Short HW1 - Dry

1. נתונים $X_1, X_2, \ldots, X_m$ משתנים מקריים i.i.d בעלי תוחלת $\mathbb{E}[X_i] = \mu$ ושונות $Var(X_i) = \sigma^2$.

ונגדיר את המשתנה המקרי $\overline{X} = \frac{1}{m}\sum_i X_i$, נרצה לחשב את $\mathbb{E}[\overline{X}],\ Var[\overline{X}]$:

$$1.\ \mathbb{E}[\overline{X}] \underbrace{=}_{\overline{X}\text{ הגדרת}} \mathbb{E}\left[\frac{1}{m}\sum_i X_i\right] \underbrace{=}_{\text{לינאריות תוחלת}} \frac{1}{m}\mathbb{E}\left[\sum_i X_i\right] = \frac{1}{m}\mathbb{E}[X_1 + \ldots + X_m]$$

$$\underbrace{=}_{\text{לינאריות תוחלת}} \frac{1}{m}(\mathbb{E}[X_1] + \ldots + \mathbb{E}[X_m]) \underbrace{=}_{\mathbb{E}[X_i]=\mu\text{ נתון}} \frac{1}{m}\big(\underbrace{\mu + \ldots + \mu}_{m\text{ פעמים}}\big) = \frac{m\mu}{m} = \mu$$

$$2.\ Var(\overline{X}) \underbrace{=}_{\overline{X}\text{ הגדרת}} Var\left(\frac{1}{m}\sum_i X_i\right) \underbrace{=}_{Var(aX)=a^2Var(X)} \frac{1}{m^2}Var\left(\sum_i X_i\right) = \frac{1}{m^2}Var(X_1 + \ldots + X_m)$$

$$= \frac{1}{m^2}(Var(X_1) + \ldots + Var(X_m)) \underbrace{=}_{Var(X_i)=\sigma^2\text{ נתון}} \frac{1}{m^2}\big(\underbrace{\sigma^2 + \ldots + \sigma^2}_{m\text{ פעמים}}\big) = \frac{m\sigma^2}{m^2} = \frac{\sigma^2}{m}$$

.2

1. ההתפלגות המתוארת בשאלה היא התפלגות בינומית, כאשר כל הטלה מתארת ניסוי ברנולי עם סיכוי הצלחה p, לכן $\theta_i \sim Bin(50, p)$.

2. נחשב את התוחלת של θ_i לכל $1 \leqslant i \leqslant m$:

$$\mathbb{E}[\theta_i] \underbrace{=}_{\text{הגדרת שונות + הגדרת התפלגות בינומיאלית}} \sum_{i=1}^{m} i \cdot \binom{m}{i} \cdot p^i \cdot (1-p)^{m-i} = \sum_{i=1}^{m} \frac{m!}{(m-i)! \cdot (i-1)!} \cdot p^i \cdot (1-p)^{m-i}$$

$$\underbrace{=}_{i \to i+1} \sum_{i=0}^{m} \frac{m!}{(m-i-1)! \cdot i!} \cdot p^{i+1} \cdot (1-p)^{m-i-1} = mp\underbrace{\sum_{i=0}^{m-1}\binom{m-1}{i} \cdot p^i \cdot (1-p)^{m-i-1}}_{\substack{=1 \\ \text{לפי הגדרת פונקציית התפלגות}}} = mp$$

3. נמצא את m המינימלי שעבורו מתקיים $P(|\overline{\theta}(m) - \mu| > 1) \leqslant 0.01$:

$$P(|\overline{\theta}(m) - \mu| > 1) \leqslant 2e^{-\frac{2m \cdot 1^2}{(50-0)^2}} = 2e^{\frac{-2m}{50^2}} = 2e^{-0.0008m}$$

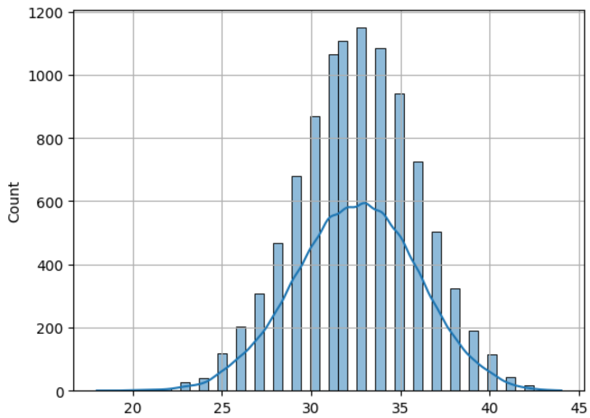
כלומר נבדוק עבור איזה ערך של m מתקיים $2e^{-2m} = 0.01$:

$$2e^{-0.0008m} = 0.01 \,/\, :2$$
$$e^{-0.0008m} = 0.005 \,/\, ln$$
$$-0.0008m = ln(0.005)$$

$$m = -\frac{ln(0.005)}{-0.0008} = 6622.89$$
$$\Longrightarrow m = 6623$$

m צריך להיות מספר שלם כי מתאר מספר הטלות, לכן ה-m המינימלי המקיים את הדרישה הוא 6623.

