

מודלים סטטיסטיים ויישומיהם 52518 תשע"ח – תרגיל 5

להגשה עד 4.12.17 בשעה 23:55

1. להלן נתוני שלוש קבוצות:

קבוצה 1	3.2	2.6	3	2.9
קבוצה 2	2.4	4	2.7	3.8
קבוצה 3	2	2.2	3.7	3.4

- א. חשבו ידנית את סטטיסטי Kruskal-Wallis.
- ב. חשבו את מובהקות הסטטיסטי בהתבסס על קירוב אסימפטוטי להתפלגות חי-בריבוע.
- ג. כתבו פונקציה ב-R אשר מקבלת אוסף נתונים ומחשבת עבורו את מובהקות מבחן Kruskal-Wallis בהתבסס על סימולציות מונטה קרלו. הריצו את הפונקציה על הנתונים לעיל.
- ד. השתמשו בפונקציה `qn.test` מתוך החבילה `KSample` ב-R כדי לבצע את מבחן Kruskal-Wallis, עם שלושת האופציות הקיימות עבור הפרמטר `method` (הריצו 3 פעמים).
- ה. השוו בין תוצאות סעיפים ב', ג' ו-ד'.

2. בהתייחס לנתוני הקובץ `ex4a.txt`:

- א. כתבו פונקציה ב-R לביצוע מבחן Levene והריצו אותה על נתוני הקובץ.
- ב. כתבו פונקציה ב-R לביצוע מבחן Welch והריצו אותה על נתוני הקובץ.
3. עבור נתוני הקובץ `ex4b.txt`, השתמשו בכלים שנלמדו בכיתה כדי לבדוק את הנחות המודל של ANOVA חד-כוונית. נסו לשפר את התאמת הנתונים להנחות באמצעות טרנספורמציה.

4. נניח כי אנחנו מבצעים ANOVA חד-כוונית ורוצים לבדוק האם השאריות מפולגות נורמלית באמצעות סטטיסטי קולמוגורוב-סמירנוב, $D = \max_{t \in [-\infty, \infty]} \left| \left\{ \hat{F}_n(t) - \Phi\left(\frac{t - \bar{X}}{s}\right) \right\} \right|$. כתבו פונקציה ב-R אשר מקבלת כקלט את הערך d ומחזירה קירוב עבור ההסתברות $P(D \geq d)$ במקרה בו השגיאות ϵ_{ij} אכן מתפלגות נורמלית עם שונות זהה. הריצו את הפונקציה במקרים הבאים:

- א. שתי קבוצות בעלות 15 תצפיות כ"א, $d = 0.5$.
- ב. שלוש קבוצות בעלות 10 תצפיות כ"א, $d = 0.5$.