תפקיד החזרות בהסקה סטטיסטית וההבדל בין חזרות טכניות לביולוגיות/

צחי בר[1]

[1]מורה לביולוגיה בתיכון יצחק שמיר, פתח תקווה, 05.08.42.42.08 ,tzachi.bar@gmail.com

גרסה אחרונה: 11 ביולי, 2022

תקציר

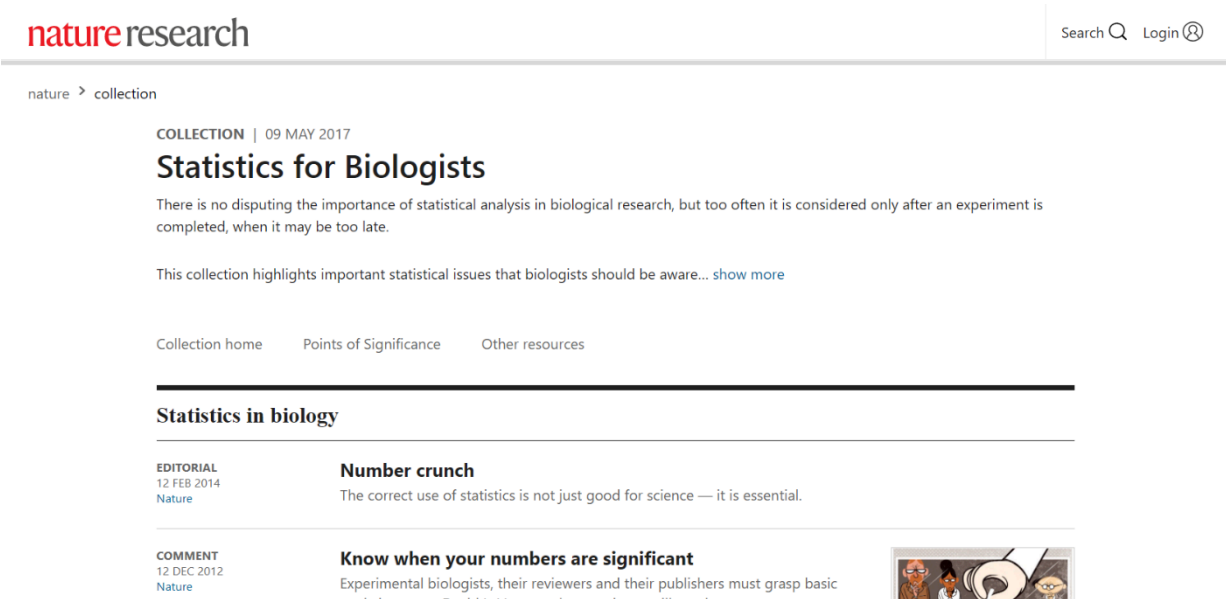
חזרות נמצאות בבסיס העשייה המדעית וקשורות באופן הדוק למדע טוב. דיונים בקבוצות מורים מהתקופה האחרונה והתבוננות בספרות הקיימת להדרכת מורים בנושא חזרות מצביעים על בלבול בקרב מורי הביולוגיה בכל הקשור להבחנה בין חזרות טכניות לביולוגיות ולתפקיד החזרות בכלל. בחינה של ספרות הדרכת המורים מצביעה על התמקדות בשני תפקידים בלבד של חזרות – שיפור דיוק (של אומדן הממוצע) ו"לוודא שהתוצאות לא מקריות". מכיוון שההסבר אינו מתייחס לתפקיד החזרות בהסקה סטטיסטית בהקשר רחב יותר, קשה להבין ממנו את מקור ההבחנה בין חזרות טכניות לחזרות ביולוגיות.

במסמך זה אני מרחיב את המבט על תפקידי החזרות בהקשרן הסטטיסטי באופן כללי, ומשם עובר למקרים הפרטיים של חזרות טכניות וחזרות ביולוגיות. לבסוף מוזכרים שני חסרונות של מספר גדול מדי של חזרות.

המסמך מתבסס על מאמרים מתוך הסדרה Statistics for Biologists שהתפרסה בשנים 2014-15 במגזין Nature, מלווה בדוגמאות מעולמות החקר המוכרים לתלמידים וכתוב תוך שימת דגש על עקרונות סטטיסטיים ללא כניסה למתמטיקה הנלווה.

מבוא

בשנת 2005 פרסם מדען יווני בשם ג'ון איונידיס מאמר עם כותרת פרובוקטיבית "למה מרבית הממצאים המחקריים המתפרסים הם שגויים?" (Ioannidis, 2005). המאמר שצוטט כ-9,000 פעמים ניער הייטב את העולם המדעי. איונידיס הסביר והדגים את הטענה שלו בצורה משכנעת, ובעקבותיו גאה גל של מאמרים שבדקו את היכולת לשחזר תוצאות של מחקרים קודמים (Reproducibility). הממצאים שעלו במאמרים היו נוקבים והובילו למספר מסקנות, כשאחת מהן – צריך לשפר את ההבנה הסטטיסטית של חוקרים רבים. בעקבות המסקנה הנ"ל מגזין Nature פרסם סדרת מאמרים בנושא תכנון ניסויים וניתוח תוצאות שנקראת Statistics For Biologists. הסדרה מציגה בצורה ברורה את העקרונות החשובים של תכנון ניסוי וניתוח תוצאות. הסידרה אינה נכנסת לעומק המתמטי של הסטטיסטיקה, היא מתמקדת בצד היישומי של סטטיסטיקה לחוקרי ביולוגיה.



