IML -EX1
Amitay Sicherman -203449004

Warm-up - Algebra Recap

Q1:

We use that

$$P_v u = \frac{\langle v, u\rangle}{|u|^2}\vec{u}$$

Calculate:

$$\mathrm{P}_v u = \frac{\left\langle \begin{pmatrix}1\\2\\3\\4\end{pmatrix}, \begin{pmatrix}0\\-1\\1\\2\end{pmatrix} \right\rangle}{\left\| \begin{pmatrix}0\\-1\\1\\2\end{pmatrix} \right\|^2} \cdot \begin{pmatrix}0\\-1\\1\\2\end{pmatrix} = \frac{9}{6} \cdot \begin{pmatrix}0\\-1\\1\\2\end{pmatrix} = \begin{pmatrix}0\\-\frac{3}{2}\\\frac{3}{2}\\3\end{pmatrix}$$

Q2:

Use the same equation,

Calculate:

$$\mathrm{P}_v u = \frac{\left\langle \begin{pmatrix}1\\2\\3\\4\end{pmatrix}, \begin{pmatrix}1\\0\\1\\-1\end{pmatrix} \right\rangle}{\left\| \begin{pmatrix}1\\0\\1\\-1\end{pmatrix} \right\|^2} \cdot \begin{pmatrix}1\\0\\1\\-1\end{pmatrix} = \frac{0}{3} \cdot \begin{pmatrix}0\\-1\\1\\2\end{pmatrix} = 0$$

Q3:

we use the definition of $\langle v, u\rangle$:

$$\langle v, u\rangle = |u| \cdot |v| \cdot \cos(\theta)$$

we will prove the two sides:

First side, $\theta = \pm 90 \to \cos(\theta) = 0 \to |u| \cdot |v| \cdot \cos(\theta) = 0 \to \langle v, u\rangle = 0$

second side: $\langle v, u\rangle = 0 \to \ |u| \cdot |v| \cdot \cos(\theta) = 0 \underbrace{\to}_{|\mathrm{u}|,|v|>0} \cos(\theta) = 0 \to \theta = \pm 90$

Q4:

we use that:

1. $for\ vector\ v: \left||v|\right|_2 = \sum v_i^2 = v \cdot v^T$
2. $for\ orthogonal\ marix\ A : AA^T = AA^{-1} = I$

And now we prove:

$$\left||Ax|\right|_2 \underbrace{=}_{1} Ax(Ax)^T = AxA^Tx^T = AA^Txx^T \underbrace{=}_{2} xx^T \underbrace{=}_{1} \left||x|\right|_2$$

SVD

Q5:

we use :

1. $for\ diagonal\ matix\ D^{n*n} = d_{ij}$

$$: calculat\ the\ invers\ matrix\ of\ D\ can\ make\ in\ O(n) \quad by - D^{-1} = \frac{1}{d_{ij}}$$

2. $for\ orthogonal\ marix\ A :: calculat\ the\ invers\ matrix\ of\ D\ can\ make\ in\ O(n^2)\ by\ A^{-1} = A^T$

proof:

$$A = UDV^T$$
$$\rightarrow A^{-1} = (V^T)^{-1}D^{-1}U^{-1}$$
$$from\ 2 \rightarrow \boxed{(VD^{-1}U^T)}$$

from (1) we can calculate D^{-1} in $O(n)$, from (2) we can calculate U,T transpose in $O(n^2)$ so the operation is more useful then the regular way that take O $\left(n^{\approx 2.8}\right)$

Q6:

Calculate:

$$c \cdot c^T = \begin{pmatrix} 50 & 30 \\ 30 & 50 \end{pmatrix}$$

We see that:

$$if\ c = UDV^T\ if\ SVD\ so\ cc^T = UDD^TU^T\ is\ EVD$$

so let say that $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $and\ DD^T = diag(d1, d2)$

calculate:

