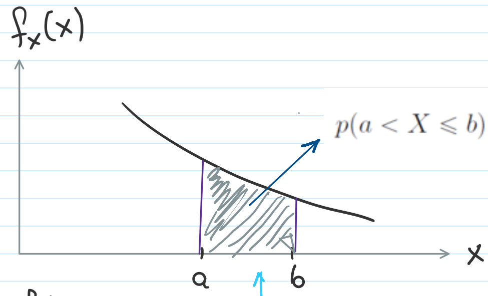
$\boxed{\text{היסטוגרמה}}$

מטרה: ייצוג מספרי/גרפי של PDF מתוך תוצאות הניסוי.

שיטה: $\boxed{\text{תאוריה}}$

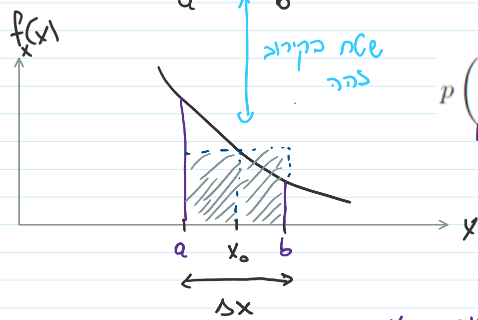
$$p(a < X \leqslant b) = \int_a^b f_X(x)dx = F_X(b) - F_X(a)$$ ← בסיס השיטה

$= p(a \le X \le b)$ ← עבור $\bar{X}$ מ"א רציף



$$p(a < X \leqslant b) = \int_a^b f_X(x)dx$$

$f_X(x)$



שטח בקירוב זהה

$$\underbrace{p\left(x_0 - \frac{\Delta x}{2}\right.}_{a} < X \leqslant \underbrace{\left.x_0 + \frac{\Delta x}{2}\right)}_{b} \approx \underbrace{f_X(x_0)}_{\text{גובה}}\underbrace{\Delta x}_{\text{רוחב}}$$

העברת אגף

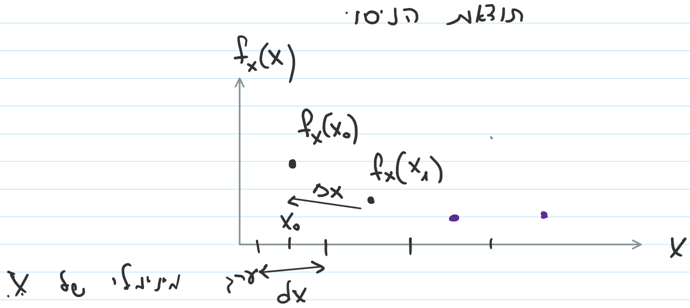
סיכום: ערך מקורב של PDF בנקודה x_0 נתון ע"י:

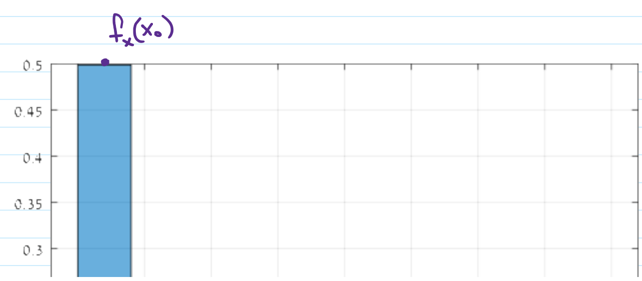
$$f_X(x_0) \approx p\left(x_0 - \frac{\Delta x}{2} < X \leqslant x_0 + \frac{\Delta x}{2}\right)\frac{1}{\Delta x}$$

שיטה: $\boxed{\text{אופן מימוש מעשי}}$

* Δx: נחלק תחום תוצאות הניסוי ל- m חלקים ⇐ Δx

* הערך של $p\left(x_0 - \frac{\Delta x}{2} < X \leqslant x_0 + \frac{\Delta x}{2}\right) = \dfrac{\text{מספר תוצאות הניסוי בתחום } \Delta x \text{ נתון}}{\text{מספר סה"כ של תוצאות הניסוי}}$





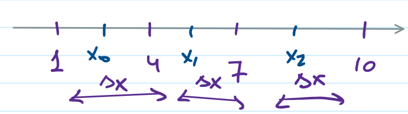
* דוגמה מספרית:

* $X = \{1, 2, 3, 1.5, 5, 7, 10, 2.5, 1.5, 3\}$

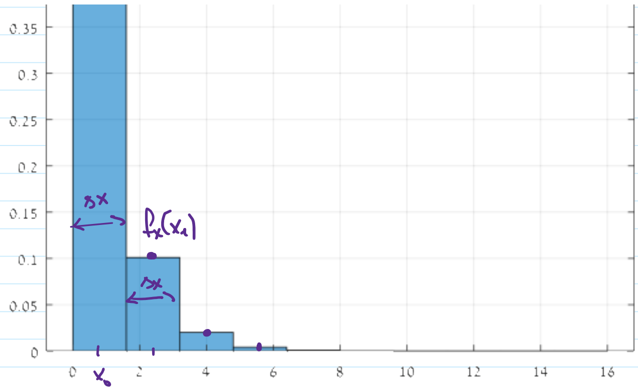
* מספר תוצאות הניסוי $N=10$;

* תחום ערכי הניסוי $[1, 10]$ - רוחב 9

* $m = 3 \Rightarrow \Delta x = \frac{9}{m}$



$x_0 = 2.5$
$x_1 = 5.5$



$X_0 = 2.5$
$X_1 = 5.5$
$X_2 = 8.5$

$$f_X(X_0) = \frac{\overset{\text{מספר תוצאות}}{\text{בין 1 4-8}} = 7}{10} \cdot \frac{1}{\Delta x} = \frac{7}{10}\cdot\frac{1}{3} = \frac{7}{30}$$

$$f_X(X_1) = \frac{2}{10}\cdot\frac{1}{3} = \frac{2}{30}$$

$$f_X(X_2) = \frac{1}{10}\cdot\frac{1}{3} = \frac{1}{30}$$

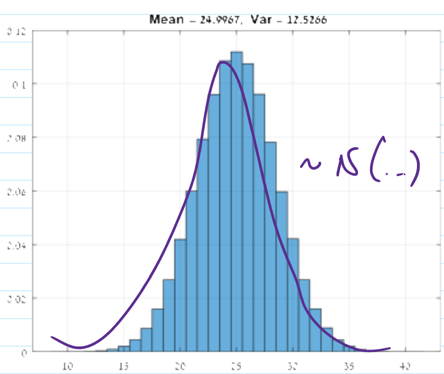
התפלגות גאוסית

הקדמה: משפט גבול המרכזי: central limit theorem

סכום n משתנים בלתי תלויים בעלי התפלגות זהה ובעלי תוחלת
μ ושונות σ^2 תתפלג גאוסית עבור n מספיק גדול:

$$\frac{1}{n}\sum_{k=1}^{n} X_n \sim N\left(\mu, \sigma^2\right)$$

הערה: זה לא משנה אם
מדובר במ"א בדיד
או רציף



$E[Y]=\mu \quad Var[Y]=\sigma^2$

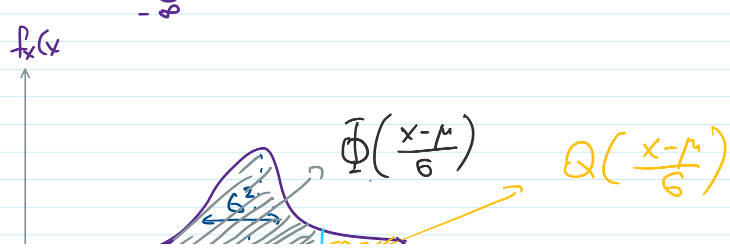
הגדרה:

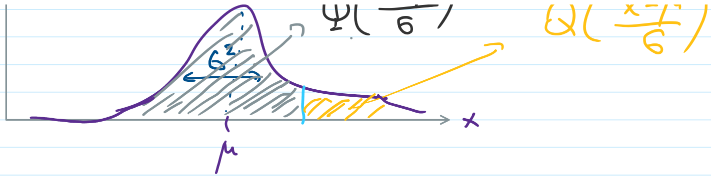
PDF (הגדרה 3.8): עבור משתנה אקראי גאוסי $Y \sim N(\underbrace{\mu}_{\text{תוחלת}}, \underbrace{\sigma^2}_{\text{שונות}})$, בעל תוחלת μ ושונות σ^2, מתקיים

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad \text{PDF} \tag{3.17}$$

CDF (הגדרה 3.9): מסומן ע"י $\Phi(x)$ וניתן לחישוב נומרי (מספרי) בלבד.