אחד המאמרים בסדרה עסק בחזרות טכניות ובחזרות ביולוגיות (Paul Blainey, 2014), ומאמר אחר עסק במקורות של שונות (Naomi Altman, 2014), יחד הם מהווים את הבסיס לטקסט הנוכחי.

להבחנה בין סוגי החזרות חשיבות רבה, שכן הבחנה שגויה בין סוגי חזרות פוגעת בתקפות המחקר, ובמקרים רבים תאיין את מסקנות המחקר מכל ערך.

"לא חוזרים על ניסוי שהצליח" אדוארד. א. מרפי, או שבעצם כן?

אחד מחוקי מרפי אומר "לא חוזרים על ניסוי שהצליח" (Bloch, 2003). הדבר העצוב בחוק המצחיק הזה הוא שהוא נכון סטטיסטית![1] אבל אין מה לעשות, מדע טוב מבוסס על הסקה סטטיסטית, אוסף כלים שעוסקים בהסקת מסקנות לגבי אוכלוסיות. מכיוון שחזרות הן אלו שבונות אוכלוסיות, חזרות הן אלו שמאפשרות הסקה סטטיסטית. בין הדברים שהסקה סטטיסטית עושה באמצעות חזרות:

- אומדן של פרמטרים סטטיסטיים כגון ממוצע וסטיית תקן (ככל שגודל המדגם עולה כך משתפר דיוק האומדן)
- שימוש באותם אומדנים כדי לכמת את הבטחון במסקנות של ניסויים
- הכללת הממצאים מהמחקר לכלל האוכלוסייה שמיוצגת במחקר
- זיהוי תוצאות חריגות כדי להחריג אותן מניתוח התוצאות

כדי להסביר כיצד חזרות מאפשרות לחוקר לעשות את שני הסעיפים הראשונים נלווה זוג תלמידים, נוגה ואלי, שמנסים לענות על השאלה "מה הקשר בין ריכוז מלח במי השרייה של זרעי מש לנביטה", מתוך התבוננות באורך השורשון של הנבטים. הדיאלוג וההסברים לאורכו ואחריו אינם כוללים ביטויים

[1] ראו תסוגה אל הממוצע בויקיפדיה

מתמטיים, אך כוללים הפניות לכאלה עבור הקוראים המתעניינים. בחלק מהמקומות נעשה עיגול פינות קל לטובת פישוט ההסבר. התפקידים האחרים של החזרות יוזכרו רק בקצרה.

הדרך למסקנה רצופה בכוונות טובות

אלי: מאוד קל לראות את הקשר בין ריכוז מלח במי השרייה ונביטה של זרעי מש. הנבטתי זרע אחד במי ברז וזרע נוסף בתמיסת מלח, השורשון של הנבט שהיה במי מלח קצר יותר, 20 מ"מ, לעומת 21 מ"מ במי ברז. מסקנה: תמיסת מלח מעכבת נביטה.

נוגה: אתה בטוח שאלו אורכי השורשונים? אולי תמדוד שוב?

אלי: צודקת. חכי רגע. [הולך למדוד, חוזר אחרי דקה]. בעצם... אורך השורשון בתמיסת מלח 22 מ"מ ולא 20 מ"מ, במדידה הראשונה לא מדדתי את הפיתול בשורשון ליד הזרע. יותר טוב! תמיסת מלח מאיצה נביטה. עלינו על משהו חדש. נוגה, מלח בישול הוא דשן! מלח בישול מאוד זול, אנחנו הולכים להציל את העולם! יש לנו סטארט אפ!

נוגה: בטוח?

אלי: את יודעת מה. חכי רגע. אמדוד את האורך של כל שורשון 30 פעמים ואחשב ממוצע. [חוזר אחרי שעה, אבל וחפוי ראש]. טוב, יש הכרעה. האורך הממוצע של השורשון במי ברז הוא 20.5 מ"מ, והאורך הממוצע של השורשון בתמיסת המלח 19.7 מ"מ. מלח מעכב נביטה. [הבהרה: אם אתם חושבים שזאת חזרה מיותרת, אתם צודקים, זה רק לצורך הדיון שלנו].

נוגה: בטוח?

אלי: למה לא? חזרות משפרות דיוק של מדידה.

נוגה: זה נכון, אבל סבלנות, אתה רץ ישר לסוף. יש לנו כמה דברים להכיר על חזרות לפני שנגיע לשיפור דיוק של מדידה.

כשאלי מדד את אורך השורשון מספר פעמים הוא גילה שלא כל המדידות היו זהות. איזה מהתוצאות מייצגת את התשובה האמיתית? הוא הרי לא יכול להמשיך ולמדוד עד אין סוף וכל פעם לשנות את התשובה לפי המדידה האחרונה. כדי לפתור את הבעיה אלי לקח קבוצת מדידות שמייצגת את כל המדידות והסיק מהן לגבי כל השאר. לפעולה הראשונה של אלי, בחירת תת-קבוצה מתוך האוכלוסייה (במקרה זה של מדידות האורך), קוראים דגימה. אלי דגם[2] את המדידות. המדגם, אם נעשה לפי הכללים (Daniel, 2012), מאפשר לאלי לראות כיצד התכונה משתנה בקרב האוכלוסייה, ובאמצעות אוסף הכלים של הסקה סטטיסטית, הוא יכול להסיק מסקנה אותה הוא יוכל להכליל מהמדגם לכלל המדידות. כלומר, אלי יוכל לומר משהו על ההבדלים בין כל האורכים שימדדו משני השורשונים האלה, גם אם הוא לא מדד אותם אין סוף פעמים.