$DD^T = diag(d1^2, d2^2)$

$$\begin{pmatrix} 50 & 30 \\ 30 & 50 \end{pmatrix} = \begin{pmatrix} a^2d1^2 + cbd2^2 & abd1^2 + dbd2^2 \\ cad1^2 + dcd2^2 & abd1^2 + dbd2^2 \end{pmatrix}$$

And the solve is :

$$d1^2 = 80\ , d2^2 = 20\ , a = b = c = -d = \frac{1}{\sqrt{2}}$$

$$\rightarrow D = \begin{pmatrix} \sqrt{80} & 0 \\ 0 & \sqrt{20} \end{pmatrix}, U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Now, $c = UDV^T$ multiplate two sides with V in the right we get $cV = UDV^TV = UD$, multiplate two sides in $c^{-1} = \begin{pmatrix} \frac{7}{40} & -\frac{1}{8} \\ \frac{1}{40} & \frac{1}{8} \end{pmatrix}$ in the left we get $c^{-1}cV = V = c^{-1}UD$

set the values of $c^{-1}UD$ and calculate the multiplate we get :

$$V = \begin{pmatrix} \sqrt{\frac{1}{10}} & \sqrt{\frac{9}{10}} \\ \sqrt{\frac{9}{10}} & -\sqrt{\frac{1}{10}} \end{pmatrix}$$

Q7:

$\boxed{\text{שאלה 1:}}$

נציג את [illegible] כך:

$$b_0 = \sum_{i=0}^{n} a_i v_i$$

נקבל:

$$A^k b_0 = \sum_{i=1}^{n} a_i A^k v_i$$

$$= \sum_{i=1}^{n} a_i \lambda_i^k v_i$$

$$= a_1 \lambda_1^k \sum_{i=1}^{n} \left(\frac{a_i}{a_1}\right)\left(\frac{\lambda_i}{\lambda_1}\right)^k v_i$$

ולכן כל [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]

[illegible] [illegible] [illegible] v_1 

Multivariate Calculus

Q8

$\boxed{\text{שאלה 8:}}$

ראשית נכתוב את הפונקציה כסכום (בעזרת הפירוק הסינגולרי)

$$f(\bar\sigma) = \sum_{i=1}^{n} \sigma_i v_i v_i^T \bar{x}$$

v_i הוא העמודה ה-i ב-V.

ולכן

$$f(\bar\sigma)_j = \left(\sum_{i=1}^{n} \sigma_i v_i v_i^T \bar{x}\right)_j$$

ולכן נחשב את הנגזרת:

$$\frac{\partial f(\bar\sigma)_j}{\partial \sigma_k} = \frac{\partial}{\partial \sigma_k}\left(\sum_{j=1}^{n} \sigma_i v_i v_i^T \bar{x}\right)_j$$

ברור כי עבור $j = k$ נקבל כי $\frac{\partial \sigma_i}{\partial \sigma_k} = 1$ ואחרת נקבל 0, כלומר של הסכום נקבל רק איבר אחד, כלומר:

$$\boxed{\frac{\partial f(\bar\sigma)_j}{\partial \sigma_k} = \left(v_k v_k^T \bar{x}\right)_j}$$

אם נכתוב את הנגזרת כולה עם ההנחה הנוכחית
ה-k העמודה תהיה $v_k v_k^T \bar{x}$

ובסה"כ נקבל ונכתוב

$$\boxed{J_\sigma(f) = V \operatorname{diag}(V^T \bar{x})}$$

Q9

$\boxed{\text{שאלה 9:}}$

נשתמש בכלל השרשרת:

נגדיר

$$x = f(\sigma) - y$$

$$y = \frac{1}{2}\|X\|^2$$

ולכן

$$h = \frac{1}{2}\|x(\sigma)\|^2$$

$$\nabla h = \frac{\partial}{\partial x}\left(\frac{1}{2}\|x(\sigma)\|^2\right) \cdot \frac{\partial x(\sigma)}{\partial \sigma}$$

$$= (\sigma) \cdot (f(\sigma) - y)^T \cdot \nabla f(\sigma)$$

Q10

שאלה 10

$$g(S)_j = \frac{e^{s_j}}{\sum_i e^{s_i}}$$

$$\frac{\partial g_i}{\partial s_j} = \frac{\partial}{\partial s_j}\left(\frac{e^{s_i}}{\sum_k e^{s_k}}\right)$$

