סמך שהתקבל? נמק.

ג. מהו גודל המדגם שעל החוקר לבדוק אם הוא מעוניין לקבל רווח סמך, באותה רמת סמך, שאורכו מחצית מאורך הרווח הנתון? נמק.



שאלה 17

התפלגות משקל התבלינים בצנצנות מתפלג נורמלית עם סטיית תקן 0.8 גרם.

על מנת לאמוד את המשקל הממוצע של תבלינים מתוצרת חברת ה"תבלין" בצנצנת, נלקח מדגם מקרי של 100 צנצנות, ונמצא כי המשקל הממוצע שלהן 9.85 גרם.

א. בנה רווח סמך לתוחלת המשקל של התבלינים בצנצנות ברמת סמך של 95%.

ב. מה צריך להיות גודל המדגם שיש לדגום על מנת שבאותה רמת סמך שבסעיף א' נקבל רווח סמך ששגיאת האמידה שלו לא תעלה על 0.05 גרם?

שאלה 1

מתקן מפעיל התראה למצב חירום על סמך החוזק של 9 אותות שהוא קולט. חוזק האותות מתפלג נורמלית. במצב רגיל, תוחלת החוזק היא 5000 יחידות וס"ת 900 יחידות. במצב חירום, תוחלת החוזק היא 6300 יחידות וס"ת נשארת 900. המתקן מפעיל את האתראה רק כאשר העוצמה הממוצעת של תשעת האותות עולה על 5600 יחידות.

א. מהי ההסתברות של אתראת שווא?

ב. מהי ההסתברות שלא תופעל אתראה במצב חירום?

ענה על הסעיפים הבאים בהסבר ללא חישוב. כיצד יושפעו הסתברויות של א' ו- ב':

ג. מתקן האתראה יופעל כאשר ממוצע החוזק יעלה על 5700 יחידות.

ד. מספר האותות הנקלטים יהיה 12.

ה. תוחלת החוזק במצב חירום תהיה 6400 יחידות.



שאלה 19

הזמן הדרוש לייצור מוצר מסוים הוא בממוצע 10 שעות וסטיית תקן של זמני הייצור היא שעה אחת. בית חרושת טוען שיש לו תהליך יצור מהיר יותר. נערכה בדיקה ונמצא ש- 100 מוצרים נוצרים במשך 960 שעות.

- א. האם מוצדקת טענת בית החרושת? בדוק ברמת מובהקות של 0.05.
 ב. מהו הטעות האפשרית במסקנתך מסעיף א' ומה ערכה? נמק.
 ג. אם ידוע כי זמן היצור בבית החרושת הנ"ל הוא 9.8, מה הסיכוי שנצליח לגלות זאת? כיצד נקראת הסתברות זו ומה ערכה.
 ד. חשב גודל מדגם מינימלי כך שיבטיח שהסיכוי לדחייה לא מוצדקת של H_0 לא יעלה על 0.01, והסיכוי לקבלה שגויה של H_0 לא יעלה על 0.05.



שאלה 20

על פי נתוני הלשכה המרכזית לסטטיסטיקה, רמת החיסכון הממוצעת למשפחה בישראל היא 1500 ₪ לחודש עם סטיית תקן 400 ₪, במדגם מיקרי של 100 משפחות התקבל חיסכון ממוצע של 1550 ₪. מנהל חברת השקעות טוען כי רמת החיסכון הממוצעת באוכלוסייה עלתה. האם תוצאות המדגם מאששות את טענתו? בדוק ברמת מובהקות של 0.05.



שאלה 21

כמות הסידן בגביעי יוגורט צריכה להיות עם ממוצע של 120 מיליגרם וסטיית תקן של 20 מיליגרם. תזונאי משרד הבריאות טוען כי ממוצע תכולת הסידן בגביעים נמוכה מהנדרש. לבדיקת הטענה הוחלט לבדוק מדגם מקרי של 64 גביעים ועל סמך תכולתם הממוצעת להחליט האם הטענה נכונה.

- א. ברמת מובהקות 0.05, מהי המסקנה אם במדגם התקבל ממוצע של 115 מיליגרם?
 ב. מהי רמת המובהקות המינימלית שעבורה תתקבל טענת התזונאי?
 ג. אם טענת התזונאי נכונה וכמות הסידן הממוצעת בגביעים היא 110 מיליגרם, מהי ההסתברות להצליח לגלות זאת עבור המבחן ברמת מובהקות 0.05?

openbook
המרכז לקידום אקדמי

שאלה 1

מהנדס רכב מציע מבנה חדש לבלמים. מרחק העצירה הממוצע של בלמים במבנה הרגיל, בנסיעה של 80 קמ"ש, הוא 50 מטר.
 הוחלט להרכיב את הבלמים החדשים ברכב החברה רק אם ניסוי יצביע על ירידה משמעותית במרחק העצירה שלהם, לעומת הבלמים הרגילים.
 הוצע להרכיב את הבלמים החדשים באם במדגם מקרי של 36 מדידות עם הבלמים החדשים, מרחק העצירה הממוצע יהיה קטן מ- 48.6 מטר.
 ידוע כי סטיית התקן של מרחק העצירה של הבלמים החדשים היא 3 מטר.
 א. נסח את ההשערות וחשב מהי רמת המובהקות של המבחן המוצע.

ב. מהי ההסתברות לכך שעל סמך הניסוי יוחלט לא להרכיב את הסוג החדש, אם למעשה מרחק העצירה הממוצע שלו הוא 48 מטר?

ג. כמה מדידות יש לבצע אם מעוניינים שאם מרחק העצירה הממוצע עם הבלמים החדשים הוא 48 מטר, הדבר יתגלה בניסוי בהסתברות של 0.99 לפחות, ואם הבלמים החדשים לא גורמים לירידה במרחק העצירה יוחלט בטעות להכניסם לשימוש בהסתברות של 0.05?



שאלה 22

חברת תרופות מפתחת תרופה חדשה להורדת לחץ דם בחולים הסובלים מיתר לחץ דם. מנכ"ל החברה טוען שאם התרופה לא תוריד את לחץ הדם ב- 5 יחידות, הפיתוח לא יהיה לו כדאי. צוות החוקרים פיתחו את התרופה ומתכוונים לבדוק אותה במדגם של 30 חולים. אם ידוע שלחץ הדם בקרב חולים הסובלים מיתר לחץ דם הוא בממוצע 160 עם סטיית תקן של 20:

א. מה תהיה עוצמת המבחן אם החוקרים משתמשים ברמת מובהקות של 0.05?

ב. החוקר פנה בבקשה לביצוע הניסוי, לארגון פיתוח התרופות האירופאי. תשובתם הייתה שהם אינם מאפשרים ניסויים שעוצמתם הסטטיסטית קטנה מ- 0.9.

מהו מספר החולים שעל החוקרים לדגום כדי לעמוד בדרישה זו ברמת מובהקות של 0.05?

ג. נבדק מדגם של חולים בגודל שנמצא בסעיף ב, שהשתמשו בתרופה החדשה והתקבל ממוצע לחץ דם של 152.

מצא רווח סמך ברמת סמך 95% לממוצע לחץ הדם בקרב משתמשי התרופה החדשה.



שאלה

במפעל מסוים זמן הלימוד של פועל על מכונה חדשה מתפלג נורמלית. מהעבר ידוע כי הוא בעל תוחלת של 90 דקות עם סטיית תקן של 7 דקות. הגיעה למפעל מכונה חדשה והמנהל החליט לבדוק האם תוחלת זמן הלימוד היא אכן מה שהיתה בעבר או שהשתנתה.

לשם כך נלקח מדגם של 20 פועלים בו התקבל זמן לימוד ממוצע של 86 דקות.

א. אמדו את תוחלת זמן הלימוד בעזרת רווח סמך ברמת סמך של 98%.

ב. מה תסיקו לגבי שאלת המנהל לפי רווח הסמך שקיבלתם בסעיף א'? נמקו.

ג. 1. בדקו, ברמת מובהקות של 0.01, את שאלת המנהל.

רשמו באופן ברור את ההשערות הנדרשות, הבדיקה ומסקנה מילולית.

2. מה רמת המובהקות המינימלית בה תקבלו את השערת האפס לפי מדגם זה?



שאלה

- במסגרת מחקר על מעמד האישה נדרש לאמוד את תוחלת מספר שעות העבודה בשבוע של נשים עובדות. ידוע כי סטיית התקן של מספר שעות העבודה השבועיות היא 5 שעות.
- א. נלקח מדגם מקרי של 225 נשים עובדות ובו התקבל ממוצע של 37.5 שעות.
- מצאו רווח סמך עבור תוחלת מספר שעות העבודה בשבוע של נשים עובדות, ברמת סמך של 95%.
 - מהו גודל המדגם שיש לקחת אם רוצים לקבל רווח סמך, באותה רמת סמך, שאורכו מחצית מהאורך שהתקבל בסעיף 1?
 - נבנה רווח סמך לפי מדגם בגודל שמצאתם בסעיף 2'. האם ניתן לטעון בביטחון רב יותר שרווח זה כולל את התוחלת (μ)? נמקו!
- ב. התקן המוצהר של מספר שעות העבודה הנדרשות מנשים עובדות הוא 37.
- אחד החוקרים טען כי בפועל תוחלת מספר שעות העבודה של נשים עובדות היא גדולה יותר. לצורך בדיקת טענתו הוצע המבחן הבא: ילקח מדגם מקרי של 225 נשים עובדות. אם ממוצע המדגם יהיה גדול מ- 37.6 שעות אז טענת החוקר תתקבל.
- רשמו את ההשערות הנבדקות
 - מה רמת המובהקות של המבחן המוצע?
 - אם החוקר צודק ותוחלת מספר שעות העבודה של נשים עובדות היא 38 שעות, מה ההסתברות שהמבחן המוצע יגלה זאת? כיצד נקראת הסתברות זאת?

openbook
המרכז לקידום אקדמי



שאלה

- מנהל חדר מיון הכניס נוהל חדש לטיפול בחולים המגיעים לחדר המיון.
- בדקו את זמן ההמתנה (בדקות) של 12 חולים שהגיעו לחדר המיון לאחר הפעלת הנוהל החדש והתוצאות שקיבלו היו:
- 25, 28, 29, 30, 33, 34, 35, 45, 50, 58, 60, 65.
- הניחו כי זמן ההמתנה מתפלג נורמלית.
- א. אמדו את תוחלת זמן ההמתנה בעזרת רווח סמך ברמת סמך 95%.
 - ב. מנהל חדר המיון טוען כי תוחלת זמן ההמתנה אחרי הכנסת הנוהל החדש היא 38 דקות.

האם תקבלו את טענתו לפי רוח הסמך שקיבלתם בסעיף א'? נמקו!

ג. תוחלת זמן ההמתנה לפני הפעלת הנוהל החדש הייתה 48.5 דקות.

מנהל בית החולים טוען כי תוחלת זמן ההמתנה קטנה לאחר הפעלת הנוהל החדש.

1. בדקו את טענתו של מנהל בית החולים ברמת מובהקות 0.05. פרטו מה המבחן בו אתם משתמשים ורשמו מסקנה מילולית ברורה.

2. משרד הבריאות דרש לבדוק את טענת מנהל בית החולים ברמת מובהקות מחמירה יותר של 0.025. האם תשתנה מסקנתכם? נמקו.



מודל 2: בדיקת השערות של ממוצע – סטיית התקן אינה ידועה

ביחידה 10 שאלנו את השאלה:

"מדוע אנו לומדים גם רווח סמך לתוחלת כשהשונות לא ידועה?"

כשאנו באים לאמוד את μ או $\mu_1 - \mu_2$, רק לעתים רחוקות נתונות השוניויות באוכלוסיות הנידונות.

גם הפעם נעזר באומד חסר הטיה לשונות באוכלוסייה - $\hat{S}^2$.

כפי שלמדנו ביחידה 10,

התפלגות הדגימה של הסטטיסטי המתקבל היא התפלגות t.

הנחות:

$$x \sim N(\mu, \sigma^2)$$

כאשר $n > 30$ נשתמש בהתפלגות נורמלית עם אמד לשונות: $S^2 = \frac{\sum_{i=1}^n x_i^2 - n\bar{x}^2}{n-1}$

$$T_x = \frac{\bar{x} - c}{S/\sqrt{n}}$$

נשתמש בסטטיסטי המבחן:

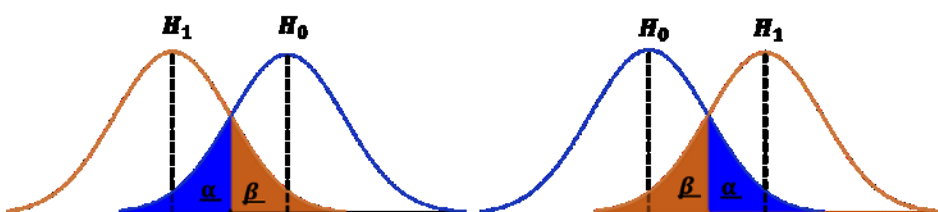
התפלגות t עם $(n-1)$ דרגות חופש

$$\begin{aligned} H_0: \mu &\geq \mu_0 \\ H_1: \mu &< \mu_0 \end{aligned}$$

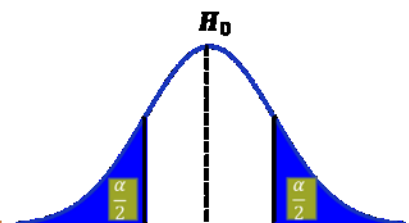
$$\begin{aligned} H_0: \mu &\leq \mu_0 \\ H_1: \mu &> \mu_0 \end{aligned}$$

$$\begin{aligned} H_0: \mu &= \mu_0 \\ H_1: \mu &\neq \mu_0 \end{aligned}$$

חד צדדי



דו צדדי





שאלה 1

יצרן מכוניות מפרסם שמכונית מסוימת שהוא מייצר צורכת ליטר דלק ל-20 ק"מ נסיעה בממוצע. לחברה מסוימת יש מספר רב של מכוניות מאותה תוצרת. בבחירה מקרית של 9 מכוניות נבדקו מספר הק"מ לליטר שצרכה כל אחת מהמכוניות בנסיעה מיוחדת של 400 ק"מ. התקבלו התוצאות הבאות (בק"מ לליטר) :

$$\sum_{i=1}^9 x_i^2 = 3271, \quad \sum_{i=1}^9 x_i = 171$$

בהנחה שמספר הק"מ לליטר מתפלג נורמלית, האם יש בנתונים אלה הוכחה לכך שיצרן המכוניות אינו נוקט בשיטת "אמת בפרסום" ? נסח את ההשערות ובדוק עבור רמת מובהקות של 0.05.



שאלה 2

ממוצע הציונים במבחן הבודק כושר ריכוז באוכלוסייה רגילה הוא 84.16 נבדקים קיבלו תרופת הרגעה ונבדקה השפעת תרופה זו על הביצוע שלהם במבחן הריכוז. החוקר משער שבהשפעת התרופה יורד כושר הריכוז.

$$\sum_{i=1}^{16} x_i^2 = 107,225$$

$$\sum_{i=1}^{16} x_i = 1299$$

לגבי ציונים שהתקבלו במדגם התוצאות היו :

א. האם התוצאות מעידות על כך שהתרופה מפחיתה את כושר הריכוז, ברמת מובהקות 0.05 ?

נסח את ההשערות ורשום את ההנחות הדרושות.

ב. מצא רווח סמך ברמת סמך 95% עבור התוחלת של הציון במבחן הריכוז של סמך תוצאות המדגם.

ג. האם מתוך הרווח סמך שמצאת בסעיף ב', ניתן להסיק כי בהשפעת התרופה חל שינוי בכושר הריכוז, ברמת מובהקות 0.01 ? הסבר.



שאלה 3

בשני מדרונות נדגם עובי הקרקע (Cm) (להלן X) :

Thichness	1	2	3	4	5	6	7	8
מדרון א	157	161	155	166	145	150	157	153
מדרון ב	143	150	149	157	156	146	159	

א. חשב רווח סמך לתוחלת X על סמך המדגם ממדרון א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת X על סמך המדגם ממדרון ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק את ההשערה שתוחלת X במדרון א שונה מ-160 ברמת מובהקות 0.05. נמק.

ה. בדוק את ההשערה שתוחלת X במדרון א שונה מתוחלת X במדרון ב ברמת מובהקות 0.05. נמק.

ו. בדוק את ההשערה ששונות X במדרון א שונה משונות X במדרון ב ברמת מובהקות 0.1, בהנחה שא מתפלג נורמלית. נמק.

ז. האם דרושה הנחה כלשהי לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.



שאלה 4

בשני רשתות חברתיות נדגם מספר החברים בקרב המשתמשים (להלן X):

Obs	1	2	3	4	5	6	7	8
רשת א	157	161	155	166	145	150	157	153
רשת ב	143	150	149	157	156	146	159	

א. חשב רווח סמך לתוחלת X על סמך המדגם מרשת א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת X על סמך המדגם מרשת ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק את ההשערה שתוחלת X ברשת א שונה מ-160 ברמת מובהקות 0.05. נמק.

ה. בדוק את ההשערה שתוחלת X ברשת א שונה מתוחלת X ברשת ב ברמת מובהקות 0.05. נמק.

ו. בדוק את ההשערה ששונות X ברשת א שונה משונות X ברשת ב ברמת מובהקות 0.1, בהנחה שא מתפלג נורמלית. נמק.

ז. האם דרושה הנחה כלשהי לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.

מודל 6: בדיקת השערות על פרופורציה

פרופורציה

$$\hat{p} = \frac{x}{n}$$

אומדן חסר הטיה לפרופורציה באוכלוסייה

באמידת פרופורציה אנו מעוניינים לחשב אחוז כלשהו מהאוכלוסייה.

p – הפרמטר האמיתי = הפרופורציה באוכלוסייה

ניתן לומר ש- p הוא ההסתברות להצלחה בהתפלגות בינומית $X \sim B(n, p)$.

לכן אמידת פרופורציה היא ניסוי בינומי בעל הסתברות להצלחה p , כאשר אנו מבצעים n חזרות ב"ת ומס' איברים בעלי התכונה המבוקשת שנקבל היא x , אז:

$$X \sim \text{Bin}(n, p) \quad \Rightarrow \quad E(X) = np, \quad V(X) = npq$$

בדיקת השערות על פרופורציה

הנחות:

התפלגות נורמלית

או אם $np \geq 10$, $nq \geq 10$

המבחן מבוסס על קירוב נורמלי למ"מ בינומי.

$$Z_p = \frac{p - c}{\sqrt{\frac{c(1-c)}{n}}}$$

הסטטיסטי במדגם:

שאלה 1

מנתוני משרד הבריאות ידוע כי התרופות המקובלות לכאבי ראש מביאות הקלה ל-70% מהמשתמשים בהן. יצרן תרופות טוען שיש תרופה חדשה יעילה יותר. לשם בדיקת הטענה נבדק מדגם מקרי של 250 איש הסובלים מכאבי ראש.

א. אם נמצא כי ההסתברות לגלות, שאכן התרופה החדשה מביאה הקלה ל-78%, היא 0.85, מהי ההסתברות לטעות מסוג ראשון של המבחן?

ב. אם במדגם נמצא כי התרופה החדשה הביאה הקלה ל-188 מהם – מהי מסקנת הבדיקה ברמת מובהקות 0.05? נמק.

ג. מהו גודל המדגם המינימלי שיבטיח, שאם אכן התרופה החדשה מביאה הקלה ל-78%, ההסתברות לגלות זאת תהיה 0.95, וההסתברות להחליט בטעות שהתרופה החדשה יעילה תהיה 0.01.



שאלה 2

בבית ספר תיכון מסוים נמצא כי 40% מהתלמידים מאחרים לפחות פעם אחת בחודש. מנהל בית הספר מציע שיטת תמריצים מיוחדת לצמצום האיחורים. השיטה נבדקה על מדגם מקרי של 200 תלמידים. נמצא כי 68 מתוכם אחרו לפחות פעם אחת מחודש.

א. האם השיטה יעילה? נסח את ההשערות ובדוק ברמת מובהקות 0.025.

ב. מהי ההסתברות לטעות האפשרית במסקנה מסעיף א', אם שיטת התמריצים מפחיתה את אחוז התלמידים המאחרים לפחות פעם אחת בחודש, ל-30%?

ג. מהי רמת המובהקות המינימלית שעבורה נחליט ששיטת התמריצים יעילה, על סמך תוצאות המדגם?



שאלה 3

בחברת התעופה "רקיע" אחוז המזוודות שהולכות לאיבוד הוא 4%.

לאחרונה שינתה החברה את נוהלי העבודה במטרה להקטין את אחוז המזוודות שאובדות.

א. מנהל התפעול בחברה החליט לבדוק 500 מזוודות מקריות, ואם לפחות 485 מהן יגיעו ליעדן הוא יסיק כי הנוהל החדש משיג את מטרתו.

מהי רמת המובהקות של המבחן שהציע מנהל התפעול?

ב. עבור מדגם בגודל 500, מהו המבחן המתאים ברמת מובהקות 0.025, כלומר עבור אילו פרופורציות מדגמיות יוחלט ברמת מובהקות 0.025, שהנוהל החדש מוריד את אחוז המזוודות שהולכות לאיבוד?

ג. בהמשך לסעיף ב', אם הנוהל החדש אכן מוריד את אחוז המזוודות האובדות ל-2%, מה ההסתברות שמבחן שבו תיבדקנה 500 מזוודות, יגלה זאת?



שאלה 4

60% מכלל הנבחנים בבחינת בגרות במתמטיקה עוברים את הבחינה בהצלחה. מורה מציע שיטת הוראה חדשה להגדלת פרופורציית המצליחים. השיטה החדשה נבדקה על מדגם מקרי של 200 תלמידים, ובמדגם נמצא כי 66% מהנבחנים עברו בהצלחה את הבחינה.

א. מהי מסקנת הבדיקה ברמת מובהקות 0.01?

ב. מהי ההסתברות לטעות במסקנה שנתקבלה, אם בשיטה החדשה לפחות 75% מהנבחנים עוברים את הבחינה בהצלחה?

ג. מהו גודל המדגם שיש לחקור כדי לבנות מבחן עבור רמת מובהקות 0.01 ועוצמה 95% לפחות?



שאלה 5

יצרן תנורים להורדת גשם טוען כי לתנורים שהוא מיצר יש 80% הצלחה בהורדת גשם. בעל תחנה חקלאית רוצה לבדוק טענה זו, החורגת מהסיכוי המקובל להורדת גשם, של 70% הצלחה, ברמת מובהקות 0.05.

א. נסח את ההשערות.

ב. כדי להבטיח עוצמה של 0.9, כמה פעמים יש לנסות את התנורים הנ"ל?

ג. לאיזו מסקנה יגיע בעל התחנה החקלאית, אם ב-125 ניסיונות ירד גשם 95 פעמים?

ד. מהי הטעות האפשרית במסקנה מסעיף ג', ומהי הסתברותה?



שאלה 6

חברה לדיוור ישיר שוקלת לצאת במבצע פרסום למוצר מסוים, באמצעות הדואר.

משיקולים כלכליים החברה מעריכה שמבצע הפרסום יהיה כדאי רק אם שיעור התגובה של הצרכנים, כלומר מספר ההזמנות ביחס למספר המכתבים שנשלחו הינו לפחות 15%.

בכדי לבדוק את כדאיות המבצע, נשלחו 100 מכתבים ל-100 צרכנים שנבחרו באופן מקרי מרשימת תפוצה, ומתוכם התקבלו 10 הזמנות.

א. מהי המלצתך לחברה ברמת מובהקות 3%?

ב. אם שיעור התגובה האמיתי של הצרכנים הוא 12%, מהי עוצמת המבחן מסעיף א'?

ג. בנה רווח סמך ברמת סמך 95% לשיעור התגובה באוכלוסייה על סמך המדגם שלעיל.

ד. לדעת מנכ"ל החברה, רווח הסמך שהתקבל גדול מדי. מהו גודל המדגם הדרוש כדי שאורך רווח הסמך יהיה מחצית מהאורך שקיבלת בסעיף ג' באותה רמת סמך?



שאלה 7

על פי נתוני משרד הבריאות 50% מהאנשים שוהים בשמש ללא אמצעי הגנה.

איגוד רופאי העור מעוניין לבדוק האם, בעקבות הפרסום על נזקי השמש,

אחוז האנשים השוהים בשמש ללא הגנה ירד.

מהו גודל המדגם המינימלי שתמליץ לרופאי העור לדגום

כדי לבדוק אם האחוז ירד ל-40% עבור רמת מובהקות 0.01 ובעצמה של 0.98 לפחות?



שאלה 8

התרופה המקובלת נגד כאבי ראש מביאה הקלה ל-70% מהמשתמשים בה.

יצרן תרופות טוען שמצא תרופה חדשה המביאה הקלה לאחוז גבוה יותר של משתמשים.

לשם בדיקת טענתו, נבדק מדגם מקרי של 121 איש הסובלים מכאבי ראש, ונמצא כי התרופה החדשה הביאה הקלה ל-88 מהם.

- א. בדוק את ההשערה שהתרופה החדשה יעילה מקודמתה ברמת מובהקות 0.01.
- ב. מהי רמת המובהקות המינימלית שבה נחליט שהתרופה אכן מקלה על כאבי ראש, על סמך נתוני המדגם?
- ג. מהי הטעות האפשרית במסקנה מסעיף א' ומהי הסתברותה אם התרופה החדשה מביאה הקלה ל- 85% מהמשתמשים בה?



שאלה 9

- בחברת התעופה "רקיע" אחוז המזוודות שהולכות לאיבוד הוא 4%.
- לאחרונה שינתה החברה את נוהלי העבודה במטרה להקטין את אחוז המזוודות שאובדות.
- א. מנהל התפעול בחברה החליט לבדוק 500 מזוודות מקריות, ואם לפחות 485 מהן יגיעו ליעדן הוא יסיק כי הנוהל החדש משיג את מטרתו.
- מהי רמת המובהקות של המבחן שהציע מנהל התפעול?
- ב. עבור מדגם בגודל 500, מהו המבחן המתאים ברמת מובהקות 0.025, כלומר עבור אילו פרופורציות מדגמיות יוחלט ברמת מובהקות 0.025, שהנוהל החדש מוריד את אחוז המזוודות שהולכות לאיבוד?
- ג. בהמשך לסעיף ב', אם הנוהל החדש אכן מוריד את אחוז המזוודות האובדות ל- 2%, מה ההסתברות שמבחן שבו תיבדקנה 500 מזוודות, יגלה זאת?
- ד. מהו גודל המדגם המינימלי שיבטיח, שאם הנוהל החדש אכן מוריד את אחוז המזוודות האובדות ל- 2%, ההסתברות לגלות זאת תהיה 0.95, וההסתברות להחליט בטעות שהנוהל החדש מוריד את אחוז המזוודות שהולכות לאיבוד תהיה 0.01.

שאלה 1

- בעיר מסוימת נמצא על-פי מפקד האוכלוסין האחרון, כי אחוז התושבים מעל גיל 24, בעלי תואר אקדמי הוא 25%.
- מנהל אגף החינוך באותה עיר מעוניין לבדוק האם חל שינוי באחוז התושבים בעלי תואר אקדמי.
- סטטיסטיקאי הרשות הציע לבדוק מדגם מקרי של 20 תושבים מעל גיל 24 ולהחליט שלא חל שינוי באחוז האקדמאים, אם מספר בעלי תואר אקדמי במדגם יהיה 6, 7 או 8.
- א. מהי ההסתברות לטעות מסוג ראשון של המבחן המוצע?
- ב. מהי ההסתברות לגלות שינוי באחוז האקדמאים בעיר, אם זה אכן השתנה ל- 40%?
- ג. חזור על סעיפים א' ו-ב' באם יוחלט שחל שינוי באחוז האקדמאים אם מספר בעלי תואר אקדמי יהיה שונה מ- 7?

שאלה 1

- ידוע משנים קודמות כי מבין כל 100 סטודנטים שהתחילו ללמוד באו"פ 20 הפסיקו את לימודיהם. משערים כי השנה אחוז הסטודנטים שהפסיקו ללמוד שונה מהעבר. לשם כך נלקח מדגם מיקרי של 150 סטודנטים ונמצא כי 35 הפסיקו את לימודיהם.

א. נסח השערות ובדוק את הטענה בר"מ 0.02.

$$H_0 : p = 0.2$$

ב. אם ההשערות היו:

$$H_1 : p > 0.2$$

האם ניתן לדעת, ללא חישוב נוסף מה תהיה המסקנה באותה ר"מ? אם לא, מדוע לא?

ג. אם ההשערה היא כי השנה הפסיקו את לימודיהם 25% מהסטודנטים, מה תהיה ההסתברות לטעות מסוג שני על פי המבחן בסעיף א'.



הסקה סטטיסטית לפי שני מדגמים

באילו מצבים החוקר יתעניין בהשוואה לפי שני מדגמים ?

כאשר רוצים להשוות בין 2 טיפולים ולהחליט מי טוב יותר מדגם הראשון יטופל בטיפול א' ומדגם ב' יטופל בטיפול ב' ונשווה בין שתי האוכלוסיות.

1. שני מדגמים בלתי תלויים

המדגמים נבחרו באופן אקראי, אין תלות בין בחירת המקרים בין מדגם א' לב'.

לדוגמא: השוואת שכר גברים לשכר נשים

יתכנו 3 מצבים:

1. שונות ידועה – מודל 3
2. שונות לא ידועה אך שוות – מודל 4
3. שונות לא ידועה ולא שוות – לא נתעסק בקורס זה
2. שני מדגמים מזווגים (תלויים) מודל 5
בין שני מדגמים מזווגים יש מתאם סטטיסטי, מדגם מזווג זה כמו בחירת אותו נחקר לשני מדגמים, השוואת לפני טיפול ואחרי טיפול וכו.

openbook
המרכז לקידום אקדמי

מודל 3+4: בדיקת השערות על הפרש תוחלות במדגמים ב"ת

מדגמים ב"ת

שני המדגמים נבחרו באופן מקרי, אין תלות בין בחירת המקרים של מדגם א' לבחירת המקרים של מדגם ב'

אינפורמציה של נבדק בקבוצה א' לא משפיעה ולא אומרת כלום על נבדק בקבוצה ב'

המדגמים יכולים להיות שונים או שווים בגודלם

כאשר חישבנו רווח סמך ביחידה 10 להפרש תוחלות כשהשונויות ידועות ראינו שהתפלגות הדגימה היא:

תוחלת: $\mu_1 - \mu_2$

שונות: $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

לכן בבדיקת השערות על הפרש תוחלות כאשר השונויות ידועות יהיה הסטטיסטי:

$$Z_{\bar{x}_1 - \bar{x}_2} = \frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

מודל 3 שונויות ידועות - בדיקת השערות על הפרש תוחלת האוכלוסייה כאשר השונות

ידועה

הנחות:

$x_2 \sim N(\mu_2, \sigma_2^2)$ ו- $x_1 \sim N(\mu_1, \sigma_1^2)$

או לפי מ.ג.מ כאשר $n_1 > 30$ ו- $n_2 > 30$.

המדגמים ב"ת

הסטטיסטי במדגם: $Z_{\bar{x}_1 - \bar{x}_2} = \frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$

אם השונויות אינן ידועות ומדובר במדגמים גדולים מספיק נציב את s_1, s_2 במקום סטיות התקן של האוכלוסיות.



שאלה 1

בחוות כבשים הגיעה עונת הגז והעובדים עושים תחרות בין שתי קבוצות: בקבוצה א' העובדים הוותיקים ובקבוצה ב' העובדים החדשים שזה להם הגז הראשון. בנתונים שנאספו נמצא כי קבוצה א' גזמה 120 ק"ג צמר ב-40 ימי עבודה שנדגמו מקרית בעוד שקבוצה ב' גזמה 70 ק"ג צמר ב-35 ימי עבודה שנדגמו מקרית.

ידוע כי סטיית התקן בקרב העובדים הוותיקים היא 1.5 ובקרב העובדים החדשים היא 2.5.

א. האם לאור התוצאות ניתן לומר, ברמת מובהקות 0.05 כי ממוצע ק"ג צמר ליממה שנאסף בקבוצה א' גבוה מממוצע ק"ג צמר ליממה שנאסף בקבוצה ב'?

ב. מהי מובהקות התוצאה, כלומר, מהי רמת המובהקות המינימלית עברה יוחלט כי ממוצע ק"ג צמר ליממה שנאסף בקבוצה א' אכן גבוה מממוצע ק"ג צמר ליממה שנאסף בקבוצה ב'?

