סימונים בתרגול

P פונק' התפלגות

p (קטן) פונק' התפלגות נק'

הסתברות 1 - תרגול 4

8 בנובמבר 2018

1 משתנים מקריים

הגדרה יהא (Ω, P) מרחב הסתברות (בדיד) . פונקציה $X : \Omega \to \mathbb{R}$ נקראת משתנה מקרי. נכתוב בקיצור ש X הוא מ"מ.

הערה 1 שימו לב שלא דרשנו שום תכונה של הפונקציה או של S בהגדרה הנ"ל. זו תכונה מיוחדת של מרחבים סופיים או בני מנייה. לו היינו מטפלים גם במרחבים גדולים יותר, היינו צריכים לשנות את ההגדרה.

הערה 2 שימו לב שאם X הוא מ"מ אז הפונקציה

פונק' התפלגות

$$\boxed{P_X(A) = P(X^{-1}(A)) = \sum_{\omega : X(\omega) \in A} P(\{\omega\})}$$

הסבר לשלב עמוד הבא

מגדירה פונקציית הסתברות על $\mathbb{R}$. פונקציה זו נקראת ההתפלגות (הבדידה) של המשתנה המקרי X.

הערה 3 נניח ש $X, Y : \Omega \to \mathbb{R}$ הם שני מ"מ שמוגדרים על אותו מרחב הסתברות (Ω, P). אז גם $X + Y, X \cdot Y$ וכן $\alpha \cdot X$ לכל $\alpha \in \mathbb{R}$ הם משתנים מקריים (כאשר הטווח שלהם יכול להיות שונה) שמוגדרים על אותו מרחב.

הגדרה יהא X משתנה מקרי. פונקציית ההתפלגות הנקודתית $p_X : Im(X) \to \mathbb{R}$ של המשתנה המקרי X היא הפונקציה שנותנת לכל ערך אפשרי של המשתנה המקרי את ההסתברות שלו. כלומר, אם $x \in \mathbb{R}$ אז

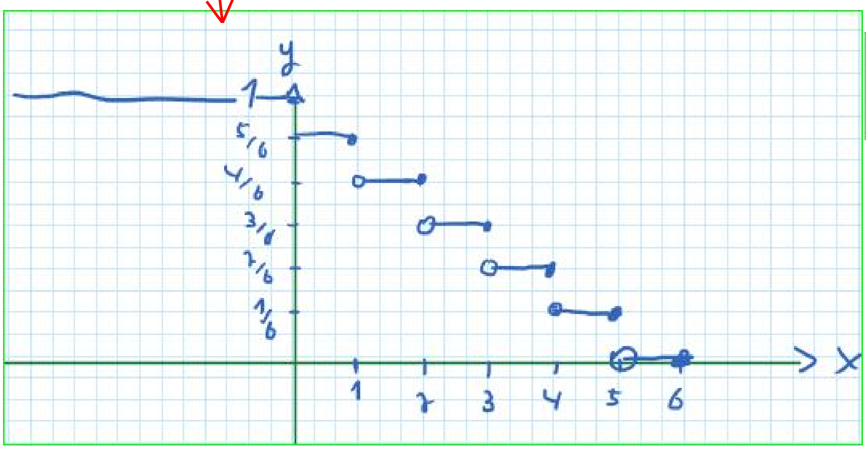
$$p_X(x) = P(\{\omega \in \Omega : X(\omega) = x\})$$

פונקציית ההתפלגות השיורית $F_X : \mathbb{R} \to \mathbb{R}$ של X מוגדרת ע"י

$$\bar{F}_X(x) = P(\{\omega \in \Omega : X(\omega) > x\})$$

.

הערה 4 כאשר עובדים עם משתנה מקרי לא תמיד מתייחסים למרחב הסתברות ברקע אלא עבור חישובים מפורשים כאשר צריך. בהרבה המקרים, כאשר נתון משתנה מקרי ופונקציית ההתפלגות המצטברת שלו, למשל, יש לנו מספיק מידע ואין צורך להתייחס למרחב הסתברות ברקע.



דוג' של מרחב התפלגות אחיד של הטלת קוביה אחת

משתנה מקרי: הוא בסה"כ פונק' $X: \Omega \to [0,1]$ אני יכול לחשוב עליו כמו **קטגוריה כללית**

כמו הגובה של כל האנשים בקבוצה ומה שמעניין אותי זה התפלגות של X מבחינת ביחס למרחב המדגם שאיתו אני עובד. כי תמיד בהגדרה כן שכחת אני מקבל התפלגות של X ואני לא תלוי במרחב המדגם שלי

