אותות ומשתנים אקראיים 67652

עבודת MATLAB #1

פרק 1

בפרק זה נשערך את פונקצית צפיפות הסתברות עבור משתנה אקראי כלשהוא X , באמצעות מדידות i.i.d. המתפלגות לפי אותו משתנה אקראי.

תיאור שיטת השערוך

נציג את שיטת היערוך ואת חוק המספרים הגדולים.

<u>תזכורת : החוק החלש של המספרים הגדולים</u>

תהיה $\{X_n\}_{n=1}^{\infty}$ סדרת משתנים אקראיים המתפלגים i.i.d. עם תוחלת סופית μ. נגדיר

$$.S_n = \sum_{k=1}^{n} X_n$$

אזי לכל $\epsilon > 0$ מתקיים

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) \xrightarrow[n\to\infty]{} 1$$

- על מנת לקבל קירוב של פונקצית צפיפות, נחלק את המרחב הערכים ל־n מקטעים (bins)

$$t_1 < t_2 < \cdots t_{n-1} < t_{n+1}$$

ועבור כל מקטע נקרב את ההסתברות

$$\mathbb{P}\left(t_i < X \le t_{i+1}\right) \quad , 1 \le i \le n$$

כאשר המרווחים קטנים, עבור $t_i < X \le t_{i+1}$ נוכל להשתמש בקירוב הבא:

$$f_X(x) \cong \frac{\mathbb{P}\left(t_i < X \le t_{i+1}\right)}{t_{i+1} - t_i} \tag{1}$$

- נגדיר את **פונקצית האינדיקטור** של מאורע A במרחב הסתברות נתון:

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & ,\omega \in A \\ 0 & ,\omega \notin A \end{cases}.$$

- שימו לב כי $\mathbb{1}_A$ הינו משתנה אקראי. הראו כי

$$\mathbb{E}\left[\mathbb{1}_A\right] = \mathbb{P}\left(A\right)$$

- נסמן את A_i להיות המאורע ש־$t_i < X \le t_{i+1}$. על פי האמור לעיל ובאמצעות חוק המספרים הגדולים, נשערך את ההסתברות של A_i באמצעות K מדידות i.i.d. של X על ידי:

$$\mathbb{P}\left(A_i\right) = \mathbb{P}\left(t_i < X \le t_{i+1}\right) = \mathbb{P}\left(\mathbb{1}_{A_i}\right) = \mathbb{E}\left[\mathbb{1}_{A_i}\right] \cong \frac{1}{K}\sum_{j=1}^{K} \mathbb{1}_{t_i < X \le t_{i+1}}(X_j)$$

דוגמת שיערוך פונקצית צפיפות הסתברות

העזרו בתוכנית MATLAB מהמודל (normal_density_estimation.m) המציגה שערוך של פונקצית צפיפות הפילוג של משתנה אקראי גאוסי בעל תוחלת 0 ושונות 1 (בעזרת משתנים אקראיים i.i.d. שאותם מגרילים). תוצאת השערוך מוצגת בדיאגרמת מלבנים (היסטוגרמה) ובאותו השרטוט מוצגת גם פונקצית צפיפות הפילוג המדויקת, לשם השוואה. (שיטת השערוך הזו מכונה גם שערוך באמצעות "סימולציית מונטה־קרלו"). הריצו את התוכנית, התרשמו מפלטיה, וענו על השאלות הבאות:

א) תארו את השיטה שבאמצעותה משערכים את פונקצית צפיפות ההסתברות והסבירו את נכונותה.

ב) מהם הפרמטרים המשפיעים על טיב השערוך ? כיצד הם משפיעים?

מעתה והלאה כאשר תתבקשו לבצע שערוך מונטה קרלו של פונקצית צפיפות ההסתברות של משתנה אקראי, בחרו את הפרמטרים המשפיעים על טיב השערוך שיתקבל שערוך סביר, ותנו פלטים רק עבור פרמטרים אלו.

שאלות

שאלה 1

הציגו שערוך מונטה־קרלו של פונקצית צפיפות פילוג של משתנה אקראי עם פילוג ברנולי $X \sim \mathrm{Ber}\left(1/4\right)$.