שאלה 1

המרכז הארצי לבחינות החליט לבדוק כמה סוגיות בקשר לתוחלת הציונים בבחינה הפסיכומטרית. ידוע כי

הציונים בבחינה הפסיכומטרית מתפלגים נורמלית עם סטיית תקן 100, ללא קשר למועד אליו ניגשים

הנבחנים. נלקח מדגם של 20 תלמידי תיכון שנגשו לבחינה הפסיכומטרית והתקבל בו ציון ממוצע בבחינה של

540. בנוסף, נלקח מדגם של 25 נבחנים שנגשו לבחינה הפסיכומטרית לאחר הצבא והתקבל בו ציון ממוצע

בבחינה של 572.

א. האם ניתן להסיק, ברמת מובהקות 0.05, כי עדיף להיבחן בבחינה הפסיכומטרית לאחר הצבא מאשר בתיכון? רשמו את המבחן בו אתם משתמשים ונמקו תשובתכם.

ב. ידוע כי תוחלת הציון בבחינה הפסיכומטרית לאורך השנים היא 535.

1. האם ניתן לומר, ברמת מובהקות 0.05, כי הציונים של הנבחנים בבחינה לאחר הצבא שונים

מהממוצע הכללי? רשמו את המבחן בו אתם משתמשים ונמקו תשובתכם.

2. אם תוחלת הציון בבחינה הפסיכומטרית, של הנבחנים בבחינה לאחר הצבא, היא באמת 560, מה

הסתברות שהמבחן בו השתמשו ב-1 יגלה זאת? איך נקראת הסתברות זו?



שאלה 3

ידוע כי סטיית התקן של מספר שעות העבודה בשבוע של נשים עובדות היא 4 שעות, במדגם מקרי של 225

נשים נמצא ממוצע 36.4 שעות עבודה.

במדגם מקרי נוסף של 100 גברים נמצא כי ממוצע שעות העבודה שלהם בשבוע הוא 38 שעות. ידוע כי סטית

התקן של מספר שעות העבודה השבועיות בקרב הגברים היא 5 שעות.

א. האם לאור התוצאות ניתן לומר ברמת מובהקות 0.01 שנשים עובדות בממוצע פחות שעות? הסבר ונמק.

ב. בהמשך לסעיף א, חשב את עצמת המבחן אם למעשה ממוצע שעות העבודה בשבוע של גברים גבוה ב-3 שעות ממוצע שעות העבודה של נשים.

מודל 4 – בדיקת השערות: הפרש תוחלות כאשר השונויות אינן ידועות

כאשר חישבנו רווח סמך ביחידה 10 להפרש תוחלות כאשר סטיות התקן לא ידועות אך n_1 וגם n_2 גדולים דיים,

הצבנו במקום סטיות התקן את האומדנים שלהן: $\hat{s}_1, \hat{s}_2$, והשתמשנו בהתפלגות z.

ביחידה 10 לא דנו במקרה בו המדגמים קטנים (n_1, n_2) והשונויות לא ידועות.

במקרה זה באה לעזרתנו התפלגות t.

אנו צריכים להניח 3 הנחות עיקריות:

✓ הנחות מתפלגים נורמלית

✓ המדגמים בלתי תלויים

✓ השונויות שאינן ידועות שוות זו לזו



שלב 1: איסוף נתונים

שלב 2: השערות

$$H_0: \mu_1 - \mu_2 \geq c$$

$$H_1: \mu_1 - \mu_2 < c$$

$$H_0: \mu_1 - \mu_2 \leq c$$

$$H_1: \mu_1 - \mu_2 > c$$

$$H_0: \mu_1 - \mu_2 = c$$

$$H_1: \mu_1 - \mu_2 \neq c$$

$$X_2 \sim N(\mu_2, \sigma_2^2) \text{ וכן } X_1 \sim N(\mu_1, \sigma_1^2)$$

$$\sigma_1^2 = \sigma_2^2 = \sigma^2$$

X_1 ו- X_2 בלתי תלויים

שלב 3: הנחות

והתפלגות הדגימה

$$\hat{S}^2 = \frac{(n_1 - 1)\hat{S}_1^2 + (n_2 - 1)\hat{S}_2^2}{n_1 + n_2 - 2}$$

$$\mu_1 - \mu_2 = c \text{ בהנחה}$$

עבור ת-ים גדולים:

$$\frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{\frac{\hat{S}^2}{n_1} + \frac{\hat{S}^2}{n_2}}} \sim N(0,1)$$

עבור ת-ים קטנים:

$$\frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{\frac{\hat{S}^2}{n_1} + \frac{\hat{S}^2}{n_2}}} \sim t(n_1 + n_2 - 2)$$

שלב 4: שרטוט וקביעת אזורי קבלה ודחייה של H_0

$$H_0: \mu_1 - \mu_2 \geq c$$

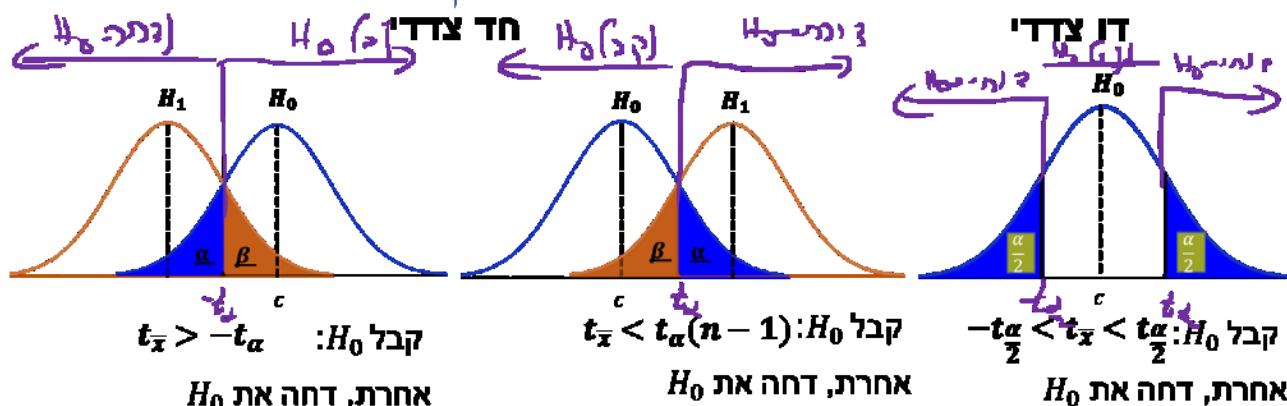
$$H_1: \mu_1 - \mu_2 < c$$

$$H_0: \mu_1 - \mu_2 \leq c$$

$$H_1: \mu_1 - \mu_2 > c$$

$$H_0: \mu_1 - \mu_2 = c$$

$$H_1: \mu_1 - \mu_2 \neq c$$



$$t_{\bar{x}_1 - \bar{x}_2} = \frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{\frac{\hat{S}^2}{n_1} + \frac{\hat{S}^2}{n_2}}}$$

שלב 5: חישוב סטטיסטי המבחן וההכרעה

הנחות:

$$X_2 \sim N(\mu_2, \sigma^2_2) \text{ ו- } X_1 \sim N(\mu_1, \sigma^2_1) \quad \checkmark$$

או לפי מ.ג.מ כאשר $n_1 > 30$ ו- $n_2 > 30$.

המדגמים ב"ת $\checkmark$

שונויות אינן ידועות אך שוות זו לזו $\checkmark$

$$\hat{S}^2 = \frac{(n_1-1) \cdot \hat{S}_1^2 + (n_2-1) \cdot \hat{S}_2^2}{n_1 + n_2 - 2} \quad \text{שלב ראשון-נחשב "שונויות משוקללות" לפי הנוסחה:}$$

$$S^2 = \frac{(n_1-1) \hat{S}_1^2 + (n_2-1) \hat{S}_2^2}{n_1 + n_2 - 2} = \frac{\sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2 + \sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2}{n_1 + n_2 - 2}$$

התפלגות t עם $n_1 + n_2 - 2$ דרגות חופש

$$t_{\bar{x}_1 - \bar{x}_2} = \frac{(\bar{x}_1 - \bar{x}_2) - c}{\sqrt{S^2/n_1 + S^2/n_2}} \quad \text{שלב שני - בדיקת השערות רגילה, הסטטיסטי במדגם:}$$



שאלה 1

בשתי שכונות נמדדו רמות הרעש [דציבלים] (להלן X) ביום העצמאות במספר מוקדים:

מוקד	1	2	3	4	5	6
שכונה א	100	80	115	95	112	
שכונה ב	110	117	102	109	98	103

א. בדוק את ההשערה שתוחלת X - בשכונה א שונה מ 116 ברמת מובהקות 0.05. נמק

ב. בדוק את ההשערה שתוחלת X בשכונה א שונה מתוחלת X בשכונה ב ברמת מובהקות 0.05. נמק.

בכל סעיף, במידה ודרושות הנחות כלשהן, רשום מהן.

שאלה 1

באוניברסיטה מסוימת קיימת הטענה כי הציון הממוצע בבחינות בסטטיסטיקה ללא חומר עזר אינו שונה מהציון הממוצע בבחינות עם חומר עזר. לבדיקת הטענה ניתנה בחינה בסטטיסטיקה ללא חומר עזר ל-16 סטודנטים

שנבחרו באופן מקרי מבין תלמידי הקורס. הציון הממוצע בבחינה היה 76 עם טעות תקן $\hat{S}_1 = 13$. כמו כן

נבחנו 15 סטודנטים אחרים שנבחרו באופן מקרי, בבחינה דומה (הבחינות היו חסויות) כאשר בבחינה ניתן

היה להיעזר בחומר עזר. הציון הממוצע בבחינה היה 73 עם טעות תקן $\hat{S}_2 = 9$.

נסח את ההשערות ורשום את ההנחות הדרושות.

בדוק את ההנחה בדבר שוויון השונות ברמת מובהקות 0.1.

האם התוצאות מאוששות את טענת האוניברסיטה ברמת מובהקות 0.05?



שאלה 3

לבדיקת הטענה כי הציון הממוצע במבחני לשכת עו"ד של סטודנטים למשפטים שלמדו באוניברסיטאות גבוה

מהציון המוצע של סטודנטים למשפטים שלמדו במכללות נבדקו ציונים של שתי קבוצות סטודנטים והתקבלו

התוצאות המוצגות בטבלה שלהלן:

מקום הלימוד	מספר הסטודנטים	ציון ממוצע	אומדן לשונות ($\hat{S}^2$)
אוניברסיטה	38	81	410
מכללה	27	72	500

בדוק את הטענה ברמת מובהקות 0.05? נמק.

בשנים הקודמות הציון הממוצע במבחני לשכת עו"ד בקרב בוגרי מכללות היה 65. האם ניתן לומר על סמך התוצאות הנ"ל ברמת מובהקות 0.05 שחלה עלייה בציונים בקרב בוגרי מכללות? נמק.



שאלה 4

לבדיקת הטענה כי הציון הממוצע במבחני לשכת עו"ד של סטודנטים למשפטים שלמדו באוניברסיטאות גבוה מהציון הממוצע של סטודנטים למשפטים שלמדו במכללות נבדקו ציונים של שתי קבוצות סטודנטים והתקבלו התוצאות המוצגות בטבלה שלהלן:

מקום הלימוד	מספר הסטודנטים	ציון ממוצע	אומדן לשונות ($\hat{s}^2$)
אוניברסיטה	38	81	410
מכללה	27	72	500

א. בדוק את הטענה ברמת מובהקות 0.05? נמק.

ב. על סמך התוצאות חשב רווח סמך ברמת סמך 98% לציון הממוצע במבחני לשכת עו"ד בקרב סטודנטים שסיימו לימודיהם באוניברסיטאות.

ג. בשנים הקודמות הציון הממוצע במבחני לשכת עו"ד בקרב בוגרי מכללות היה 65. האם ניתן לומר על סמך התוצאות הנ"ל ברמת מובהקות 0.05 שחלה עלייה בציונים בקרב בוגרי מכללות? נמק.



שאלה 5

לבדיקת הטענה כי הציון הממוצע במבחני לשכת עו"ד של סטודנטים למשפטים שלמדו באוניברסיטאות גבוה מהציון הממוצע של סטודנטים למשפטים שלמדו במכללות נבדקו ציונים של שתי קבוצות סטודנטים והתקבלו התוצאות המוצגות בטבלה שלהלן:

מקום הלימוד	מספר הסטודנטים	ציון ממוצע	אומדן לשונות ($\hat{s}^2$)
אוניברסיטה	38	81	410
מכללה	27	72	500

א. בדוק את הטענה ברמת מובהקות 0.05? נמק.

ב. על סמך התוצאות חשב רווח סמך ברמת סמך 98% לסטיית התקן של ציון במבחני לשכת עו"ד בקרב סטודנטים שסיימו לימודיהם באוניברסיטאות.

ג. בשנים הקודמות הציון הממוצע במבחני לשכת עו"ד בקרב בוגרי מכללות היה 65. האם ניתן לומר על סמך התוצאות הנ"ל ברמת מובהקות 0.05 שחלה עלייה בציונים בקרב בוגרי מכללות? נמק.



שאלה 6

לבדיקת הטענה כי הציון הממוצע במבחני לשכת עו"ד של סטודנטים למשפטים שלמדו באוניברסיטאות גבוה מהציון המוצע של סטודנטים למשפטים שלמדו במכללות נבדקו ציונים של שתי קבוצות סטודנטים והתקבלו התוצאות המוצגות בטבלה שלהלן:

מקום הלימוד	מספר הסטודנטים	ציון ממוצע	אומדן לשונות ($\hat{s}^2$)
אוניברסיטה	38	81	410
מכללה	27	72	500

א. בדוק את הטענה ברמת מובהקות 0.05? נמק.

ב. על סמך התוצאות חשב רווח סמך ברמת סמך 98% לסטיית התקן של ציון במבחני לשכת עו"ד בקרב סטודנטים שסיימו לימודיהם באוניברסיטאות.



שאלה 7

בשני מדרונות נדגם עובי הקרקע (Cm) (להלן X):

Thicknes	1	2	3	4	5	6	7	8
מדרון א	157	161	155	166	145	150	157	153
מדרון ב	143	150	149	157	156	146	159	

א. חשב רווח סמך לתוחלת X על סמך המדגם ממדרון א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת X על סמך המדגם ממדרון ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק את ההשערה שתוחלת X במדרון א שונה מ-160 ברמת מובהקות 0.05. נמק.

ה. בדוק את ההשערה שתוחלת X במדרון א שונה מתוחלת X במדרון ב ברמת מובהקות 0.05. נמק.

ו. בדוק את ההשערה ששונות X במדרון א שונה משונות X במדרון ב ברמת מובהקות 0.1, בהנחה שא מתפלג נורמלית. נמק.

ז. האם דרושה הנחה כלשהי לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.



שאלה 8

בשני רשתות חברתיות נדגם מספר החברים בקרב המשתמשים (להלן X):

Obs	1	2	3	4	5	6	7	8
רשת א	157	161	155	166	145	150	157	153
רשת ב	143	150	149	157	156	146	159	

א. חשב רווח סמך לתוחלת X על סמך המדגם מרשת א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת X על סמך המדגם מרשת ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק את ההשערה שתוחלת X ברשת א שונה מ-160 ברמת מובהקות 0.05. נמק.

ה. בדוק את ההשערה שתוחלת X ברשת א שונה מתוחלת X ברשת ב ברמת מובהקות 0.05. נמק.

ו. בדוק את ההשערה ששונות X ברשת א שונה משונות X ברשת ב ברמת מובהקות 0.1, בהנחה שא מתפלג נורמלית. נמק.

ז. האם דרושה הנחה כלשהי לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.



שאלה 9

טיפול חדש בחולים לאחר ניתוח מושווה לטיפול סטנדרטי. נלקח מדגם מקרי של 12 חולים, שחולקו באקראי ל-2 קבוצות ונמדד הזמן בימים עד להחלמתו של כל אחד ואחד מהם. להלן הנתונים:

טיפול סטנדרטי	24	48	40	36	32	30
טיפול חדש	25	37	34	29	28	26

בהנחה שהזמן להחלמה מתפלג נורמלית:

א. בדוק האם יש הבדל בין 2 הטיפולים ברמת מובהקות 0.05. נסח את ההשערות ופרט את ההנחות הדרושות.

ב. בהמשך לסעיף א', מה תהיה המסקנה ברמת מובהקות 0.05 אם האלטרנטיבה היא שהטיפול החדש עדיף? נמק.

ג. חזור לסעיף א' אם לא ניתן להניח שזמן ההחלמה מתפלג נורמלית.

שאלה 1

הליגה לשוויון זכויות האישה טוענת כי ההכנסה של נשים נמוכה מזו של גברים העובדים באותו מקצוע. לשם בדיקת הטענה נלקח מדגם של 8 נשים ו- 5 גברים העובדים במחלקה אחת באותו מפעל ונרשמו ההכנסות החודשיות (בדולרים) :

משכורת נשים	900	700	1000	925	1220	800	950	940
משכורת גברים	1350	870	1150	930	1400			

מהי מסקנתך ברמת מובהקות 2.5% ?

נסח את ההשערות, רשום את ההנחות הדרושות לביצוע המבחן הסטטיסטי המתאים .

openbook
המרכז לקידום אקדמי

מודל 5: בדיקת השערות על הפרש תוחלות/ממוצעים לפי שני במדגמים מזווגים



מדגמים תלויים/מזווגים

מדגמים מזווגים בנויים מזוגות של תצפיות כאשר לכל תצפית במדגם א' יש בן זוג מתואם במדגם ב'.

✓ המדגמים זהים בגודלם ומורכבים מזוגות של תצפיות כאשר בין זוגות התצפיות קיימת אי תלות

✓ את הזיווג ניתן ליצור ע"י:

✓ לפני אחרי טיפול מסוים

✓ חלוקה לפי קריטריון רלוונטי לניסוי. בחירת זוגות עם משתני רקע זהים.

במדגמים מזווגים נגדיר משתנה חדש $\bar{D}$ שיהיה ההפרש בין זוגות התצפיות.

במבחנים מזווגים דנים במדגם אחד של n זוגות, בכל זוג נחשב את הפרש התצפיות בתוך הזוג וכך נקבל מדגם של n הפרשים.

בדיקת השערות על הפרש תוחלות מדגמים מזווגים

המבחן מבוסס על מעבר משני מדגמים מזווגים למדגם אחד של הפרשים.

בודקים השערה על תוחלת ההפרשים $\mu_D = \mu_X - \mu_Y$

נוסחאות עזר לחישוב הסטטיסטיים:

ממוצע ההפרשים: $\bar{d} = \frac{\sum_{i=1}^n d_i}{n}$

שונות ההפרשים: $\hat{S}_d^2 = \frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1} = \frac{\sum_{i=1}^n d_i^2 - n \cdot \bar{d}^2}{n-1}$

הנחות:

✓ $x_1, x_2 \sim N(\mu, \sigma^2)$ המדגמים באים מאוכלוסייה המתפלגת נורמלית או בקירוב נורמלית (לפי מ.ג.מ.

(כאשר $n > 30$)

✓ המדגמים תלויים / מזווגים (לפני ואחרי)

כשהשונות ידועות: $D \sim N(\mu_D, \sigma^2)$

כשהשונות לא ידועות: $D \sim t\left(\mu_D = \bar{D}, \frac{\hat{S}_D^2}{n}\right)$

נחשב לכל זוג את ההפרש: $D_i = X_1 - X_2$ ונתייחס אליהם כאל מדגם אחד.

הסטטיסטי במדגם: $T_D = \frac{\bar{D} - c}{S_D / \sqrt{n}}$ התפלגות t עם (n-1) דרגות חופש

הנחות x_1, x_2 מזווגים, $D \sim N(\mu_D, \sigma^2)$, **סטטיסטי במדגם:** $T_D = \frac{\bar{D} - c}{S_D / \sqrt{n}}$ נחשב הפרשים $D_i = X_1 - X_2$

ונתייחס אליהם כאל מדגם אחד. התפלגות t עם (n-1) דרגות חופש.



שאלה 1

להלן נתונים על מספר השעות שהקדישו עשרה ילדים בגיל 14 במשך שבוע להכנת שעורים ולצפייה בטלוויזיה:

12	15	16	19	13.5	10	14.5	16.5	15	13	הכנת שעורים
16	17	16	18	14	13.5	20	15	14	17.5	צפייה בטלוויזיה

האם ניתן לומר כי ילדים מקדישים בממוצע יותר זמן לצפייה בטלוויזיה מאשר להכנת שעורים, בהנחה שמספר השעות מתפלג נורמלית? נסח את ההשערות ובדוק ברמת מובהקות 0.05.



שאלה 2

קיימת טענה שאנשים המפסיקים לעשן עולים בממוצע יותר מ-4 ק"ג. לצורך בדיקת הטענה נלקח מדגם מקרי של 10 גברים, שנשקלו פעם אחת לפני שהפסיקו לעשן, ופעם שנייה שנה לאחר הפסקת העישון. להלן התוצאות שנתקבלו:

80	84	74	77	76	68	82	78	75	62	משקל לפני הפסקת העישון
86	89	80	82	80	75	80	84	84	70	משקל אחרי הפסקת העישון

האם ניתן לקבוע שהטענה נכונה?

נסח את ההשערות, רשום את ההנחות שהנך מניח לצורך הפתרון ובדוק ברמת מובהקות 0.05.



שאלה 3

הטבלה הבאה מציגה את מספר שעות השינה לאחר קבלת תרופה מסוימת, ומספר שעות השינה לפני קבלתה, של 9 חולים שקבלו תרופה מסוימת.

8	8	4.5	5	7	6.5	6.5	5	7	לפני קבלת התרופה
9	7	8.5	8	7	8	6	7	8	אחרי קבלת התרופה

האם יש לתרופה השפעה חיובית על מספר שעות השינה, בהנחה שמספר שעות השינה מתפלג נורמלית? נסח את ההשערות, הסבר מהו המבחן הסטטיסטי המתאים ובדוק ברמת מובהקות 0.05.



שאלה 4

במחקר מסוים נבדק יבול הקפה [מספר שקים בחודש] ב-7 חוות קטנות לפני ואחרי שימוש בשיטת טיפול חדשנית נגד פטרייה מזיקה:

חוה	1	2	3	4	5	6	7
לפני	5	6	7	8	4	3	6
אחרי	10	9	9	8	13	8	9

ידוע שהיבול מתפלג נורמלית.

בדוק את ההשערה שתוחלת היבול לאחר הטיפול שונה מתוחלת היבול לפני הטיפול ברמת מובהקות 0.05. נמק



שאלה 5

במחקר מסוים נבדק ההבדל בין לחצי-האוויר (Bar) בצמיגים בקדמת כלי רכב ב-7 מכוניות נבדק לחץ האוויר בשני הצמיגים הקדמיים. להלן רמות הלחץ שנמדדו:

Pressure	1	2	3	4	5	6	7
ימין	2.5	2.6	2.7	2.8	2.4	2.3	3
שמאל	3	2.9	2.9	2.6	2.5	2.8	2.9

ידוע שלחץ האוויר מתפלג נורמלית.

בדוק את ההשערה שתוחלת לחץ האוויר בצד שמאל גבוהה יותר מתוחלת לחץ האוויר בצד ימין ברמת מובהקות 0.05. נמק



שאלה 6

במחקר מסוים נבדק ההבדל בין צריכות החשמל (Kwh) במשקי בית בימי שישי, בבוקר ואחה"צ.

להלן רמות הצריכה שנמדדו ב-7 משקי-בית:

7	6	5	4	3	2	1	Pressure
30	23	24	28	27	26	25	ימין
29	28	25	26	32	29	30	שמאל

ידוע שלחץ האוויר מתפלג נורמלית.

בדוק את ההשערה שתוחלת הצריכה אחת גבוהה יותר מתוחלת הצריכה בבוקר ברמת מובהקות 0.05. נמק



מודל 7: בדיקת השערות על הפרש פרופורציות

הנחות:

✓ $x_1 \sim N(\mu_1, \sigma_1^2)$ ו- $x_2 \sim N(\mu_2, \sigma_2^2)$ או לפי מ.ג.מ כאשר $n_1 > 30$ ו- $n_2 > 30$.

✓ המדגמים ב"ת

✓ $n_1 \hat{p}_1 \geq 10$, $n_1 \hat{q}_1 \geq 10$ ו- $n_2 \hat{p}_2 \geq 10$, $n_2 \hat{q}_2 \geq 10$

$$Z_{\hat{p}_1 - \hat{p}_2} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p} \cdot \hat{q}}{n_1} + \frac{\hat{p} \cdot \hat{q}}{n_2}}}$$

הסטטיסטי במדגם:

$$\hat{p} = \frac{\hat{p}_1 n_1 + \hat{p}_2 n_2}{n_1 + n_2}$$

כאשר:

שאלה 1

להלן תוצאות סקר שנעשה לבדיקת מספר בתי אב המנויים על כבלים באזורים שונים ברחבי הארץ:

אזור גודל המדגם מספר המנויים במדגם

ירושלים	162	124
תל אביב	250	155
חיפה	120	84
מחוץ לערים הגדולות	620	341

על סמך תוצאות הסקר ענה על השאלות הבאות:

א. בדוק ברמת מובהקות 0.05 את הטענה כי מחוץ לערים הגדולות, יותר ממחצית בתי האב מנויים על כבלים. נמק.

ב. האם ניתן לומר, ברמת מובהקות 0.05, כי אחוז המנויים בת"א שונה מאחוז המנויים בחיפה? נמק.

ג. בדוק ברמת מובהקות 0.05 את הטענה כי פרופורציית בתי האב המנויים על כבלים בירושלים נמוכה מ-0.8. נמק

שאלה 2

בסקר דעת קהל על תכנית חדשה בטלוויזיה נמצאו התוצאות הבאות :

	צופים	לא צופים
נשים	38	162
גברים	108	192

- א. בדוק את ההשערה שלמין אין כל השפעה על הצפייה בתכנית ברמת מובהקות 0.05? נמק.
- ב. האם לאור התוצאות ניתן לומר, ברמת מובהקות 5%, שפרופורציית הנשים הצופות בתכנית נמוכה מ- 20%? נמק.

שאלה 3

- יצרן משקאות קלים מעוניין לבדוק את השפעתו של מבצע פרסום שמטרתו להחדיר משקה חדש לשוק. שני מדגמים מקריים נבחרו לפני ואחרי מבצע הפרסום ובו נשאלו צרכנים האם הם שמעו על המשקה החדש. התוצאות הראו שלפני המבצע שמעו על המשקה 450 מתוך 1200 צרכנים ואחריו 700 מתוך 1250.
- א. מצא רווח סמך ברמת סמך 95% להפרש הפרופורציות של הצרכנים שמכירים את המשקה לפני מבצע הפרסום ולאחריו.
- ב. משרד הפרסום טוען כי המבצע היה אפקטיבי בהעלאת המודעות למשקה החדש. האם הטענה מוצדקת ברמת מובהקות 0.01? נמק.
- ג. מצא רווח סמך ברמת סמך 90% לפרופורציית הצרכנים שמכירים את המשקה אחרי המבצע.

שאלה 4

- על סמך נתוני הלשכה המרכזית לסטטיסטיקה ידוע כי 90% מבוגרי מדעי המחשב מתקבלים לעבודה מיד עם סיום לימודיהם.
- חוקר במשרד הכלכלה טוען כי כיום האחוז נמוך יותר. לצורך בדיקת הטענה נבדק מדגם מקרי של 150 בוגרים ונמצא כי מתוכם 132 התקבלו לעבודה מיד עם סיום לימודיהם.
- א. מהי מסקנת החוקר ברמת מובהקות 0.05? נמק.
- ב. מהי הסתברות לטעות האפשרית במסקנה מסעיף א אם למעשה רק 85% מהבוגרים מתקבלים לעבודה מיד עם סיום לימודיהם?
- ג. במדגם של 100 בוגרי הנדסת תוכנה נמצא כי 70 התקבלו לעבודה מיד עם סיום לימודיהם. האם ניתן לומר, ברמת מובהקות 0.05, כי אחוז הבוגרים המתקבלים לעבודה מיד עם סיום לימודיהם בקרב בוגרי מדעי המחשב גבוה מאחוז הבוגרים בקרב בוגרי הנדסת תוכנה? נמק.

שאלה 1

חוקר בדק את הקשר בין עישון ללחץ דם גבוה בעזרת מדגם מקרי של 800 אנשים. במדגם היו 300 מעשנים ומתוכם 8% סובלים מלחץ דם גבוה, וכן 500 איש שאינם מעשנים ומתוכם 4% סובלים מלחץ דם גבוה. א. בדוק ברמת מובהקות של 0.01, את ההשערה שבקרב מעשנים אחוז הסובלים מלחץ דם גבוה גדול יותר מאשר בקרב אלה שאינם מעשנים.

ב. אמוד את אחוז הסובלים מלחץ דם גבוה באוכלוסייה כולה, ברמת בטחון של 95%.

ג. האם תוצאות המדגם מצדיקות את הטענה כי אחוז הסובלים מלחץ דם גבוה באוכלוסייה

כולה, הוא גבוה מ-5%? נסח את ההשערות ובדוק ברמת מובהקות 0.01.

שאלה 1

במדגם מקרי של 400 תושבי גוש דן נמצא של-140 מהם יש מכשיר DVD.

במדגם מקרי של 500 תושבים משאר חלקי הארץ נמצא כי של-150 מהם יש מכשיר DVD.

א. האם ניתן לומר שפרופורצית בעלי DVD בגוש דן גבוה מפרופורצית בעלי DVD בשאר חלקי הארץ? נסח את ההשערות והסק עבור רמת מובהקות 0.05.

ב. בדוק את ההשערה שפרופורצית בעלי DVD בכל רחבי הארץ שונה מ-0.3 ברמת מובהקות 0.05.

openbook
המרכז לקידום אקדמי

מודל 8: בדיקת השערות על שונות אחת

הנחות:

$x \sim N(\mu, \sigma^2)$ או לפי מ.ג.מ. כאשר $n > 30$.

$$\chi^2 = \frac{(n-1)S^2}{c} = \frac{nS^2}{c} \quad \text{הסטטיסטי במדגם:}$$

התפלגות χ^2 עם $(n-1)$ דרגות חופש.

שאלה 1

חוקר טוען שבידיו שיטה חדשה המשפרת את הצלחת כל התלמידים אך לא באופן שווה. לצורך בדיקת העניין דגם 30 תלמידים מקרית ומצא כי ממוצע ציוניהם הוא 124 עם סטיית תקן 9. בהנחה שהציונים מתפלגים נורמלית באוכלוסייה עם ממוצע 120 ושונות 100. האם ניתן לטעון שהשיטה החדשה אכן שמפרת את שונות הציונים? בדקו בר"מ 0.05.

שאלה 2

יצרן צינורות טוען שקוטר הצינורות שהוא מייצר מתפלג נורמלית עם שונות של 1 ס"מ. בבדיקה שנערכה נבדקו 15 צינורות והשונות במדגם - S^2 , הייתה 2.25. האם על סמך מדגם זה ניתן לשלול את טענת היצרן ברמת מובהקות $\alpha = 0.02$. נמק.

שאלה 3

חברת יעוץ להשקעות במניות מייצעת למשקיע מסוים לא להסתכן, ולדחות כל השקעה שסטיית התקן של שיעור התשואה הצפוי ממנה עולה על 5%. שיעורי התשואה הכוללים של מניות חברת STOCK היו בתשע השנים האחרונות כדלקמן (באחוזים): 25, 35, 45, 10, 40, 50, 20, 35, 35. בהנחה ששיעורי התשואה מתפלגים נורמלית:

- בדוק בר"מ של 0.05 האם ישקיע במניות של חברה זו?
- משקיע שמוכן לקחת סיכון גדול יותר, מוכן לקבל כל השקעה שס"ת של שיעור התשואה הצפוי ממנה לא עולה על 10%, האם משקיע זה ישקיע ב"סטוק" באותה רמת מובהקות?

שאלה 4

סטיית התקן של השכר בשירות הציבורי באוכלוסייה היא 830 ₪. הליגה לשוויון זכויות האישה טוענת כי סטיית התקן של השכר בקרב נשים העובדות בשירות הציבורי נמוכה מסטיית התקן באוכלוסייה.

לשם בדיקת הטענה, נלקח מדגם של 20 נשים העובדות בשירות הציבורי והתקבלו התוצאות הבאות:

$$\sum_{i=1}^{20} (x_i - \bar{x})^2 = 6,475,626$$

האם הטענה מוצדקת ברמת מובהקות 0.05 ? נמק



מודל 9: הסקה על שתי שונויות של אוכלוסיות ✓

כאשר משווים שונות באוכלוסייה לנתון של מדגם, זוהי בדיקת השערות לשונות אחת ונשתמש בהתפלגות χ^2 . לעומת זאת, כאשר נתונים 2 מדגמים בלתי תלויים ומעוניינים להשוות בין נתוני 2 המדגמים, אזי משווים בין 2 אוכלוסיות ולכן זוהי בדיקת השערות ל-2 שונויות ומשתמשים בהתפלגות F.