כעת אני מגדיר את **פונק' ההתפלגות** כך: $\mathbb{P}_X : \mathcal{B}^{\mathbb{R}} \to [0,1]$

$$\mathbb{P}_X(A) = \mathbb{P}[\underbrace{X^{-1}(A)}]$$

כל תת קב' של Ω עם המרחב הישן והמוכר

פונק' התפלגות מקושר

פונקצי' **ההסתברות** הישנה והמוכרת מהמרחב מדגם הישן שלי קרי $(\Omega, F, \mathbb{P})$

הרעיון בהגדרה החדשה זה שיש תת-השמות בהתפלגות של **X** ולהשתמש מהמרחב מדגם הישן שלי

עכשיו מהרגע שהוגדרו לי המ"מ אני משתמש פחות ממרחב המדגם שלי ובהמשך נחקור **התפלגויות שונות**

פונק' התפלגות נק': מוגדרת כך: $P_X(\tilde{x}) = \{w \in \Omega \mid P(w) = \tilde{x}\}$ $\quad P_X : \mathrm{Im}X \to \mathbb{R}$

יחידון!

למעשה שוב זו אותה פונק' ההסתברות נק' של המרחב הישן היא מגדירה לי כיצד פונק' התפלגות נק' מתנהגת.

נשים לב שמהרגע שבו פונק' ההסתברות הנק' נקבעה נקבעה גם כן פונק' ההתפלגות **שכן יש:**

$$\mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A)) = \sum_{w \in X(w) \in A} P_X(\{w\})$$

פונק' התפלגות נק' שאעלה יחידון $(\{w\})$

כאמור גם כאן $P(\Omega) = 1$

פונק' התפלגות נק'

דוגמאות:

1. סכום ההטלות של שתי קוביות הוגנות.

2. פוגשים באקראיות אדם ברחוב ורוצים לדעת מה ההסתברות שגובהו מעל 1.70 מ'. במודל ההסתברותי, Ω הוא אוסף כל האפשרויות לפגישת אדם ברחוב (נאמר - כל בני האדם האפשריים בעולם), $\mathbb{P}$ תהיה איזו פונקציית הסתברות שתתאר נכון את מי יותר סביר שנפגוש ואת מי לא. בשביל הגובה, נגדיר $X:\Omega\to\mathbb{R}$ על ידי:

$$X(\omega) = \text{Height of } \omega$$

והשאלה הומרה לחישוב ההסתברות של המאורע "$X > 1.70$", כלומר: $\mathbb{P}(X > 1.70) = \mathbb{P}(\{\omega\in\Omega : X(\omega) > 1.70\})$.

3. מטילים פעמיים קוביה, מה ההסתברות שהערך של ההטלה הראשונה גדול מערך ההטלה השניה? אם $(\Omega,\mathcal{F},\mathbb{P})$ הוא מרחב ההסתברות שמתאים לשאלה, עבור $i=1,2$ נגדיר את X_i להיות הערך שיצא בהטלה ה־i:

$$X_i(\omega) = \text{Output of the } i\text{th roll in } \omega$$

והשאלה הומרה לחישוב הערך $\mathbb{P}(X_1 > X_2)$. לדוגמה, אם בחרנו את מרחב ההסתברות להיות כמו שאנו פותרים בדרך כלל שאלה מסוג זה - $\Omega = \{1,...,6\}^2$, אז נוכל להציג במפורש את הפונקציות X_i:

$$X_1(a,b) = a$$
$$X_2(a,b) = b$$

2 התפלגויות מיוחדות

2.1 התפלגות אחידה

הגדרה יהא (Ω, P) מרחב הסתברות. נאמר שמ"מ X המוגדר על Ω מתפלג אחיד אם קיים n כך שהטווח שלו הוא קבוצה בת n איברים S, וכן פונקציית ההתפלגות הנקודתית שלו מקיימת לכל s בטווח

$$p_X(s) = \frac{1}{n}$$

ונסמן $X \sim U(\{1,...,n\})$.

הערה משתנה מקרי אחיד מתאר למעשה את רוב הדוגמאות שראינו עד עכשיו בקורס. הוא מתקבל כאשר עובדים עם מרחב הסתברות סופי, ולכל יחידון יש הסתברות שווה.

דוגמא מטילים קובייה הוגנת. נסמן ב X את תוצאת ההטלה. אז הטווח של X הוא הקבוצה $\{1,..,6\}$ ולכל k בקבוצה הנ"ל מתקיים $P(X=k)=\frac{1}{6}$. לכן, $X \sim U(\{1,...,6\})$.

$p_X(1) = \underbrace{\mathbb{P}_X(X=1)}_{\text{פונק' ההתפלגות אוכלת קבוצה באגף שמאלי את הערך 1 ונשמה X}} = \mathbb{P}_X(X=\{1\})$

פונק' התפלגות נק' אוכלת (1)

סימון אחר

2.2 התפלגות ברנולי

הגדרה יהא (Ω, P) מרחב הסתברות. נאמר שמ"מ X המוגדר על Ω מתפלג ברנולי (או שפשוט X הוא מ"מ ברנולי) עם פרמטר $p \in [0,1]$, אם הטווח שלו הוא $\{0,1\}$, ופונקציית ההתפלגות הנקודתית שלו מקיימת

$$p_X(1) = P(X=1) = p$$

ואז נסמן $X \sim ber(p)$.