---

[2] כללי הדגימה, המסבירים את מה, כמה ואיך לדגום, הם בעלי חשיבות גדולה להצלחת הניסוי, אבל מחוץ לנושא המאמר הזה. קוראים המעוניינים להרחיב את ידיעתם בנושא מוזמנים לקרוא את הפרק Methods of sampling from a population בעמוד של HealthKnowledge.

פיזור אורכי השורשונים ביחס לאורך הממוצע שלהם נקרא שונות (Variance). כאשר אנחנו דוגמים את האוכלוסייה באמצעות חזרות אנחנו יודעים איך התכונה משתנה באוכלוסייה, כלומר, דגמנו את השונות. בעזרת הידע הזה אנחנו יכולים להכליל את הממצאים שלנו, תחת מגבלות מסוימות, נרחיב על זה עוד מעט.

לעת עתה נסכם את שנאמר עד כה:

כדי להכליל מסקנה לאוכלוסייה באמצעות הסקה סטטיסטית עלינו לדעת איך המשתנה שלנו משתנה בקרב האוכלוסייה, כלומר, את השונות שלו. לצורך כך עלינו לדגום מספר פעמים את האוכלוסייה, קרי, לבצע מדידות חוזרות של המשתנה אותו אנו רוצים להכליל.

נמשיך להסתכל על הדיאלוג של נוגה ואלי. הוא מזמן מבט על יסודות ההסקה סטטיסטית.

כאשר אנחנו שואלים שאלה בהסקה סטטיסטית 'האם טיפול מסוים השפיע על התוצאות?' אנחנו שואלים מה ההסתברות לקבל במקרה את התוצאות שהתקבלו, או קיצוניות מהן, בהנחה שהטיפול לא משפיע על התוצאות.

כדי לדעת מה הסיכוי לקבל את תוצאות הניסוי שלנו במקרה, צריך לדעת איך התוצאות משתנות בקרב האוכלוסייה. השונות בחזרות מראה לנו איך התוצאות משתנות. כך החזרות מאפשרות הסקה סטטיסטית והכללה של התוצאות.

נוגה: סבלנות אלי, אתה רץ ישר לסוף. יש לנו כמה דברים להכיר על חזרות לפני שנגיע לשיפור דיוק של מדידה.

אלי: חזרתי על המדידה של אורך השורשון 30 פעמים!

נוגה: מצוין, אז אתה יכול להסיק מסקנה, באופן סטטיסטי, לגבי אורך השורשונים.

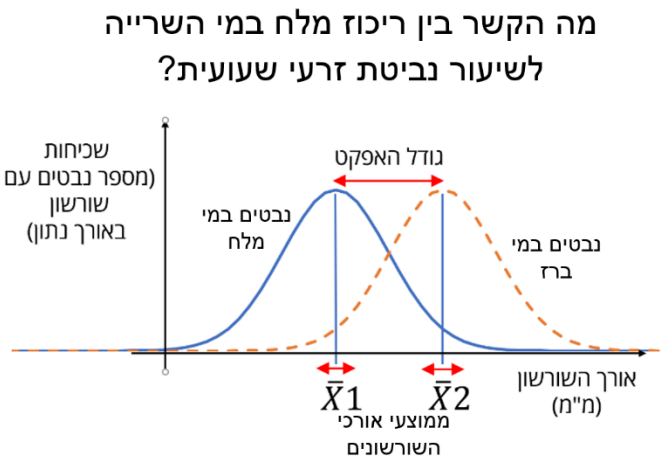
אלי: איך עושים את זה?

נוגה: כדי לבצע הסקה סטטיסטית מבחנים סטטיסטיים רבים בודקים את היחס בין גודל האפקט של הטיפול לרעש במדידת האפקט (איור 2). הכוונה ברעש היא לחוסר הדיוק במדידת האפקט. אגב, לאפקט קוראים גם סיגנל, ולכן היחס הזה נקרא Signal to Noise Ratio. אם האפקט (סיגנל) בולט מאוד על פני הרעש, כנראה שמדובר באפקט של ממש. אך אם האפקט נבלע ברעש, כנראה שמדובר ברעש (איור 3).

התפקי

מה הסיכוי שקי
התוצאה שקיבלנו
בהנחה שהטיפ
התו

איור 2: התפקיד העיקרי של חזרות בניסוי. כדי לענות על השאלה "מה הסיכוי שקיבלנו במקרה את התוצאה שקיבלנו, או קיצונית ממנה, בהנחה שהטיפול לא משפיע על התוצאות" רבים מהמבחנים הסטטיסטיים בודקים את היחס בין האות, שהוא גודל האפקט של הטיפול (מסומן בחץ מעל הגרף), לרעש, שהוא חוסר הדיוק במדידת האפקט (מסומן בחיצים מתחת לגרף). איך חזרות קשורות לבדיקה הזאת? חזרות מאפשרות לאמוד את השונות, שחיונית לחישוב הרעש.