הערות

- זכרו כי שלב ראשון בתהליך השערוך הוא ייצור של סדרת משתנים אקראיים i.i.d. המתפלגים לפי הפילוג של X.
- השתמשו במשתנה אקראי אחיד רציף על מנת לייצר את X. היעזרו בפונקציה `rand`.
- הציגו באותו השרטוט את פונקצית צפיפות הפילוג המדויקת ואת השערוך שלכם בעזרת היסטוגרמה. היעזרו בתוכנית מהחלק הקודם.

שאלה 2

נתון משתנה אקראי $Y \sim \mathcal{N}\left(0,1\right)$ ומגדירים את המשתנים האקראיים הבאים :

$$
\begin{aligned}
Z_1 &= \max\left\{Y, Y^2\right\} \\
Z_2 &= \ln\left|Y\right|
\end{aligned}
$$

א) חשבו את פונקציות צפיפות הפילוג של Z_1, Z_2.

ב) בצעו שערוך מונטה־קרלו של פונקציות צפיפות הפילוג של Z_1, Z_2 בעזרת MATLAB. הציגו באותו השרטוט את פונקצית צפיפות הפילוג המדויקת ואת השערוך שלכם בעזרת היסטוגרמה.

הערות

- זכרו כי שלב ראשון בפתרון הוא ליצור סדרות משתנים אקראיים i.i.d. המתפלגות לפי הפילוגים של Z_1, Z_2.

פרק 2

בפרק זה נייצר דגימות מפילוג גאוסי רב־מימדי עם מימד p.

<u>שאלה 1</u>

בשאלה זו נייצר אוסף וקטורים גאוסיים במשותף ולאחר מכן נבצע טרנספורמציה לשינוי התוחלת ומטריצת הקווריאנס. מימשו פונקציה ב־MATLAB:

```
function [x] = create_gaussian_points(mu,covariance,N)
```

כאשר פרמטרי הכניסה מוגדרים על ידי:

- MU: וקטור באורך p .
- COVARIANCE: מטריצה בגודל $p \times p$.
- N: מספר הדגימות (כלומר מספר הוקטורים) שיש ליצור.

המוצא x הוא מטריצה בגודל $p \times N$ כך שכל עמודה הינו וקטור אקראי גאוסי במשותף עם תוחלת MU ומטריצת קווריאנס COVARIANCE. כמו כן, העמודות השונות הן בלתי תלויות סטטיסטית זו בזו.

הערות

- יש ליצור תחילה אוסף של וקטורים גאוסיים עם תוחלת אפס וקווריאנס מטריצת יחידה. לאחר מכן יש להשתמש בטרנספורמציה ליניראית מתאימה על מנת לקבל את התוחלת וקווריאנס הרצויים.
- יש לבצע את הטרנספורמציה בשלשה שלבים:

א) טרנספורמציה אלכסונית (השקולה למתיחה בכל ציר).

ב) טרנספורמציה אורתוגונלית (השקולה לסיבוב של מערכת הצירים).

ג) הוספת קבוע וקטורי.

- פונקציות MATLAB שימושיות בשאלה זו הן EIG, RANDN.

שאלה 2

נבחר

$$p=2,\;\; \texttt{N}=100000\,,\;\; \texttt{mu}=\begin{pmatrix}0\\0\end{pmatrix},\quad \texttt{covariance}=\begin{pmatrix}4 & 1.5\\ 1.5 & 1\end{pmatrix}$$

בעזרת הפונקציה שממשתם בשאלה הקודמת, ייצרו אוסף של N וקטורים גאוסיים עם התוחלת והקווריאנס הנ"ל. שימו לב שעל ידי כך אנחנו יוצרים דגימות של

$$\begin{pmatrix}x_1\\x_2\end{pmatrix}\sim\mathcal{N}\left(\texttt{mu},\texttt{covariance}\right)$$

מהי הטרנספורציה המפורשת במקרה זה?
הציגו את הדגימות במערכת צירים (למשל על ידי שימוש בפונקצית PLOT) בשלשה שלבים:

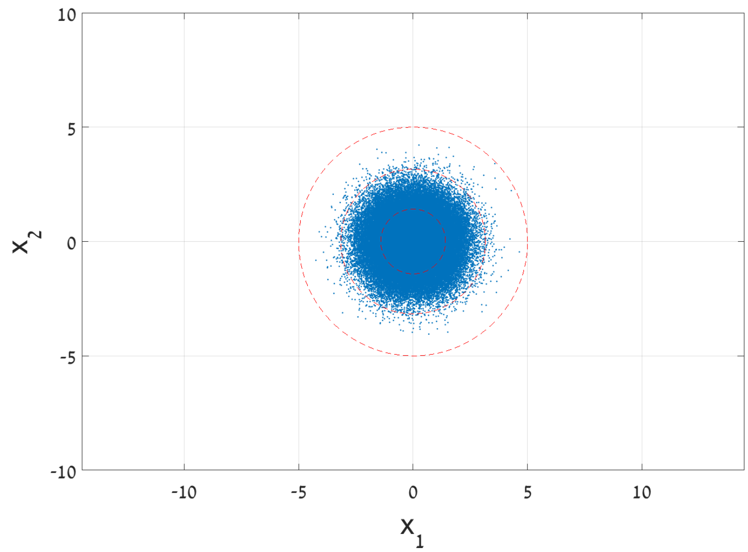
א) דגימות לפני טרנספורמציה (דגימות i.i.d.).

ב) דגימות לאחר טרנספורמציה אלכסונית.

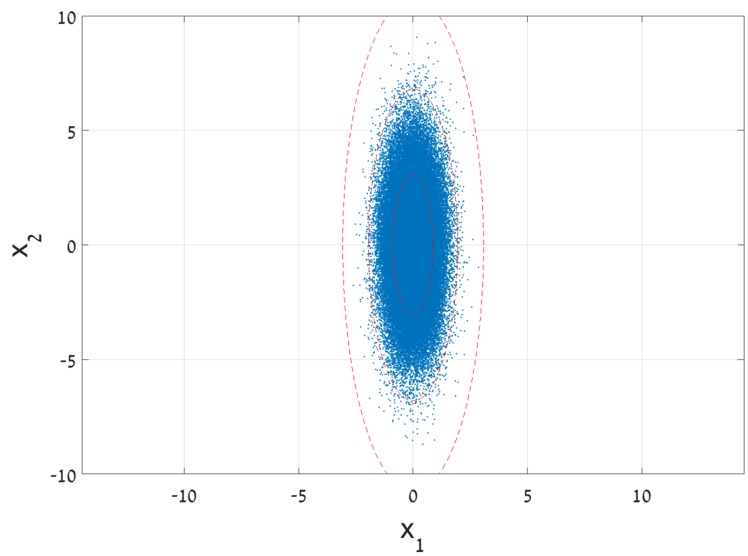
ג) דגימות לאחר טרנספורמציה אורתוגונלית.

הערות

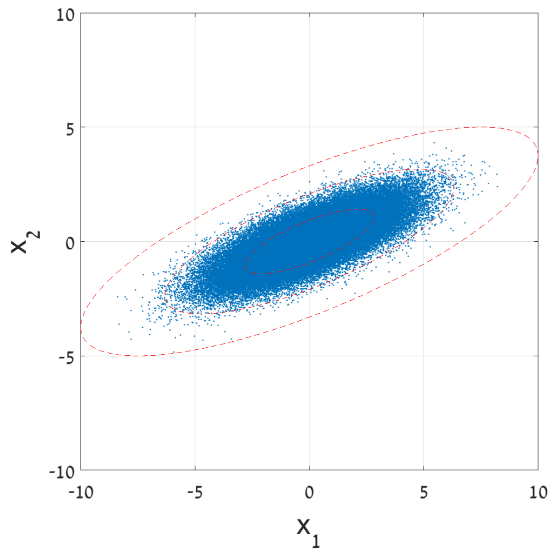
- יש לקבע את הטווח של הצירים על מנת שיהיו זהים בכל המקרים (על ידי הפונקציות XLIM, YLIM).
- ודאו כי אתם מקבלים את הפלטים הבאים (ללא האליפסות):



איור 1: דגימות i.i.d.



איור 2: דגימות לאחר מתיחה



איור 3: דגימות לאחר סיבוב

שאלה 3

בשאלה זו נכתוב פונקציה המקבלת את אוסף הדגימות ומשרטטת קירוב של פונקצית הצפיפות הסתברות המותנית. נעריך את הצפיפות המותנית $f_{X_1|X_2}(x_1|x_2)$ בעזרת הדגימות שייצרתם. נשתמש בקירוב

$$f_{X_1|X_2=\beta}(\alpha|\beta) \cong f_{X_1|-\delta\leq X_2-\beta\leq\delta}(\alpha|\beta)$$

עבור בחירה של δ. נבצע את החישוב באותו אופן כמו במשוואה (1) כאשר נשתמש רק בדגימות המקיימות את המאורע

$$|X_2-\beta|\leq\delta \tag{2}$$

יש לממש פונקציה ב־MATLAB:

```
function display_conditional_pdf(x,x2,delta,mu,covariance,...
                                  leftLimit,rightLimit,binWidth);
```

כאשר פרמטרי הכניסה מוגדרים על ידי:

- x: מטריצה בגודל $N \times p$ של אוסף המדידות (המוחזר מ־ CREATE_GAUSSIAN_POINTS).