מבחן F לשוויון שונויות

במבחן F אנו שואלים: אם שני מדגמים אכן לקוחים מהתפלגויות בעלות שונויות שוות, אנו נצפה שההבדלים שבין אומדי השונויות של שני המדגמים יהיו קטנים. ההבדל שביניהם נובע מטעות מקרית. אם ההבדל שביניהם גדול דיו, אנו נחשוש שהמדגמים לא לקוחים מהתפלגויות בעלות שונויות שוות ואז לא נוכל להניח שוויון שונויות במבחן t למדגמים בלתי תלויים.

הנחות: דגימה מקרית, התפלגות הדגימה F, x_1, x_2 בלתי תלויים, מתפלגים נורמלית.

$$\text{השערות: } H_1: \frac{\sigma_1^2}{\sigma_2^2} \neq 1, H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$$

הסטטיסטי במדגם: $F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$ יחושב כיחס שבין האומדן הגדול חלקי האומדן הקטן מבין השניים.

התפלגות F - התפלגות זו לא בודקת הפרש, אלא בודקת יחס בין שונויות ולכן הסטטיסטי שמתפלג התפלגות F הינו: $F \sim \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$

התפלגות F עבור השערה דו צדדית

$$\text{ערך קריטי ימני } F_{\frac{\alpha}{2}}(n_1 - 1, n_2 - 1) \\ \text{ערך קריטי שמאלי } \frac{1}{F_{\frac{\alpha}{2}}(n_2 - 1, n_1 - 1)}$$

התפלגות F היא כמו התפלגות χ^2 , זוהי התפלגות אסימטרית שבה כל הערכים בהכרח חיוביים, המשתנה F בנוי מיחס של שני משתנים בלתי תלויים, שלכל אחד מהם התפלגות χ^2 עם דרגות חופש מתאימות. התפלגות F תלויה בשני פרמטרים: מספר דרגות החופש של הביטוי במונה ומספר דרגות החופש של הביטוי במכנה:

דרגת חופש במונה $V_1 = (n_1 - 1)$ שורה - תלוי המספר התצפיות במדגם אחד

דרגת חופש במכנה $V_2 = (n_2 - 1)$ טור - תלוי המספר התצפיות במדגם השני

תחת הנחת H_0 מניחים שאין הבדל בין השונויות ולכן מרכז ההתפלגות $H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$ ולכן הערכים הקריטיים העליונים בהתפלגות F גדולים מ-1 והתחתונים נמצאים בין 0 ל-1.

בהשערה דו צדדית, כמו ב- χ^2 , יש להוציא שני ערכים קריטיים. בטבלה מופיע רק ערך קריטי עליון $F_{(v_1, v_2)}$

$$\text{והערך הקריטי התחתון } F_{(v_2, v_1)} = \frac{1}{F_{(v_1, v_2)}} \text{ מחושב לפי הנוסחה: } F_{1-\alpha/2} = \frac{1}{F_{\alpha/2}}$$



שאלה 1

חברה ליצור סלילי חוטים לתפירה טוענת כי אורך החוט בסליל המיוצר על ידיה עולה על זה של חברה מתחרה, אך השונות שווה. לשם בדיקת הטענה לגבי שוויון השונות נלקח מדגם מקרי של 7 סלילים מחברה א', ומדגם מיקרי של 5 סלילים מחברה ב' והתקבלו אורכי החוטים הבאים:

חברה א': 114 87 92 134 97 165 81

חברה ב': 92 109 98 86 102

א. בדוק את ההשערה שהשונות שוות בשתי החברות בר"מ 0.1. מהי מסקנתך?

ב. על מנת לקבוע תקינות מכונות התפירה, סטיית התקן של אורך חוט בסליל אינו צריך לעלות על 10 ס"מ בחברה ב'? בדוק תקינות המכונות ברמת מובהקות של 0.01.



שאלה 2

מתוך כל אחד משני עדרי פרות הוצא מדגם מקרי של 10 פרות ונמדדה תפוקת שומן החלב של כל פרה במשך חודש ימים.

הנח שתפוקת שומן החלב מפולגת נורמלית. להלן התוצאות שהתקבלו:

$\sum_{i=1}^{10} x_i^2 = 15717$	$\sum_{i=1}^{10} x_i = 365$	65	37	17	34	51	18	46	36	15	46	מדגם 1
$\sum_{i=1}^{10} y_i^2 = 17453$	$\sum_{i=1}^{10} y_i = 405$	31	40	36	49	36	40	29	47	32	65	מדגם 2

א. בדוק את ההשערה שאין הבדל בין שונות תפוקת שומן החלב של הפרות בשני העדרים על סמך תוצאות שני המדגמים ברמת מובהקות 0.1.

ב. בדוק את ההשערה שהתפוקה הממוצעת של עדר 2 גדולה מהתפוקה הממוצעת של עדר 1 ב-2 יחידות, לעומת האלטרנטיבה שההבדל גדול ביותר מ-2 יחידות, ברמת מובהקות 0.01. נמק.

ג. על סמך נתוני מדגם 1 בלבד, מצא רווח סמך לתוחלת תפוקת שומן החלב של פרות ברמת סמך 98%



שאלה 3

פעיל איכות הסביבה טוען כי תכשיר כביסה אקולוגי חדש מנקה טוב יותר מאבקת כביסה רגילה.

לצורך בדיקת טענתו נבדק מדגם מקרי של 12 בגדים מלוכלכים שחולקו לשתי קבוצות שוות.

6 בגדים כובסו בתכשיר האקולוגי החדש וה- 6 האחרים כובסו באבקת כביסה רגילה.

לאחר הכביסה נבדק כל בגד וניתן לו ציון על מידת הניקיון.

$\sum_{i=1}^n x_i^2$	$\sum_{i=1}^n x_i$								
32,774	438	89	86	74	70	60	59	תכשיר אקולוגי	
21,620	354	71	69	67	55	52	40	אבקת כביסה רגילה	

בהנחה כי ציון הניקיון מתפלג נורמלית:

א. בדוק את ההשערה ששונות הציונים בשתי הקבוצות שוות ברמת מובהקות 0.1.

ב. האם התוצאות מאששות את טענתו של פעיל איכות הסביבה ברמת מובהקות 0.05? נמק.

ג. מצא רווח סמך ברמת סמך 95% לתוחלת מידת הניקיון בכביסה עם תכשיר אקולוגי.

שאלה 1

כדי לבדוק האם קיימת השפעה של המתרגל על הצלחת סטודנטים בבחינה במבוא לכלכלה, חולקו סטודנטים הלומדים אצל מרצה מסוים לשתי קבוצות תרגול באופן מקרי. בטבלה שלהלן נרשמו תוצאותיהם בבחינת הגמר:

שם המתרגל	ציוני בחינת הגמר		
מר ידעני	77 89 67 45 88 95 68 72	$\sum_{i=1}^8 x_i = 601$	$\sum_{i=1}^8 x_i^2 = 46941$
ד"ר חכמי	83 72 95 35 50 76 74 92 88 65	$\sum_{i=1}^{10} y_i = 730$	$\sum_{i=1}^{10} y_i^2 = 56508$

האם יש הבדל בממוצע הציונים אצל שני המתרגלים, בהנחה שהציונים מתפלגים נורמלית? נסח את ההשערות, נמק והסבר מהו המבחן הסטטיסטי המתאים ובדוק ברמת מובהקות 0.05.

שאלה 1

חוקר טוען כי שונות הציונים בקורס סטטיסטיקה א' גדולה משונות הציונים בקורס סטטיסטיקה ב'. נבדקו הציונים במדגם מיקרי של 20 סטודנטים הלומדים סטטיסטיקה ב' ובמדגם מיקרי של 25 סטודנטים הלומדים סטטיסטיקה א' והתקבלו התוצאות הבאות:

Y- סטטיסטיקה ב': $\sum y = 1600$; $\sum y^2 = 128285$

X- סטטיסטיקה א': $\sum x = 1750$; $\sum x^2 = 123460$

- א. נסח השערות, רשום את ההנחות הדרושות ובדוק בר"מ של 0.05?
- ב. על סמך הנתונים, בדוק את ההשערה שיש הבדל בין השוניות בר"מ 0.1? נמק ללא חישוב.



תרגילים מסכמים



שאלה 1

בשני סוגי טלפונים נדגם זמן התגובה של המסך למגע (מילי-שניות):

Obs(Time)	1	2	3	4	5	6	7	8
טלפון א	68.2	72.4	71.5	69.6	68.4	67.8	66.7	68.3
טלפון ב	70.4	66.3	66.9	68.3	67.3	66.3	64.9	

א. חשב רווח סמך לתוחלת Time על סמך המדגם מטלפון א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת Time על סמך המדגם מטלפון ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק את ההשערה שתוחלת Time בטלפון א שונה מ-68, ברמת מובהקות 0.05:

1. באמצעות מבחן סטטיסטי מתאים.

2. באמצעות רווח הסמך שחושב בסעיף א. נמק.

ה. בדוק את ההשערה שתוחלת Time בטלפון א שונה מתוחלת Time בטלפון ב, ברמת מובהקות 0.05. נמק.

ו. בדוק את ההשערה ששוונות Time בטלפון א שונה משוונות Time בטלפון ב, ברמת מובהקות 0.1. נמק.

ז. האם דרושות הנחות כלשהן לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.



שאלה 2

נבדק ההבדל בין זמני תגובה לפרסומת קופצת (מיל-שניות) לפני ואחרי אימון. אצל 7 נערים:

Obs(Time)	1	2	3	4	5	6	7
לפני	300	294	323	261	252	284	299
אחרי	251	262	273	281	248	237	302

ידוע ש- Time מתפלג נורמלית.

ברמת מובהקות 0.05, בדוק את ההשערה שתוחלת Time לפני האימון גבוהה יותר מתוחלת Time אחרי

האימון. נמק



שאלה 3

במחקר מסוים נבדק ההבדל בין חוויית השימוש ב 7 סוגי טלפון.

7 נבדקים השתמשו פעם בטלפון מסוג א ופעם בטלפון מסוג ב. להלן רמות החווייה שדווחו:

Observation	1	2	3	4	5	6	7
טלפון א'	50	60	70	50	40	30	100
טלפון ב'	100	90	90	60	50	80	90

ידוע שחוויית השימוש מתפלגת נורמלית.

בדוק את ההשערה שחוויית השימוש בטלפון ב טובה יותר מחוויית השימוש בטלפון א ברמת מובהקות 0.05.
נמק



שאלה 4

חוקר בדק את ההשערה כי היקף הגולגולת הממוצע של יילודים הסובלים ממחלה מסוימת גדול מהיקף הגולגולת של יילודים נורמליים. ידוע כי היקף הגולגולת הממוצע של יילודים נורמליים הוא 34 ס"מ.

נבדק מדגם מקרי של 10 יילודים הסובלים מהמחלה, והתקבלו התוצאות הבאות של היקף הגולגולת שלהם:

42, 36, 35, 37, 33, 40, 38, 38, 32, 43

א. מהי מסקנת החוקר ברמת מובהקות 0.05? נסח את ההשערות ורשום את ההנחות הנדרשות.

ב. תן הערכה לרמת המובהקות הקטנה ביותר לפיה ניתן לקבל את השערת המחקר.

ג. מצא רווח סמך לסטיית התקן של היקף הגולגולת של יילודים הסובלים מהמחלה, על סמך נתוני המדגם וברמת סמך 95%.



שאלה 5

בשתי שכונות נדגם שיעור ההצבעה באזורים סטטיסטיים (אחוזים):

SA(Rate)	1	2	3	4	5	6	7	8
שכונה א	8.25	2.46	1.55	9.65	68.4	67.8	66.7	58.3
שכונה ב	70.4	66.3	66.9	68.3	67.3	66.3	64.9	

א. חשב רווח סמך לתוחלת Rate על סמך המדגם משכונה א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת Rate על סמך המדגם משכונה ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק את ההשערה שתוחלת Rate בשכונה א שונה מ- 65, ברמת מובהקות 0.05:

1. באמצעות מבחן סטטיסטי מתאים.

2. באמצעות רווח הסמך שחושב בסעיף א. נמק.

ה. בדוק את ההשערה שתוחלת Rate בשכונה א שונה מתוחלת Rate בשכונה ב, ברמת מובהקות 0.05. נמק.

ו. בדוק את ההשערה ששוונות Rate בשכונה א שונה משוונות Rate בשכונה ב, ברמת מובהקות 0.1. נמק.

ז. האם דרושות הנחות כלשהן לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.

שאלה 6

בשתי רשתות שיווק נדגם מחיר של מוצר בסניפים שונים (שקלים):

Net(Price)	1	2	3	4	5	6	7	8
רשת א	582	624	515	596	684	678	667	583
רשת ב	704	663	669	683	673	663	620	

א. חשב רווח סמך לתוחלת Price על סמך המדגם מרשת א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת Price על סמך המדגם מרשת ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק את ההשערה שתוחלת Price ברשת א שונה מ- 650, ברמת מובהקות 0.05:

1. באמצעות מבחן סטטיסטי מתאים.

2. באמצעות רווח הסמך שחושב בסעיף א. נמק.

ה. בדוק את ההשערה שתוחלת Price ברשת א שונה מתוחלת Price ברשת ב, ברמת מובהקות 0.05. נמק.

ו. בדוק את ההשערה ששוונות Price ברשת א שונה משוונות Price ברשת ב, ברמת מובהקות 0.1. נמק.

ז. האם דרושות הנחות כלשהן לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.



שאלה 7

על פי נתוני הלשכה המרכזית לסטטיסטיקה אחוז המעשנים בקרב בני נוער (גילאים 16-18) בארץ הוא 25%. במדגם מקרי של 80 בני נוער תושבי חיפה נמצאו 18 מעשנים. במדגם מקרי של 120 בני נוער תושבי ת"א נמצאו 33 מעשנים.

א. האם ניתן לומר, ברמת מובהקות 0.05, שבת"א פרופורציית בני הנוער המעשנים גדולה מהפרופורציה הארצית?

ב. מצאו רווח סמך ברמת סמך 95% לפרופורציית בני הנוער המעשנים בחיפה.

ג. האם ניתן לומר, ברמת מובהקות 0.05, שפרופורציית בני הנוער המעשנים בחיפה נמוכה מפרופורציית בני הנוער המעשנים בת"א?

ד. האם ניתן לומר, ברמת מובהקות 0.05, שבחיפה פרופורציית בני הנוער המעשנים שונה מהפרופורציה הארצית?



שאלה 8

אגודת הצרכנים ערכה בדיקה לקביעת צריכת הדלק של מכוניות בנפח 1,600 סמ"ק בנסיעות בינעירוניות. בבחירה מקרית של 9 מכוניות בנפח 1600 סמ"ק נבדקו מספר הק"מ לליטר שצרכה כל אחת מהמכוניות בנסיעה בינעירונית. התקבלו התוצאות הבאות (בק"מ לליטר):

$$\sum_{i=1}^9 x_i = 171 \quad ; \quad \sum_{i=1}^9 x_i^2 = 3271$$

בהנחה שמספר הק"מ לליטר מתפלג נורמלית:

א. מצא רווח סמך לסטיית התקן של מספר הק"מ לליטר ברמת סמך 95%.

ב. מצא רווח סמך ברמת סמך 95% עבור ממוצע הק"מ לליטר.



שאלה 9

יצרן צינורות טוען שקוטר הצינורות שהוא מייצר מתפלג נורמלית עם שונות של 1 ס"מ.

בבדיקה שנערכה נבדקו 15 צינורות והשונות במדגם - S^2 , הייתה 2.25.

האם על סמך מדגם זה ניתן לשלול את טענת היצרן ברמת מובהקות $\alpha = 0.02$. נמק.



שאלה 10

מתוך כל אחד משני עדרי פרות הוצא מדגם מקרי של 10 פרות ונמדדה תפוקת שומן החלב של כל פרה במשך חודש ימים.

הנח שתפוקת שומן החלב מפולגת נורמלית. להלן התוצאות שהתקבלו:

$\sum_{i=1}^{10} x_i^2 = 15717$	$\sum_{i=1}^{10} x_i = 365$	65	37	17	34	51	18	46	36	15	46	מדגם 1
$\sum_{i=1}^{10} y_i^2 = 17453$	$\sum_{i=1}^{10} y_i = 405$	31	40	36	49	36	40	29	47	32	65	מדגם 2

א. בדוק את ההשערה שאין הבדל בין שוניות תפוקת שומן החלב של הפרות בשני העדרים על סמך תוצאות שני המדגמים ברמת מובהקות 0.1.

ב. בדוק את ההשערה שהתפוקה הממוצעת של עדר 2 גדולה מהתפוקה הממוצעת של עדר 1 ב-2 יחידות, לעומת האלטרנטיבה שההבדל גדול ביותר מ-2 יחידות, ברמת מובהקות 0.01. נמק.

ג. על סמך נתוני מדגם 1 בלבד, מצא רווח סמך לתוחלת תפוקת שומן החלב של פרות ברמת סמך 98%



שאלה 11

בשני מושבים נדגם יבול התפוחים במטעים (טון):

Plant(Amount)	1	2	3	4	5	6	7	8
מושב א	58	62	51	59	68	67	66	58
מושב ב	70	66	66	68	67	66	62	

א. חשב רווח סמך לתוחלת Plant על סמך המדגם ממושב א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת Plant על סמך המדגם ממושב ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק את ההשערה שתוחלת Plant במושב א שונה מ-57, ברמת מובהקות 0.05:

1. באמצעות מבחן סטטיסטי מתאים.

2. באמצעות רווח הסמך שחושב בסעיף א. נמק.

- ה. בדוק את ההשערה שתוחלת Plant במושב א שונה מתוחלת Plant במושב ב, ברמת מובהקות 0.05. נמק.
- ו. בדוק את ההשערה ששונות Plant במושב א שונה משונות Plant במושב ב, ברמת מובהקות 0.1. נמק.
- ז. האם דרושות הנחות כלשהן לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.



שאלה 12

נבדק ההבדל בין ציונים בשני מועדי בחינה אצל 7 סטודנטים:

Student(Score)	1	2	3	4	5	6	7
מועד א	60	94	93	81	52	84	99
מועד ב	71	62	73	61	48	37	82

ידוע ש-Score מתפלג נורמלית.

ברמת מובהקות 0.05, בדוק את ההשערה שתוחלת Score במועד א נמוכה מתוחלת Score במועד ב.



שאלה 13

בשני כבישים נדגם זמן ההמתנה בפקקים (דקות):

Jam (Time)	1	2	3	4	5	6	7	8
כביש א	50	54	44	51	59	58	57	50
כביש ב	63	59	59	61	60	59	58	

א. חשב רווח סמך לתוחלת Time על סמך המדגם מכביש א ברמת סמך 95%.

ב. חשב רווח סמך לתוחלת Time על סמך המדגם מכביש ב ברמת סמך 95%.

ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק את ההשערה שתוחלת Time בכביש א שונה מ-50, ברמת מובהקות 0.05:

1. באמצעות מבחן סטטיסטי מתאים.

2. באמצעות רווח הסמך שחושב בסעיף א. נמק.

ה. בדוק את ההשערה שתוחלת Time בכביש א שונה מתוחלת Time בכביש ב, ברמת מובהקות 0.05. נמק.

ו. בדוק את ההשערה ששונות Time בכביש א שונה משונות Time בכביש ב, ברמת מובהקות 0.1. נמק.

ז. האם דרושות הנחות כלשהן לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.



שאלה 14

נבדק ההבדל בין ציונים בשני מועדי בחינה אצל 7 סטודנטים:

Student(Score)	1	2	3	4	5	6	7
מועד א	60	94	93	81	52	84	99
מועד ב	71	62	73	61	48	37	82

ידוע ש- Score מתפלג נורמלית.

ברמת מובהקות 0.05, בדוק את ההשערה שתוחלת Score במועד א גבוהה מתוחלת Score במועד ב.



שאלה 15

חוקר בדק את ההשערה כי היקף הגולגולת הממוצע של יילודים הסובלים ממחלה מסוימת גדול מהיקף הגולגולת של יילודים נורמליים. ידוע כי היקף הגולגולת הממוצע של יילודים נורמליים הוא 34 ס"מ.

נבדק מדגם מקרי של 10 יילודים הסובלים מהמחלה, והתקבלו התוצאות הבאות של היקף הגולגולת שלהם:

42, 36, 35, 37, 33, 40, 38, 38, 32, 43

תן הערכה לרמת המובהקות הקטנה ביותר לפיה ניתן לקבל את השערת המחקר.



שאלה 16

בעל סופרמרקט טוען, שסטיית התקן של המשקל של שקיות תפוזים הארוזים בשקיות של 10 ק"ג במוצק, גדולה מ- 0.15 ק"ג.

במדגם מקרי של 10 שקיות נתקבלו המשקלים הבאים (בק"ג):

10.2, 9.7, 10.1, 10.3, 10.1, 9.8, 9.9, 10.4, 10.3, 9.8

בהנחה שהתפלגות המשקל נורמלית, מהי המסקנה עבור רמת מובהקות 0.05? נמק.



שאלה 17

לשם השוואה בין שתי תרופות שונות, שנועדו לריפוי מחלה מסוימת, נלקח מדגם מקרי של עכברים שהודבקו במחלה.

לאחר שחלו, חולקו העכברים באופן מקרי, לשתי קבוצות ולכל קבוצה ניתנה תרופה שונה.

להלן סיכום הנתונים לגבי מספר ימי המחלה של כל אחד מהעכברים:

תרופה א – $n_1 = 10$, $\sum_{i=1}^{10} x_i = 70$, $\sum_{i=1}^{10} (x_i - \bar{x})^2 = 12$

תרופה ב - $n_2 = 11$, $\sum_{i=1}^{11} y_i = 110$, $\sum_{i=1}^{11} (y - \bar{y})^2 = 36$

א. האם קיים הבדל מובהק בין שונות שתי האוכלוסיות ברמת מובהקות 0.1?

ב. האם קיים הבדל מובהק בין שתי התרופות ברמת מובהקות 0.05?

ג. מצא רווח סמך לתוחלת מספר ימי המחלה בקרב העכברים שניתנה להם תרופה ב ברמת סמך 98%.



שאלה 18

הועלתה השערה כי הציון הממוצע בבחינות הבגרות באזור המרכז גבוה ביותר מ- 3 נקודות מהציון הממוצע בבחינות הבגרות בפריפריה.

לבדיקת ההשערה נבחרה באופן מקרי קבוצה בת 15 תלמידים מאזור המרכז ונמצא שממוצע הציונים שלהם בבחינות הבגרות היה 80

והאומדן חסר ההטיה לשונות היה 225 . כמו כן, לגבי 16 תלמידים שנבחרו באופן מקרי מהפריפריה ממוצע הציונים

בבחינות הבגרות היה 70 והאומדן חסר ההטיה לשונות היה 144.

א. האם התוצאות מאששות את ההשערה ברמת מובהקות 0.05 ? נמקו

ב. מצאו רווח סמך לסטיית התקן של הציונים באזור המרכז, ברמת סמך 95%.

ג. האם ניתן לומר שתוחלת הציון בבחינות הבגרות בפריפריה נמוך מ- 75?

נסחו את ההשערות והסק עבור רמת מובהקות 0.01.



שאלה 19

במפרט של יצרן נורות ליבון רשום כי זמן הבעירה של הנורות הוא 1250 שעות בממוצע עם סטיית תקן של 55 שעות.

המועצה לצרכנות קיבלה תלונות מצרכנים שזמן הבעירה של הנורות קצר יותר.

לצורך בדיקת התלונות בדקה המועצה לצרכנות מדגם מקרי של 121 נורות מתוצרת היצרן.

א. נסחו את ההשערות ורשמו עבור אילו תוצאות מדגם יוחלט, ברמת מובהקות 0.025, שתלונות הצרכנים מוצדקות. נמקו

ב. בהמשך למבחן מסעיף א, מה ההסתברות שהמועצה לצרכנות תצדיק את תלונות הצרכנים,

אם זמן הבעירה הממוצע האמיתי הוא 1230 שעות?

ג. זמן הבעירה הממוצע של הנורות שנבדקו היה 1242 שעות.

מהי רמת המובהקות המינימלית שעבורה המועצה לצרכנות תצדיק את תלונות הצרכנים?
 ד. בהמשך לסעיף ג, האם לאור תוצאות המדגם ניתן להסיק שתלונות הצרכנים מוצדקות ברמת מובהקות 0.02? נמקו.



שאלה 20

חוקר ביקש לבדוק את הטענה כי ההיעדרות מהעבודה שכיחה בקרב נשים יותר מאשר בקרב גברים. במדגם מקרי של 200 נשים מועסקות נמצא כי 32 נעדרו לפחות יום אחד בחודש המחקר, ואילו במדגם מקרי של 250 גברים מועסקים נעדרו 25 לפחות יום אחד באותו חודש.
 א. מהי מסקנת הבדיקה? נסח את ההשערות ובדוק ברמת מובהקות 0.05.
 ב. האם נכונה הטענה כי פרופורציית ההיעדרות של נשים עולה על 0.1?
 נסחו את ההשערות ובדקו על סמך נתוני המדגם שלעיל, ברמת מובהקות 0.02.
 ג. מצאו רווח סמך ברמת סמך 98% לפרופורציית המועסקים (גברים ונשים) הנעדרים לפחות יום אחד בחודש.



שאלה 21

בשני אזורים נדגם מרחק הנסיעה השבועי (ק"מ):

Region(Distance)	1	2	3	4	5	6	7	8
אזור א	450	486	396	459	531	522	513	450
אזור ב	567	531	531	549	540	531	522	

- חשב רווח סמך לתוחלת Distance על סמך המדגם מאזור א ברמת סמך 95%.
- חשב רווח סמך לתוחלת Distance על סמך המדגם מאזור ב ברמת סמך 95%.
- איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.
- בדוק ב-2 דרכים את ההשערה שתוחלת Distance באזור א שונה מ-510, ברמת מובהקות 0.05:
 - באמצעות מבחן סטטיסטי מתאים.
 - באמצעות רווח הסמך שחושב בסעיף א. נמק.
 - בדוק את ההשערה שתוחלת Distance באזור א שונה מתוחלת Distance באזור ב, ברמת מובהקות 0.05. נמק.

- ו. בדוק את ההשערה ששונות Distance באזור א שונה משונות Distance באזור ב, ברמת מובהקות 0.1. נמק.
- ז. האם דרושות הנחות כלשהן לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.



שאלה

חוקרי חלל רצו לבחון האם בני אדם מבצעים תפקוד מוטורי מסוים ביעילות גבוהה יותר בתנאי חוסר משקל מאשר בתנאים רגילים.

לשם כך הועברו מבחני ביצוע ל- 10 אנשים ששהו בתנאי חוסר משקל ול- 15 נבדקים אחרים ששהו בתנאים רגילים.

כל נבדק קיבל ציון והתוצאות שהתקבלו מוצגות בטבלה שלהלן :

תנאי הביצוע	מספר נבדקים	ממוצע המדגם	אומדן חסר הטיה לשונות
תנאי חוסר משקל	10	84.7	6.74
תנאים רגילים	15	82.1	5.66

א. האם התוצאות מעידות על רמת ביצוע גבוהה יותר בתנאי חוסר משקל לעומת תנאים רגילים, ברמת מובהקות 0.05 ? נמקו

ב. האם לאור התוצאות ניתן להסיק ששונות הציונים בקרב נבדקים בתנאי חוסר משקל גבוהה משונות הציונים בקרב נבדקים הנבחנים בתנאים רגילים, ברמת מובהקות 0.05 ? נמקו.

ג. האם לאור התוצאות ניתן להסיק כי ממוצע הציונים בתנאי חוסר משקל נמוך מ- 86 , ברמת מובהקות 0.01 ? נמקו.



שאלה

הועלתה הטענה כי שכרם של עובדי חינוך נמוך מהשכר הממוצע במשק.

לבדיקת הטענה נבדק מדגם מקרי של 100 עובדי חינוך ונמצא כי שכרם הממוצע 8,950 שקל בחודש.

על פי נתוני הלשכה המרכזית לסטטיסטיקה השכר הממוצע במשק כולו הוא 9,100 שקל וסטיית התקן של השכר היא 800 שקל.

א. מהי המסקנה ברמת מובהקות 0.05 ? נמקו.

ב. האם ניתן להגיע לאותה מסקנה, ללא חישוב נוסף, עבור רמת מובהקות 0.01 ? הסבירו.

ג. מהי רמת המובהקות המינימלית שעבורה השערת האפס תדחה?

ד. חשבו רווח סמך לשכר הממוצע של עובדי החינוך במשק כולו על סמך תוצאות המדגם ברמת סמך 99%.



שאלה

אורך ברגים הוא משתנה מקרי נורמלי עם סטיית תקן 0.05 ס"מ.
האורך הממוצע של הברגים μ אינו קבוע והוא תלוי באופן כיוון המכונה.
לצורך בדיקת הטענה שהמכונה כוונה לממוצע גבוה מ- 3 ס"מ נלקח מדגם מקרי של n ברגים
והוחלט לדחות את השערת האפס אם ממוצע המדגם שיתקבל יהיה גדול מ- 3.01.
מהו גודל המדגם, n , עבור רמת מובהקות 0.01?



שאלה

הלשכה המרכזית לסטטיסטיקה פרסמה ש- 50% מהתושבים ממחזרים בקבוקים.
הוחלט על מסע הסברה תקשורתית ברחבי המדינה להעלאת המודעות למחזור.
בתום מסע ההסברה נדגמו בשכונה מסוימת 150 משפחות ונמצא כי 82 משפחות ממחזרות בקבוקים.
א. האם לאור התוצאות ניתן להסיק שבשכונה זו חלה עלייה במודעות למחזור לאחר מסע ההסברה
התקשורתית ברמת מובהקות 0.05? נמקו.
ב. בהמשך לסעיף א מהי עצמת המבחן, אם אחוז המשפחות שממחזרות לאחר מסע ההסברה הוא 58%?
ג. בשכונה אחרת יותר גדולה נדגמו באופן מקרי 250 משפחות ונמצא כי 145 משפחות ממחזרות בקבוקים.
האם לאור התוצאות ניתן להסיק שיש הבדל בין השכונות מבחינת המודעות למחזור בקבוקים ברמת מובהקות
0.05? נמקו.



שאלה

בשני אזורים נדגמה הטמפרטורה היומית (מעלות):

8	7	6	5	4	3	2	1	Region(Degrees)
35	41	42	43	36	34	39	35	אזור א
	42	43	44	45	43	43	47	אזור ב

- א. חשב רווח סמך לתוחלת Degrees על סמך המדגם מאזור א ברמת סמך 95%.
ב. חשב רווח סמך לתוחלת Degrees על סמך המדגם מאזור ב ברמת סמך 95%.
ג. איזה רווח מדויק יותר? איזה רווח בטוח יותר? נמק.

ד. בדוק ב-2 דרכים את ההשערה שתוחלת Degrees באזור א שונה מ-40, ברמת מובהקות 0.05:

1. באמצעות מבחן סטטיסטי מתאים.

2. באמצעות רווח הסמך שחושב בסעיף א. נמק.

ה. בדוק את ההשערה שתוחלת Degrees באזור א שונה מתוחלת Degrees באזור ב, ברמת מובהקות 0.05. נמק.

ו. בדוק את ההשערה ששוונות Degrees באזור א שונה משוונות Degrees באזור ב, ברמת מובהקות 0.1. נמק.

ז. האם דרושות הנחות כלשהן לצורך החישובים? אם כן – רשום והסבר אותה. אם לא – נמק מדוע.



שאלה

נבדק ההבדל בין טמפרטורות במהלך היממה ב-7 אזורים:

Region(Deg.)	1	2	3	4	5	6	7
בוקר	30	32	31	27	27	37	33
ערב	35	31	37	32	34	32	41

ידוע ש- Deg. מתפלג נורמלית.

ברמת מובהקות 0.05, בדוק את ההשערה שתוחלת Deg. בערב גבוהה מתוחלת Deg. בבוקר. נמקו.



שאלה

על סמך נתוני הלשכה המרכזית לסטטיסטיקה ידוע כי 90% מבוגרי מדעי המחשב מתקבלים לעבודה מיד עם סיום לימודיהם.

חוקר במשרד הכלכלה טוען כי כיום האחוז נמוך יותר. לצורך בדיקת הטענה נבדק מדגם מקרי של 150 בוגרים ונמצא כי מתוכם 132 התקבלו לעבודה מיד עם סיום לימודיהם.

א. מהי מסקנת החוקר ברמת מובהקות 0.05? נמק.

ב. מהי הסתברות לטעות האפשרית במסקנה מסעיף א אם למעשה רק 85% מהבוגרים מתקבלים לעבודה מיד עם סיום לימודיהם?

ג. מהו גודל המדגם המינימלי שעבורו יוחלט בטעות שטענת החוקר ממשרד הכלכלה נכונה בהסתברות שלא תעלה על 0.02

וגם ההסתברות שיוחלט בטעות שנתוני הלשכה המרכזית לסטטיסטיקה נכונים
אם למעשה רק 85% מהבוגרים מתקבלים לעבודה מיד עם סיום לימודיהם, לא תעלה על 0.01.