במקרה הנוכחי, עליך לחשב את ממוצע האורך של כל שורשון, וההפרש בין הממוצעים הוא האפקט (החץ העליון באיור 2). כשחזרת על מדידת אורך כל שורשון באותם תנאים ראית איך התוצאה משתנה בין המדידות השונות. המדד המתאר את פיזור התוצאות האלה ביחס לממוצע נקרא שונות, Variance. בעזרת השונות ומספר החזרות ניתן לחשב את חוסר הדיוק במדידה של הממוצע של אורך כל שורשון, זה כל אחד משני החצים הקטנים בתחתית איור 3 (למתעניינים: שניאת תקן). אם החץ העליון גדול יותר מהחיצים התחתונים ההפרש בממוצעי האורכים הוא בסבירות גבוהה לא מקרי (למעשה זה קצת יותר מורכב. למתעניינים ראו לדוגמה מבחן t).

חזרות מאפשרות לאפיין את השונות במדידה, הנחוצה להסקה סטטיסטית.

נוגה: עכשיו שם לב לעוד משהו מעניין. אם מחשבים כמה ממוצעים, כל אחד באמצעות מספר קטן של חזרות, מקבלים ממוצעים שונים מאוד אחד מהשני. אבל אם מחשבים כמה ממוצעים, כל אחד באמצעות מספר גדול של חזרות, מקבלים ממוצעים שיותר קרובים אחד לשני. זה בגלל שבקבוצה קטנה, לתוצאה

חריגה יש השפעה יותר גדולה על הממוצע, מאשר בקבוצה גדולה. בקבוצה גדולה, תוצאות נדירות חריגות "מדוללות" בתוך יותר תוצאות לא חריגות, ולכן ההשפעה שלהן על הערך הממוצע, קטנה יותר. זה אומר שחזרות משפרות את הדיוק של אומדן הממוצע, כלומר, מקטינות את החיצים התחתונים באיור 3 ובעצם הן מבליטות את הסיגנל לעומת הרעש. המשמעות של זה היא שחזרות רבות יותר מאפשרות לגלות אפקט קטן יותר.

סיכום ביניים:

הגדלת מספר החזרות תשפר את דיוק אומדן הממוצע (ועוד פרמטרים). לכן, ככל שמספר החזרות גדול יותר ניתן לזהות הבדלים קטנים יותר בין הטיפול לבקרה.

אלי: למה אמרת 'אומדן הממוצע' ולא 'ממוצע'?

נוגה: כי ממוצע הוא תוצאת חישוב הכוללת בתוכה את כל האוכלוסייה, ואומדן של הממוצע מחושב מתוך מדגם של האוכלוסייה.

אלי: אוקיי, אז אם השורשון של הנבט שהושרה במי ברז ארוך יותר מהשורשון של הנבט שהושרה בתמיסת מלח, אז מלח מעכב נביטה!

נוגה: איך אתה יודע את זה?

אלי: השורשון של הנבט שהונבט במלח יותר קצר! מה לא ברור?

נוגה: אתה בטוח שהוא יותר קצר?

אלי: בוודאי! חזרתי על המדידה 30 פעמים.

נוגה: ומה יקרה אם תחזור על הניסוי עם זוגות נבטים נוספים?

אלי: כבר יש לי חזרות, מדדתי 30 פעמים כל שורשון.

נוגה: כשחזרת על מדידת אורך השורשונים אפיינת את השונות של מדידת האורך, לא את השונות של תגובת הזרעים לטיפול. כדי לומר משהו על השפעת הטיפול על זרעים אחרים אתה צריך לחזור על הטיפול – השרייה במי ברז ותמיסת מלח עם זרעים נוספים. אכליל ואומר:

חזרות על מרכיב בניסוי מאפשרת להסיק מסקנות סטטיסטיות רק לגבי הרכיב שחזרו עליו.

מהסיבה הזאת מקובל להבדיל בין שני סוגים של חזרות:

חזרות טכניות לעומת חזרות ביולוגיות

חזרות טכניות הן מדידות חוזרות של אותה דגימה ביולוגית**. חזרות טכניות מייצגות רעש אקראי, בלתי תלוי, הקשור לפרוטוקול או לציוד המדידה.**

חזרות ביולוגיות הן מדידות מקבילות של דגימות ביולוגיות נפרדות**. חזרות ביולוגיות כוללות את השונות הביולוגיות האקראית של המשתנה הנמדד. מכאן, שחזרות טכניות יכולות לשמש להסקה סטטיסטית לגבי הטכניקה, אבל רק חזרות ביולוגיות יכולות לשמש להסקה סטטיסטית לגבי הביולוגיה והכללת המסקנות לפרטים נוספים.**

בעוד שחזרות ביולוגיות בהכרח מחייבות גם חזרה טכנית כלשהי, חזרות טכניות יכולות, אבל אינן מחייבות, להיות כרוכות עם חזרות ביולוגיות. הדוגמה של מדידה חוזרת של אותו שורשון מציגה חזרה טכנית ללא חזרה ביולוגית. איור 4 מסכם את המאפיין והמבדיל בין שני סוגי החזרות.