ונגזור לפי כלל המנה:

$$= \left(\sum_j e^{s_j}\right)^{-2} \cdot \left(\frac{\partial}{\partial s_j} e^{s_i} - e^{s_i} \cdot \frac{\partial}{\partial s_j} \sum_k e^{s_k}\right)$$

if $i = j$: $\Rightarrow \left(\sum_k e^{s_k}\right)^{-2} \cdot \left(e^{s_i} - e^{s_j} \cdot e^{s_i}\right)$

$$= g(s)_i \cdot (1 - g(s_i))$$

if $i \neq j$: $\Rightarrow \left(\sum_k e^{s_k}\right)^{-2} \left(0 + e^{s_i} e^{s_j}\right)$

$$= g(s)_i \cdot (-) g(s_j)$$

ובסה"כ נקבל:

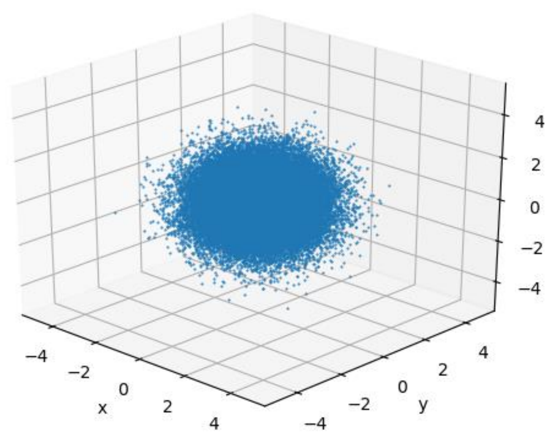
$$\frac{\partial g_i}{\partial s_j} = g(s)_i \left(\delta_{ij} - g(s)_j\right)$$

[illegible]

