



חיובי או חיובי כוזב?

הבעיה

בזלנד מתמודדים עם התפרצות של מחלה נגיפית.

כדי לבדוק הידבקות במחלה מבצעים בדיקת מעבדה.

לכל בדיקה רפואית יתכנו תוצאות שגויות הנובעות ממגבלות טכניות של המעבדות. כאשר אדם בריא נמצא בטעות חיובי, התוצאה נקראת חיובי כוזב (false positive). כאשר אדם חולה נמצא בטעות שלילי, התוצאה נקראת שלילי כוזב (false negative).

שאלה 1:

מה ההשלכות הבריאותיות והחברתיות של ריבוי מקרים של שלילי כוזב? מה ההשלכות של חיובי כוזב?

שאלה 2:

ידוע כי במעבדות של זלנד ההסתברות לשלילי כוזב היא 1.5%, כלומר, לכל חולה יש הסתברות של 1.5% לצאת בטעות שלילי בבדיקת המעבדה. ההסתברות לחיובי כוזב היא 3%, כלומר, לכל אדם בריא יש הסתברות של 3% לצאת בטעות חיובי בבדיקת המעבדה. לפניכם 2 טבלאות שמתארות תוצאות בדיקה של מדגם אקראי של 10,000 נבדקים בשני ימים שונים. איזו מהטבלאות מתאימה לנתונים על ההסתברות של חיובי/שלילי כוזב?

תוצאות הבדיקה			סה"כ
נמצאו חיוביים בבדיקה	נמצאו שליליים בבדיקה	סה"כ	
696	36	חולים	
598	8670	בריאים	
	10,000	סה"כ	

תוצאות הבדיקה			סה"כ
נמצאו חיוביים בבדיקה	נמצאו שליליים בבדיקה	סה"כ	
534	8	חולים	
266	9192	בריאים	
	10,000	סה"כ	

שאלה 3:

ביום מסוים תמר ודן נבחרו באקראי להשתתף בבדיקה מדגמית.

תמר נבדקה למחלה ונמצאה חיובית בבדיקה. דן נבדק ונמצא שלילי בבדיקה.

נתוני הבדיקות שהיו ידועים באותו היום מתוארים בטבלה הבאה.

תוצאות הבדיקה		
נמצאו חיוביים בבדיקה	נמצאו שליליים בבדיקה	
162	3	חולים
285	9550	בריאים

על פי נתוני הטבלה:

- א. ודאו כי המעבדות פעלו כצפוי ביום זה (כ- 1.5% שלילי כוזב וכ- 3% חיובי כוזב).
- ב. מה ההסתברות שתוצאת הבדיקה של תמר שגויה, כלומר שאיננה חולה למרות שיצאה חיובית בבדיקה?
- ג. מה ההסתברות שתוצאת הבדיקה של דן שגויה, כלומר שהוא חולה למרות שיצא שלילי בבדיקה?
- ד. מה דעתכם על אמינות הבדיקה? האם לדעתכם היתה תקלה במעבדות ביום שבו נאספו הנתונים?
רמז: מומלץ לחזור על החישוב של סעיפים ב-ג עם הטבלה משאלה 2, ולחשוב מדוע התוצאות משתנות מטבלה לטבלה. כמו כן אפשר לחזור לסעיף זה אחרי שאלה 4.
- ה. באותו היום נבדקו ויצאו חיוביים גם עמית ויובל. עמית נבדק כי נחשף לחולה מאומת, ויובל כיון שהופיעו אצלה סימפטומים של המחלה.
שניהם טוענים, על פי נתוני הטבלה, כי יש סיכוי גבוה מ 50% שתוצאת הבדיקה שלהם שגויה, והם מקרה של חיובי כוזב. מה דעתכם?

שאלה 4:

הקטע הבא מובא מתוך כתבה באתר גלובס. הכתבה הופיעה ביוני 2020, כשלושה חודשים לאחר התפרצות מגפת הקורונה בישראל¹.