נסתכל על שתי דוגמאות, הראשונה פשוטה והשנייה יותר מורכבת.

עוד דוגמאות

נחזור לנוגה ואלי שבדקו את הקשר בין ריכוז מלח במי השרייה של נבטי מש לבין קצב הנביטה באמצעות מדידת אורך השורשון לאחר 72 שעות.

נוגה ואלי הכינו 5 כוסות להנבטת זרעי מש, כוס אחת עם מי ברז, וארבע כוסות עם ריכוזי מלח שונים. בכל כוס השרו 30 זרעי שעועית. לאחר 72 שעות הם מדדו שלוש פעמים את אורך השורשון של כל נבט.

מה כאן חזרות טכניות ומה חזרות ביולוגיות?

- 30 זרעי שעועית בכל כוס – כל זרע הוא מרכיב ביולוגי נפרד, ולכן זאת חזרה ביולוגית
- שלוש מדידות של אורך אותו שורשון – חזרה טכנית על אותה דגימה ביולוגית.

הנקודה המרכזית בהבחנה בין חזרות טכניות לחזרות ביולוגיות נמצאת בתשובה לשאלה "האם החזרה נעשית על אותה דגימה ביולוגית או על דגימות ביולוגיות נפרדות?".

ההבדל בין

חזר

איור 4: המאפיין והמבדיל בין חזרות ביולוגיות וטכניות. חזרות ביולוגיות הן מדידות מקבילות של המשתנה הנמדד על דגימות ביולוגיות שונות, שכוללות את השונות הביולוגית, האקראית, של המשתנה הנמדד. אותה שונות הכרחית להסקת מסקנות על המשתנה הביולוגי ולזיהוי תוצאות ביולוגיות חריגות. לעומת זאת, חזרות טכניות נעשות על אותה דגימה ביולוגית, ולכן לא כוללות את השונות הביולוגית, האקראית, של המשתנה הנמדד. לעומת זאת הן כוללות רעש אקראי, בלתי תלוי, הקשור לשונות של הפרוטוקול או ציוד המדידה. השונות הזאת, הטכנית, הכרחית להסקת מסקנות לגבי המשתנה הטכני ולזיהוי תוצאות חריגות מסיבות טכניות.

נסתכל על ניסוי אחר, שיוביל אותנו למסקנה מפתיעה.

זוג תלמידים רצה לבדוק את הקשר בין ריכוז סוכרוז בתמיסת הגידול של שמרי אפיה לקצב הנשימה התאית של השמרים. התלמידים הכינו סטוק של תרחיף שמרים והוסיפו נפח שווה של הסטוק ל-15 מבחנות עם נפחים שווים של תמיסות סוכרוז ב-5 ריכוזים שונים של סוכרוז, שלוש מבחנות לכל ריכוז סוכרוז. לאחר 30 דקות הם עצרו את הניסוי ומדדו את כמות הפד"ח שפלטו השמרים בכל בקבוק.

האם שלוש המבחנות עבור כל ריכוז סוכרוז הן חזרות טכניות או ביולוגיות?

כדי לענות על השאלה ננסח את השאלה מעט אחרת "האם בשלוש המבחנות דגימות זהות מבחינה ביולוגית? האם השמרים במבחנות זהים?".

מכיוון שהמדידה נעשית על כל השמרים שבמבחנה יחד, יש להתייחס אל כל השמרים במבחנה כדגימה ביולוגית אחת. מאחר שמקור כל המבחנות מאותו סטוק, ובכל מבחנה יש הרבה מאוד שמרים (מאות מיליונים?) חוק המספרים הגדולים יאמר שגם אם יש מידה מסוימת של שונות בין השמרים, הם התפזרו באופן אחיד בין כל המבחנות בזמן החלוקה. על כן, תוכן המבחנות זהה מבחינה ביולוגית, ואלו חזרות טכניות בלבד.

משמעות הדבר שמבחינה מדעית, ניסויים כמו הניסוי שתואר בפסקה הקודמת, בגלל שאין בהם חזרות ביולוגיות אי אפשר להכליל אותם והם יהיו רלוונטיים רק לסטוק שממנו הוכנו. כדי להסיק מסקנות לגבי הקשר בין ריכוז סוכרוז לקצב נשימה תאית בשמרים צריך לחזור על הניסוי עם שמר בודד בכל פעם. ניתן לגדל ממנו תרחיף עם שמרים רבים, וכפי שצויין, אם הם נמדדים יחד הם נחשבים ליחידה ביולוגית אחת, וכך לייצר חזרות ביולוגיות. הדבר נכון גם לחיידקים. הניסוי שתואר כאן בקצרה והמסקנה שלו מתוארים בפרוטרוט במאמר של Naomi Altman, רק עם תאי מכרסם במקום שמרים (Naomi Altman, 2014).

נחזור לנוגה ואלי.

אלי: את יודעת מה, אני אמדוד את הנשימה התאית של תאי שמר אפייה בודדים, אצור חזרות ביולוגיות, וכך אוכל לאפיין את השונות של הנשימה התאית באוכלוסיית שמרי האפייה ולבדוק האם יש הבדל בנשימה התאית שלהם עם סוכרוז ובלעדיו. מכאן נוכל להסיק מסקנות לכל סוגי השמרים.

נוגה: אתה בטוח?