Multivariate Gaussian- practical question

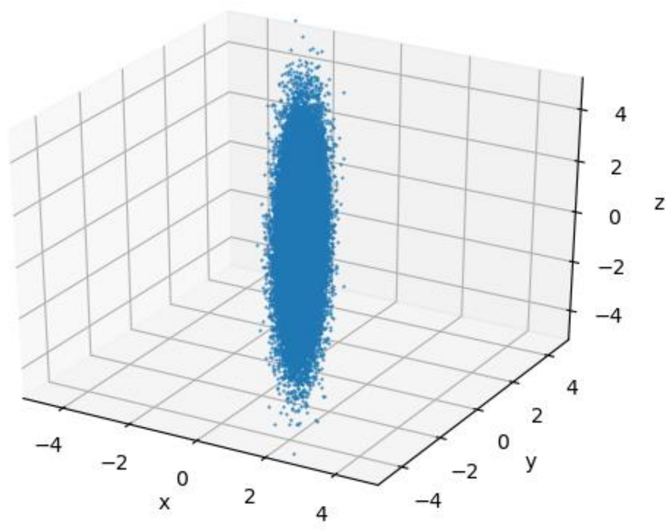
Q11

3D gaussian



Q12

Scale 3D gaussian



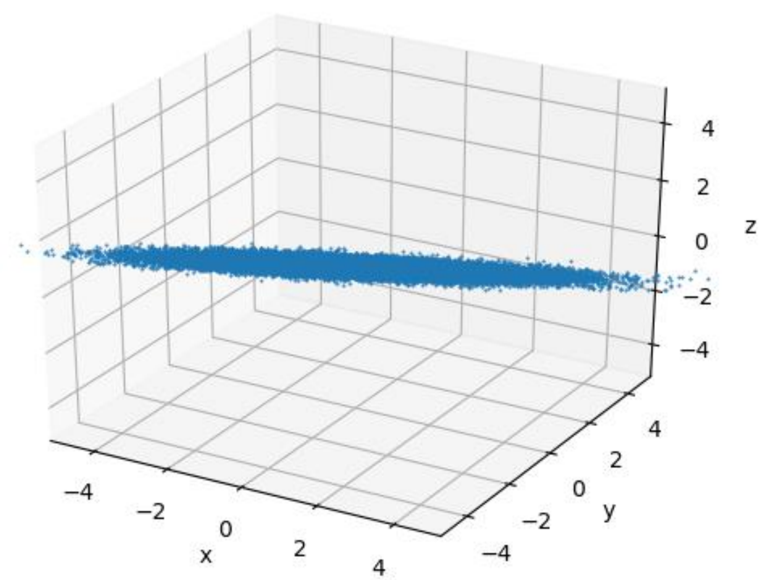
מטריצת ה COV

[[0.01 0.   0.]
 [0.   0.25 0.]
 [0.   0.   4.]]

ניתן לראות כי עדיין אין קורלציה בין הצירים (רק האיברים על האלכסון קיימים) אבל יש מתיחה – כמו שניתן לראות בסרטוט. יש מתיחה אבל אין סיבוב. עדיין מקביל לצירים.

Q13

Scale and rotate 3D gaussian



מטריצת הCOV:

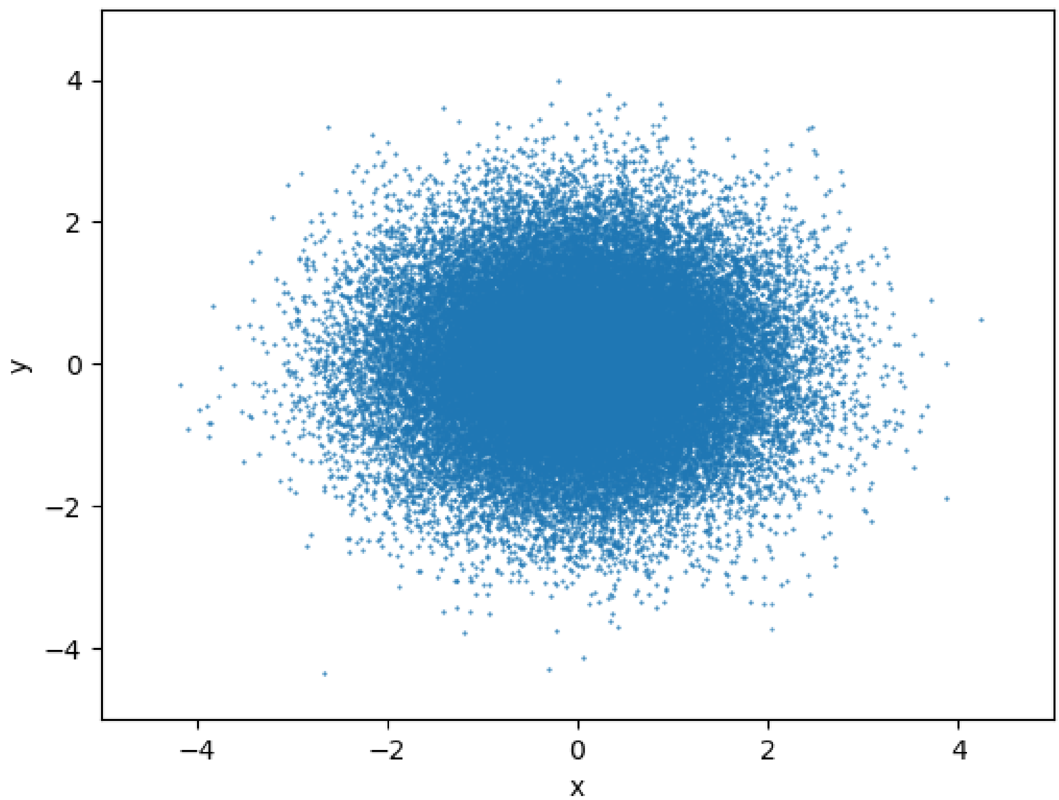
[[0.2012325 0.77954016 0.22139559]
 [0.77954016 3.38267323 1.18514263]
 [0.22139559 1.18514263 0.67609427]]

