

עבודה 2 עקרונות ויישומים בניתוח סטטיסטי

בעבודה זו אתם נדרשים לבצע סטטיסטיקה היסקית על קובץ הנתונים עליו ביצעתם סטטיסטיקה תיאורית. יש להסביר בשתי מילים או בנוסחה קצרה את השיטה בה השתמשתם ובעיקר להציג את התוצאות של הפעלת השיטה. את העבודה יש להגיש כקובץ PDF למודל (אין צורך לצרף את הקוד).

יהיו X ו- Y המשתנים בהם השתמשתם ברגרסיה בעבודה הראשונה (X המשתנה המסביר ו- Y המשתנה המוסבר).

1.

- הניחו כי ל- Y התפלגות נורמלית $N(\mu, \sigma^2)$ ואמדו את התוחלת והשונות.
- נסמן ב- m את הערך המינימלי מבין ערכי X ונגדיר $W = X - m$. הניחו כי ל- W התפלגות $Gamma(\alpha, \lambda)$ ואמדו את שני הפרמטרים של ההתפלגות.
- על פי ההתפלגויות שאמדתם בסעיפים א ו ב חשבו את האחוזונים 10, 50, 75, 90 של X ו- Y, W .
- חשבו את האחוזונים האמפיריים (אלה המתקבלים מהנתונים כפי שלמדנו בשבוע הראשון) המתאימים לאלו שהתקבלו בסעיף ג והשוו אותם לאלה המבוססים על המודלים (השוואה בטבלה פשוטה, לא השוואה פורמלית). האם האחוזונים על פי המודל דומים לאלה האמפיריים?
- ה. (סעיף רשות). חזרו על סעיף ג וד לכל האחוזונים 1-99. ציירו בגרף את האחוזונים על פי המודל לעומת האחוזונים האמפיריים. האם המודלים נראים מתאימים לנתונים?

2. חשבו רווחי סמך ברמת סמך של 97% לתוחלת של Y וברמת סמך 92% לשונות של Y , בהנחה כי ל- Y התפלגות נורמלית.

3. בשאלה זו נרצה לבדוק האם יש קשר בין Y לבין X . יהיה $med(X)$ החציון של ערכי X . חלקו את הזוגות (Y, X) על פי אלה בהן ערך X קטן מהחציון $\{(X_i, Y_i) | X_i < med(X)\}$ ואלה עבורם הוא גדול מהחציון או שווה לו $\{(X_i, Y_i) | X_i \geq med(X)\}$. נרצה לבדוק את ההשערה כי התוחלת של Y שווה בשתי הקבוצות.

- נסחו את ההשערה במילים – דהיינו, נסחו את שאלת המחקר (בניסוח זה אסור להשתמש באותיות Y, X אלא חייבים להשתמש בתיאור המשתנים עצמם). הגדירו את ההשערה האלטרנטיבית (גדול/קטן/שווה) על פי מה שנראה לכם נכון מבחינת אופי הנתונים.
- נסחו את ההשערות בצורה סטטיסטית/מתמטית בעזרת הסימונים Y, X ועל פי ההשערה האלטרנטיבית שהגדרתם בסעיף א.
- בצעו מבחן סטטיסטי: רשמו את סטטיסטי המבחן, ציינו כיצד הוא מתפלג תחת השערת האפס וחשבו את ערך ה- P .
- השתמשו ברמת מובהקות של 3% על מנת להכריע האם לדחות או לקבל את השערת האפס. סכמו את הממצאים במילים לפי השערת המחקר שהגדרתם בסעיף א.