

מעבדה לסטטיסטיקה 2024, מטלה 4
הצגה ב 11.6

הקשר בין GC לבין הכיסוי: רגרסיה לא פרמטרית וניתוח שאריות.

א. אמדו רגרסיה בעזרת בסיס פולינומית למקטעים (עד מעלה 3 לכל היותר). בחרו את ה knots ואת הדרגה כך שקו הרגרסיה יתאים בצורה הטובה ביותר. לפני ההתאמה, הורידו תאים ש"מפריעים" לגרף לתפוס היטב את גרף המציג.

הציגו גרף שמראה הן את הנתונים והן את קו המגמה.

ב. הציגו את גרף השארית כנגד GC. האם יש עדיין הטיה עבור ערכי GC מסוימים?

ג. מצאו מדד מספרי לאיכות החיזוי. פי כמה השיפור בהתאמה לעומת מודל לינארי שמתאים לנתונים? (שימו לב ששני הקווים לא יושפעו יותר מדי מתאים יוצאי דופן).

ד. האם יש הבדלים באיכות החיזוי (גודל השארית) בין ערכי GC שונים? ענו בעזרת חיווי ויזואלי או מספרי.

.