- x2: הבחירה של β עבור המאורע במשוואה (2).
- DELTA: הבחירה של δ עבור המאורע במשוואה (2).
- MU: וקטור באורך p , זהה לשאלה הקודמת.
- COVARIANCE: מטריצה בגודל $p \times p$, זהה לשאלה הקודמת.
- LEFTLIMIT: גבול שמאלי בציר האופקי.
- RIGHTLIMIT: גבול ימני בציר האופקי.
- BINWIDTH: רוחב מקטע בציר האופקי עבור חישוב היסטוגרמה.

הפונקציה DISPLAY_CONDITIONAL_PDF מבצעת את הפעולות הבאות:

א) בוחרת מתוך אוסף הדגימות את אלו שמקיימים את המאורע במשוואה (2).

ב) מקרבת את הצפיפות המותנית מתוך דגימות אלו על ידי היסטוגרמה.

ג) מחשבת את התוחלת המותנית והשונות המותנית באמצעות נוסחה.

ד) מציירת על אותה מערכת צירים את ההיסטוגרמה ואת הצפיפות המותנית האמיתית (שמחושבת באמצעות הנוסחה לצפיפות גאוסית).

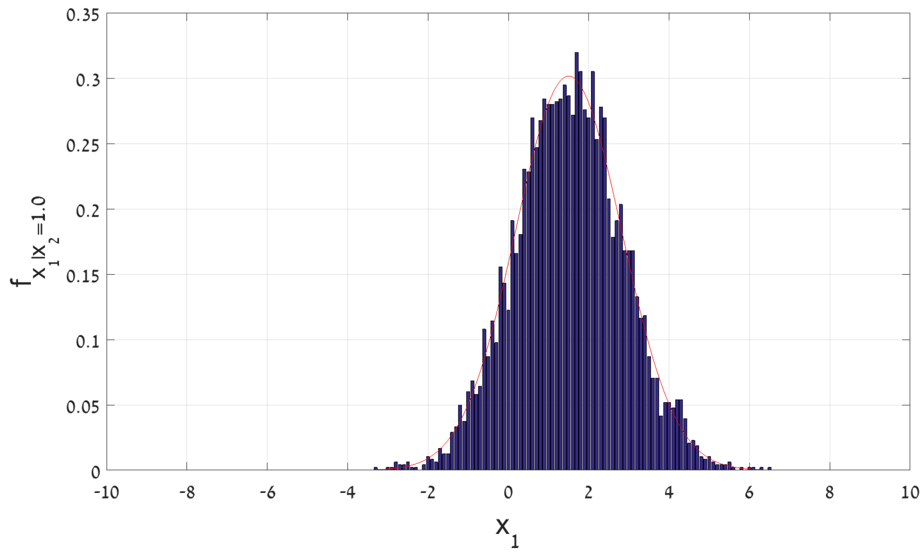
הערות

- פונקציות MATLAB שימושיות בשאלה זו הן PLOT, BAR, HIST.
- על מנת שההיסטוגרמה תתלכד עם השרטוט התיאורטי, יש לבחור באופן מושכל את DELTA ואת BINWIDTH.
- מה המשמעות של בחירת DELTA גדול מידי או קטן מידי?
- נסו להגדיל את N ולראות שמתקבלת קירוב טוב יותר לצפיפות המותנית האמיתית.
- נסו ערכים שונים של x2 וודאו כי ההיסטוגרמה תואמת לצפיפות המותנית האמיתית (כאשר x2 לא גדול מידי).

קבעו את הערכים:

```
x2 = 1.0, leftLimit = -8, rightLimit = -8
```

והפעילו את הפנקציה. להלן פלט להשוואה:



0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-10
-8
-6
-4
-2
0
2
4
6
8
10
$f_{X_1|X_2=1.0}$
x_1