אלי: בוודאי! יש לי חזרות ביולוגיות. את אמרת שמדידה על תאי שמר נפרדים תהייה חזרה ביולוגית.

נוגה: כשמדדת את קצב הנשימה התאית על תאי שמר בודדים יצרת חזרות ביולוגיות עבור אותו מין שמר, אבל לא דגמת מיני שמרים שונים, ולכן אין לך חזרות ביולוגיות עבור מיני שמרים אחרים.

אלי: נראה לי שאני מזהה תכונה מעניינת של חזרות. מקודם אמרת ש" חזרות על מרכיב בניסוי מאפשרת להסיק מסקנות סטטיסטיות רק לגבי הרכיב שחזרו עליו". הכלל הזה תופס גם לרמות ארגון! חזרה ברמת ארגון נמוכה לא מכילה חזרות ברמת ארגון גבוהה ממנה, ולכן לא ניתן להסיק מסקנות סטטיסטיות מרמת ארגון נמוכה לגבוהה ממנה.

נוגה: אתה צודק אבל אתה גם נכנס לנושא קצת מורכב. בוא נשאיר משהו גם ללימודים באוניברסיטה?

ומה עם "תפקיד החזרות לוודא שהתוצאות לא מקריות"?

אם תשאלו תלמיד ביולוגיה "מה תפקיד החזרות בניסוי?" התשובה השכיחה תהייה "לוודא שהתוצאות שקיבלנו לא מקריות ושהן חוזרות על עצמן". זה מה שהמורה שלו, שנעזרה במונחון של משרד החינוך, לימדה אותו. שכן, על פי המונחון, חזרות, ללא הבחנה בין חזרות טכניות לחזרות ביולוגיות, הן "פעולות שמבצעים במהלך תכנון מערך החקר וביצועו, במטרה למנוע שגיאות מדידה והפרעות לא צפויות, ובמטרה להגביר את מהימנות התוצאות (לוודא שהתוצאות אינן מקריות)". (משרד החינוך, המזכירות הפדגוגית, אגף מדעים, הפיקוח על הוראת המדע והטכנולוגיה).

ההגדרה הזאת קרובה לאמת אבל מערבבת תפקידים שונים של החזרות. בשלב הראשון, מספר החזרות וההתפלגות שלהן, מאפשרים לאפיין את השונות במשתנה וכך לזהות באופן הסתברותי חריגות ממנה (איור 5). תוצאות חריגות יכולות להיות מאוד מעניינות, כמו זרע עמיד למלח בין זרעים רגילים, או שגיאת מדידה. אבל, תוצאות חריגות, בהגדרתן, הן תוצאות שבסבירות גבוהה לא מייצגות את אוכלוסיית המחקר ולכן נרצה להחריג אותן מניתוח הנתונים בהמשך (ראו את הערך Anomaly detection בויקיפדיה). זיהוי תוצאות חריגות ונטרולן, מטייבים את ההסקה הסטטיסטית, שאומנם באה אחריהם באופן כרונולוגי, אך עולה עליהם בחשיבותה ובתרומתה לגילוי תופעות מדעיות והכללתן.

איור 5: מספר החזרות ואופן התפלגות שלהן, מאפשרים לזהות תוצאות חריגות.

מבולבלים? אתם ממש לא לבד

אם עניין תפקיד החזרות בכלל וההבחנה בין חזרות טכניות לחזרות ביולוגיות בפרט, נשמע לכם מבלבל אז הרגישו בנוח, אתם ממש לא לבד. הבלבול במושגים סטטיסטיים, ובפרט בין חזרות טכניות לביולוגיות לא התחיל אצל מורי הביולוגיה. מחקר שבדק לאחרונה 200 מחקרים באופן אקראי מצא שרק כ 22% מהם השתמשו נכון בחזרות טכניות וביולוגיות (Stanley E. Lazic, 2018). כלומר, רוב החוקרים התייחסו בניתוח הנתונים לדגימות כנפרדות במובן הביולוגי שלהן, כשהן לא היו כאלו. חזרות כאלו נקראות חזרות מדומות (Pseudo-replicates). מחקר אחר, שדגם גיליון של Nature Neuroscience, מגזין מכובד מאוד לכל הדעות, הראה שרק בחצי מהמאמרים ניתן היה למצוא שימוש נכון בחזרות טכניות וביולוגיות (Lazic, 2011). אלו רק שתיים מתוך דוגמאות רבות לבעיה קיימת. בקיצור, אנחנו לא לבד.

הבלבול בספרות המחקרית כנראה הוביל לניסוחים לא ברורים ולא מדוייקים גם בספרות למורים. בפרק שעוסק בחזרות ביולוגיות וטכניות וחזרות בספר 'ביולוגיה חוקרת' של עמינדב ברנשטיין (ברנשטיין, 1991) אין הגדרה כלשהי של המונחים, אלא רק ציון שמם באנגלית. בפרק אומנם מצוינים בערבוביה מספר תפקידים של חזרות כולל בהסקה סטטיסטית, לרבות שיפור הדיוק של מדידת הממוצע. אך חוסר הדיוק במדידה ("רעש") אינו מקושר למדידת היחס בין אות לרעש. לכן, מקור תפקיד החזרות בהסקה סטטיסטית נישאר לא ברור. בנוסף, אין התייחסות מפורשת לכך שחזרות מאפשרות הסקה סטטיסטית רק על מה שחזרו עליו, ומכאן גם לא ברורה החשיבות של סוגי החזרות השונים. אולי בגלל כל אלה, ברנשטיין מגדיר את החזרות בניסוי עם בקבוקי השמרים שהוזכרו למעלה, באופן שגוי, כשילוב של חזרות טכניות וחזרות ביולוגיות, כאשר מדובר בחזרה טכנית בלבד.

