סטטיסטיקה למדעי המחשב – תרגיל בית שבוע 2

שאלה 1 (50 נקודות)

משרד החינוך רצה לבדוק האם יש צורך ברפורמה במתמטיקה ובחינוך הגופני. לשם כך נאספו ציוני המבחנים של תלמידי כיתות ז' עד ט' ב- שלושה בתי ספר שונים. הורידו את הקובץ grades.csv מהמודל. הטבלה מכילה, לכל תלמיד, את בית הספר בו הוא לומד (A,B,C), את ציון סוף השנה במתמטיקה (math) ואת ציון סוף השנה בחינוך גופני (gym).

א. כתבו פונקציה בשם ()kernel_density המקבלת וקטור x ורוחב חלון h. הפונקציה תצייר היסטוגרמה חלקה של x בעזרת הגרעין *Cosine*.

הדרכה:

1. בעזרת הפונקציה ()np.linspace צרו וקטור t עם ערכים סדורים בין min(x)-h לבין max(x)+h בקפיצות קטנות זהות.
2. חשבו לכל ערך ב- t את ערך הפונקציה J(t) לפי הנוסחה שנלמדה בשיעור ובתרגול.
3. ציירו גרף של J כפונקציה של t.

ב. באמצעות הפונקציה שיצרתם, ציירו את הצפיפות של המשתנה math ומצאו את רוחב החלון *h שמציג את ההתפלגות בצורה הטובה ביותר (לדעתכם). חפשו בין כמה שיותר ערכי h שונים**.**

ג. כמה שיאים לדעתכם יש להתפלגות?

ד. כעת ציירו **על גרף אחד** שלוש היסטוגרמות חלקות של ציוני המתמטיקה **-** היסטוגרמה נפרדת לכל בית ספר. השתמשו בצבע שונה לכל בית ספר. מה ניתן להגיד על רמת הידע במתמטיקה בבתי הספר השונים? (אם הגרפים חורגים מהצירים, ניתן להיעזר ב- ,plt.ylim plt.xlim)

ה. ציירו את פונקציית הצפיפות של המשתנה gym עם רוחב החלון *h שמצאתם בסעיף ב', עבור כל הנתונים ועבור כל בית ספר בנפרד. האם ניתן להגיע לאותן מסקנות על היחס בין בתי הספר השונים כמו בסעיף ד'?

שאלה 2 (30 נקודות)

א. נתונה הפונקציה

$$f(x) = \begin{cases} \frac{1}{4}x^3 & 0 \le x \le b \\ 0 & else \end{cases}$$

מצאו את b עבורו f היא פונקציית צפיפות.

יהי מ״מ רציף X המתפלג לפי פונקציית הצפיפות מסעיף א׳.

ב. מצאו את $P(X > 1)$.

ג. מצאו את הערך x_0 כך ש- $P(X \le x_0) = 0.25$.

שאלה 3 (20 נקודות)

א. נזכיר את פונקציית הגרעין הגאוסיאנית:

$$w(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$$

הוכיחו כי הפונקציה

$$J(a) = \frac{1}{n \cdot h} \underbrace{\sum_i w\left(\frac{X_i - a}{h}\right)}_{\text{סופר כמה נתונים בחלון } w}$$

w היא פונקציית החלון
אפשר לקחת חלון
שנרצה - Wc ,Wr

n - מספר הנתונים
מחלקים ב n
לקבלת שכיחות
יחסית

h - רוחב החלון
מחלקים ב h
לקבלת יחידת
המדידה

עם פונקציית גרעין גאוסיאנית (כלומר עם $w(u)$), היא פונקציית צפיפות תקנית.

תוכלו להשתמש ללא הוכחה בכך שהפונקציה:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

היא פונקציית צפיפות תקנית לכל μ, $\sigma > 0$.

(זו פונקציית הצפיפות של משתנה מקרי נורמלי עם תוחלת μ ושונות σ^2 ונדבר עליה עוד בהמשך)