שאלה

- משרד הכלכלה מעריך, על סמך נתוני העבר, כי אחוז הבוגרים במדעי המחשב המתקבלים לעבודה מיד עם סיום לימודיהם הוא 90. חוקר במשרד טוען כי הערכה זו אינה נכונה למצב כיום. לצורך בדיקת טענתו נלקח מדגם מקרי של 100 בוגרים. מתוכם 82 התקבלו לעבודה מיד עם סיום לימודיהם.
- א. בדקו את טענת החוקר ברמת מובהקות 0.05.
- ב. אם אחוז הבוגרים במדעי המחשב המתקבלים לעבודה מיד עם סיום לימודיהם הוא באמת 85, מה ההסתברות שהמבחן מסעיף א לא יגלה זאת? כיצד נקראת הסתברות זו?
- ג. בנוסף, מעוניינים לבדוק האם אחוז בוגרי מדעי המחשב המתקבלים לעבודה מיד עם סיום לימודיהם גבוה מאחוז בוגרי הנדסת תוכנה המתקבלים לעבודה מיד עם סיום לימודיהם. לשם כך נלקח מדגם מקרי של 50 בוגרי הנדסת תוכנה והתברר כי 35 מתוכם התקבלו מיד לעבודה.
1. מצאו רווח סמך להפרש בין פרופורציה בוגרי מדעי המחשב המתקבלים לעבודה מיד עם סיום לימודיהם לבין פרופורציה בוגרי הנדסת תוכנה המתקבלים לעבודה מיד עם סיום לימודיהם, ברמת סמך 90%.
2. מה ניתן להסיק מרווח הסמך שקיבלתם בתת סעיף 1?



שאלה

משך זמן השינה של ילדים מתפלג נורמלית עם סטיית תקן של שעה אחת.

חוקר החליט לבדוק האם זמן השינה מושפע מצפייה בסדרת אימה. לשם כך הוא בדק עבור 10 ילדים מקריים את זמן השינה בלילה לפני שצפו בסדרת אימה ואת זמן השינה בלילה אחרי שצפו בסדרת אימה.

הנתונים מוצגים בטבלה הבאה

ילד	1	2	3	4	5	6	7	8	9	10	סכום
זמן שינה לפני	7.9	8.2	10.1	9.3	9.2	9.6	8.6	9.8	9.4	10.4	92.5
זמן שינה אחרי	7.2	8.7	8.9	8.4	8.8	9.1	8.4	9.2	9.2	10.5	88.4

א. בדקו, ברמת מובהקות 0.05, האם משך זמן השינה של ילדים מתקצר לאחר צפיה בסדרת אימה. רשמו מה המבחן בו אתם משתמשים.

ב. אמדו את תוחלת ההפרש בין זמן השינה לפני צפיה בסדרת אימה לבין זמן השינה לאחר צפיה בסדרת אימה בעזרת רווח סמך ברמת סמך 95%.

ג. מה ניתן להסיק מרווח הסמך שקיבלתם בסעיף ב'?
האם יש סתירה למסקנה מסעיף א'?



שאלה

ברשת חנויות בגדים רצו לבדוק האם יש הבדל בין גברים ונשים בהוצאה החודשית על ביגוד. לשם כך החליטו לקחת מדגם מקרי של 100 גברים ו-200 נשים.

ידוע כי סטיית התקן של ההוצאה החודשית היא 50 ₪ אצל גברים ו-100 ₪ אצל נשים.

א. מנהל הרשת טוען כי תוחלת ההוצאה החודשית של נשים על ביגוד גדולה יותר.

1. רשמו את ההשערות שיש לבדוק לצורך בדיקת טענת המנהל.

2. השלימו: במבחן ברמת מובהקות 0.05 נטען כי המנהל צודק אם _____

3. אם אכן תוחלת ההוצאה החודשית של נשים על ביגוד גבוהה ב-30 ₪ מתוחלת ההוצאה החודשית של גברים על ביגוד,

מה ההסתברות שהמבחן מתת סעיף 2 יגלה זאת? כיצד נקראת הסתברות זו?

ב. במדגם שנלקח, לפי הגדלים הרשומים בפתיחת השאלה, התקבלה הוצאה חודשית ממוצעת על ביגוד של 295 ₪ אצל הגברים ו-350 ₪ אצל הנשים.

1. מצאו רווח סמך, ברמת סמך 96%, להפרש בין תוחלת ההוצאה החודשית על ביגוד של נשים לבין תוחלת ההוצאה החודשית על ביגוד של גברים.

2. ישנה טענה כי תוחלת ההוצאה החודשית של נשים על ביגוד גבוהה ב-40 לפחות מתוחלת ההוצאה החודשית של גברים על ביגוד. האם תקבלו טענה זו על פי רווח הסמך מתת סעיף 1? נמקו.



שאלה

חברת תרופות הוציאה לשוק תרופה חדשה להקלת כאבים. החברה טוענת בפרסומיה כי התרופה בעלת זמן השפעה ארוך בשעתיים לפחות מהתרופות המקבילות בשוק.

לצורך בדיקת טענת החברה נלקחו 18 נבדקים הסובלים מכאבים וחולקו באופן מקרי לשתי קבוצות. קבוצה אחת קיבלה את אחת התרופות המקבילות הקיימות בשוק והקבוצה השניה קיבלה את התרופה החדשה. כל נבדק דיווח על הזמן שעבר עד שהשפעת התרופה פגה.

התוצאות נתונות בטבלה הבאה:

קבוצה א – תרופה קיימת	קבוצה ב – תרופה חדשה
3	7
4	8
6	9
7	10
7	11
7	12
7	13
9	14
13	15

א. רשמו את השערות שיש לבדוק.

ב. מה ההנחות הדרושות לביצוע המבחן המתאים?

ג. בדקו את אחת ההנחות הניתנת לבדיקה עפ"י נתוני השאלה. ברמת מובהקות 2%.

ד. בדקו והסיקו מסקנות לגבי טענת החברה ברמת מובהקות 5%.



שאלה

במפעל מסוים זמן הלימוד של פועל על מכונה חדשה מתפלג נורמלית. מהעבר ידוע כי הוא בעל תוחלת של 90 דקות עם סטיית תקן של 7 דקות.

הגיעה למפעל מכונה חדשה והמנהל החליט לבדוק האם תוחלת זמן הלימוד היא אכן מה שהיתה בעבר או שהשתנתה.

לשם כך נלקח מדגם של 20 פועלים בו התקבל זמן לימוד ממוצע של 86 דקות.

א. אמדו את תוחלת זמן הלימוד בעזרת רווח סמך ברמת סמך של 98%.

ב. מה תסיקו לגבי שאלת המנהל לפי רווח הסמך שקיבלתם בסעיף א'? נמקו.

ג. 1. בדקו, ברמת מובהקות 0.01, את שאלת המנהל.

רשמו באופן ברור את ההשערות הנדרשות, הבדיקה ומסקנה מילולית.

2. מה רמת המובהקות המינימלית בה תקבלו את השערת האפס לפי מדגם זה?



שאלה

במסגרת מחקר על מעמד האישה נדרש לאמוד את תוחלת מספר שעות העבודה בשבוע של נשים עובדות.

ידוע כי סטיית התקן של מספר שעות העבודה השבועיות היא 5 שעות.

א. נלקח מדגם מקרי של 225 נשים עובדות ובו התקבל ממוצע של 37.5 שעות.

1. מצאו רווח סמך עבור תוחלת מספר שעות העבודה בשבוע של נשים עובדות, ברמת סמך של 95%.

2. מהו גודל המדגם שיש לקחת אם רוצים לקבל רווח סמך, באותה רמת סמך, שאורכו מחצית מהאורך שהתקבל בסעיף 1?

3. נבנה רווח סמך לפי מדגם בגודל שמצאתם בסעיף 2'. האם ניתן לטעון בביטחון רב יותר שרווח זה כולל את התוחלת (μ)? נמקו!

ב. התקן המוצהר של מספר שעות העבודה הנדרשות מנשים עובדות הוא 37.

אחד החוקרים טען כי בפועל תוחלת מספר שעות העבודה של נשים עובדות היא גדולה יותר.

לצורך בדיקת טענתו הוצע המבחן הבא: ילקח מדגם מקרי של 225 נשים עובדות.

אם ממוצע המדגם יהיה גדול מ- 37.6 שעות אז טענת החוקר תתקבל.

1. רשמו את ההשערות הנבדקות

2. מה רמת המובהקות של המבחן המוצע?

3. אם החוקר צודק ותוחלת מספר שעות העבודה של נשים עובדות היא 38 שעות,

מה ההסתברות שהמבחן המוצע יגלה זאת? כיצד נקראת הסתברות זאת?



שאלה

מנהל חדר מיון הכניס נוהל חדש לטיפול בחולים המגיעים לחדר המיון.

בדקו את זמן ההמתנה (בדקות) של 12 חולים שהגיעו לחדר המיון לאחר הפעלת הנוהל החדש והתוצאות שקיבלו היו:

25, 28, 29, 30, 33, 34, 35, 45, 50, 58, 60, 65.

הניחו כי זמן ההמתנה מתפלג נורמלית.

א. אמדו את תוחלת זמן ההמתנה בעזרת רווח סמך ברמת סמך 95%.

ב. מנהל חדר המיון טוען כי תוחלת זמן ההמתנה אחרי הכנסת הנוהל החדש היא 38 דקות.

האם תקבלו את טענתו לפי רווח הסמך שקיבלתם בסעיף א'? נמקו!

ג. תוחלת זמן ההמתנה לפני הפעלת הנוהל החדש הייתה 48.5 דקות.

מנהל בית החולים טוען כי תוחלת זמן ההמתנה קטנה לאחר הפעלת הנוהל החדש.

1. בדקו את טענתו של מנהל בית החולים ברמת מובהקות 0.05. פרטו מה המבחן בו אתם משתמשים ורשמו מסקנה מילולית ברורה.

2. משרד הבריאות דרש לבדוק את טענת מנהל בית החולים ברמת מובהקות מחמירה יותר של 0.025. האם תשתנה מסקנתכם? נמקו.

המרכז לקידום אקדמי

מבחנים א- פרמטריים

עסקנו עד פרק זה במבחנים על z - ו t - המבוססים על הנחות לגבי האוכלוסייה ממנה נדגמו הפריטים.

משפחת המבחנים הא-פרמטריים נועדו לתת מענה למצבים בהם לא ניתן להניח הנחות על האוכלוסייה. רוב המבחנים הא-פרמטריים מבוססים דירוגים, דירוג מספרים מתמקד במיקום כל ערך יחסית לערכים האחרים ולא בעוצמתו (הערך של המספר). דירוגים מתייחסים אך ורק לתכונת הסדר של המספרים, לכן דורשים סולם סדר לפחות.

המבחנים הא-פרמטריים מחולקים ל:

1. מבחנים לבדיקת השערות על **משתנה מקרי אחד** (מבחן הבינום, מבחן טיב ההתאמה)
2. מבחנים לבדיקת השערות על **שני משתנים מקריים (2 מדגמים):**
 - א. **בלתי תלויים** (מבחן וילקוקסון למדגמים בלתי תלויים, מבחן פישר)
 - ב. **תלויים – מזווגים** (מבחן הסימן, מבחן וילקוקסון למדגמים מזווגים, מבחן מקנמר).

מדגם אחד – מבחן הבינום או מבחן טיב התאמה

מבחן הבינום

אנו נשתמש במבחן זה כאשר המשתנה בינומי באוכלוסייה, המדגם קטן ולא עומד בתנאי הקירוב לנורמלי $np, nq \leq 10$. לכן נשתמש בנוסחת הבינום.

תנאים:

- (1) מדגם אחד
- (2) משתנה בינומי באוכלוסייה
- (3) תנאי קירוב לנורמלי לא מתקיים $np < 10, nq < 10$

חישוב

נחשב את אלפא מינימום: $p.v = \alpha_{\min} = P(X \geq k)$

כאשר k התוצאה שהתקבלה במדגם, ע"י טבלת ההסתברות הבינומית או נוסחת הבינום.

נשים לב שבטבלה נתון ערך מצטבר ולכן נתון: $P(X \leq k)$

תזכורת: הסיכוי של שחקן מסוים לקלוע לסל בזריקת עונשין הוא 0.7. מה הסיכוי שמתוך 5 זריקות עונשין הוא יקלע 3 פעמים בדיוק? $0.3087 = \frac{5!}{3!2!} \cdot 0.7^3 \cdot 0.3^2$

$$\binom{5}{3} \cdot 0.7^3 \cdot 0.3^2 = \frac{5!}{3!2!} \cdot 0.7^3 \cdot 0.3^2 = 0.3087$$

באופן כללי:

במצב שבו ההסתברות להצלחה בניסיון מסוים היא p

וההסתברות לכישלון היא $q = 1 - p$.

הסיכוי להצליח k פעמים מתוך n ניסיונות ב"ת שווה ל:

$$\binom{n}{k} p^k q^{n-k}$$

אלה הם ניסיונות Bernoulli: ניסיונות ב"ת, כאשר לכל ניסיון 2 תוצאות אפשריות.

אם נחשב פונקציית ההסתברות של k (מספר הצלחות מתוך n ניסיונות), נקבל התפלגות בינומית.

הכרעה:

נדחה את H_0 אם $p.v < \alpha$

או כאשר $p.v < \frac{\alpha}{2}$



✓ מבחן חי בריבוע לטיב התאמה (χ^2) Chi-square for goodness of fit

מטרתנו לבדוק האם משתנה באוכלוסייה מסוימת מתפלג לפי התפלגות מסוימת.

$-O_i$ ערך בטבלה הנצפית, $-E_i$ מה אנו מצפים שיהיה

בדיקה זו נעשית באמצעות χ^2

כאשר H_0 תמיד יניח שלא קיים הבדל בין O_i ל- E_i , כלומר שיש התאמה

כאשר H_1 תמיד יניח שלא קיימת התאמה.

השערות במבחן טיב התאמה:

$$H_0: p_1 = p_1^0, p_2 = p_2^0, \dots, p_k = p_k^0$$

$H_1: \text{other}$

כלומר ההשערה היא תמיד דו צדדית כאשר בודקים באמצעות χ^2 את הפער בין O_i ל- E_i .

$$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad \text{חישוב הסטטיסטי:}$$

הכרעה: דחה את H_0 אם: $\chi^2 > \chi_{\alpha}^2(k-1)$

אם הסטטיסטי הנ"ל **קטן** מהערך הקריטי אזי נאמר שהפער קטן ונקבל את H_0

אם הסטטיסטי הנ"ל **גדול** מהערך הקריטי אזי נדחה את H_0 ונאמר שלא קיימת התאמה למודל שהוצא ע"י החוקר.

התפלגות סטטיסטית זה היא חי בריבוע עם $k-1$ דרגות חופש, כאשר k יוגדר כמספר התאים, והוא יקרא ערך קריטי.

ככל שההבדל בין הטבלה הצפויה לנצפית גבוהה יותר, ניטה לדחות את השערת האפס, לגבי צורת ההתפלגות המקורית.

אופן החישוב

א- **נסח השערות:** התפלגות מסוימת: H_0 , אחרת: H_1

ב- בנה טבלת שכיחויות המכילה:

i. O_i - שכיחות נצפית של ערך i

ii. E_i שכיחות צפויה תחת הנחת H_0 של ערך i כאשר $E_i = np$

ג- חשב את סטטיסטי המבחן: $\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$

ד- כלל ההכרעה: דחה H_0 עבור ערכים מספיק גדולים של סטטיסטי המבחן.

דחה את H_0 אם $\chi^2 < \chi^2_{1-\alpha}(k-1)$, קריטי, $\chi^2 > \chi^2_{1-\alpha}(k-1)$

שאלה 1

שאלה 3

מרצה בקורס "מבוא למקור" טוען כי הציונים של 500 תלמידיו במבחן הסיום מתפלגים נורמלית עם תוחלת 70 וסטית תקן 5.

הציונים שחתקבלו הם:

שכיחות	ציון
1	40-55
67	56-65
344	66-75
85	76-85
3	86-100

בדוק את טענת המרצה ברמת מובהקות 0.05.

openbook
המרכז לקידום אקדמי

מבחנים א – פרמטריים לשני מדגמים תלויים

במדגמים תלויים אנו נחשב הפרש בין 2 המדגמים.

מדגמים תלויים – לפני/אחרי שימוש בתרופה וכו

דיכוטומיים – טבלה 2×2 . תשובות כן/לא

במבחן א – פרמטריים למדגמים תלויים מתעלמים מחישוב זוגות שלא חל שינוי – בהם ההפרש=0.

✓ מבחן הסימן

כאשר מדובר בשני מדגמים תלויים (מזוגים) אך לא ניתן לחשב הפרשים או שלא נתונים הפרשים אלא רק מספר העליות או הירידות (או ללא שינוי)

א. בודקים בכמה מקרים יש עליה (+) ירידה (-) או ללא שינוי (0).

ב. מתעלמים בחישוב מזוגות שלא חל בהם שינוי (לא נכללים בחישוב).

ג. שיטת הבדיקה זהה לחלוטין למבחן הבינום כאשר תמיד תחת $p=0.5$.

השערות:

לא קיים הבדל בין (+) ל (-) : H_0

$$H_0: p(+) = p(-) = 0.5$$

נדחה את H_0 אם $p.v < \alpha$

שלבי מבחן הסימן

א- להבחין במקרה

ב- לנסח את ההשערות (מבחן שמאלי, ימני, דו צדדי)

ג- לבנות טבלת הסימנים

ד- לקבל את ההחלטה: דחה H_0 אם:

a. חד צדדי: $P(X \geq k) = 1 - P(X \leq k - 1) \leq \alpha$

i. מבחן ימני – k הוא מספר ה "+"

ii. מבחן שמאלי – k הוא מספר ה "-"

b. דו-צדדי: $2 \cdot P(X \leq k) \leq \alpha$ (k הנמוך מבין ה- "-" או "+")

✓ מבחן וילקוקסון למדגמים מזווגים תלויים

זהו מבחן המקביל למבחן t למדגמים תלויים.

כאשר מדובר בשני מדגמים תלויים (מזווגים), ואין הנחת התפלגות נורמלית, מדגם קטן, וניתן לחשב הפרשים. במבחן הסימן- לא שמנו לב לגודל המוחלט של ההפרש כדי להתגבר על בעיה זו הציע וילקוקסון בשנת 1945 את המבחן הזה.

מתחשבים בגודל המוחלט של ההפרש ולא רק בכוונו.

לצורך חישוב מבחן זה יש לערוך טבלת דירוגים.

כאשר שני המשתנים תלויים, המידע הוא n זוגות של תצפיות, אין הנחת נורמליות, מדרגים את הערך המוחלט של ההפרשים.

$D = X_2 - X_1$, מתעלמים מזוגות שהפרשם אפס.

אלגוריתם לחישוב מבחן וילקוקסון למזווגים

1. כמו במבחן t למדגמים תלויים, עבור כל נבדק/זוג מחשבים את ההפרש בין שתי התצפיות $d_i = x_i - y_i$

(כאשר קיבלנו הפרש=0 הנבדקים בעלי $d=0$ נזרקים מהניתוח ומתקנים את גודל n בהתאם

למספר זוגות להם קיים הפרש בין שתי תצפיות).

2. מדרגים את הערכים המוחלטים $|d_i|$, צריך לנטרל אפס

3. נדרג את הערך מוחלט מ-1 (הנמוך) אם יש ערכים שווים הדורג יהיה ממוצע הדורגים

4. נסמן בנפרד דירוגים שבמקור שליליים מדירוגים שבמקור חיוביים ע"י הוספת 2 שורות נוספות w^-

המצוין דירוגים שבמקור שליליים, ו w^+ שמצוין דירוגים שבמקור חיוביים.

5. סוכמים בנפרד את הדירוגים שמקורם בהפרש חיובי, ובנפרד את הדירוגים שמקורם בהפרש שלילי.

6. בוחרים את הקטן שמבין שני הסוכמים ומגדירים אותו כ- w זהו הסטטיסטי של מבחן זה

7. בודקים אם הסטטיסטי w קטן או שווה לערך הקריטי המופיע בטבלה המתאימה עבור α נתונה, ואז

נדחה/נקבל את H_0 (אנו דוחים כאשר ערך הסטטיסטי קטן או שווה לערך הקריטי המופיע בטבלה)

C- ערך קריטי בטבלה ערכים קריטיים של W למבחן וילקוקסון למדגמים מזווגים.

הכרעה: דחה את H_0 אם $w \leq c$ לפי הטבלה של חד / דו צדדי

בדיקה: $W_+ + W_- = \frac{n(n+1)}{2}$

מבחן מקנמר ✓

נשתמש במבחן זה כאשר:

מדגמים תלויים מזווגים, שמיים (בינומיים), דיכטומיים (כן/לא), טבלה דו ממדית (2×2) לנתונים שמיים, מבחן דו צדדי.

המבחן עושה שימוש בטיב התאמה עם 2 קטגוריות.

הסטטיסטי נקבע ע"י השכיחויות של התאים/מקרים ששינו את דעתם $B+C$.

מתחלק ל 2 חלקים:

1. כשההשערה **דו צדדית** ו- $B+C > 20$ משתמשים בהתפלגות **חי בריבוע** עם דרגת חופש אחת. סטטיסטי

$$\chi^2 = \frac{(B-C)^2}{B+C} \sim \chi^2_{(1)} \quad \text{נדחה את } H_0 \text{ אם } \chi^2 > \chi^2_{\alpha}$$

2. כשההשערה **חד צדדית** או $B+C \leq 20$ משתמשים בהתפלגות **בינומית**. משתמשים במבחן הסימן

$$B+C=n, \quad \text{דחה } H_0 \text{ אם } p.v. < \alpha$$

אמרנו שאם אין השפעה – מה זה השפעה? הפרש 0 אז מתעלמים לכן אם לפני אין שינוי ואחרי אין שינוי – נתעלם? ואם לפני יש שינוי ואחרי יש שינוי מתעלם? לכן נמחק את התוצאות הלא רלוונטיות.

טבלת שכיחויות 2×2 :

	ערך 1	ערך 2	נסמן ב-B ו-C את 2 התאים שחל בהם שינוי
ערך 1	אין שינוי	יש שינוי B	סך הנבדקים בעלי ערך 1 אחרי
ערך 2	יש שינוי C	אין שינוי	סך הנבדקים בעלי ערך 2 אחרי
	סך הנבדקים בעלי ערך 1 לפני (בעבר)	סך הנבדקים בעלי ערך 2 לפני (בעבר)	$N =$ סך הנבדקים הכולל

לאחר מכן בודקים באמצעות התפלגות χ^2 האם קיים פער בין B ל-C.

✓ מבחן וילקוקסון לשני מדגמים בלתי תלויים

מבחן זה מקביל למבחן t למדגמים ב"ת. המשתנים סידוריים או מנה ולא ניתן להניח נורמליות. הבסיס התיאורטי למבחן: נניח שיש לנו שני מדגמים (ערכי x וערכי y) הלקוחים משתי התפלגויות.

H_0 : אין הבדל

H_1 : יש הבדל

שלבים בחישוב סטטיסטי המבחן:

(1) **מדרגים** את כל הערכים בסדר עולה מהנמוך לגבוה ללא הבחנה לקבוצה אליה שייכים ($n = n_1 + n_2$ מס' הנבדקים הוא סכום מספר הנבדקים בשתי הקבוצות).

(2) **מסכמים** את הדירוגים של כל קבוצה **בנפרד**, הסכום של הקבוצה הקטנה יסומן W_A ואילו של הקבוצה הגדולה W_B (מספיק לחשב רק את W_A). אם $n_1 = n_2$, לא משנה איזה קבוצה בוחרים.

(3) ההתפלגות של W_A תחת H_0 ($p(x > y) = 1/2$) ניתנת לחישוב עבור n_1 ו- n_2 נתונים.

שוב מישו עשה את העבודה בשבילנו וריכז את הערכים הקריטיים בטבלה.

הטבלת מציגה את הערכים הקריטיים של W_A עבור α -ות נבחרות ובהתאם לגודלם של n_1 ו- n_2 .

ע"מ לדחות את H_0 הסטטיסטי W_A צריך להיות קטן או שווה ל- W_L (הערך הקריטי בזנב התחתון) או לחלופין גדול או שווה ל- W_U (הערך הקריטי בזנב העליון).

$$\begin{aligned} U_1 &= W_1 - \frac{n_1(n_1+1)}{2} \\ U &= \min(U_1, U_2) \\ U_2 &= W_2 - \frac{n_2(n_2+1)}{2} \end{aligned}$$

הסטטיסטי U מחושב:

המרכז לקידום אקדמי

עקרונות: לשם קביעת כלל הכרעה משתמשים ב-2 מצבים:

(1) כאשר $n_1, n_2 \leq 8$ משתמשים בטבלאות $P_{H_0}(U \leq u)$ הנותנת α מינימלית.

במבחן חד צדדי דחה את H_0 כאשר $P.v = P(U \leq u) < \alpha$

במבחן דו צדדי דחה את H_0 כאשר $P(U \leq u) < \frac{\alpha}{2}$ או $2P(U \leq u) < \alpha$

(2) כאשר $n_1, n_2 > 8$ משתמשים בטבלאות הנותנות ערכים קריטיים ל- U .

דחה את H_0 כאשר $u < u_c$

שאלה 1

במחקר על היכולות הקוגניטיביות של קופים, לקחו 19 קופים בני אותו גיל וחלקו אותם באופן מקרי לשתי קבוצות. אחת של 10 קופים והשנייה של 9 קופים. שתי הקבוצות אומנו להגיב לתמריצים קוליים, כל קבוצה בשיטה אחרת. לאחר מכן בחנו את הקופים והתוצאות נתונות בטבלה הבאה:

שיטה	ממוצע	$\hat{\sigma}^2$
1	102	103
2	98	100
1	158	157
2	140	132
1	156	155
2	116	106
1	149	131
2	105	104
1	129	120
2	102	103
1	136	492.22
2	114	227.75

בדוק האם יש הבדל בין שתי שיטות האימון ברמת מובהקות 0.05, כאשר:

- א. ניתן להניח כי תוצאות הבחינה מתפלגות נורמלית. (לאיזו הנחה נוספת הינך נזקק?)
- ב. לא ניתן להניח כי תוצאות הבחינה מתפלגות נורמלית.



✓ מבחן פישר

שני משתנים ב"ת, שמיים (בינומיים) ודיכוטומיים (2×2 בעד ונגד).

מחליף את מבחן Z להפרש בין פרופורציות.

מבחן שנכנס לשימוש כאשר המדגמים בלתי תלויים, והמשתנה בינומי ב-2 האוכלוסיות

$$\begin{aligned} n \cdot \hat{p}_1, n \cdot \hat{q}_1 &< 10 \\ n \cdot \hat{p}_2, n \cdot \hat{q}_2 &< 10 \end{aligned} \quad \text{המדגמים קטנים ולא עומדים בתנאי הקירוב לנורמלי (לפחות 1 מהם):}$$

(אם לפחות אחד מהם קטן מ-10 נשתמש בפישר).

הסטטיסטי נקבע ע"י השכיחות A שנצפה שערכה צריך להיות נמוך מבין התאים בכישלון (A הוא התא בו אמור להיות שיפור אך אין),

נחשב את רמת המובהקות המינימלית $p.v = \alpha' = \alpha_{min}$

שהיא ההסתברות לקבל את תוצאת המדגם ב-A או קיצונית לה (קטנה יותר)

$$p.v = P(A \leq \alpha)$$

דחה את H_0 כאשר $P.v < \frac{\alpha}{2}$ או $P.v < \alpha$

טבלת שכיחות 2×2 :

	כישלון	הצלחה	
סך הנבדקים בקבוצה 1-ה			קבוצה 1
סך הנבדקים בקבוצה 2-ה			קבוצה 2
$N =$ סך הנבדקים בשתי הקבוצות	סך הכישלונות	סך ההצלחות	

שאלה 2 ממבחן

- העמותה, למניעת התמכרות למשחקי מחשב בקרב בני נוער, החליטה לבדוק את השפעת מספר השעות ביום שבהן נער נחשף למחשב, על הצורך במשקפיים. לצורך הבדיקה נלקח מדגם מקרי של 14 נערים, "מכורים", כלומר, נחשפים למחשב לפחות 4 שעות ביום, ומדגם מקרי של 10 נערים, לא "מכורים", אשר נחשפים פחות מ-4 שעות ביום למחשב.
- לאחר מעקב של שנתיים, התברר כי מבין הנערים ה"מכורים" 6 נזקקו למשקפים. מבין הנערים הלא "מכורים" 1 נזקק למשקפיים.
- א. בדוק, ברמת מובהקות 0.1, האם חשיפה גדולה למחשב מגדילה את הצורך במשקפיים. (שים לב לגודלי המדגם). רשום את ההשערות אותן אתה בודק ואת סוג המבחן בו אתה משתמש.
- ב. בדוק, ברמת מובהקות 0.1, האם לאור תוצאות המדגם ניתן לומר כי אחוז הנזקקים למשקפיים, בקרב הנערים ה"מכורים" למחשב, הוא 50%.

שאלה 3 ממבחן

להלן ציוני הגמר בקורס מסוים באוניברסיטה הפתוחה.

ציון	מין	אישה	גבר	סה"כ
		100	60	160
נכשל		20	20	40
סה"כ		120	80	200

- בדוק ברמת מובהקות 0.05 את ההשערה שנשים מצליחות בקורס זה יותר מגברים.
- א. עי"י מבחן אי-פרמטרי.
- רשום את הנחותיך ונמק את הבחירה במבחן בו אתה משתמש.
- ב. עי"י מבחן פרמטרי.

אססבוק
המרכז לקידום אקדמי

שאלות מבחנים א-פרמטרים



שאלה

משווק ביצים טוען שרק 10% מהביצים שהוא משווק סדוקות. הקונים טוענים שאחוז הסדוקות גבוה יותר. נקלח מדגם של 15 ביצים ונמצאו 5 ביצים סדוקות. מהי מסקנתך ברמת מובהקות 0.05?



שאלה

בסקר קודם נמצא כי 40% מהציבור בישראל מרוצה מפעולות הממשלה. כדי לבדוק אם חלה ירידה בשביעות רצון של הציבור מפעולות הממשלה, נדגמו 20 נשאלים ונמצא כי 6 מתוכם היו מרוצים. בחן את ההשערה בר"מ 0.05?



שאלה

נתוני משרד הבריאות מראים כי 80% מהמעשנים שמנסים להיגמל מעישון חוזרים לעשן. מרכז לגמילה מעישון טוען שפיתח שיטת גמילה חדשה ולפיה אחוז החוזרים לעשן נמוך מ-80%. לצורך בדיקת הטענה נלקח מדגם מקרי של 15 מעשנים שעברו את שיטת הגמילה החדשה. לאחר חצי שנה בדקו כמה מבין ה-15 חזרו לעשן. הוחלט לקבל את טענת המרכז לגמילה אם מספר האנשים שחזרו לעשן הוא לכל היותר 8.

1. מהי רמת המובהקות של המבחן המוצע?
2. מהי עצמת המבחן המוצע אם ההסתברות לחזור לעשן לאחר גמילה בשיטה החדשה היא 0.5?



שאלה

כדי לבדוק אם מטבע מסוים הוא תקין מטילים אותו 10 פעמים.

אם הוא נופל על "ראש" 2,3,4,5 או 6 פעמים מחליטים שהמטבע תקין, אחרת מזויף.

א. מהי רמת המובהקות של המבחן?

ב. אם האלטרנטיבה שההסתברות ל"ראש" במטבע זה היא 0.4, מהי עצמת המבחן?



שאלה

כדי לבדוק אם מטבע מסוים הוא תקין מטילים אותו 8 פעמים.

אם הוא נופל על "ראש" 2,3,4,5 או 6 פעמים מחליטים שהמטבע תקין, אחרת מזויף.

א. מהי רמת המובהקות של המבחן?

ב. אם האלטרנטיבה שההסתברות ל"ראש" במטבע זה היא 0.4, מהי עצמת המבחן?



שאלה

על פי נתוני הלשכה המרכזית לסטטיסטיקה אחוז המעשנים בישראל הוא 20% .
לבדיקת הטענה שלאור המצב הכלכלי הקשה חלה עלייה באחוז המעשנים נשאלו 20 אזרחים שנבחרו באופן מקרי, והוחלט שאם במדגם יהיו יותר מ- 6 מעשנים תתקבל הטענה.
חשב את רמת המובהקות של המבחן המוצע.



שאלה

התרופה המקובלת לריפוי מחלה מסוימת יעילה ב- 60% מהמקרים.
תוצאות ניסוי של תרופה חדשה שניתנה למדגם של 20 חולים הראו כי 15 מהחולים החלימו.
האם התרופה החדשה עדיפה על זו שמקובלת היום, ברמת מובהקות 0.1? נמק.