בימים האחרונים מסתובבות ברשתות החברתיות, עדויות מטרידות לפיהן פרופ' איתמר גרוטו, משנה למנכ"ל משרד הבריאות, טוען כי רוב התוצאות החיוביות בבדיקות הקורונה הן "חיובי כוזב".

...

לבדיקות הקורונה אכן יכולה להיות תוצאה חיובית כוזבת. בתנאים אידאליים מדובר בכ-4% אחוז מהבדיקות שיוצאות חיוביות למרות שהנבדק אינו חולה בפועל, ובתנאים לא אידאליים יכול המספר להגיע אף למעל 10%. כדי להבין את התשובה כדאי להכיר לעומק את המשמעות של תוצאה חיובית כוזבת. בתנאי בדיקה אידאליים, אם נדגום 100 אנשים בריאים, 4 מהם יופיעו בטעות כחולים.

...

אם נדגום 2,000 איש בקבוצה שאין בה שום חולה, עדיין יהיו לנו המון בדיקות חיוביות - על פי נתון של 4% חיובי כוזב לבדיקת PCR אידאלית, הרי שמתוך אותם 2,000 בריאים, 80 אנשים יקבלו תשובה שהם חולים. אם נדגום 2,000 איש ויש שם חולה אחד, הרי שכנראה שנוזה את החולה היחיד (הרגישות של הבדיקה בתנאים אידאליים היא 100%), אבל יחד איתו נאבחן 79 בריאים כחולים. לכן משרד הבריאות נזהר מלבצע בדיקות סקר רחבות. לעומת זאת אם נדגום אוכלוסייה של 2,000 איש שיש בה 80 חולים, כנראה נאבחן כמה עשרות בריאים כחולים ועוד 80 חולים כחולים - יחס יותר טוב, שבו יש רק חיובי כוזב אחד עבור כל חיובי אמיתי, ואז כבר היה כדאי לבצע את הבדיקה.

¹ מצורף [קישור](#) לכתבה המלאה

קראו את הקטעים הבאים מתוך הכתבה ונסו להסביר במילים שלכם מדוע ביוני 2020 משרד הבריאות לא ביצע בדיקות סקר בכלל האוכלוסייה.



TOP 15
מרחיבים את מעגל המצוינות
نوسع دائرة التميز



תוכן עניינים

1.....	הבעיה
5.....	רקע לבעיה
5.....	ניתוח מיומנויות נדרשות
5.....	המלצות הפעלה
6.....	רמזים אפשריים
6.....	רעיונות נבחרים לפתרון

רקע לבעיה

הבעיה נוצרה בזמן שישראל והעולם התמודדו עם מגפת הקורונה. הדחף לכתיבת הבעיה נבע מדיון במדיה חברתית שבה נטען שבדיקות קורונה אינן אמינות, כיוון שיש הסתברות גבוהה שאדם שנמצא חיובי בבדיקה איננו חולה באמת. הרהור על טענה בעייתית זו הציף את ההבדל בין אמינות של בדיקה ואמינות התוצאה. מתברר שגם אם הבדיקה אמינה מאד (הסתברות נמוכה שבדיקה אקראית תניב תוצאה שגויה) עדיין יתכן שרוב התוצאות החיוביות הן שגויות! את זה אפשר לראות בדוגמה הקיצונית הבאה: נניח שהבדיקה "טועה" בהסתברות של 1/1000 (זו נחשבת בדיקה אמינה מאד), ונניח שאנחנו בודקים את האוכלוסייה באופן מדגמי. אם המחלה נדירה (נניח חולה 1 לכל 10,000 נבדקים), הרי שבמדגם של 10,000 אנשים אנחנו צופים שיהיה בערך חולה אמת אחד, וכ- 10 תוצאות חיוביות שגויות!

