

השוואה בין שתי שיטות (Bootstrap BCa והשלמות מרובות) לאמידת רווח סמך למקדם מתאם שחושב במדגם שעבר ברירה מוקדמת

אילה כהן, אטי דובא, טטיאנה אומנסקי
הטכניון - מכון טכנולוגי לישראל

תמר קנת-כהן
מרכז ארצי לבחינות ולהערכה (מאל"ו)

תקציר

מחקר זה משווה בין שתי שיטות לאמידת רווח הסמך למקדם מתאם בין חזאי לקריטריון באוכלוסייה ("אוכלוסיית המועמדים") שחושב במדגם שעבר ברירה מוקדמת ("מדגם המתקבלים"). תהליך הברירה שיושם במחקר הוא המקרה הדו-משתני שבו החזאי שבתוקף הניבוי שלו מתעניינים הוא זה ששימש בברירה. שתי השיטות לחישוב רווח הסמך – Bootstrap BCa והשלמות מרובות – נבדלות זו מזו הן בחישוב האומד הנקודתי של מקדם המתאם באוכלוסייה והן בחישוב רווח הסמך סביבו. בשיטת ה-Bootstrap BCa האומד הנקודתי מחושב על פי הנוסחה המוכרת לתיקון לקיצוץ תחום ורווח הסמך מתבסס על האחוזונים הקיצוניים של המתאם המתוקן שהתקבלו במדגמי Bootstrap. בשיטת ההשלמות המרובות יוצרים עבור כל ערך חסר מספר ערכים משלימים. כך מתקבלים מספר אוספים של נתונים ללא ערכים חסרים. על בסיסם מחשבים את האומד הנקודתי, את שגיאת התקן של האומד ואת רווח הסמך. רווח הסמך בשיטת ה-Bootstrap BCa הושווה בעבר לשיטות אחרות ונמצא עדיף עליהן. חישוב רווח הסמך בשיטת ההשלמות המרובות ובפרט השוואתו לרווח המתקבל בשיטת ה-Bootstrap BCa הוא בבחינת חידוש של מחקר זה. הקריטריון העיקרי לבדיקת איכות רווח הסמך היה אחוז המקרים שהיו מחוץ לרווח הסמך המיועד (רמת סמך של 0.95).

איכותם של רווחי הסמך שהתקבלו בשתי השיטות נבדקה באמצעות סימולציות שנבדלו זו מזו בגודל אוכלוסיית המועמדים, ביחס הברירה ובמקדם המתאם באוכלוסייה. המדגמים לסימולציות התקבלו על ידי דגימה עם חזרות משלוש התפלגויות אמפיריות שונות, שהתקבלו עבור משתנה הברירה בטכניון – התפלגות הדומה לנורמלית, התפלגות עם אסימטריה שלילית והתפלגות הנוטה לאחידה.

ממצאי המחקר מעידים שהאומדים הנקודתיים, בשתי השיטות, מוטים כלפי מטה. ההטיה היא בדרך כלל הקטנה ביותר עבור ההתפלגות הדומה לנורמלית והגדולה ביותר עבור ההתפלגות הנוטה לאחידה, בפרט כאשר גודל מדגם המתקבלים קטן. כמו-כן נמצא שבשתי השיטות שונות האומדים גדולה, בפרט כאשר גודל מדגם המתקבלים קטן. הממצא העיקרי של המחקר, שנוגע לאיכות רווחי הסמך, מלמד שבכל ההתפלגויות, בשיטת ההשלמות המרובות רווח הסמך המתקבל משקף רמת ביטחון נמוכה מהרמה המיועדת. ממצא זה קיים גם בשיטת ה-Bootstrap BCa כאשר ההתפלגות אינה נורמלית, אולם הפער בין הרצוי למתקבל אינו כה גדול כמו בשיטת ההשלמות המרובות. מסקנת המחקר היא ששיטת ה-Bootstrap BCa לאמידת רווח סמך עדיפה על שיטת ההשלמות המרובות.