

הסתברות למתמטיקאים – הרצאה 5

14.03.11

איך אפשר לתאר מידה ב- $\mathbb{R}$? למשל ע"י פונקצית התפלגות מצטברת. מידת הסתברות μ ב- $\mathbb{R}$.
פונקציית התפלגות המצטברת:

$$F_\mu(x) = \mu((-\infty, x])$$

לכן, $F_\mu : \mathbb{R} \rightarrow [0, 1]$ פונקציה עולה. $F_\mu(-\infty) = 0$ כמובן הכוונה לגבול. מדוע? כי $A_n \downarrow A \Rightarrow \mu(A_n) \downarrow \mu(A)$ במקרה שלנו חיתוך הקטעים ישאף לקבוצה ריקה. באופן דומה $F_\mu(+\infty) = 1$, ע"י נימוק דומה לפונקציות עולות. $A_n \uparrow A \Rightarrow \mu(A_n) \uparrow \mu(A)$.
כמו כן,

$F_\mu(x+) = F_\mu(x)$ כי $x_n \downarrow x \Rightarrow \mu((-\infty, x_n]) \downarrow \mu((-\infty, x])$ מה לגבי הצד השני? הבעיה שבמקום $x_n \uparrow x \Rightarrow \mu((-\infty, x_n]) \uparrow \mu((-\infty, x])$ נקבל שכאשר עולים ממש אז $x_n \uparrow x \Rightarrow \mu((-\infty, x_n]) \uparrow \mu((-\infty, x))$ כלומר הקטע הוא פתוח וזה לא יעבוד. אז יש רק רציפות מצד ימין. כלומר ייתכנו קפיצות. מה זה קפיצות? זה אטומים, כלומר יחידונים עם מידה חיובית ממש. במידת לבג כמובן לא הרשנו דברים כאלה.

דוגמאות.

$$1. \mu = \delta_x \text{ כאשר } \delta_x(B) = \begin{cases} 1, & x \in B \\ 0 & x \notin B \end{cases} \text{ זה נקרא אטום.}$$

$$2. \text{ אפשר לקחת גם מולקולות: } \mu = \sum_{k=1}^n c_k \delta_{x_k} \text{ ואז } \mu(B) = \sum_{k: x_k \in B} c_k$$

$$3. \text{ אפשר גם כמובן, } \mu = \sum_{k=1}^{\infty} c_k \delta_{x_k}$$

$$4. \text{ אין סיבה לקחת סכומים ולא לקחת אינטגרלים: } \mu(B) = \int_B c(x) dx = \int_{\mathbb{R}} c(x) \mathbb{1}_B(x) dx \\ \text{Leb}(B) = 0 \Rightarrow \mu(B) = 0. \text{ } c(x) \text{ מדידה אי שלילית ו-} \int_{\mathbb{R}} c(x) dx = 1$$

$$5. \mu, \nu \text{ מידות הסתברות ו-} a \geq 0, b \geq 0, a+b=1 \text{ אז גם } a\mu + b\nu \text{ מידת הסתברות.}$$

אינטגל לבג מוגדר למידה כללית ולא רק למידת לבג. אז איך זה יראה:

$$\int_{\mathbb{R}} f d\delta_a = f(a)$$

$$\int_{\mathbb{R}} f d(\sum_k c_k \delta_{x_k}) = \sum_k c_k f(x_k)$$

$$\int_{\mathbb{R}} f d\mu = \int f(x) c(x) dx$$

איך נחשב אינטגרל של מידת הסתברות סינגולרית? מקסימום נוכל לעשות:

$$\int f d\mu = \int f dF_\mu = \int f(x) dF_\mu(x)$$

דוגמא. $\Omega = (0, 1), P = Leb(0, 1), \omega = \sum_{n=1}^{\infty} \frac{\beta_n(\omega)}{2^n}$ ספרות בינאריות.

$$X(\omega) = \sum_{n=1}^{10} \frac{\beta_n(\omega)}{2^n}; X : \Omega \rightarrow \mathbb{R}$$

איך X מתפלג? $P_X = \sum_{k=0}^{1023} \frac{1}{1024} \delta_{\frac{k}{1024}}$ כלומר מידה אטומים עם 2^{10} אטומים וכל האטומים בעלי משקל שווה.