ברנשטיין שואל "האם חזרות רבות "מחפות" על ניסויים מיעוטי פריטים?" ועונה "לא בהכרח! הרי לא ניתן לחזור על אותו ניסוי על כל פרטיו". התשובה של ברנשטיין כללית מדי ולא ברור מתי חזרות טכניות כן מחפות ועד כמה, על חזרות ביולוגיות.

התשובה לשאלה הזאת היא "כן, עד גבול מסוים". השונות הכוללת בניסוי היא סכום השונות הביולוגית והשונות הטכנית. חזרות טכניות מקטינות רק את האפקט של השונות הטכנית על דיוק המדידה של הממוצע. כדי להקטין את ההשפעה של השונות מעבר לכך, אין ברירה אלא לבצע חזרות ביולוגיות כדי לשפר את יכולת ההסקה הסטטיסטית (Naomi Altman, 2014).

סיכום

לחזרות תפקיד מרכזי במדע. חוסר היכרות עם הגדרות סוגי החזרות השונים, ותפקיד החזרות עלולים להוביל לבלבול בין סוגי החזרות וכתוצאה מכך לפגיעה בתקפות המסקנות, עד כדי איון מלא שלהן.

לחזרות מספר תפקידים. התפקיד העיקרי של החזרות הוא לאפשר הסקה סטטיסטית, שבעיקרה היא כלי לכמת את הביטחון שלנו במסקנות הניסויים. להעביר אותנו משפה של "נראה לי ש" לשפה של "בהסתברות של". ואם חישבנו נכון את ההסתברות, אנחנו יכולים לסמוך עליה. למעשה אנחנו סומכים עליה כל יום, כל היום. קשה לדמיין את המדע המודרני, בלעדי הסתברות, ומה שמייצר את ההסתברות אלו הן החזרות. תפקיד נוסף הוא לזהות תוצאות חריגות, שכנראה לא מייצגות את אוכלוסייה שממנה נדגמו הפרטים שהשתתפו בניסוי, ולכן צריך להחריג אותן מניתוח הנתונים. מאפיין נוסף של החזרות הוא שהתפקידים שלהן, שהכרנו כאן, רלוונטיים רק למה שחוזרים עליו.

אז על מה כדאי לחזור? על שלבים טכניים (חזרות טכניות) או על דגימות ביולוגיות נפרדות (חזרות ביולוגיות)? נתחיל ממה שחייבים לחזור עליו –חזרות ביולוגיות. חזרות ביולוגיות מאפשרות לנו לאפיין את השונות

הביולוגית, שהיא הכרחית להסקת מסקנות ביולוגיות, מאפשרת להכליל את התוצאות באופן ביולוגי, לדייק את אומדן הממוצע וגם לזהות תוצאות חריגות. בחלק מהמקרים השונות הביולוגית מעניינת בפני עצמה. בנוסף, חזרות ביולוגיות בהכרח כוללות גם חזרות טכניות.

אך מה בנוגע לחזרות טכניות? חזרות טכניות יאפשרו לאפיין את השונות בשלב הטכני, ויקטינו את האפקט של השונות הטכנית על הסקה סטטיסטית. זה יכול להיות מועיל, כי בתכנון ניסוי נכון, חזרות טכניות מאפשרות לזהות ולבודד תוצאות חריגות, והחקירה שלהן יכולה לגלות את מקור החריגה, כך שלעיתים ניתן לטפל במקור החריגה, להפחית אותו, ולשפר את דיוק המדידה. זה בתורו ישפר את הזיהוי של אפקט ביולוגי.

על איזה שלב טכני כדאי לחזור? זאת כבר שאלה של עלות מול תועלת. אם יש שלב טכני עם שונות גבוהה ביחס לאפקט שאנחנו רוצים למדוד, כדאי לבדוק אפשרות לחזרות בשלב הזה. אם השונות היא נמוכה ביחס לאפקט שאנחנו רוצים למדוד, כנראה שאפשר לוותר על חזרות בשלב הזה.

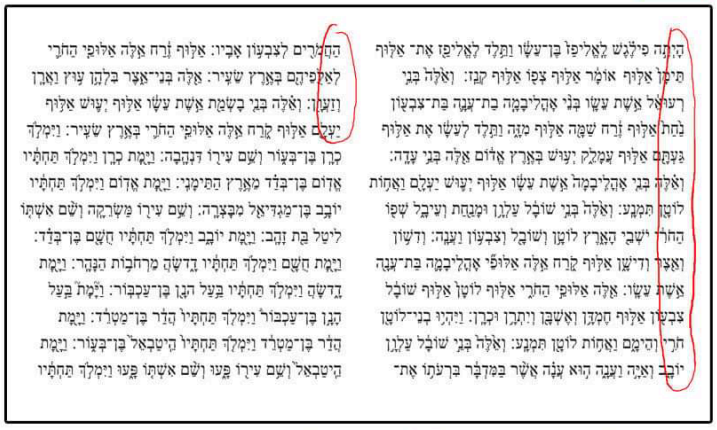
הערה לסיום, שאומנם פחות רלוונטית לביוחקר, אבל אם כבר נגענו בנושא שווה להזכיר אותה. בניגוד לאינטואיציה, לא כל הגדלה של חזרות ביולוגיות בהכרח טובה יותר. כאשר המדגם גדול מאוד אנחנו עלולים לגלות אפקטים קטנים, שאומנם יהיו מובהקים מבחינה סטטיסטית אבל חסרי משמעות ביולוגית (loannidis, 2005).

