

להיות צודק זה נחמד האם אפשר להיות בטוח? מושג העוצמה במבחן סטטיסטי

ד"ר מיכל כץ-לוירר

ד"ר כץ-לוירר, ראש החוג לפיזיותרפיה, אונ' תל אביב

להיות צודק זה נחמד, האם אפשר להיות בטוח?

אפשר לנסות להעריך **עד כמה** אנו בטוחים שאנו צודקים (תזכורת: הכוונה ב"צודקים" היא; אם נשנה את השיטה המקובלת ונטפל על פי הרעיון הפנטסטי האנשים שסובלים מ-Y **יתפקדו טוב יותר**).

לשם כך, נשתמש בשני הערכים שהגדרנו:

- הגדרנו ערך כמותי ל"טוב יותר".

- הגדרנו ערך סף להחלטה.

אנו משתמשים בשני ערכים אלו כדי להגדיר את מידת הביטחון שלנו שאנו צודקים.

מה הסיכוי שאנו צודקים? הסיכוי להיות מעל ערך הסף אם אכן הטיפול הפנטסטי טוב יותר. הסיכוי הזה במונחים סטטיסטיים מוגדר **כעוצמה**.

ככל שהעוצמה גדולה יותר, אנו בטוחים יותר שאנו צודקים בהנחתנו שהטיפול החדש טוב יותר.

אם כך, נפלא. אפשר לא רק להיות צודק אפשר **לכמת בסיכויים כמה אנו בטוחים** שאנו צודקים.

מתי לאחרונה קראנו מאמר שדיווח על עוצמה? על הסיכוי שחוקר צודק בהנחתו המתבססת על מנגנון?

אנו קוראים הרבה על כך ששיטה זו או אחרת לא מתאימה למוכר ולידוע... אבל היינו רוצים לראות גם את העוצמה, שכן, ברור לכולנו שקיים הבדל בביטחון שלנו שהחוקר צודק אם העוצמה המדווחת היא לדוגמה של 20% או של 80%.

באופן מסורתי² עוצמה העולה על 80% - היא עוצמה, תרתי משמע.

לסיכום, **להיות צודק זה נחמד, מידת הביטחון של ההחלטה חשובה! יש לדווח על העוצמה.**

בואו נניח שכבר שנים כולנו מטפלים באנשים עם Y בשיטת X. שיטת X זו שיטת הטיפול המקובלת.

בואו נדמיין שיש לנו רעיון פנטסטי (לא סתם רעיון, אלא רעיון על סמך מנגנון....) שאם נשנה את השיטה המקובלת ונטפל על פי הרעיון הפנטסטי האנשים שסובלים מ Y יתפקדו טוב יותר. (הערה: מה פירושו טוב יותר? בכמה טוב יותר?)

אף אחד לא ניסה את הרעיון הפנטסטי לפנינו, לפחות לא ידוע לנו שמישהו ניסה. אז אנו מנסים, כלומר אנו בודקים שהרעיון הפנטסטי שלנו עובד (ומחזיקים אצבעות שאנו צודקים).

לשם כך, דבר ראשון אנו חייבים להגדיר מה פירושו "טוב יותר". להגדיר, הכוונה באופן כמותי, בכמה טוב יותר משיטת X. עד כאן הכול הגיוני וברור.

- נערוך ניסוי (על פי כללים ידועים - יש רשימת כללים כזו¹ ואנו פועלים על פיה).

- ננתח את התוצאות. ו....

רגע, כאן נכנס טוויסט בעלילה:

הטוויסט הוא שראשית אנו בודקים האם התוצאות של הטיפול החדשני, על פי הרעיון הפנטסטי, "דומות" לתוצאות הטיפול על פי השיטה המקובלת.

מה פירושו "דומות"?

אנו מגדירים ערך סף. אם התוצאה של הניסוי תהייה מתחת לערך הסף - (זו אכזבה לנו - הוגי הרעיון הפנטסטי) נסכם שהטיפול החדש דומה בתוצאותיו לטיפול המוכר.

אולם, אם התוצאה שהתקבלה היא מעבר לערך הסף, (שמחה וששון לנו - הוגי הרעיון הפנטסטי), כנראה שהטיפול החדש טוב מהטיפול המוכר.

כלומר, האם אפשר לומר שצדקנו? כך נראה, כך אנו מדווחים בשמחה (עייני ערך p-value) אבל השאלה היא, עד כמה אנו בטוחים שאנו צודקים?

2. Sink, Christopher A., and Nyaradzo H. Mvududu. "Statistical power, sampling, and effect sizes: Three keys to research relevancy." Counseling Outcome Research and Evaluation 1.2 (2010): 1-18.

1. Portney, Leslie Gross, and Mary P. Watkins. Foundations of clinical research: applications to practice. Vol. 2. Upper Saddle River, NJ: Prentice Hall, 2000.