5.



הגרף מתאר בקירוב הפתלגות נורמלית, כאשר קיבלנו את קירוב זה לפי משפט הגבול המרכזי - לפיו תחת תנאים מסויימים, התפלגות הממוצע של סדרת משתנים מקריים בלתי תלויים תתנהג בקירוב כמו התפלגות נורמלית לאחר נירמול.

3.

1. תהי מטריצה $A \in \mathbb{R}^{n \times n}$, כך ש-$A \succ 0$, כלומר PD.
נראה כי כל הערכים העצמיים של A הם חיוביים ממש:

$$\forall x \in \mathbb{R}^n : \; x^T A x > 0$$

עבור ערך עצמי כלשהו λ ישנם 3 אפשרויות:

$$1. \; \lambda = 0 \quad \underbrace{\Longrightarrow}_{\text{עבור הוקטור העצמי המתאים } x} \quad Ax = \lambda x = 0 \Longrightarrow \; x^T A x = 0$$

בסתירה להגדרת מטריצה PD.

$$2. \; \lambda < 0 \quad \underbrace{\Longrightarrow}_{\text{עבור הוקטור העצמי המתאים } x} \quad Ax = \lambda x \Longrightarrow x^T A x = x^T \lambda x = \underbrace{\lambda}_{<0} \underbrace{|x|^2}_{\geqslant 0} < 0$$

בסתירה להגדרת מטריצה PD.
לכן כל ערכיה העצמיים של A הם חיוביים ממש.

2. לא נכון - דוגמה נגדית:
יהיו $A, B \in R^{n \times n}$ מטריצות סימטריות ו-PSD, נראה כי ייתכן והביטוי $2A - B$ אינו PSD:
לפי הגדרת מטריצה PSD, קיימים וקטור v וסקלרים a, b כך שמתקיים:

$$v^T A v = a \geqslant 0$$
$$v^T B v = b \geqslant 0$$

כאשר בה"כ מתקיים $2a < b$, לכן:

$$v^T(2A-B)v = 2v^TAv - v^TBv = 2a-b<0$$

ולכן המטריצה $2A-B$ אינה PSD.

4. נתון כי $f:\mathbb{R}^d \to \mathbb{R},\ f(w) = w^Tx+b$

1. נראה כי $\nabla_w f = x$:

תחילה נגדיר את x, w:

$$w^T = [w_1,..., w_d]$$
$$x^T = [x_1,..., x_d]$$

וכעת נמצא את וקטור הגרדיאנט:

$$\nabla_w f \underbrace{=}_{\text{הגדרת וקטור גרדיאנט}} \left[\frac{\partial f}{\partial w_1},..., \frac{\partial f}{\partial w_d}\right]^T \underbrace{=}_{f \text{ הגדרת}} \left[\frac{\partial\left(w^Tx+b\right)}{\partial w_1},..., \frac{\partial\left(w^Tx+b\right)}{\partial w_d}\right]^T$$

$$\underbrace{=}_{\text{לפי מכפלת וקטורים}} \left[\frac{\partial(w_1x_1+...\ +w_dx_d+b)}{\partial w_1},..., \frac{\partial(w_1x_1+...\ +w_dx_d+b)}{\partial w_d}\right]^T \underbrace{=}_{\text{לפי הגדרת נגזרות חלקיות}} [x_1,..., x_d]^T = x$$