למעשה, משתנה ברנולי מתאר כל ניסוי (מורכב ככל שיהיה) שאנחנו מסתכלים על התוצאה שלו רק כהצלחה או כשלון.

דוגמא יהא (Ω, P) מרחב ההסתברות האחידה שמתאר הטלת מטבע הוגן, כלומר $\Omega = \{H,T\}$. יהא X המשתנה המקרי שנותן 1 אם יצא ראש ו־0 אחרת. אז X מתפלג ברנולי עם פרמטר $\frac{1}{2}$. אם המטבע לא בהכרח הוגן, אלא נופל על ראש בהסתברות p, הרי ש $X \sim ber(p)$.

דוגמא סטודנט ניגש למבחן אמריקאי בלי ללמוד את החומר, ובכוונתו לנחש את כל התשובות באקראי. מרחב המדגם הוא כל הסדרות $\Omega = \{1,2,3,4\}^n$ כש־n מספר השאלות. המשתנה

$$X(\omega) = \begin{cases} 0 & \#\text{of correct answer } < 0.6\text{n} \\ 1 & \text{Otherwise} \end{cases}$$

הוא משתנה ברנולי עם פרמטר p שמחושב ממרחב המדגם המתואר. הוא מתאים, כי לסטודנט אכפת רק אם הוא עבר או לא.

2.3 התפלגות גיאומטרית

נסתכל על ניסוי חוזר בעל הסתברות הצלחה p שאינה תלויה בתוצאות הניסויים הקודמים. אנו מבקשים לדעת מה הוא מספר הניסויים עד להצלחה ראשונה. התפלגות זו נקראת התפלגות גיאומטרית.

הגדרה יהא (Ω, P) מרחב הסתברות. מ"מ X המוגדר על Ω מתפלג גיאומטרית עם פרמטר $p \in [0,1]$, $X \sim Geo(p)$, אם הוא מקבל ערכים בקבוצה $\mathbb{N}$ וכן פונקציית ההתפלגות הנקודתית שלו מקיימת

$k=$ הצלחה 1

$$p_X(k) = (1-p)^{k-1} \cdot p$$

למשל: מטילים מטבע שנופל על ראש בהסתברות p באופן ב"ת עד שמקבלים ראש. X מסמן את כמות הנסיונות שעשינו עד שקיבלנו H.

בניסוח קצת שונה,

טענה: עבור סדרה $(Z_n)_{n=1}^{\infty}$ של מ"מ ב"ת עם התפלגות $Z_n \sim ber(p)$ לכל n, יהי

$$X = \min\{n : Z_n = 1\}$$

הערה אם $X \sim Geo(p)$נקבל נוסחה יפה לפונקצית ההתפלגות הצוברת. לכל $k \in \mathbb{N}$ מתקיים

$$F_X(k) = P(X > k) = 1 - P(X \leq k) = 1 - \sum_{i=1}^{k} P(\{X = i\})$$
$$= 1 - \sum_{i=1}^{k} (1-p)^{i-1} \cdot p = 1 - p\frac{1-(1-p)^k}{1-(1-p)} = (1-p)^k$$

דוגמא מסקרים שנערכו התברר ש־40% מתושבי העיר משתמשים במשחת השיניים "לובן". איש חברת פרסום רוצה לראיין משתמש כזה. נניח שהוא פונה לאנשים ברחוב באופן מקרי. נסמן ב־X את מספר האנשים שאליהם הוא פונה עד שהוא מגיע לאדם שמשתמש במשחה "לובן". ההסתברות שהוא יראיין לכל היותר שלושה צרכנים היא

$$P(X \leq 3) = 1 - P(X > 3) = 1 - 0.6^3$$

כאשר נעזרנו בהערה בשוויון האחרון.

סגולתה המיוחדת של ההתפלגות הגיאומטרית היא תכונת חוסר הזכרון. לדוגמא, מהמר כרוני משחק במכונת מזל פעם אחר פעם בציפייה לזכייה הנחשקת. הוא מקווה שיש למכונה זכרון ושהיא תתחשב בכל נסיונותיו הכושלים עד כה. אבל למכונת המזל כמו גם למשתנה הגיאומטרי אין זכרון. אחרי כל כשלון ההתפלגות (החדשה־מותנית) נראית כמו ההתפלגות שהתחיל איתה!