שאלה

במבחן גמר בכלכלה 15 שאלות אמריקאיות עם 5 תשובות אפשריות לכל אחת.
סטודנט שהיה רוב הסמסטר במילואים ניגש לבחינה והצליח לענות נכון על 6 שאלות.
מהי רמת המובהקות הקטנה ביותר עבורה ניתן יהיה להסיק שהסטודנט השלים את החומר ולא ניחש את התשובות?



שאלה

במבחן גמר בסטטיסטיקה יש 15 שאלות אמריקאיות עם 5 תשובות אפשריות לכל אחת.
סטודנט שהיה רוב הסמסטר במילואים ניגש לבחינה והצליח לענות נכון על 7 שאלות.
האם ניתן יהיה להסיק, ברמת מובהקות 0.05, שהסטודנט השלים את החומר ולא ניחש את התשובות?



שאלה

פסיכולוג ב' ביצע סקר, שהקיף 65 חיילי מילואים. הוא בדק את מידת ההבנה החינוכית שלהם לפני שירות המילואים ולאחריו, וקיבל את התוצאות האלה:

הבנה לפני מילואים			הבנה אחרי מילואים
מעטה	רבה	סה"כ	
10	10	20	מעטה
30	15	45	רבה
40	25	65	סה"כ

מה אפשר להסיק מתוצאות אלה בר"מ 2.5%?



שאלה

חוקר ביקש לבדוק האם מצב משפחתי משפיע על עבירות תנועה. הוא דגם מקרית 80 אנשים רווקים/ ות העומדים להינשא בשנה הקרובה ומצא כי 30 מתוכם ביצעו עבירות תנועה בשנה האחרונה.

כעבור שנה, כאשר היו כבר נשואים, הוא גילה כי מתוך ה-30, 7 המשיכו לבצע עבירות, וכמו כן נוספו 9 עבריינים נוספים (שבעבר לא ביצעו עבירות תנועה).

בדוק את הטענה ברמת מובהקות 0.01 ונמק בחירתך במבחן הסטטיסטי המתאים.



שאלה

כדי לבדוק אם להורמון מסוים יש השפעה כלשהי על יכולת החשיבה, נבחרו 50 אנשים באופן מקרי ונערכה להם בחינה. לאחר מכן קבלו כל הנחקרים את ההורמון ונערך להם מבחן נוסף. התוצאות שהתקבלו הראו כי מתוך 36 האנשים שעברו את בהצלחה את המבחן הראשון, 20 עברו גם את השני בהצלחה. כמו כן מחצית האנשים שנכשלו במבחן הראשון נכשלו גם בשני.

האם ניתן לומר, בר"מ 0.025 שלהורמון יש השפעה כלשהי על יכולת החשיבה? הסבר ונמק.

שאלה

לפני עימות רדיופוני שבא לקבוע מי מבין שני מועמדים פופולרי יותר בקרב בני הנוער, נדרשו 300 בני נוער להכריע בין שניהם. לאחר העימות נשאלו אותם בני נוער בשנית. להלן התוצאות:

לפני	אחרי	
	מועמד 1	מועמד 2
מועמד 1	102	18
מועמד 2	23	157

האם שדרן רדיו יכול לקבוע ברמת מובהקות של 5% שקיים הבדל בין שני המועמדים במספר הקולות שסחפו לטובתם?



שאלה

חברה מייצרת אבקת כביסה רוצה לבדוק אם פרסום באמצעות חלוקת דוגמיות משפיע על קניית מוצרים בקרב הלקוחות. לשם כך נדגמו מקרית 50 אנשים שנשאלו אם הם משתמשים באבקת כביסה מסוימת לפני ואחרי החלוקה. נמצא כי 21 מתוכם השתמשו באבקת הכביסה בעבר, כש-15 מתוך ה-21 המשיכו להשתמש גם לאחר חלוקת הדוגמיות, ובנוסף הצטרפו 19 משתמשים חדשים באבקה לאחר חלוקת הדוגמיות.

א. בדוק בר"מ 0.01 את ההשפעה שיש לשיטת פרסום זו על השימוש במוצר.

ב. בדוק בר"מ 0.01 אם שיטת הפרסום מעלה את השימוש במוצר.



שאלה

גזבר ראשות מקומית מסוימת רצה לעודד תושבים לשלם את הארנונה בתשלום אחד בתחילת השנה. לטענתו הענקת תו חנייה המאפשר חנייה במקומות מועדפים, חינוך, למשלמים בתשלום אחד, תעלה את מספר התושבים המשלמים את הארנונה בתשלום אחד. לבדיקת הטענה, נשאלו 50 תושבים על עמדותיהם לפני הצעת המבצע ולאחריה.

התוצאות שהתקבלו היו:

23 תושבים שילמו בתשלום אחד בשנה שעברה וגם השנה בעקבות המבצע יעשו כן.

5 תושבים שילמו בתשלום אחד בשנה שעברה אך השנה למרות המבצע לא יעשו כן.

14 תושבים לא שילמו בתשלום אחד בשנה שעברה אך השנה בעקבות המבצע יעשו כן.

8 תושבים לא שילמו בתשלום אחד בשנה שעברה וגם השנה בעקבות המבצע לא יעשו כן

בדוק את טענת הגזבר ברמת מובהקות 0.05. נמק.

openbook
המרכז לקידום אקדמי



שאלה

הטבלה הבאה מציגה את מספר שעות השינה לאחר קבלת תרופה מסוימת, ומספר שעות השינה לפני קבלתה, של 9 חולים שקיבלו תרופה מסוימת.

מספר שעות השינה לפני קבלת התרופה	7	5	6.5	6.5	7	5	4.5	8	8
מספר שעות השינה אחרי קבלת התרופה	8	7	6	8	7	8	8.5	7	9

האם לתרופה השפעה חיובית על מספר שעות השינה, כאשר לא ניתן להניח התפלגות נורמלית של שעות השינה? נסח את ההשערות, הסבר מהו המבחן המתאים ובדוק ברמת מובהקות 0.05.

שאלה

מהנדס מזון מעוניין לבדוק האם הוספת חיידקים בתהליך הכנת יוגורט מחלב רזה משפרת את איכותו. לשם כך נלקח חלב רזה מ-8 רפתות שונות. החלב, מכל אחת מהרפתות, חולק לשניים. חצי אחד הועשר בחיידקים והשני לא. לאחר מכן הוכן יוגורט מכל אחד מהם בנפרד. בכל אחד מ-16 היוגורטים שהוכנו נמדדה רמת המוצקות (ככל שיוגורט מוצק יותר איכותו טובה יותר) והתוצאות מובאות בטבלה הבאה:

מספר הרפת	1	2	3	4	5	6	7	8
X- עם החיידקים	68	75	62	86	52	46	72	50
Y- ללא החיידקים	62	70	63	79	52	38	69	54

בדקו את טענת החוקר, ברמת מובהקות 0.05,

א. כאשר ניתן להניח כי רמת המוצקות מתפלגת נורמלית.

(ניתן להיעזר בחישובים הבאים: $\sum_{i=1}^8 d_i^2 = 200$, $\sum_{i=1}^8 d_i = 24$ כאשר $d_i = x_i - y_i$)

ב. כאשר לא ניתן להניח כי רמת המוצקות מתפלגת נורמלית.

1. בעזרת מבחן הסימן.

2. בעזרת מבחן וילקוסון מתאים.



שאלה

במחקר מסוים נבדק יבול הקפה [מספר שקים בחודש] ב-7 חוות קטנות לפני ואחרי שימוש בשיטת טיפול חדשנית נגד פטרייה מזיקה:

חוה	1	2	3	4	5	6	7
לפני	5	6	7	8	4	3	6
אחרי	10	9	9	8	13	8	9

ידוע שהיבול אינו מתפלג נורמלית.

בדוק את ההשערה שתוחלת היבול לאחר הטיפול שונה מתוחלת היבול לפני הטיפול ברמת מובהקות 0.05. נמק



שאלה

מרכז הוראה שבקורס שלו קיימת הנחיה דרך הליווי ("אופק") החליט לבדוק האם יש הבדל בין הנחיה זו והנחיה רגילה בכיתה. לצורך כך הוא בחר 20 סטודנטים, 10 מהם למדו ב"אופק" ו-10 בהנחיה רגילה. ציוניהם בסוף הקורס נתונים בטבלה הבאה: (לא ניתן להניח התפלגות נורמלית של הציונים)

מספר תלמיד בקבוצתו	1	2	3	4	5	6	7	8	9	10
"אופק"	35	6	18	25	37	48	49	53	81	89
רגילה	34	57	9	64	61	75	91	100	98	93

בדוק את השערת המרכז ברמת מובהקות 0.02,

א. בהנחה כי אלו 10 זוגות של סטודנטים שהותאמו לפי ביצועיהם בקורסים קודמים.

ב. בהנחה כי הסטודנטים נבחרו וחולקו באופן מקרי לשתי קבוצות בכל סעיף ציין את המבחן בו בחרת להשתמש ונמק בחירה זו.



שאלה

כדי לבדוק את הטענה שכמות השתן המופרשת ביום גדולה מזו המופרשת בלילה נמדדו כמויות אלה אצל 8 אנשים והתקבלו התוצאות הבאות:

כמות השתן ביום	37	40	73	56	71	73	96	35
----------------	----	----	----	----	----	----	----	----

39	59	71	59	53	73	42	63	כמות השתן בלילה	
----	----	----	----	----	----	----	----	-----------------	--

מהי מסקנתך ברמת מובהקות 0.05 ?

נמק בעזרת המבחן הסטטיסטי המתאים ביותר כאשר כמות השתן לא מתפלגת נורמלית.

שאלה

לילדים עם בעיות לימודיות יש לעיתים סימפטומים הדומים לאלה של ילדים עם מחלה מסוימת. עובדה זו גרמה לחוקרים לחשוב שתרופה למחלה תעזור גם לילדים עם הבעיות הלימודיות. לצורך בדיקת ההשערה נבדקו 10 ילדים עם בעיות לימודיות כאלו. במשך שלושה שבועות הם קיבלו תרופת דמי (פלסבו) ובמשך שלושה שבועות נוספים קיבלו את התרופה האמיתית. הם נבחנו במבחני IQ פעמים, פעם לאחר שלושת השבועות הראשונים (מבחן I) פעם לאחר שלושת השבועות שבהם קיבלו את התרופה האמיתית (מבחן II)

תוצאות המבחנים מוצגות בטבלה הבאה:

96	90	104	115	111	102	95	106	106	97	מבחן I
126	110	102	121	122	111	119	101	113	113	מבחן II

האם לאור התוצאות ניתן לומר שהתרופה עוזרת ?

נסח את ההשערות ובדוק ברמת מובהקות 0.05 בעזרת המבחן הסטטיסטי המתאים ביותר כאשר לא ניתן להניח התפלגות נורמלית של הציונים

שאלה

לילדים עם בעיות לימודיות יש לעיתים סימפטומים הדומים לאלה של ילדים עם מחלה מסוימת. עובדה זו גרמה לחוקרים לחשוב שתרופה למחלה תעזור גם לילדים עם הבעיות הלימודיות. לצורך בדיקת ההשערה נבדקו 10 ילדים עם בעיות לימודיות כאלו. במשך שלושה שבועות הם קיבלו תרופת דמי (פלסבו) ובמשך שלושה שבועות נוספים קיבלו את התרופה האמיתית. הם נבחנו במבחני IQ פעמים, פעם לאחר שלושת השבועות הראשונים (מבחן I) פעם לאחר שלושת השבועות שבהם קיבלו את התרופה האמיתית (מבחן II)

תוצאות המבחנים מוצגות בטבלה הבאה:

96	90	104	115	111	102	95	106	106	97	מבחן I
126	110	102	121	122	111	119	101	113	113	מבחן II

א. מהי רמת המובהקות המינימלית שעבורה יוחלט שהתרופה עוזרת על סמך מבחן הסימן ? נמק

ב. בהמשך לסעיף א' מהי המסקנה ברמת מובהקות 0.05 ?

ג. האם לאור התוצאות ניתן לומר שהתרופה עוזרת ?

נסח את ההשערות ובדוק ברמת מובהקות 0.05 בעזרת המבחן הסטטיסטי המתאים ביותר כאשר לא ניתן להניח התפלגות נורמלית של הציונים



שאלה

במפעל מסוים ביקשו לבדוק את ההשערה שתפוקת העובדים הממוצעת גדולה יותר בשעות הבוקר מאשר בשעות אחה"צ. לשם כך נלקח מדגם מקרי אשר כלל 8 מעובדי המפעל והתקבלו התוצאות הבאות

עובד	1	2	3	4	5	6	7	8
תפוקה בבוקר	5.1	5.3	6	5.8	5.5	5.1	5.4	5.7
תפוקה אחה"צ	4.1	4.8	5	6.3	4.5	5	5.2	5.8

מהי מסקנתך ברמת מובהקות 0.01 ? נמק.

בעזרת מבחן ווילקוקסון המתאים



שאלה

במפעל מסוים ביקשו לבדוק את ההשערה שתפוקת העובדים הממוצעת גדולה יותר בשעות הבוקר מאשר בשעות אחה"צ. לשם כך נלקח מדגם מקרי אשר כלל 8 מעובדי המפעל והתקבלו התוצאות הבאות

עובד	1	2	3	4	5	6	7	8
תפוקה בבוקר	5.1	5.3	6	5.8	5.5	5.1	5.4	5.7
תפוקה אחה"צ	4.1	4.8	5	6.3	4.5	5	5.2	5.8

א. מהי מסקנתך ברמת מובהקות 0.01 ? נמק.

1. בעזרת מבחן הסימן.

2. בעזרת מבחן ווילקוקסון המתאים

ב. חזור על סעיף א' כאשר ניתן להניח שהתפוקה מתפלגת נורמלית



שאלה

נבדק ההבדל בין זמני תגובה לפרסומת קופצת (מילי-שניות) לפני ואחרי אימון. אצל 7 נערים:

bs(Time)	1	2	3	4	5	6	7
לפני	300	294	323	261	252	284	299
אחרי	251	262	273	281	248	237	302

ידוע ש- Time אינו מתפלג נורמלית.

- א. ברמת מובהקות 0.05, בדוק את ההשערה ש Time לפני האימון גבוה יותר מ Time אחרי האימון. נמק
 ב. חזור על סעיף א באמצעות מבחן הסימן.



שאלה

יצרן של משחות שיזוף מעוניין לבדוק אם קרם חדש שיצא מגדיל את ההגנה מפני קרני השמש. לשם כך בחר 7 מתנדבים וביקש מהם לחשוף את גבם לשמש כשעל צד אחד מרוח הקרם הישן ובצד השני החדש. תוצאות המדידה של דרגת הכוויה הן אלו: (על פי סולם ערכים מ-31 עד 70 = כווירה חומרה ביותר)

מספר מתנדב	1	2	3	4	5	6	7
קרם ישן	42	51	31	61	44	55	48
קרם חדש	38	53	36	52	33	49	36

מה תהיה מסקנת היצרן בר"מ 0.05 ו 0.025 אם לא ניתן להניח התפלגות נורמלית של דרגות הכוויה?

נסח את ההשערות והסבר מהו המבחן הסטטיסטי המתאים לבדיקתן



שאלה

יצרן של משחות שיזוף מעוניין לבדוק אם קרם חדש שיצא מגדיל את ההגנה מפני קרני השמש. לשם כך בחר 7 מתנדבים וביקש מהם לחשוף את גבם לשמש כשעל צד אחד מרוח הקרם הישן ובצד השני החדש. תוצאות המדידה של דרגת הכוויה הן אלו: (על פי סולם ערכים מ-31 עד 70 = כווירה חומרה ביותר)

מספר מתנדב	1	2	3	4	5	6	7
קרם ישן	42	51	31	61	44	55	48

ק"מ חדש	38	53	36	52	33	49	36
---------	----	----	----	----	----	----	----

מה תהיה מסקנת היצרן על פי מבחן הסימן בר"מ 0.05 , 0.025 ו 0.1?

שאלה

יצרן של משחות שיזוף מעוניין לבדוק אם קרם חדש שיצא מגדיל את ההגנה מפני קרני השמש. לשם כך בחר 7 מתנדבים וביקש מהם לחשוף את גבם לשמש כשעל צד אחד מרוח הקרם הישן ובצד השני החדש. תוצאות המדידה של דרגת הכוויה הן אלו: (על פי סולם ערכים מ-31 עד 70 = כווירה חומרה ביותר)

מספר מתנדב	1	2	3	4	5	6	7
קרם ישן	42	51	31	61	44	55	48
קרם חדש	38	53	36	52	33	49	36

א. מה תהיה מסקנת היצרן בר"מ 0.05 ו 0.025 אם לא ניתן להניח התפלגות נורמאלית של דרגות הכוויה?

נסח את ההשערות והסבר מהו המבחן הסטטיסטי המתאים לבדיקתן

ב. מה תהיה מסקנת היצרן על פי מבחן הסימן בר"מ 0.05 , 0.025 ו 0.1?

שאלה

בבית ספר מסוים הועלתה טענה כי ההישגים בספרות טובים יותר מההישגים בלשון.

מנהל בית הספר טוען שהפער עולה על 10 נקודות בממוצע, ואם יתברר כי טענתו נכונה יצטרכו לתגבר את לימודי הלשון.

לצורך בדיקת טענת המנהל נבדק מדגם מקרי של 9 תלמידים שציוניהם בלשון ובספרות מופיעים בטבלה שלהלן:

ציון בלשון	70	78	81	67	90	57	68	84	78
ציון בספרות	90	88	93	89	100	69	80	75	90

א. האם ניתן לטעון ברמת מובהקות 0.05 שהמנהל צודק, בהנחה שהציונים מתפלגים נורמלית? נמק.

ב. בהמשך לסעיף א, תן הערכה לרמת המובהקות המינימלית שעבורה השערת האפס תידחה על סמך תוצאות המדגם.



שאלה

בבית ספר מסוים הועלתה טענה כי ההישגים בספרות טובים יותר מההישגים בלשון.

מנהל בית הספר טוען שהפער עולה על 10 נקודות בממוצע, ואם יתברר כי טענתו נכונה יצטרכו לתגבר את לימודי הלשון.

לצורך בדיקת טענת המנהל נבדק מדגם מקרי של 9 תלמידים שציוניהם בלשון ובספרות מופיעים בטבלה שלהלן:

ציון בלשון	70	78	81	67	90	57	68	84	78
ציון בספרות	90	88	93	89	100	69	80	75	90

א. האם ניתן לטעון ברמת מובהקות 0.05 שהמנהל צודק, בהנחה שהציונים מתפלגים נורמלית? נמק.

ב. בהמשך לסעיף א, תן הערכה לרמת המובהקות המינימלית שעבורה השערת האפס תידחה על סמך תוצאות המדגם.

ג. בדוק את ההשערה שיש הבדל בין הציון בספרות לבין הציון בלשון ברמת מובהקות 0.05 בהנחה שהציונים אינם מתפלגים נורמלית. נמק.



שאלה

להלן ציוני חרדה של 8 סטודנטים לפני ואחרי בחינה בסטטיסטיקה:

	8	7	6	5	4	3	2	1
לפני הבחינה	6	9	8	5	10	7	4	8
אחרי הבחינה	7	7	6	5	8	8	3	7

לאור התוצאות האם ניתן לומר, ברמת מובהקות 0.025 כי רמת החרדה יורדת לאחר הבחינה כאשר לא ניתן להניח שהציונים מתפלגים נורמלית? נמק.



שאלה

נבדק ההבדל בין ציונים בשני מועדי בחינה אצל 7 סטודנטים:

Student(Score)	1	2	3	4	5	6	7
מועד א	60	94	93	81	52	84	99
מועד ב	71	62	73	61	48	37	82

ידוע ש-Score אינו מתפלג נורמלית.

א. ברמת מובהקות 0.05, בדוק את ההשערה ש-Score במועד א נמוך מ-Score במועד ב.

ב. חזור על סעיף א באמצעות מבחן הסימן.



שאלה

חוקרים בחינוך פיתחו תוכנית חדשה בהוראה מתקנת בקריאה, לתלמידים המתקשים בקריאה. התוכנית נוסתה על 8 תלמידים שנבחנו במבחן

קריאה לפני שלמדו בתוכנית החדשה (מבחן א'), ובמבחן נוסף (מבחן ב') לאחר שלמדו לפי התוכנית החדשה. הציונים שהתקבלו בשני המבחנים הם:

	1	2	3	4	5	6	7	8
ציון במבחן א	31	33	30	38	34	28	27	36
ציון במבחן ב	40	39	30	45	32	29	38	42

בדוק את ההשערה שהתוכנית משפרת את הציונים בקריאה ברמת מובהקות 0.05 בהנחה שלא ניתן להניח התפלגות נורמלית.



שאלה

נבדק ההבדל בין ציונים בשני מועדי בחינה אצל 7 סטודנטים:

Student(Score)	1	2	3	4	5	6	7
מועד א	60	94	93	81	52	84	99

ידוע ש-Score אינו מתפלג נורמלית.

א. ברמת מובהקות 0.05, בדוק את ההשערה ש-Score במועד א גבוהה מ-Score במועד ב.

ב. חזור על סעיף א באמצעות מבחן הסימן.



openbook
המרכז לקידום אקדמי



שאלה

מרכז הוראה שבקורס שלו קיימת הנחיה דרך הלוויין ("אופק") החליט לבדוק האם יש הבדל בין הנחיה זו והנחיה רגילה בכיתה. לצורך כך הוא בחר 20 סטודנטים, 10 מהם למדו ב"אופק" ו-10 בהנחיה רגילה. ציוניהם בסוף הקורס נתונים בטבלה הבאה: (לא ניתן להניח התפלגות נורמלית של הציונים)

מספר תלמיד בקבוצתו	1	2	3	4	5	6	7	8	9	10
"אופק"	35	6	18	25	37	48	49	53	81	89
רגילה	34	57	9	64	61	75	91	100	98	93

בדוק את השערת המרכז ברמת מובהקות 0.02,

א. בהנחה כי אלו 10 זוגות של סטודנטים שהותאמו לפי ביצועיהם בקורסים קודמים.

ב. בהנחה כי הסטודנטים נבחרו וחולקו באופן מקרי לשתי קבוצות בכל סעיף ציין את המבחן בו בחרת להשתמש ונמק בחירה זו.



שאלה

בשתי חנויות של רשת גדולה נרשמו המכירות של מעילי רוח (להלן X) במשך מס' ימים

יום	1	2	3	4	5	6	7	8
חנות א	80	78	60	90	64	58	76	84
חנות ב	62	66	60	76	68	56	54	

א. בדוק את ההשערה ששוונות X בחנות א' שונה משוונות X בחנות ב' ברמת מובהקות 0.1, בהנחה ש X מתפלג נורמלית. נמק.

ב. בדוק את ההשערה שהתפלגות X בחנות א' שונה מהתפלגות X בחנות ב' ברמת מובהקות 0.05, כאשר אי אפשר להניח ש X מתפלג נורמלית.



שאלה

במחקר על משך הזמן של ביצוע חישובים מתמטיים בעקבות אכילת ממתקים, הנבדקים חולקו אקראית לשתי קבוצות (קבוצה אחת קיבלה ממתקים והקבוצה השנייה לא קיבלה) לכל אחד מהמשתתפים נמדד משך הזמן של פתרון שאלות מתמטיות. התוצאות מוצגות להלן:

3.7	7.8	4.1	6.23	2.51	5.1	2.7	ללא ממתקים
	2.3	4.45	2.9	3.5	5.3	2.52	עם ממתקים

בדוק ברמת מובהקות 0.05, האם לאכילת ממתקים יש השפעה כלשהי על משך זמן הביצוע של חישובים מתמטיים, כאשר לא ניתן להניח שמשך הזמן מתפלג נורמלית? נמק.



שאלה

קבוצת צרכנים החליטה לבדוק האם יש הבדל בכמות הקפאין בין שני מותגים מובילים של קפה נמס מגורען. לצורך כך נבחרו 8 צנצנות מקריות מכל אחד משני המותגים ונמדדה כמות הקפאין בכל אחת מהצנצנות. הנתונים נמצאו מופיעים בטבלה הבאה:

כמות הקפאין								
82	81	79	84	73	85	77	82	מותג 1
78	73	74	72	81	76	80	75	מותג 2

א. לא ניתן להניח כי כמות הקפאין מתפלגת נורמלית. באיזה מבחן תשתמשו בכדי לבדוק את שאלת קבוצת הצרכנים?

ב. בצעו את המבחן ורשמו את מסקנתכם לגבי שאלת קבוצת הצרכנים, ברמת מובהקות 0.05.



שאלה

נערך ניסוי במטרה לבחון את היעילות של שיטה ליצירת גשמים מלאכותיים באמצעות הזרעת עננים. בניסוי הוזרעו 8 סופות, ואילו 6 סופות אחרות לא הוזרעו. כדי לבדוק את יעילות ההזרעה הושוו כמות המשקעים שהתקבלו משתי קבוצות הסופות.

התוצאות שהתקבלו מוצגות בטבלה הבאה:

7	16	19	28	29	31	31	35	עם הזרעה
		3	5	11	13	25	30	ללא הזרעה

מהי רמת המובהקות הקטנה ביותר שעבורה תתקבל ההשערה שהשיטה מעלה את רמת המשקעים, כאשר לא ניתן להניח התפלגות נורמלית



שאלה

בשני סוגי טלפונים נדגם זמן התגובה של המסך למגע (מילי-שניות):

Obs(Time)	1	2	3	4	5	6	7	8
טלפון א	68.2	72.4	71.5	69.6	68.4	67.8	66.7	68.3
טלפון ב	70.4	66.3	66.9	68.3	67.3	66.3	64.9	

בדוק את ההשערה שהתפלגות Time בטלפון א שונה מהתפלגות Time בטלפון ב ברמת מובהקות 0.05 ,
 כאשר אי אפשר להניח ש- Time מתפלג נורמלית. נמק.



שאלה

בשני כנסים נדגם:

	1	2	3	4	5	6	7	8
כנס א	57	61	55	66	45	50	57	53
כנס ב	43	50	49	57	56	46	59	

בדוק את ההשערה שהתפלגות X בכנס א שונה מהתפלגות X בכנס ב ברמת מובהקות 0.05 ,
 כאשר אי אפשר להניח ש- X מתפלג נורמלית. נמק.



שאלה

בשני מדדים נדגם:

	1	2	3	4	5	6	7	8
מד א	1.15	1.22	1.09	1.12	0.9	1	1.14	1.05
מד ב	1.07	1.01	0.98	0.91	1.13	0.92	1.19	

בדוק את ההשערה שהתפלגות X במד א שונה מהתפלגות X במד ב ברמת מובהקות 0.05 ,
 כאשר אי אפשר להניח ש- X מתפלג נורמלית. נמק.



שאלה

בשני מדרונות נדגם עובי הקרקע (Cm) (להלן X) :

8	7	6	5	4	3	2	1	Thickness
153	157	150	145	166	155	161	157	מדרון א
	159	146	156	157	149	150	143	מדרון ב

בדוק את ההשערה שהתפלגות X במדרון א שונה מהתפלגות X במדרון ב ברמת מובהקות 0.05 ,

כאשר אי אפשר להניח ש-X מתפלג נורמלית. נמק.



שאלה

נערך ניסוי במטרה לבחון את היעילות של שיטה ליצירת גשמים מלאכותיים על-ידי הזרעת עננים.

בניסוי הוזרעו 8 סופות, ואילו 6 סופות אחרות לא הוזרעו

טיפול	1.14	1.02	0.89	0.73	0.68	0.29	0.26	0.07
ביקורת	0.81	0.67	0.23	0.11	0.05	0.03		

בהנחה שלא ניתן להניח התפלגות נורמלית,

בדוק את ההשערה שהשיטה מעלה את רמת המשקעים לאור התוצאות הנ"ל, ברמת מובהקות 0.05.



שאלה

מאמן כושר טוען כי אכילת שוקולד שעה לפני תחילת ריצה מגבירה את מהירות הריצה.

לצורך בדיקה ואישוש טענתו נבחרו שישה עשר אצנים.

תשעה מהם שנבחרו באופן מקרי אכלו שוקולד שעה לפני הריצה.

כל שישה עשר האצנים התמודדו בתחרות ריצה.

מבין התשעה שאכלו שוקולד לפני הריצה - חמישה סיימו בחמשת המקומות הראשונים, אחד סיים במקום השביעי,

שניים במקומות ה- 10 וה- 11 ואחד סיים מקום אחד לפני האחרון (מקום 15).

כל זמני הריצה של שישה עשר המתמודדים היו שונים זה מזה.

מהי מסקנתך? נסח את ההשערות ובדוק ברמת מובהקות 0.05



שאלה

פסיכולוגים טוענים שאנשים הנבחנים במבחנים פסיכומטריים פעם שנייה מגיעים לתוצאות טובות יותר מאלו שלא נבחנו מעולם. לצורך בדיקת הטענה נדגמו באופן מקרי 8 אנשים שמעולם לא נבחנו במבחן פסיכומטרי, וכן 7 אנשים אחרים שנבחנו כבר פעם בעבר. כל 15 הנדגמים נבחנו במבחן פסיכומטרי וקיבלו את הציונים הבאים:

ציוני הנבחנים לראשונה	664	600	575	540	580	680	510	595
ציוני הנבחנים בעבר	675	640	724	570	700	680	680	

האם התוצאות מאששות את טענת הפסיכולוגים, כשלא ניתן להניח התפלגות נורמלית, ברמת מובהקות 0.05 ? נמק



שאלה

לבדיקת השפעתן של דרגות חוסר שינה שונות על מספר טעויות בביצוע תפקיד זכירה מסוים נדגמו שתי קבוצות בלתי תלויות של סטודנטים. קבוצה א' סבלה מחוסר שינה של 24 שעות וקבוצה ב' מחוסר שינה של 72 שעות. כל נבדק ביצע את תפקיד הזכירה ומספר הטעויות נרשם בטבלה הנתונה:

קבוצה א'	7	5	3	4	1	2	6	4	6
קבוצה ב'	9	3	5	6	2	9	9	6	8

כאשר לא ניתן להניח שמספר הטעויות מתפלג נורמלית, האם ניתן לומר שיש לדרגות חוסר השינה השפעה על מספר הטעויות בביצוע תפקיד זכירה?
נסח את ההשערות, רשום את ההנחות הנדרשות לביצוע המבחן הסטטיסטי המתאים ובדוק ברמת מובהקות 0.1



שאלה

בשני שכונות נדגם שיעור ההצבעה באזורים סטטיסטיים (אחוזים):

SA(Rate)	1	2	3	4	5	6	7	8
שכונה א	58.2	62.4	51.5	59.6	68.4	67.8	66.7	68.3
שכונה ב	70.4	66.3	66.9	68.3	67.3	66.3	64.9	

בדוק את ההשערה שהתפלגות Rate בשכונה א שונה מהתפלגות Rate בשכונה ב ברמת מובהקות 0.05 ,
כאשר אי אפשר להניח ש-Rate מתפלג נורמלית. נמק.



שאלה

פסיכולוג ביקש לברר אם תרופה מסוימת יכולה לעזור בשליטה על התנהגות תוקפנית אצל חולים.
9 חולים קיבלו את התרופה ו-9 אחרים קיבלו פלצבו.
לאחר מכן כל הנבדקים עברו מבחן שבדק שליטה על התנהגות תוקפנית וקיבלו את הציונים הבאים
(ציון גבוה מעיד על התנהגות תוקפנית רבה):

תרופה	8	7	10	13	1	14	18	17	16
פלצבו	11	16	13	10	16	12	20	21	25

האם לאור התוצאות ניתן לומר, ברמת מובהקות 0.05, שהתרופה עוזרת בשליטה על התנהגות תוקפנית אצל חולים,
בהנחה שלא ניתן להניח התפלגות נורמלית של הציונים? נמק.



שאלה

בשתי רשתות שיווק נדגם מחיר של מוצר בסניפים שונים (שקלים):

Net(Price)	1	2	3	4	5	6	7	8
רשת א	582	624	515	596	684	678	667	583
רשת ב	704	663	669	683	673	663	620	

בדקו את ההשערה שתוחלת Price ברשת א שונה מתוחלת Price ברשת ב ברמת מובהקות 0.05,
כאשר לא ניתן להניח ש-Price מתפלג נורמלית. נמקו.



שאלה

טיפול חדש בחולים לאחר ניתוח מושווה לטיפול סטנדרטי. נלקח מדגם מקרי של 12 חולים, שחולקו באקראי
ל-2 קבוצות ונמדד הזמן בימים עד להחלמתו של כל אחד ואחד מהם. להלן הנתונים:

טיפול סטנדרטי	24	48	40	36	32	30
טיפול חדש	25	37	34	29	28	26

בהנחה שהזמן להחלמה מתפלג נורמלית:

א. בדוק האם יש הבדל בין 2 הטיפולים ברמת מובהקות 0.05. נסח את ההשערות ופרט את ההנחות הדרושות.