המטריצה לא אלכסונית- יש גם מתיחה וגם סיבוב (כמו שרואים בסרטוט), אבל היא סימטרית

Q14

נעריך כי הסרטוט יראה כמו גאוסיאן דו מימד סביב אותם ערכים, וזה אכן מה שקיבלנו:

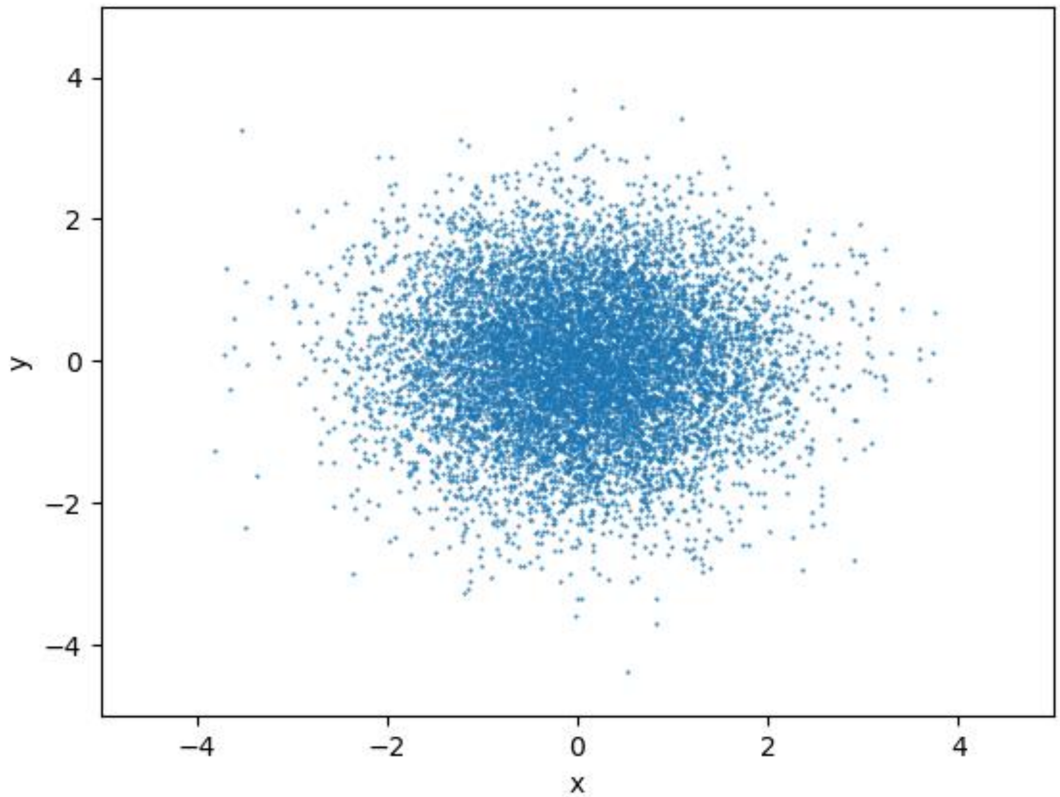
marginal distribution- 2D gaussian



Q15

נצפה לראות גאוסיאן בעל אותם ערכים, אבל בצפיפות נמוכה יותר. כלומר צפיפות הנקודות דלילה יותר אבל הצורה והערכים זהים- וזה אכן מה שקיבלנו:

__marginal distribution and conditional distribution - 2D gaussian__



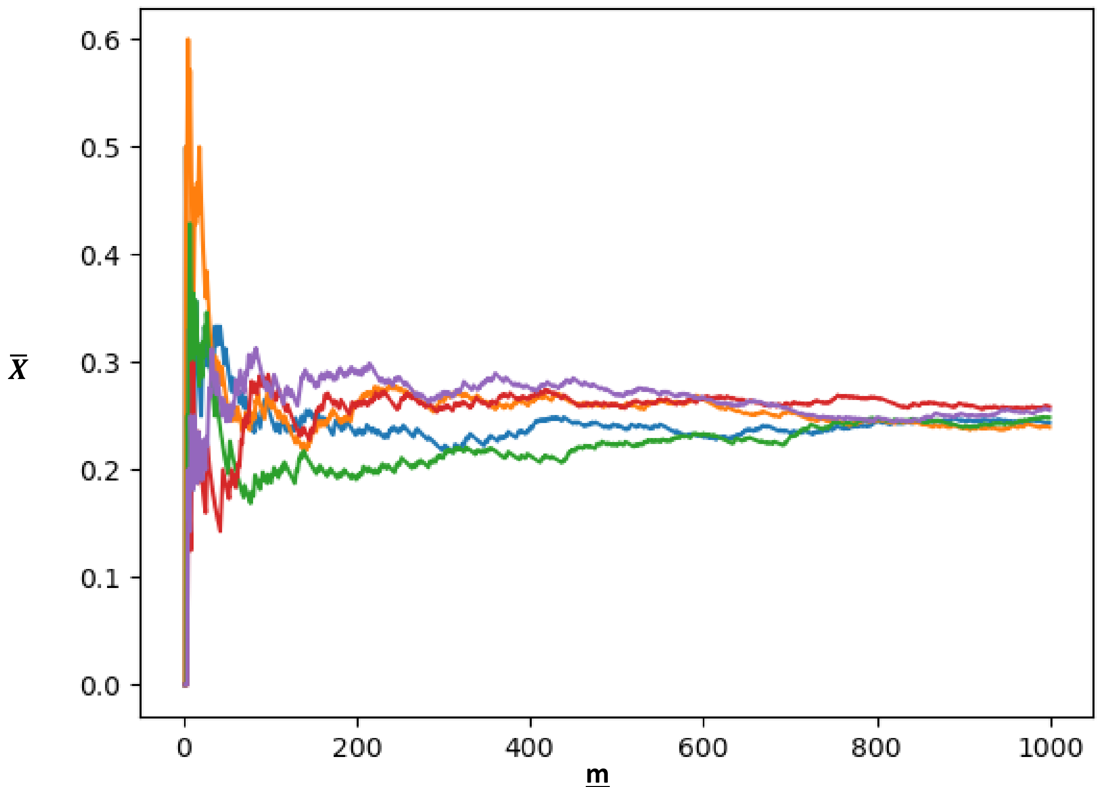
Concentration inequalities - practical question

Q16

שאלה 16 A:

נצפה לראות ככל שM גדל, שאיפה אסימפטוטמטית לממוצע האמיתי (אשר מוערך להיות 0.25)