כדומר, אם ידוע לנו ש k הנסיונות הראשונים נכשלו. הסיכוי שנצליח תוך m נסיונות נוספים הוא בדיוק הסיכוי שנצליח ב m נסיונות בלבד. אחרי כל כשלון אנחנו למעשה מתחילים מהתחלה.

לפני ניסוח הטענה, נזכר בהגדרה הבאה:

הגדרה 2.1 (התפלגות מותנית) יהי X מ"מ על המרחב הבדיד $(\Omega, \mathcal{F}, P)$ ויהי $A \in \mathcal{F}$ מאורע המקיים $P(A) > 0$. נסמן ב $(X|A)$ את המ"מ X על מרחב ההסתברות $(\Omega, \mathcal{F}, P_A)$. כאשר P_A ההסתברות המותנית.

טענה יהא X מ"מ הנתמך על $\mathbb{N}$. אזי התנאים הבאים שקולים:

1. $X \sim Geo(p)$ עבור $0 < p < 1$ כלשהו.
2. X ו $(X-1|X>1)$ שווי התפלגות.
3. X ו $(X-k|X>k)$ שווי התפלגות לכל $k \in \mathbb{N}$.

שימו לב שמספיק להראות שיוון
בהינתן התפלגות הנ' [illegible] [illegible] משתנה
להראות שהם שווי התפלגות

הוכחה:

(1) **גורר** (2):

נניח ש X מתפלג גיאומטרית. לכל n מתקיים ע"פ הגדרה (אנו מתנים במאורע $A = \{X > 1\}$)

$$P(X-1 = n|X > 1) = \frac{P(X = n+1)}{P(X > 1)} = \frac{(1-p)^n p}{1-p} = (1-p)^{n-1}p = P(X = n)$$

וסיימנו!

(2) **גורר** (3)**:**

נניח כי X ו $(X-1|X>1)$ שווי התפלגות. לכל $k \in \mathbb{N}$ נגדיר מ"מ $Y_k = (X-k)$ נניח באינדוקציה שהמשתנים $(Y_{k-1}|Y_{k-1}>0)$ ו X שווי התפלגות. אזי

$$(Y_{k-1}|Y_{k-1}>0) \overset{d}{=} X \overset{d}{=} (X-1|X>1) \overset{d}{=} (Y_{k-1}-1|Y_{k-1}>1, Y_{k-1}>0)$$
$$\overset{d}{=} (Y_{k-1}-1|Y_{k-1}>1) = (Y_k, Y_k>0) = (X-k|X>k$$

כאשר השוויון הראשון הוא הנחת האינדוקציה, השני הוא בדיוק הנחה (2), השלישי שוב מהנחת האינדוקציה ($(Y_{k-1}|Y_{k-1}>0)$ מתפלג כמו X) הרביעי כי ההתניה Y_{k-1} לא מוסיפה דבר (לוקחים את חיתוך ההתניות) ושני האחרונים מתקבלים ישירות מההגדרה .

(3) **גורר** (1)**:** **(החלק המעניין)**

(1) יהי X מ"מ עם ערכים ב $\mathbb{N}$ המקיים את תכונה (3). נזכר בכלל השרשרת לסדרה יורדת של מאורעות:

אם $\cancel{A_1 \in A_2 \in \cdots \in A_n}$ עם הסתברות חיובית

$$\boxed{P(A_n) = P(A_1 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_2)\cdots P(A_n|A_{n-1})}$$

(שנובע ישירות באינדוקציה מ $P(A \cap B) = P(B)P(A|B)$) נבחר $A_i = \{X > i\}$ (מדוע הם בעלי הסתברות חיובית?) ונקבל

$$\boxed{P(X>k) = P(X>k|X>k-1)P(X>k-1|X>k-2)\cdots P(X>1) = P(X>1)^k}$$

כאשר השוויון האחרון נובע מהנחה (3), כלומר, מכך שלכל $s \in \mathbb{N}$

$$P(X>1) = P(X-s>1|X>s) = P(X>s+1|X>s)$$

נבחר את p ע"י הדרישה $1-p = P(X>1)$. ונקבל שפונקצית ההתפלגות השיורית של X היא

$$F_X(k) = P(X>k) = (1-p)^k$$

בדיוק כמו התפלגות גיאומטרית. הטענה אפוא נובעת מכך שפונקצית ההתפלגות השיורית קובעת את פונקצית ההתפלגות באופן יחיד (נובע משאלה 1 בתרגיל הבית הקרוב).

(1) נשים לב שאם $A_k = \{X > k \mid k \in N\}$ אז

$A_1 \supset A_2 \supset A_3 \supset — \supset A_k$

כי אם אני לוקח

מ-2 אז קל להראות את כולם מ1 לכן

אם $w \in A_2 = \{x > 2\}$

בהכרח $w \in A_1 = \{x > 1\}$

כלומר $A_1 \supset A_2$