ב. בהמשך לסעיף א', מה תהיה המסקנה ברמת מובהקות 0.05 אם האלטרנטיבה היא שהטיפול החדש עדיף? נמק.

ג. חזור לסעיף א' אם לא ניתן להניח שזמן ההחלמה מתפלג נורמלית.



שאלה

בשני מושבים נדגם יכול התפוחים במטעים (טון):

Plant(Amount)	1	2	3	4	5	6	7	8
מושב א	58	62	51	59	68	67	66	58
מושב ב	70	66	66	68	67	66	62	

בדקו את ההשערה שהתפלגות Amount במושב א שונה שהתפלגות Amount במושב ב ברמת מובהקות 0.05, כאשר אי אפשר להניח ש- Amount מתפלג נורמלית. נמקו.



שאלה

בשני כבישים נדגם זמן ההמתנה בפקקים (דקות):

Jam (Time)	1	2	3	4	5	6	7	8
כביש א	50	54	44	51	59	58	57	50
כביש ב	63	59	59	61	60	59	58	

בדקו את ההשערה שהתפלגות Time בכביש א שונה מהתפלגות Time בכביש ב ברמת מובהקות 0.05, כאשר אי אפשר להניח ש- Time מתפלג נורמלית. נמק.



שאלה

בשני אזורים נדגם מרחק הנסיעה השבועי (ק"מ):

8	7	6	5	4	3	2	1	Region(Distance)
450	513	522	531	459	396	486	450	אזור א
	522	531	540	549	531	531	567	אזור ב

א. בדוק את ההשערה שהתפלגות Distance באזור א שונה מהתפלגות Distance באזור ב ברמת

מובהקות 0.05, כאשר אי אפשר להניח ש-Distance מתפלג נורמלית. נמק.

ב. אם Distance קטן מ- 530 ק"מ אז המרחק נחשב קצר. בדקו את ההשערה ששיעור המרחקים הקצרים גדול יותר באזור א בהשוואה לאזור ב.

openbook
המרכז לקידום אקדמי



שאלה

פסיכולוג א' ביצע סקר, שהקיף 14 משפחות של חיילים, ששבוע קודם לכן השתחררו משירות מילואים ממושך. הוא בדק את מידת ההבנה החינוכית של האב לבנו בהשוואה לזו של האם. התוצאות הראו, שב- 11 משפחות עלתה ההבנה החינוכית של האם על זו של האב, וב- 3 משפחות – ההפך. מה אפשר להסיק מתוצאות אלה ברמת מובהקות 2.5%?



שאלה

15 זוגות נשואים טופלו בשני סוגי טיפולים:

טיפול פרטני לאחד מבני הזוג וטיפול קבוצתי לבן הזוג השני.

פסיכולוגים נדרשו לבצע הערכות להצלחת הטיפול לכל אחד מבני הזוג בנפרד ולציין ב- (+) אם הטיפול הפרטני היה מוצלח יותר וב- (-) אם הטיפול הקבוצתי היה מוצלח יותר.

לבדיקת טענת הפסיכולוגים, שהטיפול הפרטני עדיף, נקבע כי אם יהיו לפחות 11 זוגות, שבהם הטיפול הפרטני מוצלח יותר, השערת האפס תידחה.

א. מהי רמת המובהקות של המבחן המוצע?

ב. מהי עצמת המבחן אם ידוע כי טיפול פרטני יעיל יותר מטיפול קבוצתי ב- 80% מהמקרים.

ג. אם נמצא כי רק אצל 3 זוגות, הטיפול הקבוצתי היה מוצלח יותר מהטיפול הפרטני, מה תהיה מסקנתך ברמת מובהקות 0.05?



שאלה

חוקר במשרד התחבורה טוען כי נהגים ותיקים מבצעים פחות עבירות תנועה מנהגים חדשים.

לצורך בדיקת הטענה נלקח מדגם מקרי של 16 נהגים, ונספרו מספר עבירות התנועה שלהם בשנה הראשונה לקבלת הרישיון וכעבור עשר שנים.

התוצאות הראו כי אצל 11 נהגים מספר העבירות בשנה העשירית היה קטן ממספר העבירות בשנה הראשונה

ואילו אצל 5 נהגים מספר העבירות בשנה העשירית היה גדול ממספר העבירות בשנה הראשונה לקבלת

הרישיון. האם התוצאות מאששות את טענת החוקר ברמת מובהקות של 5%?

ב. חוקר במשרד התחבורה טוען, שנהגים ותיקים מבצעים פחות עבירות תנועה מנהגים חדשים. לצורך בדיקת הטענה נלקח מדגם מקרי של 16 נהגים, ונספרו מספר עבירות התנועה שלהם בשנה הראשונה לקבלת הרישיון (שבה היו נהגים חדשים) וכעבור עשר שנים כשהיו כבר נהגים ותיקים. התוצאות הראו שאצל 12 נהגים מספר העבירות בשנה העשירית היה קטן ממספר העבירות בשנה הראשונה ואילו אצל 4 נהגים מספר העבירות בשנה העשירית היה גדול ממספר העבירות בשנה הראשונה לקבלת רישיון הנהיגה.

1. מה המבחן בו תשתמש לבדיקת טענת החוקר?

2. האם תקבל את טענת החוקר ברמת מובהקות 0.05?



שאלה

לבדיקת הטענה, שהחדרה לשוק של נוזל כביסה חדש באמצעות חלוקת דוגמיות מעלה את מידת הצריכה במוצר,

נשאלו 50 נבדקים על מידת השימוש במוצר לפני חלוקת הדוגמיות, ושבוע לאחר השימוש בדוגמיות. התוצאות שהתקבלו היו:

21 אנשים השתמשו במוצר לפני קבלת הדוגמית וגם לאחר החלוקה.

7 אנשים השתמשו במוצר לפני קבלת הדוגמית אך לא השתמשו בו לאחר החלוקה.

14 אנשים לא השתמשו במוצר לפני קבלת הדוגמית אך השתמשו בו לאחר החלוקה.

ו- 8 אנשים לא השתמשו במוצר לפני קבלת הדוגמית וגם לא לאחר מכן

מהי מסקנת המחקר ברמת מובהקות 0.05? נמק.



שאלה

כדי לבדוק אם תרגול נוסף מעלה את סיכויי ההצלחה בקורס בסטטיסטיקה נדגמו באופן מקרי סטודנטים מבין המשתתפים בקורס ונמצא כי:

9 סטודנטים השתתפו בתרגול הנוסף ועברו את הבחינה בהצלחה,

סטודנט אחד השתתף בתרגול הנוסף ונכשל,

2 סטודנטים לא השתתפו בתרגול הנוסף ועברו את הבחינה בהצלחה

ו- 3 סטודנטים לא השתתפו בתרגול הנוסף ונכשלו בבחינה.

האם לאור התוצאות ניתן לומר ברמת מובהקות 0.025 שתרגול נוסף מעלה את סיכויי ההצלחה בקורס? נמק



שאלה

לילדים עם בעיות לימודיות יש לעיתים סימפטומים הדומים לאלה של ילדים עם מחלה מסוימת. עובדה זו גרמה לחוקרים לחשוב שתרופה למחלה תעזור גם לילדים עם הבעיות הלימודיות. לצורך בדיקת ההשערה נבדקו 10 ילדים עם בעיות לימודיות כאלו. במשך שלושה שבועות הם קיבלו תרופת דמי (פלסבו) ובמשך שלושה שבועות נוספים קיבלו את התרופה האמיתית. הם נבחנו במבחני IQ פעמים, פעם לאחר שלושת השבועות הראשונים (מבחן I) פעם לאחר שלושת השבועות שבהם קיבלו את התרופה האמיתית (מבחן II)

תוצאות המבחנים מוצגות בטבלה הבאה:

96	90	104	115	111	102	95	106	106	97	מבחן I
----	----	-----	-----	-----	-----	----	-----	-----	----	--------

מבחן II	113	113	101	119	111	122	121	102	110	126
---------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

א. מהי רמת המובהקות המינימלית שעבורה יוחלט שהתרופה עוזרת על סמך מבחן הסימן ? נמק

ב. בהמשך לסעיף א' מהי המסקנה ברמת מובהקות 0.05 ?



שאלה

במפעל מסוים ביקשו לבדוק את ההשערה שתפוקת העובדים הממוצעת גדולה יותר בשעות הבוקר מאשר בשעות אחה"צ. לשם כך נלקח מדגם מקרי אשר כלל 8 מעובדי המפעל והתקבלו התוצאות הבאות

עובד	1	2	3	4	5	6	7	8
תפוקה בבוקר	5.1	5.3	6	5.8	5.5	5.1	5.4	5.7
תפוקה אחה"צ	4.1	4.8	5	6.3	4.5	5	5.2	5.8

מהי מסקנתך ברמת מובהקות 0.01 ? נמק.

בעזרת מבחן הסימן



שאלה

לבחינת השפעת ייבוש במייבש כביסה על איכות בדים נבחנו 12 זוגות של מגבות, כל זוג מגבות היה מאותו דגם ובאותו צבע.

עבור כל זוג מגבות - אחת יובשה במייבש והשנייה באוויר החופשי. לאחר הייבוש נבחנה איכות הבד.

בשני זוגות של מגבות הייתה איכות הבד גבוהה יותר לאחר ייבוש במייבש

בעוד שבעשרה זוגות של מגבות הייתה איכות הבד גבוהה יותר לאחר ייבוש באוויר החופשי.

בדוק את השפעת הייבוש במייבש ברמת מובהקות 0.05 בעזרת מבחן סטטיסטי מתאים.



שאלה

כדי לבדוק אם חומר חדש המוצע לייצור עדשות מגע מקטין את גירוי העין על ידי העדשה,

מרכיבים ל- 23 נבדקים עדשה סטנדרטית בעין אחת, ועדשה העשויה מהחומר החדש בעין השנייה.

אצל 14 מהנבדקים, נמצא שעצמת הגירוי של העין שקיבלה את העדשה הסטנדרטית הייתה גדולה יותר,

אצל 4 נבדקים לא היה הבדל בעוצמת הגירוי ואצל השאר נמצא שעוצמת הגירוי של העין שקיבלה את העדשה הסטנדרטית הייתה קטנה יותר.

מהי המסקנה ברמת מובהקות 0.05? נמק.



שאלה

כדי לבדוק שביעות רצון משירות בבנק מסוים הועבר שאלון שביעות רצון ללקוחות קבועים.

לאור התוצאות החליטה הנהלת הבנק להעביר סדנת שיפור איכות השירות לעובדים. לאחר הסדנה הועברו

שאלוני שביעות רצון לאותם הלקוחות

ונמצא כי 17 טענו שרמת השירות השתפרה,

8 טענו שרמת השירות ירדה

ו- 10 טענו שלא חלק שינוי ברמת השירות.

האם לאור התוצאות ניתן לומר שסדנת שיפור איכות השירות עזרה, ברמת מובהקות 0.025? נמק.

openbook
המרכז לקידום אקדמי



שאלה

דוגמא: בניסוי הוטלה קובייה 300 פעמים ונספרו שכיחות התוצאות. התקבלה הטבלה הבאה:

תוצאה	1	2	3	4	5	6
שכיחות	54	36	30	50	70	60

האם ניתן לומר שהקובייה הוגנת בדוק ר"מ 0.05



שאלה

שתי קוביות הוטלו 180 פעם. בכל הטלה נרשם ההפרש בין התוצאות בערכו המוחלט. לאחר מכן נרשמה התפלגות השכיחויות של הפרשים מוחלטים אלו. התקבלו הנתונים הבאים:

הפרש	0	1	2	3	4	5	סה"כ
מספר הטלות	28	45	55	32	15	5	180

א. האם התפלגות התוצאות הנתונה מתיישבת עם ההשערה, שהקוביות הוגנות? בדוק ברמת מובהקות 0.05.

ב. האם ניתן להגיע לאותה מסקנה גם עבור רמת מובהקות הנמוכה מ-0.05? הסבר.



שאלה

מעונינים לבדוק האם גובה הוא משתנה מקרי המתפלג נורמלית עם תוחלת של 178 וס"ת 4. לשם כך נדגמו 1000 אנשים והתקבלו התוצאות הבאות:

גובה בס"מ	מתחת 170	170-174	174-178	178-182	182-186	מעל 186
מספר נבדקים	25	138	315	385	102	35

בדוק את ההשערה בר"מ 0.05?



שאלה

רצו לבדוק האם סיכוי ללידת בן הוא 0.5

בקרב 300 משפחות עם 2 ילדים התקבלו התוצאות הבאות:

	שני בנים	בן אחד ובת אחת	שתי בנות
מספר משפחות	80	152	68

האם ניתן להגיד ברמת מובהקות 0.01 כי התוצאות מעידות על כך שהסיכוי ללידת בן הוא 0.5? נמק.



שאלה

למשחק קופסא מצורפת קוביית משחק הצבועה בשלושה צבעים לפי הפירוט הבא:

שלוש פאות צבועות באדום, שתיים בצהוב ואחת בכחול.

עפר רצה לבדוק האם הקובייה הוגנת ולשם כך הטיל אותה 60 פעמים ורשם את הצבע שהתקבל בכל הטלה.

סיכום הנתונים שהתקבלו הוא:

הצבע	אדום	צהוב	כחול
מספר הפעמים שהתקבל	33	12	15

בדקו האם הקובייה הוגנת ברמת מובהקות 0.05.

ציינו מהו המבחן בו אתם משתמשים



שאלה

בעקבות המצב הכלכלי והמחאה החברתית התרחב מאוד נושא המותגים הפרטיים של רשתות השיווק בסקר נשאלו 1000 בוגרים האם הם קונים את המותגים הפרטיים של הרשת או את המותג הכללי. הנתונים מובאים בטבלה הבאה:

	מותג כללי	מותג פרטי
נשים	200	250
גברים	200	350

בדקו, ברמת מובהקות 0.025,

האם אין קשר בין המגדר לבין הבחירה במותג פרטי או כללי. ציינו מהו המבחן בו אתם משתמשים



שאלה

101 שופטים נתבקשו לבחור 3 נציגים לתחרות בינלאומית מתוך קבוצת זמרים שכוללת 7 מקצוענים ו- 5 חובבים.

מנהל תאגיד השידור קיבל דו"ח אודות התפלגות מספר המקצוענים שנבחרו (להלן X) ע"י השופטים:

x	0	1	2	3
$f(x)$	10	35	45	11

א. ברמת מובהקות 0.05, בדוק את השערת המנהל, שהנציגים נבחרו עם החזרה. (*)

ב. ברמת מובהקות 0.05, בדוק את השערת המנהל, שהנציגים נבחרו ללא החזרה.



שאלה

100 עובדים נתבקשו לבחור 3 נציגים מתוך מועצת העובדים שכוללת 7 אקדמאים ו- 5 שאינם אקדמאים.

x	0	1	2	3
f(x)	8	37	45	10

מנהלת הארגון קיבלה דו"ח אודות התפלגות מספר האקדמאים שנבחרו (להלן X) ע"י העובדים:

א. בדוק את השערת המנהלת, שהנציגים נבחרו עם החזרה (*), ברמת מובהקות 0.05.

ב. בדוק את השערת המנהלת, שהנציגים נבחרו ללא החזרה, ברמת מובהקות 0.05.



שאלה

103 צופים נתבקשו לבחור באמצעות הגרלה 3 נציגים לתחרות בינלאומית מתוך קבוצת שחקנים שכוללת 8 מקצוענים ו-4 חובבנים. מנהל תאגיד השידור קיבל דו"ח אודות התפלגות מספר המקצוענים שנבחרו (להלן X) ע"י הצופים:

x	0	1	2	3
f(x)	2	24	60	17

א. ברמת מובהקות 0.05, בדוק את השערת המנהל, שהנציגים נבחרו עם החזרה (*).

ב. ברמת מובהקות 0.05, בדוק את השערת המנהל, שהנציגים נבחרו ללא החזרה.



שאלה

הטילו זוג קוביות משחק 72 פעמים והתקבלו התוצאות הבאות:

סכום התוצאות בכל הטלה	2-5	6-7	8-9	10-12
מספר ההטלות	15	23	26	8

האם התוצאות מעידות על קוביות הוגנות ברמת מובהקות 0.05? רשום מהו המבחן בו אתה משתמש.



שאלה

סדרת הגמר של ה-N.B.A מתנהלת בשיטה "הטוב מבין שבע". כלומר, מתקיימים משחקים עד אשר אחת מן הקבוצות מנצחת ב-4 משחקים. להלן נתונים על מספר המשחקים ששוחקו בסדרת הגמר במשך 50 שנה.

מספר המשחקים	מספר השנים
4	9
5	11

6	8
7	22

בדוק את ההשערה כי כל משחק בסדרת הגמר הוא ניסוי ברנולי בלתי תלוי עם הסתברות שווה לניצחון של כל אחת מהקבוצות, ברמת מובהקות 0.05.

שאלה

בבית חולים אספו את הנתונים הבאים על מספר הלידות בארבעה רבעונים עוקבים:

רבעון:	ינואר-מרץ	אפריל-יוני	יולי-ספטמבר	אוקטובר-דצמבר
מס' לידות	110	67	53	40

א. בחנו את ההשערה שיחס השכיחויות הוא 1:1:1:2 בהתאמה, בר"מ של 0.05?

ב. בחנו את ההשערה כי ברבעון הראשון מתקיימות 30% מהלידות, כנגד האלטרנטיבה שאין זה כך ברמת מובהקות של 0.01? (הסבר אם ניתן לבדוק זאת לפי מבחן סטטיסטי אחר).

שאלה

בבית חולים נאספו נתונים על מספר הלידות בשנה ונרשמו בטבלה שלהלן:

חודש:	ינואר-מרץ	אפריל-יוני	יולי-ספטמבר	אוקטובר-דצמבר
מס' לידות	110	57	53	80

לאור התוצאות, האם ניתן לומר שמספר הלידות בחודשים ינואר-מרץ כפול ממספר הלידות בכל רבעון אחר, ברמת מובהקות 0.05? נמק.

שאלה

בלוח ספרות אקראיות בן 250 ספרות נמצאה ההתפלגות הבאה של הספרות:

ספרה	0	1	2	3	4	5	6	7	8	9
שכיחות	17	31	29	18	14	20	35	30	20	36

בדקו ברמת מובהקות 1% אם אמנם הספרות מופיעות באופן אקראי.

שאלה

קופסא מכילה כדורים שחורים, לבנים, אדומים וירוקים. דגמו מהקופסא עם החזרה באופן מקרי 100 כדורים והתקבלו התוצאות הבאות:

צבע	שחור	לבן	אדום	ירוק
מספר כדורים	14	20	15	51

בדקו את ההשערה שבקופסא כדור שחור, כדור לבן, כדור אדום ו- 2 ירוקים, ברמת מובהקות 0.025.



שאלה

כדי לבדוק אם סביבון מדגם חדשני הוא סימטרי סובבו את הסביבון 400 פעמים. התקבלו התוצאות הבאות: 110 פעמים התוצאה "נס", 140 פעמים התוצאה "גדול", 75 פעמים התוצאה "היה" ו- 75 פעמים התוצאה "פה".

האם לאור התוצאות ניתן לומר ברמת מובהקות 0.025 שהסביבון הוגן? נמק.



שאלה

הועלה חשד שסביבון חדשני איננו הוגן. לבדיקת הטענה הסביבון הוטל 80 פעם ונמצא כי 24 פעם התקבלה התוצאה "נס", 8 פעם התוצאה "גדול", 30 פעם התוצאה "היה" ו- 18 פעם התוצאה "פה".
בדוק את ההשערה בר"מ 0.05



שאלה

עיתון יצא במבצע שמטרתו להגדיל את מספר המודעות המתפרסמות בו.
על פי תנאי המבצע, לקוח שלא הצליח למכור את מכוניתו/דירתו, מודעתו תתפרסם שוב- בחינם!
הפרסום הנוסף מוגבל ל-2 פעמים. כלומר בסה"כ המודעה תופיע לכל היותר 3 פעמים.
בבדיקה של 150 לקוחות מקריים שפרסמו מודעות למכירת מכונית/דירה התקבלו הנתונים הבאים:

52 ✓ לקוחות מכרו לאחר מודעה אחת בלבד.
38 ✓ לקוחות מכרו לאחר פרסום 2 מודעות.
27 ✓ לקוחות מכרו לאחר פרסום 3 מודעות.
שאר הלקוחות לא הצליחו למכור. ✓

בדוק ברמת מובהקות 0.05 את ההשערה שסיכוי ההצלחה של מודעה (כלומר מכירה) הוא 0.3 ואינו תלוי במספר הפעמים שהמודעה כבר פורסמה בעיתון. נמק.



שאלה

בסקר של 200 לקוחות של בנק מסוים התקבלו התוצאות הבאות:

לא משתמש בכרטיס אשראי	משתמש בכרטיס אשראי אחד	משתמש ביותר מכרטיס אשראי אחד

שכיר	60	50	20
עצמאי	25	25	20

בדוק ברמת מובהקות 0.05 את ההשערה שאין קשר בין אופי השימוש בכרטיס אשראי לבין היות הלקוח שכיר או עצמאי.



שאלה

במפקד שכלל את כל האוכלוסייה, ובו הפרטים נשאלו לגבי התוכנית הכלכלית החדשה של שר האוצר, עלה כי 80% מסך האוכלוסייה תמכו בתוכנית, 8% התנגדו, ול-12% לא הייתה דעה מגובשת. כעבור חודשיים נבחר מדגם של 600 נשאלים, ובו הביעו 466 את תמיכתם בתוכנית, 51 הביעו התנגדות, ולשאר לא הייתה דעה מגובשת. האם יש יסוד להסיק, שחל שינוי משמעותי בתמיכת הציבור בתוכנית הכלכלית, ברמת מובהקות 0.05? נמק.



שאלה

110 גולשים נתבקשו לבחור באמצעות הגרלה 3 נציגים לתחרות בינלאומית מתוך קבוצת מועמדים שכוללת 8 בנות ו-4 בנים. מנהל האתר קיבל דו"ח אודות התפלגות מספר הבנות שנבחרו (להלן X) ע"י הגולשים:

X	0	1	2	3
$f(x)$	4	28	61	17

- ברמת מובהקות 0.05, בדוק את השערת המנהל, שהנציגים נבחרו עם החזרה. (*)
- ברמת מובהקות 0.05, בדוק את השערת המנהל, שהנציגים נבחרו ללא החזרה.
- במקרה זה אפשר שנציג ישחק יותר מפעם אחת.



שאלה

עיתונאי פוליטי טוען כי בבחירות הבאות תהיה פרופורציה המצביעים למפלגות השונות כדלהלן:

0.4 יצביעו עבור מפלגה A, 0.3 יצביעו עבור מפלגה B וכל השאר יתחלקו באופן שווה בין מפלגות C ו-D.

כדי לבדוק את השערת העיתונאי נלקח מדגם של 200 אנשים.

כל משתתף במדגם נשאל על אופן הצבעתו והתקבלו התוצאות הבאות: 85 טענו כי יצביעו עבור מפלגה A,

50 טענו כי יצביעו עבור מפלגה B, 30 עבור מפלגה C והשאר עבור מפלגה D.

האם לאור התוצאות השערת העיתונאי מתקבלת ברמת מובהקות 0.05? נמקו.



שאלה

עיתונאי פוליטי טוען כי בבחירות הבאות תהיה פרופורציה המצביעים למפלגות השונות כדלהלן:

0.4 יצביעו עבור מפלגה A, 0.3 יצביעו עבור מפלגה B וכל השאר יתחלקו באופן שווה בין מפלגות C ו-D.

כדי לבדוק את השערת העיתונאי נלקח מדגם של 200 אנשים.

כל משתתף במדגם נשאל על אופן הצבעתו והתקבלו התוצאות הבאות: 85 טענו כי יצביעו עבור מפלגה A,

50 טענו כי יצביעו עבור מפלגה B, 30 עבור מפלגה C והשאר עבור מפלגה D.

א. האם לאור התוצאות השערת העיתונאי מתקבלת ברמת מובהקות 0.05? נמקו.

ב. עיתונאי אחר טוען כי שיעור ההצבעה עבור מפלגה B יהיה שונה מ-0.3.

האם לאור תוצאות המדגם ניתן לקבל את טענתו ברמת מובהקות 0.01?

נמק בעזרת מבחן טיב התאמה.

האם ניתן לבדוק את ההשערה בעזרת מבחן סטטיסטי אחר שלמדת?



שאלה

מכירות השוקולד התחלקו עד כה בין החברות השונות בשוק באופן הבא:

לחברה A נתח שוק של 67%, לחברה B 24% ולכל שאר החברות נתח שוק של 9%.

עם כניסתה של חברה חדשה לשוק נבדקה התפלגות המכירות במדגם מקרי של 120 חפיסות נמכרות.

נמצא כי מתוך 120 החפיסות הנמכרות, 78 היו מתוצרת חברה A, 24 מתוצרת חברה B והיתר משאר החברות שבשוק.

האם ניתן לומר, כי חל שינוי מובהק בחלוקת שוק השוקולד בין החברות השונות, ברמת מובהקות 0.05? נמקו.



שאלה

110 דיירים נתבקשו לבחור באמצעות הגרלה 3 נציגים לוועד הבית מתוך קבוצת מועמדים שכוללת 8 בעלי בתים ו- 4 שוכרים.

x	0	1	2	3
f(x)	5	28	61	16

יו"ר הוועד היוצא קיבל דו"ח אודות התפלגות מספר בעלי הבתים שנבחרו (להלן X) ע"י הדיירים:

א. ברמת מובהקות 0.05, בדוק את השערת היו"ר, שהנציגים נבחרו עם החזרה. (*)

* במקרה זה אפשר שנציג ישחק יותר מפעם אחת.

ב. ברמת מובהקות 0.05, בדוק את השערת היו"ר, שהנציגים נבחרו ללא החזרה.



שאלה

חברת כוח אדם טוענת שלמודעת "דרושים" פונים אנשים במשך ימי השבוע לפי היחס הבא:

ביום א' פונים פי 5 מאשר ביום ה', ביום ב' פי 4 מאשר ביום ה', ביום ג' פי 3 מאשר ביום ה' וביום ד' פי 2 מאשר ביום ה'.

(אין פניות בימי שישי ושבת)

בשבוע מסוים ענו למודעה אחת 180 איש: 84 מהם פנו ביום א', 62 ביום ב', 22 ביום ג', 10 ביום ד' ו- 2 ביום ה'.

האם לאור התוצאות, ברמת מובהקות 0.05, ניתן לומר כי טענת החברה אינה נכונה?

נמקו



שאלה

רוצים לבדוק האם הנחיה מוגברת מעלה את הסיכוי להצליח בבחינה בקורס סטטיסטיקה ב'. לשם כך נבחרו 6 תלמידים שלמדו בהנחיה מוגברת ומתוכם 5 עברו את הבחינה. כמו כן, 5 תלמידים שלא עברו הנחיה מוגברת נבחרו, כשמתוכם 3 עברו את הבחינה. בחן את ההשערה שהנחיה מוגברת משפרת את הסיכוי להצליח בבחינה בר"מ 0.05.



שאלה

כדי לבדוק אם חיסון חדש לכלבים מסייע בהפחתת ההדבקות במחלה מסוימת, נערך ניסוי שבו 10 כלבים קיבלו חיסון ו-8 לא קיבלו חיסון. מבין הכלבים שחוסנו כלב אחד נדבק במחלה, ומבין הכלבים שלא חוסנו 5 נדבקו במחלה. מהו המבחן הסטטיסטי המתאים ביותר לבדיקת השערת המחקר, נסח את ההשערות והסק מסקנה ברמת מובהקות 0.05? נמק



שאלה

לבדיקת יעילותו של מסיר כתמים חדש, השתמשו בו להסרת כתמים קשים מ-10 חולצות. את התוצאות השוו מול מסיר כתמים ישן, בו השתמשו להסרת כתמים מ-15 חולצות אחרות. התוצאות הראו שב-2 חולצות שטופלו במסיר הכתמים החדש וב-7 חולצות שטופלו במסיר הכתמים הישן, נותרו הכתמים.

האם ניתן לומר שמסיר הכתמים החדש יעיל יותר מהישן בר"מ 0.05?



שאלה

לבדיקת יעילותו של מסיר כתמים חדש, השתמשו בו להסרת כתמים קשים מ-10 חולצות. את התוצאות השוו מול מסיר כתמים ישן, בו השתמשו להסרת כתמים מ-15 חולצות אחרות. התוצאות הראו שב-2 חולצות שטופלו במסיר הכתמים החדש וב-7 חולצות שטופלו במסיר הכתמים הישן, נותרו הכתמים.

א. בדקו האם מסיר הכתמים החדש יעיל יותר מהישן, ברמת מובהקות 0.05.

ב. מהו הערך המקסימלי של מספר החולצות, שטופלו במסיר הכתמים החדש ונותרו בהן כתמים, שעבורו נסיק שהמסיר החדש יעיל יותר, בר"מ 0.05?



שאלה

לבדיקת יעילותו של טיפול למניעת אבנית, ניתן טיפול ל-10 קומקומים שנבחרו באופן מקרי מיצרנים שונים, ו-15 קומקומים אחרים, שגם הם נבחרו אקראית, שימשו כקבוצת ביקורת.

בכל קומקום הרתיחו מים 10 פעמים.

התוצאות הראו, שב-2 קומקומים מבין הקומקומים שעברו את הטיפול, ו ב-7 קומקומים שלא עברו את הטיפול, נוצרה אבנית.

א. נסח את ההשערות וחשב את ההסתברות לקבל תוצאה כפי שנתקבלה או קיצונית ממנה תחת השערת האפס.

ב. האם, לאור התוצאות, ניתן לומר שהטיפול יעיל, ברמת מובהקות 0.05?



שאלה

כדי לבדוק אם תרגול נוסף משפר את הצלחת הסטודנטים בקורס "מבוא לסטטיסטיקה" נבחר מדגם מקרי של סטודנטים המשתתפים בקורס ובו נמצא כי :

8 סטודנטים השתתפו בתרגול ועברו את הבחינה

2 סטודנטים השתתפו בתרגול ונכשלו בבחינה

2 סטודנטים לא השתתפו בתרגול ועברו את הבחינה

3 סטודנטים לא השתתפו בתרגול ונכשלו בבחינה

א. נסח את ההשערות וחשב את ההסתברות לקבל תוצאה כפי שנתקבלה או קיצונית ממנה תחת השערת האפס.

ב. האם, לאור התוצאות, ניתן לומר שתרגול נוסף משפר את הצלחת הסטודנטים בקורס, ברמת מובהקות 0.05?

ב. התוצאה גדולה מ-0.05 ולכן לא ניתן לדחות את השערת האפס ולא ניתן לומר שהטיפול יעיל.



שאלה

בכדי לבדוק, האם חיסון חדש למחלת השפעת מסייע להפחתת ההדבקות בשפעת נערך הניסוי הבא:

20 אנשים נבחרו באופן מקרי וחולקו אקראית לשתי קבוצות.

10 קיבלו חיסון ו-10 לא קיבלו חיסון.

מבין האנשים שחוסנו, 1 נדבקו במחלה,

ומבין האנשים שלא חוסנו 7 נדבקו במחלה.

האם ניתן לטעון ברמת מובהקות 0.05 שהחיסון מסייע במניעת ההדבקות במחלה? נמק.



שאלה

טיפול רפואי מסוים נוסה על קבוצה של 7 מתוך 12 חולים, שנבחרו לניסוי. מתוך אלו שקיבלו טיפול 5 הברואו לאחר שנה, ואילו מתוך אלו שלא קיבלו טיפול הברואו רק שניים.

בדוק את ההשערה, שלטיפול אין כל השפעה, לעומת ההשערה האלטרנטיבית, שהטיפול משפר ברמת מובהקות של 0.1.



שאלה

נערך ניסוי במטרה לבחון את היעילות של שיטה ליצירת גשמים מלאכותיים על-ידי הזרעת עננים. בניסוי הוזרעו 8 סופות, ואילו 6 סופות אחרות לא הוזרעו. התוצאות הראו שב- 7 מבין הסופות שהוזרעו וב- 3 מבין הסופות שלא הוזרעו כמות המשקעים עלתה. האם לאור התוצאות ניתן לומר שהשיטה מעלה את רמת המשקעים? נסח את ההשערות ובדוק ברמת מובהקות 0.1.