כתבה שהופיעה באתר גלובס דנה בתופעה זו, ומסבירה שמסיבה זו לא נעשו בישראל בדיקות סקר נרחבות לגילוי נגיף הקורונה. מטרת העיסוק בבעיה היא לא לבקר או להצדיק את המדיניות שננקטה, אשר נשענת על שיקולים רבים שלא מובאים בבעיה. מטרתה להכיר לתלמידים מושגים של חיובי/שלילי כוזב באופן כללי, לעסוק באופן אינטואיטיבי במצבים של הסתברות מותנית, להיחשף לתופעה הסתברותית שרבים טועים בה, וכל זאת בהקשר מוכר של מחלה ששכיחותה באוכלוסייה משתנה מעת לעת.

למי הבעיה מיועדת: לתלמידי כיתה ח' או ט', הקבצה א או מצוינות, או לתלמידי כיתה י'.

ידע נדרש: יש צורך לחשב שכיחות יחסית. אין צורך להכיר את המושג של הסתברות מותנית, למרות שהבעיה עוסקת בה בעקיפין.

ניתוח מיומנויות נדרשות

הבנת הבעיה והנתונים: רמה 5-6

נדרש מעבר בין ייצוגים שונים של הסתברות – מספרי (הסתברות של 1.5%) וייצוג באמצעות טבלת שכיחויות.

בחירת אסטרטגיה לפתרון ויישומה: רמה 4-5

כדי לחשב הסתברויות מתוך טבלת שכיחויות נדרשת הפעלה זהירה של רצף פעולות, והשוואה בין אסטרטגיות שונות (התבוננות בשורות הטבלה או בעמודות שלה). בסעיף שמבקש להשוות בין נבדק אקראי, נבדק שנחשף לחולה ומאומת ונבדק המראה סימפטומים נדרשת גישה חדשנית ושיקול דעת.

הערכת הפתרון, רפלקציה על הדרך: רמה 5-6

יש צורך לתקשר פרשנות ושיקול דעת, ולפרש את הממצאים ההסתברותיים ביחס לסיטואציה המתוארת.

המלצות הפעלה

מומלץ לחלק לתלמידים את הבעיה ולתת להם להתחיל לעבוד בקבוצות קטנות. מומלץ לוודא תוך כדי עבודת התלמידים שהם מבינים את המושגים ומחשבים את שכיחות היחסית הנכונה.

רמזים אפשריים

- לכוון תלמידים להתבונן בשכיחות היחסית של קבוצת החיובי כוזב (בריאים שיצאו חיוביים בבדיקה), פעם בהסתכלות בשורות (אחוז החיוביים מבין הבריאים) ופעם בהסתכלות בעמודות (אחוז הבריאים מבין החיוביים), לזהות שיש הבדל, ולקבוע על איזה מהם שואלת שאלה 2.
- לספר על בדיקות גנטיות, שבגלל השכיחות הנמוכה של מחלה, כמעט כל התוצאות החיוביות בבדיקות סקר הן שגויות. אפשר להדגים על מקרה קיצוני, למשל זה המופיע בתחילת מדריך זה.

רעיונות נבחרים לפתרון

שאלה 1:

זו לא שאלה מתמטית, אלא נועדה לדון בהשלכות של תוצאות שגויות.

שאלה 2:

תוצאות הבדיקה			
סה"כ	נמצאו שליליים בבדיקה	נמצאו חיוביים בבדיקה	
732	36	696	חולים
9268	8670	598	בריאים
10,000	8706	1294	סה"כ

תוצאות הבדיקה			
סה"כ	נמצאו שליליים בבדיקה	נמצאו חיוביים בבדיקה	
542	8	534	חולים
9458	9192	266	בריאים
10,000	9200	800	סה"כ

$$\frac{36}{732} \approx 5\%, \frac{598}{8670} \approx 6.4\% \quad \text{בעוד} \quad \frac{8}{542} = 0.01476 \approx 1.5\%, \frac{266}{9458} = 0.02812 \approx 3\%$$

יש לשים לב כי למשל עבור החישוב של החיובי הכוזב, ה-3% הם מתוך הבריאים, ולא מתוך כלל אלו שנמצאו חיוביים או מתוך כלל הנבדקים.