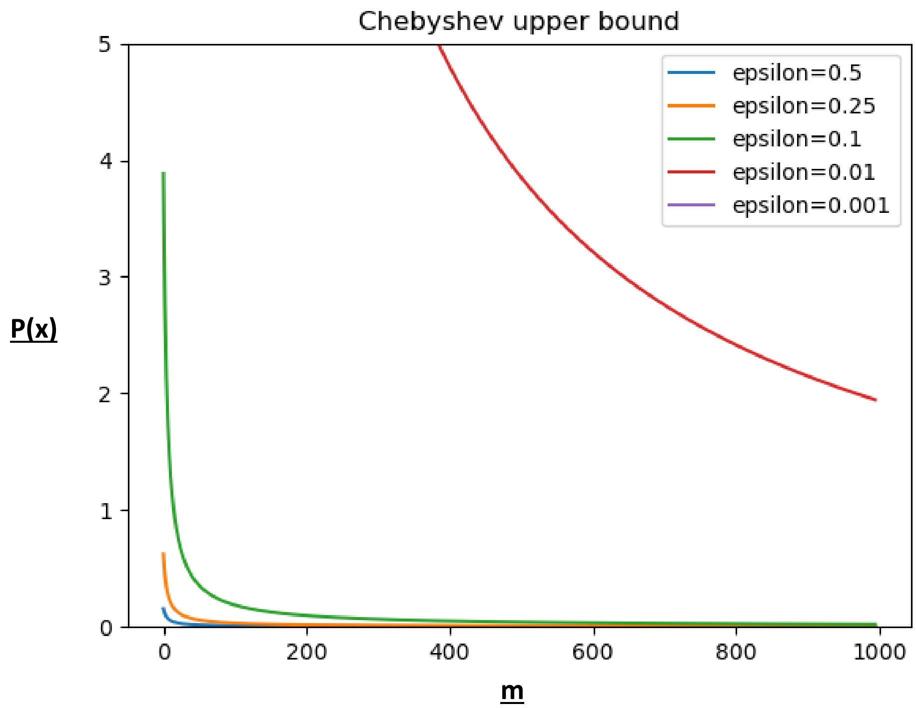
Estimate Everage of X

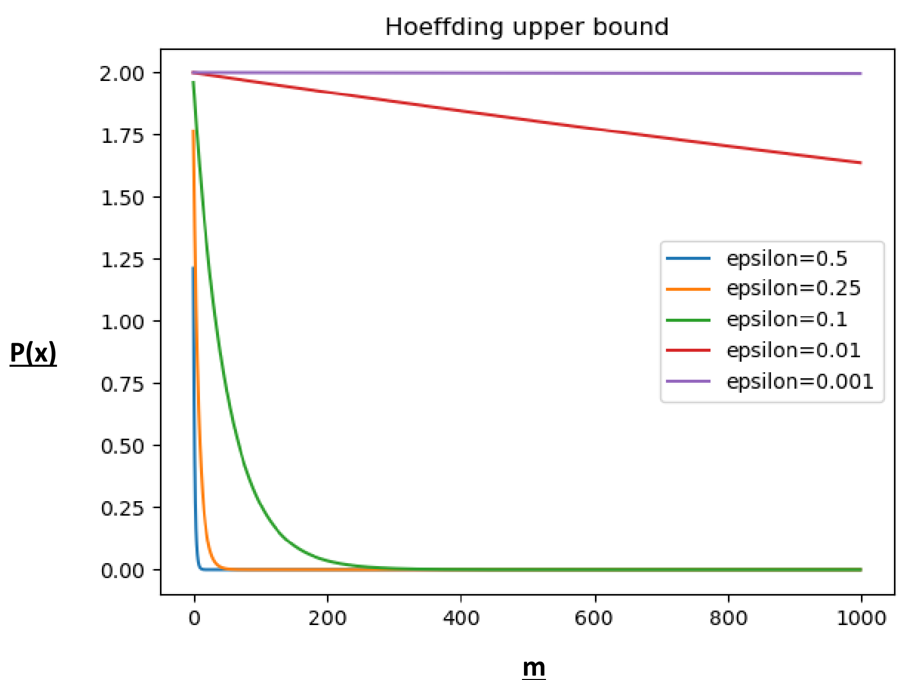


ניתן לראות כי כל השורות מתכנסות לערך מסוים ושכל הערכים הם "קרובים" לערך שציפינו לראות. כולם מתכנסים לאזור ה0.25 ("מתקרבים" זה לא מוגדר – ובהמשך השאלה נדייק יותר)

שאלה 16 B:

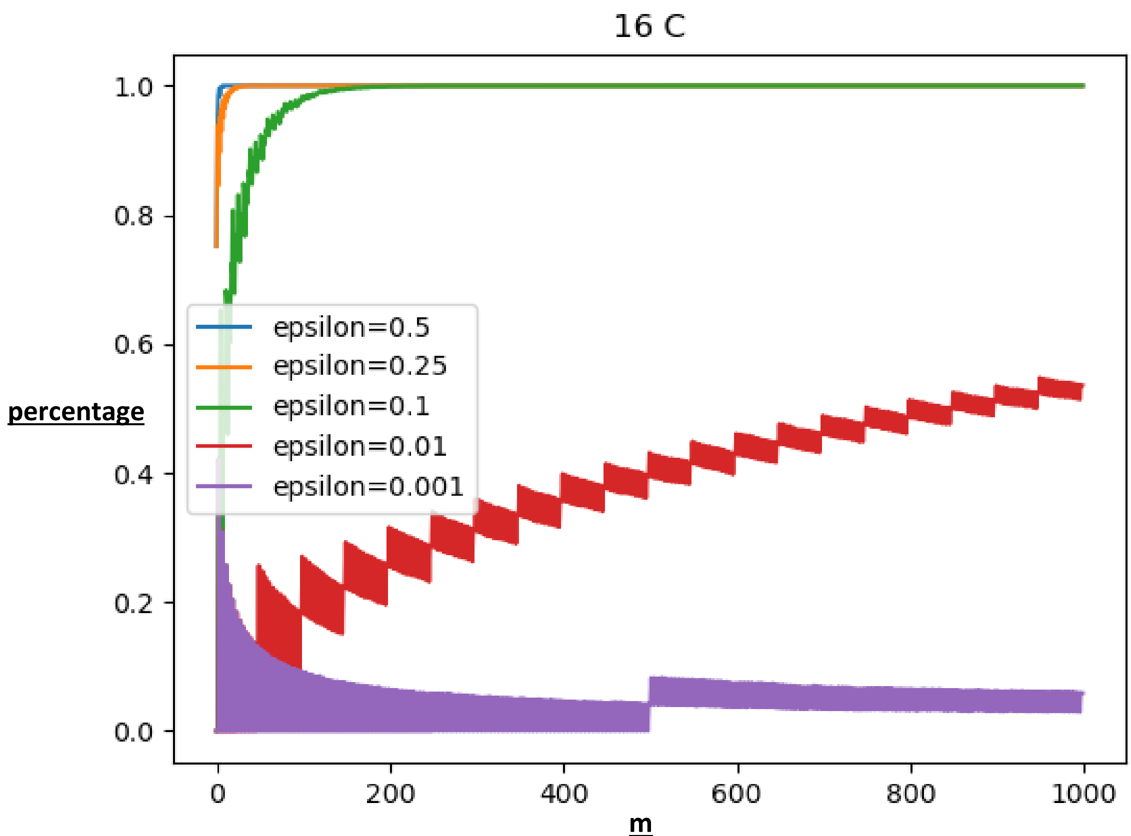
* בגרף הראשון חתכתי את ציר הY בנקודה 5 כאשר הקו הסגול מעליה על מנת לקבל גרף ברור יותר (כאשר באופן אמיתי אין משמעות לערכים ג


Chebyshev upper bound
P(x)
m
epsilon=0.5
epsilon=0.25
epsilon=0.1
epsilon=0.01
epsilon=0.001
0
1
2
3
4
5
0
200
400
600
800
1000


Hoeffding upper bound
P(x)
m
epsilon=0.5
epsilon=0.25
epsilon=0.1
epsilon=0.01
epsilon=0.001
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
0
200
400
600
800
1000

שאלה C16:

נצפה לראות שככל שM גדל נשאף למאה אחוז (1) – כאשר ככל שאפסילון גדול יותר נשאף מהר יותר



וזה אכן מה שקיבלנו- כאשר בעבור אפסילון קטן מאוד מספר הדגימות לא מספיק- המודל עדיין לא מייצג כראוי את ההתפלגות ממנה לקחנו את הנתונים