שאלה

בכדי לבדוק האם טיפול חדש מסייע בהפחתת חרדת בחינות, נבחרו באקראי 22 סטודנטים. 10 מהם טופלו והשאר לא טופלו. מבין המטופלים אצל 2 לא חלה הפחתה של החרדה, ומבין הבלתי-מטופלים אצל 7 לא חלה הפחתה של החרדה. ברמת מובהקות 0.05, בדקו את הטענה שהטיפול יעיל



שאלה

בכדי לבדוק האם טיפול חדש מסייע בהפחתת אלימות, נבחרו באקראי 20 נערים. 10 מהם טופלו והשאר לא טופלו. בקרב המטופלים הושגו 8 הצלחות. ומבין הבלתי מטופלים הושגו 3 הצלחות. ברמת מובהקות 0.05, בדקו את הטענה שהטיפול יעיל.



שאלה

בכדי לבדוק האם שיטה חדשה מסייעת בהפחתת גלישה ברשת, נבחרו באקראי 21 נערים. 10 מהם התנסו בשיטה החדשה והשאר לא התנסו בה. בקרב המתנסים הושגו 7 הצלחות. בקרב האחרים הושגו 4 הצלחות. ברמת מובהקות 0.05, בדקו את הטענה שהשיטה יעילה.



שאלה

הועלתה טענה כי פעילות גופנית מורידה את רמת הטריגליצרידים בדם.

לצורך בדיקת הטענה נמדדה רמת הטריגליצרידים בדם של שמונה אנשים פעם אחת לפני שהחלו בפעילות גופנית מסודרת ופעם שנייה לאחר שלושה חודשים של

פעילות גופנית מסודרת. להלן התוצאות שהתקבלו:

לפני פעילות גופנית	170	315	168	119	104	85	145	366
אחרי פעילות גופנית	172	304	159	114	96	88	145	360

בדקו את הטענה ברמת מובהקות 0.05,

כאשר לא ניתן להניח שרמת הטריגליצרידים מתפלגת נורמלית. נמקו.



שאלה

כדי לבדוק אם לאופן התצוגה של מוצר יש השפעה על כמות המכירות מהמוצר נבחרו באופן מקרי 16 חנויות מרשת סופרמרקטים גדולה וחולקו אקראית לשתי קבוצות.

ב- 8 מהחנויות המוצר הוצג במדף הרגיל וב- 8 חנויות המוצר הוצג במדף קדמי ליד המעבר.

לאר שבוע נרשמו המכירות בכל החנויות. התוצאות נתונות בטבלה הבאה:

תצוגה במדף רגיל	22	34	52	62	30	40	64	84
תצוגה במדף קדמי	52	71	76	34	67	83	66	90

האם קיים הבדל מובהק, ברמת מובהקות 0.05,

בממוצעי המכירות השבועיים, בהנחה שכמות המכירות לא מתפלגות נורמלית? נמקו.



שאלה

כדי לבדוק אם לאופן התצוגה של מוצר יש השפעה על כמות המכירות מהמוצר נבחרו באופן מקרי 18 חנויות מרשת סופרמרקטים גדולה

וחולקו אקראית לשתי קבוצות.

ב- 10 מהחנויות המוצר הוצג במדף הרגיל וב- 8 חנויות המוצר הוצג במדף קדמי ליד המעבר.

לאר שבוע נרשמו המכירות בכל החנויות. התוצאות נתונות בטבלה הבאה:

תצוגה במדף רגיל	22	34	52	62	30	40	64	84	56	59
תצוגה במדף קדמי	77	71	76	54	67	83	66	90		

בהנחה שלא ניתן להניח שכמות המכירות מתפלגת נורמלית,

האם קיים הבדל בממוצעי המכירות השבועיים ברמת מובהקות 0.05?



שאלה

הלשכה המרכזית לסטטיסטיקה פרסמה ש- 50% מהתושבים אינם ממחזרים בקבוקים. הוחלט על מסע הסברה תקשורתית ברחבי המדינה להעלאת המודעות למחזור. בתום מסע ההסברה נדגמו בשכונה מסוימת 15 משפחות באופן מקרי ונקבע שאם במדגם יימצאו לכל היותר 4 משפחות שאינן ממחזרות, יוסק שמסע ההסברה הצליח.

א. מהי רמת המובהקות של המבחן המוצע?

ב. מהי עצמת המבחן המוצע, אם אחוז המשפחות שאינן ממחזרות לאחר מסע ההסברה הוא 30%?

ג. בשכונה אחרת יותר גדולה נדגמו באופן מקרי 150 משפחות ונמצא כי 82 משפחות ממחזרות בקבוקים. האם לאור התוצאות ניתן להסיק שבשכונה זו חל שינוי בפרופורציית הממחזרים לאחר מסע ההסברה התקשורתית ברמת מובהקות 0.05? נמקו.



שאלה

בשני אזורים נדגמה הטמפרטורה היומית (מעלות):

Region(Degress)	1	2	3	4	5	6	7	8
אזור א	35	39	34	36	43	42	41	35
אזור ב	47	43	43	45	44	43	42	

א. בדוק את ההשערה שהתפלגות Degrees באזור א שונה מהתפלגות Degrees באזור ב ברמת מובהקות 0.05, כאשר אי אפשר להניח ש- Degrees מתפלג נורמלית. נמק.

ב. אם Degrees גבוה מ- 42 מעלות אז היום נחשב חם.

בדקו את ההשערה ששיעור הימים החמים גדול יותר באזור ב בהשוואה לאזור א. רמת מובהקות 0.05.



שאלה

נבדק ההבדל בין טמפרטורות במהלך היממה ב- 7 אזורים:

Region(Deg.)	1	2	3	4	5	6	7
בוקר	30	32	31	27	27	37	33
ערב	35	31	37	32	34	32	41

ידוע ש- Deg. אינו מתפלג נורמלית.

- א. ברמת מובהקות 0.05, בדוק את ההשערה שדרגות Deg. בערב גבוהות מדרגות Deg. בבוקר.
- ב. כאשר הטמפרטורות בערב גבוהה בהשוואה לבוקר, נרשמת החמרה בעומס החום.
- ברמת מובהקות 0.05, בדקו את ההשערה שברוב האזורים חלה החמרה.



שאלה

110 דיירים נתבקשו לבחור באמצעות הגרלה 3 נציגים לוועד הבית מתוך קבוצת מועמדים שכוללת 8 בעלי בתים ו- 4 שוכרים.

יו"ר הוועד היוצא קיבל דו"ח אודות התפלגות מספר בעלי-הבתים שנבחרו (להלן X) ע"י הדיירים:

x	0	1	2	3
$f(x)$	5	28	56	21

- א. ברמת מובהקות 0.05, בדוק את השערת היו"ר, שהנציגים נבחרו עם החזרה. (*).
- ב. ברמת מובהקות 0.05, בדוק את השערת היו"ר, שהנציגים נבחרו ללא החזרה.
- * במקרה זה אפשר שנציג יתמנה ליותר מתפקיד אחד.



שאלה

בכדי לבדוק האם טיפול חדש מסייע בהפחתת אלימות, נבחרו באקראי 20 נערים.

10 מהם טופלו והשאר לא טופלו.

בקרב המטופלים הושגו 9 הצלחות.

ומבין הבלתי מטופלים הושגו 3 הצלחות.

ברמת מובהקות 0.05, בדקו את הטענה שהטיפול יעיל.



שאלה

תזונאית טוענת כי אחוז השומן באבוקדו מזן א' קטן מאחוז השומן באבוקדו מזן ב'.

לצורך בדיקת טענתה נלקח מדגם של 5 אבוקדו מקריים מכל זן ונבדק אחוז השומן בכל אחד. התוצאות שהתקבלו מוצגות בטבלה הבאה:

אחוזי שומן							סכום	סכום הריבועים
1516171819							85	1455
1719202122							99	1975

בדוק ברמת מובהקות 0.05 את טענת התזונאית (בכל סעיף ציינו מה המבחן בו אתם משתמשים)

- א. כאשר ניתן להניח כי אחוז השומן באבוקדו מתפלג נורמלית.
 ב. כאשר לא ניתן להניח כי אחוז השומן באבוקדו מתפלג נורמלית.



שאלה

ילדים עם אוטיזם נוטים לחזור על פעולות מספר רב של פעמים. פותחה תרופה חדשה שאמורה להפחית תופעה זו.

במטרה לבדוק את יעילות התרופה נבחרו 8 ילדים עם אוטיזם. על כל אחד מהילדים בוצעה תצפית של 3 שעות לפני לקיחת התרופה

ותצפית נוספת של 3 שעות לאחר שבוע שבו קיבל הילד את התרופה החדשה.

בכל אחת מהתצפיות נרשם הזמן הכולל בו עסק הילד בפעילות חוזרת. הנתונים שנאספו מופיעים בטבלה הבאה:

הילד	1	2	3	4	5	6	7	8
לפני הטיפול	85	70	40	65	80	75	55	20
אחרי טיפול של שבוע	75	50	50	40	20	65	40	25

לא ניתן להניח כי הזמן מתפלג נורמלית.

בדקו האם התרופה אכן יעילה ברמת מובהקות 0.05,

א. בעזרת מבחן הסימן.

ב. בעזרת מבחן וילקוסון מתאים.



שאלה

במשחק מזל בקזינו מטילים 3 קוביות. הזכיה היא ביחס ישר למספר הקוביות, מבין השלוש, שבהן התקבלו התוצאות "5" או "6".

מהמר השתתף במשחק 135 פעמים. התוצאות במשחקים אלו היו:

מספר הקוביות בהן התקבלו "5" או "6"	0	1	2	3
מספר המשחקים	34	54	39	8

הנהלת הקזינו חשדה במהמר וביקשה לבדוק האם קוביות המשחק

בהן הוא השתמש הן הוגנות.

1. באיזה מבחן תשתמשו בכדי לבדוק את החשד של הנהלת הקזינו, על סמך המידע הנתון?

2. בצעו את המבחן ברמת מובהקות 0.05 ורשמו מה מסקנתכם לגבי החשד של הנהלת הקזינו.



שאלה

בעקבות בעיות האלימות וחוסר המשמעת בבתי ספר החליטו לבדוק את עמדות הציבור ביחס לשאלה: "מה דעתך ביחס למתן חופש גדול יותר למורים בהענשת תלמידים בנושא אלימות וחוסר משמעת?".

שאלו שאלה זו 300 מבוגרים מקריים והתוצאות מפורטות בטבלה הבאה:

	נמנע	נגד	בעד	
גברים	12	70	93	
נשים	6	32	87	

האם ניתן, ברמת מובהקות 0.025, לקבוע כי אין תלות בין מין הבוגר לבין עמדתו בשאלה הנדונה? ציינו מה המבחן בו אתם משתמשים.



שאלה

העמותה למניעת התמכרות למשחקי מחשב בקרב בני נוער החליטה לבדוק את השפעת מספר השעות ביום שבהן נער נחשף למחשב, על הצורך במשקפיים.

לצורך הבדיקה נלקח מדגם מקרי של 14 נערים, "מכורים", כלומר, נחשפים למחשב לפחות 4 שעות ביום,

ומדגם מקרי של 10 נערים, לא "מכורים", אשר נחשפים פחות מ-4 שעות ביום למחשב.

לאחר מעקב של שנתיים, התברר כי מבין הנערים ה"מכורים" 6 נזקקו למשקפיים. מבין הנערים הלא "מכורים" 1 נזקק למשקפיים.

א. בדקו, ברמת מובהקות 0.1, האם חשיפה גדולה למחשב מגדילה את הצורך במשקפיים. (שימו לב לגדלי המדגמים)

רשמו את ההשערות אותן אתם בודקים ואת סוג המבחן בו אתם משתמשים.

ב. בדקו, ברמת מובהקות 0.1, האם לאור תוצאות המדגם ניתן לומר כי אחוז הנזקקים למשקפיים, בקרב הנערים ה"מכורים" למחשב, הוא 50%.



openbook

המרכז לקידום אקדמי

טענות / ממ"ח 01



שאלה 1

יהי $X_1, \dots, X_{100}$ מדגם מקרי מאוכלוסייה עם תוחלת 15 ושונות 100.

הסתברות המאורע $\{\sum_{i=1}^{100} X_i < 1300\}$ היא בקירוב:



שאלה 2

מאוכלוסייה המתפלגת נורמלית עם תוחלת μ ושונות σ^2 נלקח מדגם של 100 תצפיות.

ההסתברות שממוצע המדגם יקבל ערך הגדול מ- μ היא:



שאלה 3

מאוכלוסייה המתפלגת נורמלית עם תוחלת μ ושונות σ^2 נלקח מדגם של 25 תצפיות.

ההסתברות שממוצע המדגם יקבל ערך הגדול מ- μ היא:



שאלה 4

רצו להגדיל את ההסתברות שממוצע מדגם מקרי מהאוכלוסייה המתוארת בשאלה הקודמת יקבל ערך גדול מ-

μ . ע"י שינוי גודל המדגם. אחד מהמשפטים שלהלן נכון:

א. יש לקחת מדגם יותר גדול מהמוזכר בשאלה הקודמת

ב. יש לקחת מדגם יותר קטן מהמוזכר בשאלה הקודמת

ג. גודל המדגם לא ישפיע על ההסתברות הנ"ל

(האוכלוסייה בשאלה הקודמת: מאוכלוסייה המתפלגת נורמלית עם תוחלת μ ושונות σ^2 נלקח מדגם של 100 תצפיות.)



שאלה 5

נלקח מדגם מקרי של 1000 תצפיות מהתפלגות בעלת תוחלת μ ושונות σ^2 .

לפי משפט הגבול המרכזי :

א. אפשר להניח שהתוחלת מתפלגת $N(\mu, \sigma^2/1000)$

ב. אפשר להניח שהתוחלת מתפלגת $N(\mu, \sigma^2)$

ג. אפשר להניח שממוצע המדגם מתפלג $N(0,1)$

ד. אפשר להניח שממוצע המדגם מתפלג $N(\mu, \sigma^2/1000)$



שאלה 6

יהי $X_1, X_2, \dots, X_{16}$ מדגם מקרי מתוך אוכלוסייה בעלת ממוצע μ לא ידוע ושונות $\sigma^2 = 64$.
סטיית התקן של האומד ל - μ היא:



שאלה 7

כשאומדים את μ ע"י ממוצע המדגם $\bar{X}$:

- μ גודל קבוע ו- $\bar{X}$ לפני שהוצא המדגם הוא משתנה מקרי.
- $\bar{X}$ גודל קבוע ו- μ משתנה מקרי.
- $\bar{X}$ וגם μ ידועים בתום המדגם, אך לא לפניו.
- אף לא אחת מבין התשובות א-ג נכונה.



שאלה 8

לאוכלוסייה נתונה שונות 2500.

מתוך האוכלוסייה הוצאו שני מדגמים מקריים: מדגם א' בגודל 100, ומדגם ב' בגודל 125.
סטיית התקן של ממוצע המדגם נקראת **טעות התקן (Standart Error)**

מהי טעות התקן של מדגם א' ומהי טעות התקן של מדגם ב'?



שאלה 9

במשרד התיירות רצו לאמוד את אחוז האנשים שיצאו לחופש באילת בשנת 2015,
ע"י רווח סמך ברמת סמך 95% שאורכו לא יעלה על 0.04.

גודל המדגם המינימלי הדרוש לכך הוא n_1 .

בסקר מכין התברר כי האחוז הנ"ל לא עולה על 30.

גודל המדגם המינימלי הנדרש, לאור מידע זה, הוא n_2 . הטענה הנכונה לגבי n_1 ו- n_2 היא:

א. $n_1 = n_2$

ב. $n_1 - n_2 = 384$

ג. $n_2 - n_1 = 384$

ד. $n_1 - n_2 = 97$

ה. $n_2 - n_1 = 97$



שאלה 10

אם אורך רווח סמך לתוחלת של אוכלוסיה נורמלית עם שונות ידועה, ברמת סמך 95% הוא 19.6.
אזי על סמך אותו מדגם, אורך רווח הסמך ברמת סמך 99% יהיה:



שאלה 11

במדגם מקרי של תינוקות יפניים היו 5125 בנים ו-4875 בנות. לאור התוצאות פורסם כי הסיכוי ללידת בן ביפן הוא בין 0.502 לבין 0.523.

רמת הסמך שבה השתמשו לקביעה זו היא:



שאלה 12

לאמידת ממוצע האוכלוסייה, μ_X , כאשר שונות האוכלוסייה ידועה, הוחלט להוציא מדגם מקרי ולחשב רווח סמך ברמת סמך 95%.

יהיו A ו-B, בהתאמה, הגבול התחתון והגבול העליון של רווח הסמך.

איזה מהמשפטים הבאים אינו נכון?

א. ממוצע המדגם, $\bar{X}$, הוא האמצע בין A לבין B.

ב. ההסתברות שהרווח בין A לבין B יכיל את ממוצע האוכלוסיה היא 0.95.

ג. ככל שהמדגם גדול יותר, יקטן המרחק B-A.

ד. ההסתברות שממוצע האוכלוסייה, μ_X , הוא באמצע בין A לבין B, היא 0.95.



שאלה 13

חושב רווח סמך לפרמטר μ על סמך מדגם מקרי בגודל 100 שהוצא מאוכלוסייה בעלת סטיית תקן 500, והתקבל: $988 < \mu < 1256$.

השלם:

ממוצע המדגם שווה ל: _____ ורמת הסמך שבה חושב הרווח היא: _____



שאלה 14

חושב רווח סמך לפרמטר μ על סמך מדגם מקרי בגודל 100 שהוצא מאוכלוסייה בעלת סטיית תקן 500, והתקבל: $988 < \mu < 1256$.

באותה רמת סמך, אם מעוניינים שאורך רווח הסמך יהיה לכל היותר 67 גודל המדגם המינימלי הדרוש הוא:



שאלה 15

חושב רווח סמך לפרמטר μ על סמך מדגם מקרי בגודל 400 שהוצא מאוכלוסייה בעלת שונות 625, והתקבל:

$$122.5 < \mu < 127.5$$

השלם:

ממוצע המדגם שווה ל: _____ ורמת הסמך שבה חושב הרווח היא: _____



שאלה 16

חושב רווח סמך לפרמטר μ על סמך מדגם מקרי בגודל 400 שהוצא מאוכלוסייה בעלת שונות 625, והתקבל:

$$122.5 < \mu < 127.5$$

באותה רמת סמך, אם מעוניינים שאורך רווח הסמך יהיה לכל היותר 10, גודל המדגם המינימלי הדרוש הוא:



שאלה 17

נערך שינוי בכלל ההחלטה של בדיקת השערה מסוימת.

אזור הדחייה של השערת האפס הוגדל וכל שאר הנתונים נשארו קבועים.

השינוי בערכי רמת המובהקות ועוצמת המבחן יהיו:

רמת המובהקות תגדל/ללא שינוי/תקטן

עוצמת המבחן תגדל/ללא שינוי/תקטן



שאלה 18

כדי לבדוק אם מטבע הוא תקין מטילים אותו 8 פעמים. אם הוא נופל על "ראש" שלוש, ארבע או חמש פעמים מחליטים שהוא תקין, אחרת מחליטים שהוא מזויף.

רמת המובהקות של המבחן הנתון היא



שאלה 19

כדי לבדוק אם מטבע הוא תקין מטילים אותו 10 פעמים. אם הוא נופל על "ראש" שלוש, ארבע, חמש, שש או שבע פעמים מחליטים שהוא תקין, אחרת מחליטים שהוא מזויף.

השלם:

רמת המובהקות של המבחן הנתון היא: _____ ועוצמת המבחן הנתון היא: _____



שאלה 20

לפניך שני צמדים של השערות:

B	A
$H_0: \mu = 100$	$H_0: \mu = 100$
$H_1: \mu = 80$	$H_1: \mu = 90$

כל אחד מהצמדים נבדק ע"י מדגם באותו גודל מאותה אוכלוסיה עם שונות ידועה ובאותה רמת מובהקות α .

האם הטענה: "עצמת המבחן גבוהה יותר עבור צמד ההשערות B" נכונה?



שאלה 21

לפניך שני צמדים של השערות:

B	A
$H_0: \mu = 100$	$H_0: \mu = 100$
$H_1: \mu = 80$	$H_1: \mu = 90$

כל אחד מהצמדים נבדק ע"י מדגם באותו גודל מאותה אוכלוסיה עם שונות ידועה ובאותה רמת מובהקות α .

האם הטענה: "ההסתברות לטעות מסוג שני גבוהה יותר עבור צמד ההשערות B" נכונה?



שאלה 22

לפניך שני צמדים של השערות:

B	A

$H_0: \mu = 100$	$H_0: \mu = 100$
$H_1: \mu = 80$	$H_1: \mu = 90$

כל אחד מהצמדים נבדק ע"י מדגם באותו גודל מאותה אוכלוסיה עם שונות ידועה ובאותה רמת מובהקות α .

עבור $\bar{x}$ נתון אם לא נדחה את H_0 בצמד A, מה תהיה מסקנתך לגבי צמד B?

נדחה/לא נדחה/ללא נתונים נוספים לא ניתן לדעת מה תהיה המסקנה עבור צמד B



שאלה 23

חוקר ביצע מבחן דו צדדי לבדיקת השערת האפס $H_0: \mu = \mu_0$

על סמך מדגם מקרי בגודל 25, מתוך אוכלוסיה המתפלגת נורמלית עם שונות לא ידועה.

סטטיסטי המבחן שהתקבל היה: $t_{\bar{x}} = -2.611$. לכן המסקנה היא:

א. התוצאה מובהקת ברמת מובהקות 0.1 אך לא כן ברמת מובהקות 0.05.

ב. התוצאה מובהקת ברמת מובהקות 0.05 אך לא כן ברמת מובהקות 0.02.

ג. התוצאה מובהקת ברמת מובהקות 0.02 אך לא כן ברמת מובהקות 0.01.

ד. התוצאה אינה מובהקת ברמת מובהקות 0.1.



שאלה 24

קופסה מכילה 100 כדורים. חלקם שחורים וחלקם לבנים.

רוצים לבדוק את ההשערה שהקופסה מכילה 10 כדורים שחורים ו-90 לבנים,

כנגד האלטרנטיבה שהקופסה מכילה 30 כדורים שחורים ו-70 לבנים.

לשם בדיקת ההשערה מוציאים באופן מקרי כדור אחד, בוחנים את צבעו ומחזירים אותו לקופסה.

מבצעים ניסוי זה 10 פעמים. הוחלט לדחות את השערת האפס אם מספר הכדורים השחורים במדגם הוא לפחות 2.

רמת המובהקות של המבחן הנתון היא: _____ ועצמת המבחן הנתון היא: _____.



שאלה 25

מטילים שתי קוביות משחק.

ב-72 הטלות של שתי הקוביות התקבלו התוצאות הבאות:

סכום התוצאה בכל הטלה	5-2	8-6	12-9
מספר הטלות	24	30	18

האם התוצאות מעידות על קוביות הוגנות?

- ברמת מובהקות 0.05 התוצאות מעידות על קוביות הוגנות.
- ברמת מובהקות 0.025 התוצאות מעידות על קוביות הוגנות, אך לא כן ברמת מובהקות 0.05.
- ברמת מובהקות 0.01 התוצאות מעידות על קוביות הוגנות, אך לא כן ברמת מובהקות 0.025.
- ברמת מובהקות 0.005 התוצאות מעידות על קוביות הוגנות, אך לא כן ברמת מובהקות 0.01.
- ברמת מובהקות 0.005 התוצאות מעידות על כך שקוביות אינן הוגנות.



שאלה 26

בבדיקת הטענה, שהחלוקה המינית של בוגרי התכנית ללימודי תקשורת מראה על עדיפות בולטת לגברים, נמצא כי מבין 18 בוגרים שנדגמו באופן מקרי היו 13 מהם גברים.

האם לאור תוצאות המדגם ניתן לאשש את הטענה, ברמת מובהקות 0.05:

- $P_v = 0.048 < 0.05$ ולכן ניתן לאשש את הטענה.
- $P_v = 0.952 > 0.05$ ולכן ניתן לאשש את הטענה.
- $P_v = 0.048 < 0.05$ ולכן לא ניתן לאשש את הטענה.
- $P_v = 0.952 > 0.05$ ולכן לא ניתן לאשש את הטענה.
- אף לא אחת מבין התשובות א-ד נכונה.

כאשר $P\text{-value}(P_v)$ – מובהקות התוצאה – היא ההסתברות לקבלת תוצאה קיצונית לפחות כמו שהתקבלה במדגם בהנחה שהשערת האפס נכונה.



שאלה 27

נערך ניסוי במטרה לבחון את היעילות של שיטה ליצירת גשמים מלאכותיים ע"י הזרעת עננים.

בניסוי הוזרעו 8 סופות בעוד ש-6 סופות אחרות לא הוזרעו. כדי לבדוק את יעילות ההזרעה הושוּוּ כמויות המשקעים הממוצעות שהתקבלו משתי קבוצות הסופות.

המבחן שבו תשתמש לבדיקת ההשערה שהשיטה מעלה את רמת המשקעים, אם האינפורמציה הידועה לך היא כי ב- 3 מבין הסופות שלא הוזרעו וב- 7 מבין הסופות שהוזרעו כמות המשקעים הממוצעת היתה מעל 0.2 מ"מ, הוא:

א. מבחן הסימן.

ב. מבחן פישר.

ג. מבחן מקנמר.

ד. מבחן וילקוקסון למדגמים מזווגים.

ה. מבחן וילקוקסון למדגמים בלתי תלויים.



שאלה 28

כדי לבדוק אם יש הבדל ברמת החרדה של בכורים לעומת לא בכורים, בדקו את רמת החרדה של 20 נבדקים שנבחרו באופן מקרי, מהם היו 8 בכורים ו- 12 לא-בכורים.

רמת החרדה נמדדה על ידי שאלון שבו הציון 7 ביטא רמת חרדה גבוהה ביותר והציון 1 ביטא רמת חרדה נמוכה ביותר.

המבחן הסטטיסטי המתאים לבעיה הוא:

א. מבחן t למדגמים בלתי תלויים, עם שונויות לא ידועות אך שוות.

ב. מבחן וילקוקסון למדגמים בלתי תלויים.

ג. מבחן פישר.

ד. מבחן הסימן.



שאלה 29

במדגם מקרי, מתי סכום ריבועי הסטיות מהממוצע, $\sum_{i=1}^n (x_i - \bar{x})^2$, מחולק ב- $n-1$?



שאלה 30

חוקר בנה שני רווחי סמך ל- μ באותה רמת סמך, כאשר σ ידועה.

רווח סמך אחד נבנה על סמך מדגם בגודל n , ורווח סמך שני נבנה על סמך מדגם בגודל $2n$.

אחד המשפטים הבאים נכון, איזהו?



שאלה 31

חברת תעופה מעונינת לאמוד את הפרופורציה p של ישראלים העומדים לנוסע לחו"ל בשנה הקרובה.

החברה חישה ומצאה רוח סמך ל- p על סמך מדגם של 400 ישראלים, וקיבלה את הרווח (0.488, 0.612)
האומדן לפרופורציה p הוא:



שאלה 32

בבדיקת השערת אפס מסוימת כנגד אלטרנטיבה חד צדדית הוחלט ברמת מובהקות 0.05 לדחות את השערת האפס. לכן ברמת מובהקות 0.01:
מהי מסקנתך?



שאלה 33

שני סטטיסטיקאים בודקים השערות ברמת מובהקות 0.05 על סמך אותו מדגם.
סטטיסטיקאי א' בודק את ההשערה $H_0: \mu = \mu_0$ כנגד האלטרנטיבה $H_0: \mu \neq \mu_0$
ומחליט לדחות את השערת האפס.
סטטיסטיקאי ב' בודק את ההשערה $H_0: \mu = \mu_0$ כנגד האלטרנטיבה $H_0: \mu > \mu_0$.
מה יחליט?



שאלה 34

בכדי לבדוק האם יש השפעה לעמדת הזינוק על המנצח במירון זוזים,
נאספו תוצאות של 144 מרוצים שבכל אחד מהם זינקו 8 סוסים.

בכל מרוץ נרשמה עמדת הזינוק של הסוס שניצח והתקבלו התוצאות הבאות:

עמדת הזינוק	1	2	3	4	5	6	7	8
מספר המנצחים	32	21	19	20	16	11	14	11

השלם את המשפט הבא:

המבחן בו נשתמש כדי לבדוק האם יש השפעה לעמדת הזינוק על המנצח הוא:

_____ ורמת המובהקות עבורה נדחה את השערת האפס היא: _____



שאלה 35

אם $X \sim F(29, 24)$, ההסתברות $P(X < 0.526)$ היא:



שאלה 36

מאמן כושר טוען כי אכילת שוקולד שעה לפני תחילת ריצה מגבירה את מהירות הריצה.

לצורך בדיקה ואישוש טענתו נבחרו שישה עשר אצנים.

תשעה מהם שנבחרו באופן מקרי אכלו שוקולד שעה לפני הריצה.

כל שישה עשר האצנים התמודדו בתחרות ריצה.

מבין התשעה שאכלו שוקולד לפני הריצה - חמישה סיימו בחמשת המקומות הראשונים, אחד סיים במקום השביעי,

שניים במקומות ה- 10 וה- 11 ואחד סיים מקום אחד לפני האחרון (מקום 15).

כל זמני הריצה של שישה עשר המתמודדים היו שונים זה מזה.

האם מתקבלת טענת המאמן?

א. ברמת מובהקות 0.05 טענת המאמן אינה מתקבלת

ב. ברמת מובהקות 0.05 טענת המאמן מתקבלת, אך לא כן ברמת מובהקות 0.025

ג. ברמת מובהקות 0.025 טענת המאמן מתקבלת, אך לא כן ברמת מובהקות 0.01

ד. ברמת מובהקות 0.01 טענת המאמן מתקבלת, אך לא כן ברמת מובהקות 0.025

ה. ברמת מובהקות 0.001 טענת המאמן מתקבלת

openbook



שאלה

בבית חרושת לייצור מוצרי טריקו התברר כי רוב התופרות חסרות הכשרה מקצועית. בעל בית החרושת החליט לבחור מדגם של 12 תופרות,

לערוך להן מבחנים בתפירה ולאחר מכן ללמדן בקורס מיוחד ולערוך להן מבחן נוסף.

המבחן בו תשתמשו לבדיקת ההשערה שהקורס משפר את רמת התפירה, אם האינפורמציה הידועה לכם היא כי 7 תופרות תפרו טוב יותר

לאחר הקורס, 2 תפרו גרוע יותר לאחר הקורס ועל יתר התופרות הקורס לא השפיע, הוא:

מבחן הסימן / מבחן וילקוקסון למדגמים מזווגים / מבחן וילקוקסון למדגמים בלתי-תלויים / מבחן אי-פרמטרי אחר.



שאלה

בבית חרושת לייצור מוצרי טריקו התברר כי רוב התופרות חסרות הכשרה מקצועית. בעל בית החרושת החליט לבחור מדגם של 12 תופרות,

לערוך להן מבחנים בתפירה ולאחר מכן ללמדן בקורס מיוחד ולערוך להן מבחן נוסף.

רמת המובהקות הקטנה ביותר עבורה תידחה השערת האפס, ע"ס האינפורמציה הנתונה בשאלה 1 היא:

0.98 / 0.91 / 0.806 / 0.09 / 0.019

(שאלה 1: המבחן בו תשתמשו לבדיקת ההשערה שהקורס משפר את רמת התפירה, אם האינפורמציה הידועה לכם היא כי 7 תופרות תפרו טוב יותר לאחר הקורס, 2 תפרו גרוע יותר לאחר הקורס ועל יתר התופרות הקורס לא השפיע, הוא)



שאלה

בבית חרושת לייצור מוצרי טריקו התברר כי רוב התופרות חסרות הכשרה מקצועית. בעל בית החרושת החליט לבחור מדגם של 12 תופרות,

לערוך להן מבחנים בתפירה ולאחר מכן ללמדן בקורס מיוחד ולערוך להן מבחן נוסף.

בעל בית החרושת בדק גם את הישגי 12 התופרות שעברו את הקורס כנגד הישגי 6 תופרות שלא עברו את הקורס,

וגילה כי 10 התופרות הטובות ביותר מבין ה-18 שנבדקו הן מאלה שעברו את הקורס, וכן שתי התופרות הגרועות ביותר עברו את הקורס.

המבחן בו תשתמשו לבדיקת ההשערה, שהקורס משפר את רמת התפירה לאור האינפורמציה הנ"ל, הוא:

מבחן הבינום / הסימן / וילקוקסון למדגמים מזווגים / וילקוקסון למדגמים בלתי-תלויים / מבחן אי-פרמטרי אחר.



שאלה

בבית חרושת לייצור מוצרי טריקו התברר כי רוב התופרות חסרות הכשרה מקצועית. בעל בית החרושת החליט לבחור מדגם של 12 תופרות,

לערוך להן מבחנים בתפירה ולאחר מכן ללמדן בקורס מיוחד ולערוך להן מבחן נוסף.