Wako | 100% דיסאינפורמציה

מדהים: בפרשת השבוע כבר נכתב שהתרנגול הוא צחי הלוי

הצופים נדהמו לגלות אמש (מוצ"ש) כי התרנגול ב"הזמר במסכה" הוא השחקן צחי הלוי, אולם מי שקרא בפרשת השבוע של אותה שבת לא היה מופתע: בפרשת וישלח, פרק ל"ו, כתוב "התרנגול הוא צחי הלוי" לפי צירוף האותיות הראשונות בכל שורה. "זה רק מוכיח שאלוהים הוא יודע מראש מי כל המתמודדים ב'זמר המסכה'"

מערכת Wako | מדור סאטירי | פורסם 10:25 06/12/20



בתמונה: קטע מתוך פרשת השבוע, "וישלח", עם סימון של האותיות הראשונות בכל שורה. "בהפטרה היה כתוב בדילוגי אותיות שהתרנגול הוא יהודה לוי אבל זה לא נחשב" (Wako)

מדגמים גדולים מדי, עם הרבה מאוד פריטים, הם בעייתיים גם במובן אחר, כאשר המדגם גדול מספיק עולה הסבירות שיופיעו בו גם אירועים אקראיים נדירים. מכיוון שלבני אדם יש נטייה ליחס משמעות גם לאירועים אקראיים, חסרי משמעות, תוצאות המחקר עלולות לקבל פרשנות המייחסת להם משמעות לא אקראית (Sophie Fyfe, 2008). ראו לדוגמה דילוג אותיות בתנ"ך או במקומות אחרים, כמו באיור 6.

אחרי שאמרנו את כל זה חשוב לזכור, שכדי לקבל תשובה נכונה צריך להתחיל משאלה נכונה, מה שרחוק להיות מובן מאליו כי "בשביל לשאול את השאלה הנכונה צריך לדעת לפחות חצי מהתשובה" (סטניסלב לם, סופר מדע בדיוני), ואחר כך לתכנן את הניסוי המתאים.

אני רק שאלה!

מתלבטים אם חזרה כלשהי בניסוי של התלמידים שלכם היא חזרה טכנית או ביולוגית? אתם לא לבד. אתם מוזמנים להעלות את השאלה בקבוצת הפייסבוק של מורי הביולוגיה. אם לא הגבתי שם, כיתבו לי ל tzachi.bar@gmail.com .

ביבליוגרפיה

Daniel, J. (2012). *Sampling Essentials: Practical Guidelines for Making Sampling Choices.* SAGE Publications, Inc. Retrieved from https://www.doi.org/10.4135/9781452272047

משרד החינוך, המזכירות הפדגוגית, אגף מדעים, הפיקוח על הוראת המדע והטכנולוגיה. (אין תאריך). מונחון לתהליך החקר המדעי - לתלמיד. ב- נ' ק' יהבית לוריא, *תלקיט חקר מדעי לתלמיד – לחטיבת ביניים.*

.Bloch, A. (2003). *Merphy's Law, 25th anniversary edition.* Berkley Publishing Group

Fonseca, J. F. (2014). How sample size influences research outcomes. *Orthodontics, 19*(4), /27-29. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4296634

Ioannidis, J. P. (2005). Why Most Published Research Findings Are False. *PLOS Medicine , 2*(8), 696-701. Retrieved from https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0020124&xid=17259,15700019,15700186,15700190,15700248

Lazic, S. E. (2011). The problem of pseudoreplication in neuroscientific studies: is it affecting your analysis? *BMC Neuroscience, 11*. Retrieved from https://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-11-5

Longhurst, l. (2002). Murphy's law revisited: longevity as a factor in recruitment to fish .populations. *Fisheries Research, 56*(2), 125-131

Naomi Altman, M. K. (2014). Sources of variation. *Nature Methods, 12*(1), 5-6. doi:https://doi.org/10.1038/nmeth.3224

.Paul Blainey, M. K. (2014). Replication. *Nature Methods, 11*, 879–880

Sophie Fyfe, C. W. (2008). Apophenia, theory of mind and schizotypy: Perceiving .meaning and intentionality in randomness. *Cortex, 44*(10), 1316-1325

Stanley E. Lazic, C. J.-W. (2018). What exactly is ‘N’ in cell culture and animal experiments? *PLOS Biology*. doi:https://doi.org/10.1371/journal.pbio.2005282

ברנשטיין, ע'. (1991). *ביולוגיה חוקרת : איסוף, הצגה וניתוח של נתונים ביולוגיים לבתי-הספר התיכוניים*. תל אביב: עם עובד.