2. נמצא את המטריצה המתאימה לפונקציה f המוגדרת בשאלה:

$$\nabla_w^2 f \underbrace{=}_{\text{לפי הגדרה}} \begin{bmatrix} \frac{\partial^2 f}{\partial w_1^2} & \frac{\partial^2 f}{\partial w_1 \partial w_2} & \cdots & \frac{\partial^2 f}{\partial w_1 \partial w_d} \\ \frac{\partial^2 f}{\partial w_2 \partial w_1} & \frac{\partial^2 f}{\partial w_2^2} & & \frac{\partial^2 f}{\partial w_2 \partial w_d} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial w_d \partial w_1} & \frac{\partial^2 f}{\partial w_d \partial w_2} & \cdots & \frac{\partial^2 f}{\partial w_d^2} \end{bmatrix} \underbrace{=}_{\text{לפי סעיף א'}} \begin{bmatrix} \frac{\partial x_1}{\partial w_1} & \frac{\partial x_1}{\partial w_2} & \cdots & \frac{\partial x_1}{\partial w_d} \\ \frac{\partial x_2}{\partial w_1} & \frac{\partial x_2}{\partial w_2} & & \frac{\partial x_2}{\partial w_d} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_d}{\partial w_1} & \frac{\partial x_d}{\partial w_2} & \cdots & \frac{\partial x_d}{\partial w_d} \end{bmatrix} \underbrace{=}_{\text{לפי הגדרת נגזרת חלקית}} \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

3. המטריצה שנמצאה היא מטריצת האפס, שלה d ערכים עצמיים שהם 0 - כלומר כל ערכי העצמיים מקיימים $\lambda \geqslant 0$ ולכן המטריצה PSD לפי הגדרה.

4. כעת מוגדרת הפונקציה $g:\mathbb{R}^d \to \mathbb{R},\ g(w) = \lambda||w||^2,\ \lambda > 0$, נמצא את וקטור הגרדיאנט שלה:

$$\nabla_w g \underbrace{=}_{\text{הגדרת וקטור גרדיאנט}} \left[\frac{\partial g}{\partial w_1},..., \frac{\partial g}{\partial w_d}\right]^T \underbrace{=}_{g \text{ הגדרת}} \left[\frac{\partial\left(\lambda||w||^2\right)}{\partial w_1},..., \frac{\partial\left(\lambda||w||^2\right)}{\partial w_d}\right]^T$$

$$\underbrace{=}_{\text{הגדרת נורמת 2}} \left[\frac{\partial(\lambda\left(w_1^2+\ldots\ +w_d^2\right)}{\partial w_1},\ldots,\frac{\partial(\lambda\left(w_1^2+\ldots\ +w_d^2\right)}{\partial w_d}\right]^T = [\lambda w_1,\ldots,\lambda w_d]^T = \lambda[w_1,\ldots,w_d]^T$$

$$= \lambda w$$

5. נמצא את מטריצת ההסיאן של g:

$$\nabla_w^2 g \underbrace{=}_{\text{לפי הגדרה}} \begin{bmatrix} \frac{\partial^2 g}{\partial w_1^2} & \frac{\partial^2 g}{\partial w_1 \partial w_2} & \cdots & \frac{\partial^2 g}{\partial w_1 \partial w_d} \\ \frac{\partial^2 g}{\partial w_2 \partial w_1} & \frac{\partial^2 g}{\partial w_2^2} & & \frac{\partial^2 g}{\partial w_2 \partial w_d} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 g}{\partial w_d \partial w_1} & \frac{\partial^2 g}{\partial w_d \partial w_2} & \cdots & \frac{\partial^2 g}{\partial w_d^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial(2\lambda w_1)}{\partial w_1} & \frac{\partial(2\lambda w_1)}{\partial w_2} & \cdots & \frac{\partial(2\lambda w_1)}{\partial w_d} \\ \frac{\partial(2\lambda w_2)}{\partial w_1} & \frac{\partial(2\lambda w_2)}{\partial w_2} & & \frac{\partial(2\lambda w_2)}{\partial w_d} \\ \vdots & \vdots & & \vdots \\ \frac{\partial(2\lambda w_d)}{\partial w_1} & \frac{\partial(2\lambda w_d)}{\partial w_2} & \cdots & \frac{\partial(2\lambda w_d)}{\partial w_d} \end{bmatrix}$$

$$= \begin{bmatrix} 2\lambda & 0 & \ldots & 0 \\ 0 & 2\lambda & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 2\lambda \end{bmatrix}$$

6. לפי משפט, ערכיה העצמיים של מטריצה אלכסונית הם איברי האלכסון, במטריצה שנמצאה בסעיף 5 קיבלנו מטריצה אלכסונית שאיברי האלכסון בה כולם 2λ, כאשר בשאלה נתון $\lambda > 0$ ולכן כל ערכיה העצמיים של המטריצה מקיימים $\lambda > 0$ ולכן המטריצה PD לפי הגדרה.