$$\Phi(x) = \int_{-\infty}^{x} f_x(x)\,dx = P(X \le x)$$





תכונות: כהינתן $X(0,1) \Leftarrow \mu=0,\ \sigma^2=1$

תזכורת: $Var[aX+b] = a^2 Var[X]$ $\rightarrow$ $aX \sim N(0, a^2)$

$E[b+X] = b + E[X]$ $\quad$ $X+b \sim N(b,1)$

$aX+b \sim N(b, a^2)$

כהינתן $Y(\mu,\sigma^2) \Leftarrow \frac{Y-\mu}{\sigma} = X \sim N(0,1)$

Q-function: $X \sim N(0,1^2)$

$$Q(x) = \int_x^{\infty} f_x(s)\,ds$$

תכונה חשובה: עבור $Y \sim N(\mu,\sigma^2)$

$$p(Y>y) = Q\left(\frac{y-\mu}{\sigma}\right)$$

$Q(x) = 1 - \Phi(x)$ ← ראה בדף לאל

זוג משתנים אקראיים רציפים

CDF (הגדרה 4.1):

$$F_{XY}(x,y) = p(X \leqslant x, Y \leqslant y)$$

וגם

PDF (הגדרה 4.2):

$$f_{XY}(x,y) = \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y}$$

קשר בין PDF ל-CDF (תכונה 4.1):

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(s,p)\,dp\,ds$$

תכונות PDF (תכונה 4.2): תחום ערכים ו"סכום" ערכים

$$f_{XY}(x,y) \geqslant 0$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1$$

תכונות:

תכונות:

$g(X,Y) = XY$ → מקרה פרטי.

$$E[XY] = \iint xy f_{XY}(x,y)dxdy$$

$$E[g(X,Y)] = \iint g(x,y) f_{XY}(x,y)dxdy$$

משתנים בלתי תלויים סטטיסטית (הגדרה 4.4): משתנים נקראים בלתי תלויים סטטיסטית אם ורק אם מתקיים

$$f_{XY}(x,y) = f_X(x)f_Y(y) \quad \text{(4.8א)}$$

$$F_{XY}(x,y) = F_X(x)F_Y(y) \quad \text{(4.8ב)}$$

במקרה הזה בלבד!!! ניתן לפצל חישוב של תוחלת:

$E[XY] = E[X]E[Y]$

התפלגות שולית (הגדרה 4.3):

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$$

$$F_X(x) = F_{XY}(x,\infty)$$

$$F_Y(y) = F_{XY}(\infty,y)$$

משתנים גאוסים במשותף (התפלגות גאוסית דו-ממדית)

הקדמה:

משפט. משתנים אקראיים $X_1, X_2, \dots, X_N$ הם גאוסיים במשותף (בעלי התפלגות גאוסית N-ממדית משותפת) אם ורק אם הקומבינציה לינארית $\sum_{m=1}^{N} a_m X_m$ היא מתלפגת גאוסית עבור $\forall a_m \in \mathbb{R}$.

$$a_1X_1 + a_2X_2 + \dots + a_NX_N \sim N(\ )$$

$$\Longleftrightarrow$$

חישוב מקדים: $W_2 \sim N(0,1)$ $W_1 \sim N(0,1)$ W_1, W_2 בלתי תלוי"ם

$$X_1 = \mu_1 + \sigma_{11}W_1 + \sigma_{12}W_2$$

$$X_2 = \mu_2 + \sigma_{21}W_1 + \sigma_{22}W_2$$

μ_1, σ_{ij}, μ_2 } קבועים ממשיים

מטרה של חישוב מקדים: ניתוח X_1, X_2 וקשרים ביניהם

ניתוח X_1: - $X \sim N(\ )$ בהתאם למשפט

תוספת של קבוע משאירה התפלגות גאוסית

$$- \ E[X_1] = E[\mu_1] + \sigma_{11}E[\cancel{W_1}]^{0} + \sigma_{12}E[\cancel{W_2}]^{0} = \mu_1$$

$$- \ Var[X_1] = \underbrace{\sigma_{11}^2 + \sigma_{12}^2}_{=\sigma_1^2}$$

$Var[W_1]=1$ (נתון)
$Var[W_2]=1$

תזכורת:

תזכורת: עבור משתנים בלתי תלוי"ם
$Var[aX+bY] = a^2 Var[X] + b^2 Var[Y]$

התפלגות { $X_1 \sim N(\ \ ,\ \sigma^2 + \sigma^2)$

תזכורת: $-\ \mathrm{Var}[X_1] = \sigma_{11}^2 + \sigma_{12}^2$

$\mathrm{Var}[W_1]=1$ (נתון)
$\mathrm{Var}[W_2]=1$

> תזכורת: עבור משתנים בלתי תלויים
> $\mathrm{Var}[aX + bY] = a^2\,\mathrm{Var}[X] + b^2\,\mathrm{Var}[Y]$

התפלגויות שוליות

סיכום: $X_1 \sim N(\underbrace{\mu_1}_{\text{תוחלת}}, \underbrace{\sigma_{11}^2 + \sigma_{12}^2}_{\text{שונות}})$, $= \sigma_1^2$