רמת המובהקות הקטנה ביותר עבורה תידחה השערת האפס, ע"ס הנתונים של שאלה 3, היא:

0.05 / 0.025 / 0.02 / 0.01 / 0.001

(שאלה 3: בעל בית החרושת בדק גם את הישגי 12 התופרות שעברו את הקורס כנגד הישגי 6 תופרות שלא עברו את הקורס, וגילה כי 10 התופרות הטובות ביותר מבין ה-18 שנבדקו הן מאלה שעברו את הקורס, וכן שתי התופרות הגרועות ביותר עברו את הקורס.)



שאלה

במטרה לבדוק אם מטבע מסוים תקין או לא, הטילו את המטבע 100 פעם וספרו את מספר הפעמים שהתקבל "עץ".

סטטיסטיקאי א' מציע לבדוק זאת בעזרת בדיקת ההשערות

$$H_0: p = \frac{1}{2}$$

$$H_1: p \neq \frac{1}{2}$$

ואילו סטטיסטיקאי ב' מציע להשתמש במבחן לטיב התאמה. כל אחד מהסטטיסטיקאים הללו בודק את ההשערה על פי הצעתו, אך שניהם בודקים על סמך אותו מדגם ובאותה רמת מובהקות.

איזו קביעה נכונה?

א. שני הסטטיסטיקאים יחליטו **תמיד** אותן החלטות.
ב. שני הסטטיסטיקאים יחליטו **תמיד** החלטות שונות.
ג. שני הסטטיסטיקאים לעיתים יחליטו החלטות שונות ולעיתים יחליטו החלטות זהות.



שאלה

כדי לקבוע אם תוספת לשמן מנועים במשאיות מקטינה את הסיכוי לכך שיהיה צורך לשפץ את מנועי המשאיות תוך שנתיים,

בוחרת חברת משאיות, בעלת 25 משאיות חדשות, 10 משאיות באופן מקרי לקבלת התוספת.

מבין המשאיות שקיבלו את התוספת, 3 היו זקוקות לשיפוץ מנוע תוך שנתיים,

ובין ה-15 שלא קיבלו את התוספת, 6 היו זקוקות לשיפוץ מנוע תוך שנתיים.

המבחן בו תשתמשו לבדיקת הטענה הוא:

מבחן מקנמר / פישר / הסימן / מבחן פרמטרי להפרש פרופורציות / מבחן אי פרמטרי אחר.



openbook
המרכז לקידום אקדמי



SPSS

מונח עברי	SPSS	
שכיחות	Frequency, Count	
שכיחות יחסית	Percent	
שכיחות יחסית מצטברת	Cumulative Percent	
מאון, אחוזון	Percentile	
רבעון	Quartile	
ממוצע	Mean	
חציון	Median	
שכיח	Mode	
סטיית-תקן	Std. Deviation	
מדד	Measure	
מתאם	Correlation	
טבלת שכיחות משותפת	Cross tabulation	
תלוי	Dependent	
בלתי-תלוי	Independent	
השערת האפס	Null Hypothesis	
רווח סמך	Confidence Interval	
מבחן T לתוחלת אחת	One Sample T-Test	
ממוצע	Mean	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
אומד סטיית-התקן	Std. Deviation	$\hat{s} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$
טעות-התקן	Std. Error Mean	$\frac{\hat{s}}{\sqrt{n}}$
מבחן T להפרש תוחלות במדגמים מזווגים	Paired-Samples T-Test	
ממוצע ההפרשים	Mean Difference	$\bar{d} = \bar{x}_1 - \bar{x}_2$
טעות-התקן	Std. Error Mean	$\frac{\hat{s}_d}{\sqrt{n}}$
מבחן T להפרש תוחלות במדגמים ב"ת	Independent Samples T-Test	
הפרש הממוצעים	Mean Difference	$\bar{d} = \bar{x}_1 - \bar{x}_2$
טעות-התקן	Std. Error Mean	$\sqrt{\hat{s}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$
שיוויון שונויות	Equality of Variances	
p-value במבחן דו-צדדי	Sig. (2-tailed)	



סיכום המצבים האפשריים

אפשרות (1)

הסקת מסקנות על בדיקת השערות ע"פ כלל ה-PV אפשרי רק במצבים הבאים:

- ✓ מבחן לתוחלת אחת (בכל מצב)
- ✓ מבחן הפרש תוחלות – רק כאשר הערך הנקוב תחת H_0 הינו 0.

כלל ההכרעה במצבים אלו הוא:

- א. במבחן דו זנבי כאשר $p.v > \alpha$ נקבל H_0
- ב. במבחן חד זנבי כאשר $p.v = \frac{\text{sig}(2\text{tailed})}{2} > \alpha$ נקבל H_0
- ג. במבחן חד זנבי, כאשר תוצאת המדגם רחוקה / מנוגדת לכיוון של השערת H_1 אזי נקבל H_0 אם $p.v = 1 - \frac{\text{sig}(2\text{tailed})}{2} > \alpha$

אפשרות (2)

בבדיקת השערות על הפרש תוחלות כאשר **הערך הנקוב ב H_0 שונה מ 0** יש להחליט על קבלה / דחייה של H_0 ע"פ תוצאות רווח הסמך (במבחן דו זנבי) או על פי השיטה ה"רגילה" מיחידה 13

אפשרות (3)

בהפרש תוחלות, כאשר מקבלים את H_1 יש לוודא שהממוצעים זהים לכיוון של ההשערה

אפשרות (4)

כאשר ההשערות אינן כתובות במילים יש להתייחס אל ההשערה כאל השערה דו זנבית.

לסיכום:

✓ במבחן דו צדדי $p.v = \text{sig}(2\text{tailed})$

✓ במבחן חד צדדי שבו התוצאה במדגם נמצאת בצד שבו נמצא אזור הדחייה $p.v = \frac{\text{sig}(2\text{tailed})}{2}$

✓ במבחן חד צדדי שבו תוצאת המדגם נמצאת בצד המנוגד לאזור הדחייה $p.v = 1 - \frac{\text{sig}(2\text{tailed})}{2}$

אם $p.v < \alpha$ דוחים את H_0



מבחן T לתוחלת אחת One Sample T-Test

	One Sample T-Test	מבחן T לתוחלת אחת
$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	Mean	ממוצע
$\hat{s} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$	Std. Deviation	אומד סטיית-התקן
$\frac{\hat{s}}{\sqrt{n}}$	Std. Error Mean	טעות-התקן

בדיקת השערות על תוחלת אחת

שאלה 1

נתון הפלט הבא:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
ציון בקורס סטטיסטיקה ב	36	73.50	20.708	3.451

One-Sample Test

	Test Value = 80					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
ציון בקורס סטטיסטיקה ב	-1.883	35	.068	-6.500	-13.51	.51

א. בדוק את ההשערה שציון 2 (ציון בקורס סטטיסטיקה ב) שווה ל 80 $\alpha = 0.05$

ב. בדוק את ההשערה שציון 2 (ציון בקורס סטטיסטיקה ב) קטן מ 80 $\alpha = 0.05$

ג. בדוק את ההשערה שציון 2 (ציון בקורס סטטיסטיקה ב) גדול מ 80 $\alpha = 0.05$



בדיקת השערות על שני מדגמים מזווגים Paired-Samples T-Test

	Paired-Samples T-Test	מבחן T להפרש תוחלות במדגמים מזווגים
$\bar{d} = \overline{x_1 - x_2}$	Mean Difference	ממוצע ההפרשים
$\frac{\hat{s}_d}{\sqrt{n}}$	Std. Error Mean	טעות-התקן



שאלה 2

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	ציון בקורס סטטיסטיקה א - ציון בקורס סטטיסטיקה ב	74.75	36	21.941	3.657
		73.50	36	20.708	3.451

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	ציון בקורס סטטיסטיקה א - ציון בקורס סטטיסטיקה ב	1.250	6.129	1.021	-.824	3.324	1.224	35	.229

א. בדוק את ההשערה שציון 1 גדול מציון 2 $\alpha = 0.05$

ב. בדוק את ההשערה שציון 1 גדול מציון 2 ב 0.2 יחידות $\alpha = 0.05$

ג. בדוק את ההשערה שציון 1 גדול מציון 2 ביותר מ 0.2 יחידות $\alpha = 0.05$



מבחן T להפרש תוחלות במדגמים ב"ת Independent Samples T-Test

מבחן T להפרש תוחלות במדגמים ב"ת	Independent Samples T-Test	
הפרש הממוצעים	Mean Difference	$\bar{x}_1 - \bar{x}_2$
טעות-התקן	Std. Error Mean	$\sqrt{\hat{s}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$
שיוויון שונויות	Equality of Variances	
p-value במבחן דו-צדדי	Sig. (2-tailed)	



שאלה 3

Group Statistics

מין הסטודנט	N	Mean	Std. Deviation	Std. Error Mean
זכר	16	73.81	22.364	5.591
נקבה	20	75.50	22.149	4.953

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.003	.959	-.226	34	.822	-1.688	7.461	-16.850	13.475
Equal variances not assumed			-.226	32.149	.823	-1.688	7.469	-16.899	13.524

א. בדוק האם שונות הציונים בציון 1 (בקורס סטטיסטיקה א) שוות? $\alpha = 0.05$

ב. בדוק את ההשערה שציוני הבנות גבוהים משל הבנים במשתנה ציון 1 (ציון בסטטיסטיקה א $\alpha = 0.05$)

ג. בדוק את ההשערה שציוני הבנות גבוהים משל הבנים במשתנה ציון 1 ביותר מ 0.5 יחידות,

$\alpha = 0.05$



שאלה 4

לפניך תוצאות ניתוח סטטיסטי שבדק השפעת צבע בפרסום. אצל 70 נבדקים נרשמה רמת הזכירה של מותגים דומים בפרסומות בצבע ובפרסומות בשחור-לבן. כל נבדק קיבל שני ציונים המבטאים את רמת הזכירה (לכל אחת משני סוגי הפרסומות). ידוע שהציונים מתפלגים נורמלית

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Pair 1 BW	19.94	70	4.571	.546
COLOR	23.50	70	4.286	.512

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
					Mean	Lower			
		Mean	Std. Deviation	Std. Error Mean			t	df	Sig. (2-tailed)
Pair 1	BW - COLOR	-3.557	6.469	.773	-5.100	-2.015	-4.601	69	.000

- א. בדוק ברמת-מובהקות 0.05 את ההשערה שאנשים זוכרים טוב יותר פרסומות צבעוניות בהשוואה לפרסומות בשחור-לבן. נסח את ההשערות ונמק
- ב. בדוק ברמת-מובהקות 0.05 את ההשערה שציוני הזכירה של פרסומות בצבע גבוהים ביותר מ-2 נקודות מציוני הזכירה של פרסומות בשחור-לבן. נסח את ההשערות ונמק
- ג. מתוך הפלט חשב את הערך $t_{0.025}^{69}$



שאלה 5

לפניך ניתוח סטטיסטי של נתוני שכר (באלפי שקלים) בתעשייה העילית, בשנים 2016-2017:

		Ranks	
		N	Sum of Ranks
Sal.17 – Sal.16	Negative Ranks	6 ^a	46.00
	Positive Ranks	5 ^b	W+
	Ties	0 ^c	
	Total	11	

a. Sal.17 < Sal.16

b. Sal.17 > Sal.16

c. Sal.17 = Sal.16

Test Statistics^a

	Sal.17 - Sal.16
Z	-1.156 ^b
Asymp. Sig. (2-tailed)	.248
Exact Sig. (2-tailed)	.278
Exact Sig. (1-tailed)	.139
Point Probability	.019

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks

א. מהו המבחן הסטטיסטי שבוצע? מה ערכו של הסטטיסטי $W+$?

ב. מהי רמת המובהקות המינימלית לבדיקת ההשערה שהשכר ב- 2017 שונה מן השכר ב- 2016?

ג. בהמשך לסעיף ב, מהי מסקנת המבחן, ברמת מובהקות 0.05?



שאלה 6

לפניך ניתוח סטטיסטי של שטפי אור במנורות פלואורוסצנטיות (באלפי יחידות אור, אלפי Lumen)

Explore

Descriptives

			Statistic	Std. Error
Lumen	Mean		$\bar{X}$	0.038790
	95% Confidence	Lower Bound	4.88156	
	Interval for Mean	Upper Bound	5.04103	
	Variance		0.041	
	Std. Deviation		0.201558	

על סמך הפלט ענה על הסעיפים הבאים: בהנחה ששטף האור מתפלג נורמלית:

א. הסבר מהו הערך החסר X . חשב את X .

ב. בדוק ברמת מובהקות 0.05 את ההשערה שתוחלת שטף האור שונה מ-4.5. נסח את ההשערות ונמק.

ג. בדוק ברמת מובהקות 0.05 את ההשערה שתוחלת שטף האור קטנה מ-5.6. נסח את ההשערות ונמק.



שאלה 7

לפניך פלט של ניתוח סטטיסטי שבדק מספר מילים בהודעות SMS בקרב ילדים:

Descriptives

			Statistic	Std. Error
Words	Mean		21.49	.6865
	90% Confidence Interval for Mean	Lower Bound	20.31	
		Upper Bound	22.66	
	Std. Deviation		4.119	

על סמך הפלט ענה על השאלות הבאות:

א. מהו גודל המדגם שעליו התבססו החישובים? נמק.

ב. בדוק ברמת מובהקות 0.15 את ההשערה שממוצע מספר המילים קטן מ-23. נסח את ההשערות ונמק.

ג. בדוק ברמת מובהקות 0.05 את ההשערה שממוצע מספר המילים שונה מ-21. נסח את ההשערות ונמק.



שאלה 8

לפניך תוצאות מבחן סטטיסטי שבוצע על המשתנים המקריים A ו- B.

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference			
					Lower Upper			
Pair 1	B - A	2.43276	3.19861	.34097	1.75504 3.11048	7.135	87	.000

על סמך הפלט ענה על השאלות הבאות:

א. רשום מהו המבחן הסטטיסטי ונסח את ההשערות אליהם מתייחס הפלט.

ב. בהסתמך על הפלט, בדוק את הטענה שתוחלת B גבוהה מתוחלת A ברמת-מובהקות 0.05. הסבר ונמק.

ג. בהסתמך על הפלט, בדוק את הטענה שההפרש בין תוחלת B לתוחלת A שונה מ- 4 ברמת-מובהקות 0.1. הסבר ונמק.

openbook
המרכז לקידום אקדמי



שאלה 9

לפניך פלט של תוצאות 2 מבחנים סטטיסטיים שנערכו בהסתמך על המשתנים Class ו-Measure .
למשתנה Class שני ערכים (1 ו-2) שמציינים שייכות של תצפית לקבוצה.

Group Statistics

	Class	N	Mean	Std. Deviation	Std. Error Mean
Measure	1	37	58.8777	5.70766	.93833
	2	53	54.6818	4.68179	.64309

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Measure	Equal variances assumed	2.732	.102	3.821	88	.000	4.19584	1.09822	2.01336	6.37832
	Equal variances not assumed			3.688	67.458	.000	4.19584	1.13756	1.92555	6.46613

על סמך הפלט ענה על השאלות הבאות:

א. בדוק את ההשערה שהשונויות המשתנה Measure שווה בקבוצות 1 ו-2 ברמת מובהקות 0.05.

ב. בדוק את ההשערה שתוחלת Measure בקבוצה 1 גבוהה מתוחלת Measure בקבוצה 2 ברמת מובהקות 0.05.

ג. בדוק את ההשערה שההפרש בין תוחלת Measure בקבוצה 1 לבין תוחלת Measure בקבוצה 2 שונה מ-3 יחידות, ברמת מובהקות 0.05.



שאלה 10

לפניך תוצאות ניתוח סטטיסטי שמתייחס לאתר תיירות מסוים. במשך 35 ימים נרשמו (בכל יום) מספר המבקרים באתר.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
VISITORS	35	100.0%	0	.0%	35	100.0%

Descriptives

			Statistic	Std. Error
VISITORS	Mean		53.77	1.365
	95% Confidence Interval for Mean	Lower Bound	51.00	
		Upper Bound	56.55	

על סמך הפלט ענה על השאלות הבאות:

א. בדוק ברמת-מובהקות 0.03 את ההשערה שהתוחלת היומית של מספר המבקרים שווה 52. נסח את ההשערות ונמק

ב. בדוק ברמת-מובהקות 0.07 את ההשערה שהתוחלת היומית של מספר המבקרים גבוהה מ- 50. נסח את ההשערות ונמק.

ג. חשב את הערך $t_{0.025}^{34}$



שאלה 11

להשוואת האיכויות של 2 סוגי מצמדים (קלאץ') נערך בוחן ביצועים. במסגרת הבוחן נרשם אורך החיים של כל מצמד עד הבלאי (באלפי ק"מ).

להלן התוצאות שהתקבלו:

Ranks

		N	Mean Rank	Sum of Ranks
Distance A – Distance B	Negative Ranks	17 ^a	16.88	W
	Positive Ranks	23 ^b	23.17	533.00
	Ties	0 ^c		
	Total	40		

a. Distance A < Distance B

b. Distance A > Distance B

c. Distance A = Distance B

Test Statistics^a

	Distance A - Distance B
Z	-1.653 ^b
Exact Sig. (2-tailed)	0.100
Exact Sig. (1-tailed)	0.050
Point Probability	.001

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks

על פי הפלט ובהנחה שאורך החיים אינו מתפלג נורמלית, ענה על הסעיפים הבאים:

א. מהו ערכו של הסטטיסטי W?

ב. מהי רמת המובהקות המינימלית שעבורה יוחלט שיש הבדל בין 2 סוגי המצמדים?

ג. בדוק ברמת מובהקות 0.07, את הטענה שאורך חיי מצמד A גדול מאורך חיי מצמד B.



שאלה 12

לפניך פלט שבו נותחו 80 מדידות של כמות הגשם השנתית (Rain, במילימטרים) באזור גיאוגרפי מסוים.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Rain	80	271.16	54.220	6.062

One-Sample Test

Test Value = 280						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Rain	-1.458	79	.149	-8.840	-20.91	3.23

על-סמך הפלט ענה על השאלות הבאות: בתשובה לכל סעיף יש לרשום את ההשערות,

את סטטיסטי המבחן, את כלל ההחלטה המתאים ואת המסקנה

א. רשום מהו המבחן הסטטיסטי ונסח את ההשערות אליהם מתייחס הפלט.

ב. בדוק ברמת-מובהקות 0.05 את ההשערה שתוחלת כמות הגשם שונה מ-280 מילימטר

ג. בדוק ברמת-מובהקות 0.05 את ההשערה שתוחלת כמות הגשם שונה מ-270 מילימטר

ד. בדוק ברמת-מובהקות 0.05 את ההשערה שתוחלת כמות הגשם קטנה מ-285 מילימטר.



שאלה 13

לפניך תוצאות ניתוח סטטיסטי שמתייחס לנתוני המכירה של אגרות ברכה צבעוניות ב- 70 חנויות ברחבי הארץ. ידוע שנתוני המכירות מתפלגים נורמלית.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
COLOR	70	23.5	4.286	.512

One-Sample Test

	Test Value = 22.5					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
COLOR	1.952	69	.0555	1.000	?	2.02

על-סמך הפלט ענה על השאלות הבאות:

א. בדוק ברמת-מובהקות 0.05 את ההשערה שתוחלת המכירות (לחנות) גבוהה מ- 22.5. נסח את ההשערות ונמק.

ב. בדוק ברמת-מובהקות 0.05 את ההשערה שתוחלת המכירות (לחנות) קטנה מ- 25. נסח את ההשערות ונמק.

ג. חשב את הגבול התחתון של רווח הסמך המופיע בטבלה התחתונה



שאלה 14

לפניך פלט של ניתוח סטטיסטי שבדק מספר מילים בהודעות SMS בקרב ילדים.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Words	35	21.49	4.119	.696

One-Sample Test

	Test Value = 23					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Words	-2.175	34	.037	-1.514	-2.93	-.10

על סמך הפלט ענה על השאלות הבאות:

- מהי רמת המובהקות המינימלית לבדיקת ההשערה שתוחלת מספר המילים גדולה מ-23?
- בדוק ברמת מובהקות 0.04 את ההשערה שתוחלת מספר המילים שונה מ-23. נסח את ההשערות ונמק
- בדוק ברמת מובהקות 0.025 את ההשערה שתוחלת מספר המילים קטנה מ-24. נסח את ההשערות ונמק



שאלה 15

לבדיקת תקינות משאבת לחץ אוויר, נופחו באמצעותה 33 צמיגים ללחץ נקוב של PSI 35.45.

לאחר מכן נמדד לחץ האוויר בכל צמיג ובוצע ניתוח סטטיסטי.

ממצאי הניתוח הסטטיסטי מתוארים להלן:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
AirPressure	33	Mean	1.52168	.26489

One-Sample Test

	Test Value = 35.45					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
AirPressure	2.225	32	.033	.58939	.0498	1.1290

א. מהו ממוצע המדגם?

ב. האם המשאבה תקינה? בדוק ברמת מובהקות 0.05, נסח את ההשערות ונמק.

ג. ברמת מובהקות 0.05, בדוק את ההשערה שתוחלת לחץ הניפוח שונה מ- PSI 34.4.



שאלה 16

לפניך פלט של תוצאות מבחן סטטיסטי המתייחס למשתנה המקרי X

One-Sample Test

	Test Value = 49					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
X	2.643	199	.009	1.73895	.4417	3.0362

על סמך הפלט ענה על השאלות הבאות:

- א. רשום מהו המבחן הסטטיסטי ונסח את ההשערות אליהם מתייחס הפלט.
- ב. בהסתמך על הפלט, חשב רווח-סמך לתוחלת של X ברמת סמך 0.95. הסבר ונמק.
- ג. על סמך תוצאות סעיף ב, בדוק את הטענה שהתוחלת שונה מ- 50 ברמת-מובהקות 0.02. הסבר ונמק.

openbook
המרכז לקידום אקדמי



שאלה 17

לפניך פלט של ניתוח סטטיסטי בו הושאו שכיחויות מבקרים יומיות בשני אתרי תיירות A ו- B:

Group Statistics

	SITE	N	Mean	Std. Deviation	Std. Error Mean
VISITORS	B	17	60.24	4.944	1.199
	A	18	47.67	5.122	1.207

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
VISITORS	Equal variances assumed	.040	.843	7.379	33	.000	12.569	1.703	9.103	16.034
	Equal variances not assumed			7.387	32.982	.000	12.569	1.702	9.107	16.031

על סמך הפלט ענה על השאלות הבאות:

א. בדוק ברמת מובהקות 0.05 את ההשערה שהשונויות של שכיחויות המבקרים בשני האתרים שוות.

נסח את ההשערות ונמק.

ב. בדוק ברמת מובהקות 0.05 את ההשערה שתוחלת שכיחות המבקרים באתר B גבוהה מתוחלת המבקרים באתר A.

נסח את ההשערות ונמק

ג. בדוק ברמת מובהקות 0.05 את ההשערה שתוחלת שכיחות המבקרים באתר B גבוהה מתוחלת המבקרים באתר A בפחות מ-18.

נסח את ההשערות ונמק



שאלה 18

לפניך ניתוח סטטיסטי של מספר גרגירי חול (במיליונים) בשעוני חול זעירים (בנפח של 1 סמ"ק).

Explore
Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Grains	33	100.0%	0	0.0%	33	100.0%

Descriptives

		Statistic	Std Error
Grains	Mean	.9894	X
	95% Confidence interval for Mean	Lower Bound Upper Bound	.9468 1.0320
	5% Trimmed Mean	.9892	
	Median	.9600	
	Variance	.014	
	Std. Deviation	.12002	
	Minimum	.76	
	Maximum	1.23	
	Range	.47	
	Interquartile Range	.19	
	Skewness	-.032	.409
	Kurtosis	-.680	.798

על סמך הפלט ענה על הסעיפים הבאים: בהנחה שמספר הגרגירים מתפלג נורמלית,

א. הסבר מהו הערך החסר X. חשב את X.

ב. בדוק, ברמת מובהקות 0.05, את ההשערה שתוחלת מספר גרגירי החול שונה מ- 0.96.

נסח את ההשערות ונמק.

ג. בדוק, ברמת מובהקות 0.05, את ההשערה שתוחלת מספר גרגירי החול גדולה מ- 0.94.

נסח את ההשערות ונמק.



שאלה 19

לבדיקת כמות הגשם השנתית (מ"מ) באזור מסוים נאספו נתונים מ- 26 נקודות מדידה.

Explore
Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Rain	26	100.0%	0	0.0%	26	100.0%

Descriptives

		Statistic	Std Error
Rain	Mean	283.54	A
	95% Confidence interval for Mean	Lower Bound Upper Bound	241.35 B
	5% Trimmed Mean	283.04	
	Median	270.5	
	Variance	10909.698	
	Std. Deviation	104.450	
	Minimum	114	
	Maximum	463	
	Range	393	
	Interquartile Range	162	

ענה על הסעיפים על סמך ממצאי הבדיקה.

א. מהם ערכי A ו- B ?

ב. בדוק, ברמת מובהקות 0.05, את ההשערה שתוחלת כמות הגשם שונה מ- 260 מ"מ.

נסח את ההשערות ונמק.

ג. מתוך הפלט חשב את הערך $t_{0.025}^{25}$.



שאלה 20

באתר מסוים נמדדו כיווני הרוחות (במעלות) בשעות 1 ו- 4 אחה"צ בחודש מסוים במשך 50 שנה. להלן תוצאות הניתוח הסטטיסטי

Paired Samples Statistics

	Mean	N	Std. Deviation	Std. Error Mean
Wind Direction 1pm	180.7737	50	12.36755	1.74904
Pair 1 Wind Direction 4pm	199.6735	50	14.53210	2.05515

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Wind Direction 1pm- Wind Direction 4pm	-18.89990	19.20710	2.71629	-24.35849	-13.44130	-6.958	49	.000

א. בדוק את ההשערה שממוצע כיווני הרוח שווים ברמת-מובהקות 0.05.

ב. בדוק את ההשערה שממוצע כיוון הרוח בשעה 4 אחה"צ גבוה מממוצע כיוון הרוח בשעה 1 אחה"צ ברמת-מובהקות 0.05.

ג. בדוק את ההשערה שממוצע כיוון הרוח בשעה 4 אחה"צ גבוה מממוצע כיוון הרוח בשעה 1 אחה"צ ביותר מ-20 מעלות ברמת-מובהקות 0.05.



שאלה 21

מדדו את זמני הריצה בשתי קבוצות. האחת – של מתבגרים והאחרת – של ילדים.

לפניך תוצאות הניתוח הסטטיסטי שבוצע על הנתונים שהתקבלו

Group Statistics

	Age	N	Mean	Std. Deviation	Std. Error Mean
Time [sec]	kid	45	12.6862	2.97779	.44390
	teenager	35	10.2643	1.83878	.31081

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Time [sec]	Equal variances assumed	5.526	.021	4.223	78	.000	2.42190	.57353	1.28010	3.56370
	Equal variances not assumed			4.469	74.534	.000	2.42190	.54190	1.34227	3.50153

על סמך הפלט ענה על השאלות הבאות:

א. בדוק את ההשערה ששונויות זמן הריצה בקבוצות אינן שוות ברמת מובהקות 0.05.

ב. לאור תוצאת סעיף א, בדוק את ההשערה שזמן הריצה הממוצע של הילדים גבוה מזמן הריצה הממוצע של המתבגרים, ברמת-מובהקות 0.05.

ג. לאור תוצאת סעיף א, בדוק את ההשערה שזמן הריצה הממוצע של הילדים גבוה ביותר מ- 3 שניות מזמן הריצה הממוצע של המתבגרים, ברמת-מובהקות 0.05.



שאלה 22

נתחו את ההתפלגות של שיעור חלקיקי חומר מסוים באוויר (למטר מעוקב) וקיבלו את התוצאות הבאות:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
LEVEL	201	5.0603	9.90468	.69862

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
LEVEL	7.243	200	.000	5.06029	3.6827	6.4379

One-Sample Test

	Test Value = 4					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
LEVEL	1.518	200	.131	1.06029	-.3173	2.4379

א. בדוק את ההשערה שתוחלת שיעור החלקיקים למטר מעוקב גבוהה מ-0, ברמת-מובהקות 5%.

ב. מהי רמת המובהקות המינימלית לדחיית ההשערה שתוחלת שיעור החלקיקים למטר מעוקב אינה גבוהה מ-0.05?

ג. חשב את ערכו של הערך הקריטי $t_{0.025}^{200}$.



שאלה 23

לפניך תוצאות ניתוח סטטיסטי שבדק את מידת הביטחון שחשים ילדים בבית הספר, בקרב שתי שכבות גיל: צעירים (סומנו ב-0) ובוגרים (סומנו ב-1)

Group Statistics

	Age	N	Mean	Std. Deviation	Std. Error Mean
Security	1	10	4.53	2.107	.666
	0	6	3.26	1.571	.641

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Security	Equal variances assumed	.052	.823	1.265	14	.227	1.262	.998	-.878	3.403
	Equal variances not assumed			1.365	13.125	.195	1.262	.925	-.734	3.259

על סמך הפלט ענה על הסעיפים הבאים:

- ברמת מובהקות 0.05, בדוק את ההשערה ששוניות מידת הביטחון בשתי הקבוצות שוות.
- ברמת מובהקות 0.05, בדוק את ההשערה שמידת הביטחון של הבוגרים גבוהה משל הצעירים, נמק.
- ברמת מובהקות 0.05, בדוק את ההשערה שההפרש במידת הביטחון בין הבוגרים לבין הצעירים שונה מ-2 יחידות. נמק.



שאלה 24

לפניך פלט של ניתוח סטטיסטי של בדיקות איכות בפס ייצור של אריזות מלח שמשקלן הנקוב 1000 גרם.

במהלך הבדיקה נשקלו 35 אריזות והתקבלו התוצאות הבאות:

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Salt	35	1003.06	36.445	6.160

One-Sample Test

	Test Value = 980					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Salt	3.743	34	.001	23.057	10.54	35.58

Explore Descriptives

			Statistic	Std Error
Salt	Mean		1003.06	6.160
	95% Confidence interval for Mean		Lower Bound Upper Bound	990.54 1015.58

על סמך הפלט ענה על השאלות הבאות:

בהנחה שהמשקל מתפלג נורמלי:

א. בדוק ברמת מובהקות 0.05 את ההשערה שתוחלת המשקל גבוהה מ- 980 גרם.

נסח את ההשערות ונמק.

ב. הראה שרווח הסמך בטבלה שנייה שקול לרווח הסמך בטבלה שלישית.

ג. בדוק ברמת מובהקות 0.05 את ההשערה שתוחלת המשקל שונה מ- 1000 גרם.

נסח את ההשערות ונמק.



שאלה 25

לבדיקת הטענה שסביבון מסוים הוגן בוצע ניסוי מתאים. ענה על הסעיפים הבאים על סמך ממצאי הבדיקה.

Chi – Square Test

Frequencies

Spin

	Observed N	Expected N	Residual
NES	30	25.0	5.0
GADOL	15	25.0	-10.0
HAYA	20	25.0	-5.0
PO	35	25.0	10.0
Total	100		

Test Statistics

	Spin
Chi-Square	A
df	B
Asymp. Sig.	.019
Exact Sig.	.018
Point Probability	.001

א. מהו המבחן הסטטיסטי שבוצע? נסח את ההשערות.

ב. מהם ערכי A ו-B? נמק.

ג. מהי מסקנת הבדיקה ברמת מובהקות 0.05? נמק.



שאלה 26

לפניכם ניתוח סטטיסטי של מספר חברים ברשת חברתית.

Descriptives

			Statistic	Std.Error
Friends		Mean	9.50	0.651
	95% Confidence interval for Mean	Lower Bound	8.14	
		Upper Bound	10.86	
		Median	9.00	
		Std. Deviation	2.911	
		Minimum	5	
		Maximum	15	

בהנחה שהמשתנה מתפלג נורמלית ענו על הסעיפים הבאים:

א. מהו הניתוח הסטטיסטי? מה גודל המדגם n ?

ב. ברמת מובהקות 0.05, בדוק את ההשערה שתוחלת מספר החברים שונה מ-10. נמקו

ג. ברמת מובהקות 0.05, בדוק את ההשערה שתוחלת מספר החברים קטנה מ-11. נמקו



שאלה 27

לפניכם ניתוח סטטיסטי של שיעורי הצבעה ל-3 גושי מפלגות בבחירות.

להלן הגדרות המשתנים:

Block – גוש (כחול, צהוב, ירוק)

Block

	Observed N	Expected N	Residual
Blue	40		6.7
Yellow	35		1.7
Green	25		-8.3
Total	100		

Test Statistics

	Block
Chi-Square	3.500 ^a
df	2
Asymp. Sig.	.174
Exact Sig.	.172
Point Probability	.008

א. מהו המבחן הסטטיסטי שבוצע? השלימו את הערכים החסרים בטבלה העליונה.

ב. מהן ההשערות שעבורן חושב הגודל Exact Sig.?

ג. בהמשך לסעיף ב, מהי מסקנת המבחן ברמת מובהקות 0.05?

openbook
המרכז לקידום אקדמי