שאלה 3:

נתוני הטבלה אכן מתאימים לסטטיסטיקה: $\frac{285}{285+9550} = 0.028978$ וכן $\frac{3}{162+3} = 0.018$ (אפשר לתת מקום בכיתה לויכוח האם מספר אלה די קרובים ל-3% ול-1.5% בהתאמה).

יש לשים לב שבשני הסעיפים ב-ג אנו מחשבים מתוך "שלם" אחר, בסעיף ב' הסה"כ הוא מספר האנשים שיצאו חיוביים בבדיקה ולכן מחשבים את שיעור הטעות: $\frac{285}{447} = 0.64$. בסעיף ג' הסה"כ הוא מספר האנשים

שקיבלו תוצאה שלילית בבדיקה. שיעור הטעות מבין אלו שנמצאו שליליים: $\frac{3}{9553} = 0.00031$

בסעיף ד' על התלמידים לשים לב שיש הבדל בטבלאות, וככל ששכיחות המחלה עולה, כך יורד אחוז החיוביים הכוזבים מבין החיוביים. כשהמחלה נדירה, רוב התוצאות החיוביות במדגם אקראי הן כוזבות.

מטרת סעיף ה' היא להוסיף חשיבה ביקורתית לפרשנות הנתונים. אינו דומה נבדק שיש לו סימפטומים ויצא חיובי בבדיקה לאדם שנבדק באופן אקראי ויצא חיובי. את ההבדלים האלו, בסיבות לבדיקה, אנחנו לא רואים בטבלה.

שאלה 4:

בכתבה מנסים להסביר את מה שהתלמידים אמורים להסיק מהעיסוק בשאלות הקודמות. כאשר השכיחות של מחלה מסוימת נמוכה מאד, אז אחוז הבריאים מתוך אלו שיצאו חיובי גבוה מאד. המטרה היא שיראו את החישובים הללו בהקשר שהשפיע מאד על חיי היום-יום שלהם.

הערה נוספת

בבעיה זו אנחנו עומדים על התפר בין חישובים הסתברותיים והבנה סטטיסטית. זהו נושא גדול ורחב. מבחינת חישוב הסתברותי והסקה סטטיסטית השאלה המופיעה להלן לא מספיק מבוססת, אך מבחינה דיסקטית היא יכולה לתרום להבנת ההסתברות המותנית. ניתן להוסיף אותה כשאלה נפרדת (למשל שאלה 2 סעיף ב) או להציע לתלמידים לגוון את הנתונים בשאלה, ולהחליף את הטבלאות בשאלה בכאלו שבהן משנים את הנתון על אחוז החולים בזדלנד. לדוגמה

ביום מסוים כ- 10% מתושבי זדלנד היו חולים במחלה.

מלאו טבלה משלכם המתאימה לדגימה אקראית של 10,000 תושבי זדלנד, בהנחה שהמעבדות פעלו כצפוי (הסתברות של כ- 1.5% לשלילי כוזב והסתברות של כ- 3% לחיובי כוזב). בסעיף זה יש הרבה מאד אפשרויות, הנה תוצאה אפשרית אחת:

תוצאות הבדיקה			
סה"כ	תוצאות הבדיקה		סה"כ
	נמצאו חיוביים בבדיקה	נמצאו שליליים בבדיקה	
1015	15	1000	חולים
8985	8725	260	בריאים
10,000	8740	1260	סה"כ

טבלה זו נבנתה כלהלן:

תחילה מספר החולים נבחר להיות קרוב ל- 1000 (10% מסך הנבדקים, בהתאם לנתון על תפוצת המחלה). מספר הבריאים צריך להשלים ל- 10,000. אח"כ נבחרו מספר החיובים והשליליים הכוזבים (15, 260) כך שיתאימו לנתונים (כ- 3% מתוך הבריאים וכ- 1.5% מתוך החולים בהתאמה).

$$\text{ואכן מתקיים ש } 0.015 = 1.5\% \approx \frac{\text{חולים שנמצאו חיוביים}}{\text{סה"כ החולים}}, \quad 0.03 = 3\% \approx \frac{\text{בריאים שנמצאו שליליים}}{\text{סה"כ הבריאים}}$$