* X_2: באופן דומה ל-X_2: $X_2 \sim N(\mu_2, \overbrace{\sigma_{21}^2 + \sigma_{22}^2}^{\sigma_2^2})$

* קשר בין $X_1 X_2$: לינארי

$$\mathrm{Cov}[X_1, X_2] = E[X_1 X_2] - \underbrace{E[X_1]}_{\mu_1}\underbrace{E[X_2]}_{\mu_2}$$

$$\rho_{X_1X_2} = \frac{\mathrm{Cov}[X_1, X_2]}{\sqrt{\underbrace{\mathrm{Var}[X_1]}_{\sigma_1^2}\ \underbrace{\mathrm{Var}[X_2]}_{\sigma_2^2}}}$$

$$\sigma_1^2 = \sigma_{11}^2 + \sigma_{12}^2$$
$$\sigma_2^2 = \sigma_{21}^2 + \sigma_{22}^2$$

$$E[X_1X_2] = E\left[\underbrace{(\mu_1 + \sigma_{11}W_1 + \sigma_{12}W_2)}_{X_1}\overbrace{(\mu_2 + \sigma_{21}W_1 + \sigma_{22}W_2)}^{X_2}\right]$$

$$= E\Big[\underbrace{\mu_1\mu_2}_{\text{קבוע}} + \mu_1\sigma_{21}\cancel{W_1} + \mu_1\sigma_{22}\cancel{W_2}$$
$$+\ \mu_2\sigma_{11}\cancel{W_1} + \sigma_{11}\sigma_{21}W_1^2 + \sigma_{11}\sigma_{22}\cancel{W_1W_2}$$
$$+\ \mu_2\sigma_{12}\cancel{W_2} + \sigma_{12}\sigma_{21}\cancel{W_1W_2} + \sigma_{12}\sigma_{22}W_2^2\Big]$$

* $E[W_1] = E[W_2] = 0$

* $E[W_1W_2] = E[W_1]\cdot E[W_2] = 0$ (משתנים בלתי תלויים)

* $E[W_1^2] = \quad \Rightarrow$

$= \underbrace{\mathrm{Var}[W_1]}_{=1} + \cancel{E^2[W_1]}^{0} = 1$

$= E[W_2^2]$

$$= \mu_1\mu_2 + \sigma_{11}\sigma_{21}(1) + \sigma_{12}\sigma_{22}(1)$$

$$\mathrm{Cov}[X_1, X_2] = \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}$$

> $\boxed{\mathrm{Var}[X] = E[X^2] - E^2[X]}$
> תזכורת

$$\rho = \frac{\mathrm{Cov}[X_1, X_2]}{\sqrt{\underbrace{\mathrm{Cov}[X_1, X_1]}_{\mathrm{Var}[X_1]}\ \underbrace{\mathrm{Cov}[X_2, X_2]}_{\mathrm{Var}[X_2]}}} = \frac{\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22}}{\sigma_1\sigma_2}$$

מטרה: התפלגות משותפת

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N\left(\underbrace{\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}}_{\text{תוחלת (וקטור)}}, \underbrace{\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}}_{\text{מטריצת Cov}}\right)$$

רישום קצר יותר

> תזכורת
> $\mathrm{Cov}[X, X] = Var[X]$
> $\mathrm{Cov}[X, Y] = \mathrm{Cov}[Y, X]$

הגדרה: מטריצת Cov

$$C_{\mathbf{X}} = \begin{bmatrix} \mathrm{Cov}[X_1, X_1] & \mathrm{Cov}[X_1, X_2] \\ \mathrm{Cov}[X_2, X_1] & \mathrm{Cov}[X_2, X_2] \end{bmatrix} = \begin{bmatrix} \mathrm{Var}[X_1] & \mathrm{Cov}[X_1, X_2] \\ \mathrm{Cov}[X_2, X_1] & \mathrm{Var}[X_2] \end{bmatrix}$$

זוג משתנים גאוסיים במשותף PDF (הגדרה 4.8): ה-PDF המשותף של זוג משתנים גאוסיים X_1, X_2 נתון ע"י

$$(4.16)\quad f_{X_1X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \times\right.$$

זוג משתנים גאוסיים במשותף PDF (הגדרה 4.8): ה-PDF המשותף של זוג משתנים גאוסיים X_1, X_2 נתון ע"י

$$f_{X_1X_2}(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}\times\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2}+\frac{(x_2-\mu_2)^2}{\sigma_2^2}-\frac{2\rho(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right]\right) \quad (4.16)$$