$$0, \beta_1 \beta_2 \beta_3 \dots$$

למעשה זה כמו להטיל מטבע 10 פעמים ולכתוב מספר עשרוני בהתאם. ברור שלכל מספר אותה הסתברות. התמונה ההפוכה של כל נקודה היא קטע. מתכנסת במ"ש. גם שאר הדוגמאות. מה לגבי

$$X'(\omega) = \sum_{n=1}^{10} \frac{\beta_{2n}(\omega)}{2^n}$$

הפונקציה היא לא מונוטונית יותר. אבל ההתפלגות המתאימה לא משתנה. אז X, X' הם שווי התפלגות. התמונה ההפוכה של כל נקודה הוא 1024 קטעים.

$$0, \beta_2 \beta_4 \beta_6 \dots$$

מה לגבי

$$X''(\omega) = \sum_{n=1}^{\infty} \frac{\beta_n(\omega)}{2^n}$$

עכשיו $P_{X''} = Leb(0, 1)$ כלומר $x(\omega) = \omega$ אז מה לגבי

$$X'''(\omega) = \sum_{n=1}^{\infty} \frac{\beta_{2n}(\omega)}{2^n}$$

זאת פונקציה נורא לא מונוטונית. אבל עדיין ההתפלגות נשארת אותו דבר. התמונה ההפוכה היא רצף של נקודות.

$$0, \beta_2 \beta_4 \beta_6 \dots$$

למה זאת בכלל פונקציה מדידה? כי $X_n : \Omega \rightarrow \mathbb{R}$ מדידות ו- $X_n \rightarrow X$ כ"ב אז X פונקציה מדידה. זה יודעים ממשיות. זה בסדר. איך יודעים שזה באמת מידת לבג? איך מוכיחים שזוג משתנים הם שווי התפלגות? הגדרנו:

$$\mathbb{P}(X \in (a, b)) = \mathbb{P}(X' \in (a, b)) \quad \forall a < b$$

או

$$P_x((a, b)) = P_{x'}((a, b))$$

כדי שקבוצה יוצרת תגדיר המשך מידה יחיד היא צריך להיות סגורה תחת חיתוך. כמו קטעים. ואז נשתמש בקרורורורו.
אז $B \in \mathcal{B}$ $P_X(B) = P_{X'}(B)$ וגם זאת הגדרה שקולה של ש"ה.
החדשות הרעות הן שאין נוסחא מפורשת להציג כל קבוצת בורל כקומבינציה של קטעים. החדשות הטובות הן שבכל זאת יש המשך יחיד.
מה עם קטעים חצי סגורים, סגורים? גם זה בסדר כי קטע סגור הוא גבול. להפך, קטע פתוח הוא גבול של סגורים.
מה עם לא כל קטעים אלה $[\frac{k}{2^n}, \frac{k+1}{2^n}]$? גם זה בסדר. נחבר ונקרב.
עכשיו נוכיח שזה שווה למידת לבג:

$$\mathbb{P}(X \in [\frac{k}{2^n}, \frac{k+1}{2^n}) = \frac{1}{2^n}$$

ועכשיו האינטואיציה הקומבינטורית נכונה כי שייכות לקטע נקבעת ע"י הספרות הראשונות. מה עם

$$X'''(\omega) = \sum_{n=1}^{\infty} \frac{\beta_n(\omega)}{2^{2n}}$$

עדיין פונקציה מדידה. כלומר אם נכתוב פיתוח עשרוני אז בין כל שתי ספרות שהגרלנו יש 0:

$$C = (0.0\beta_1 0\beta_2 0\beta_3 \dots)$$

P_x לא רציפה בהחלט לפי לבג. כל ההתפלגות מתרחשת בקבוצה זניחה לפי לבג. האם יש אטומים? לא. הסיכוי לקבל מה שרוצים הוא 0. לכן זאת מידה סינגולרית (לבג) בלי אטומים שמתרכזת בקבוצה זניחה. $\mathbb{P}(X \in A)$ מדידה לפי לבג. האם הביטוי בהכרח מוגדר? כן. $\mathbb{P}(X \in C) = 1, Leb(C) = 0$ $A \subseteq C$ מדידת לבג תמיד! כי C זניחה. לא. כי אחרת היה אפשר להרחיב את מידת לבג על כל הקבוצות.
ב- $\mathbb{R}$ לא נרחיב את הסיגמא אלגברה לבורל והקבוצות הזניחות כי אין לנו מוגש מה המידה שנבחר ומי תהיינה זניחות. אם אנחנו ב- $(\Omega, \mathcal{F}, P)$ אין בעיה לעשות את זה כי המידה קבועה.

X, Y ש"ה, $f: \mathbb{R} \rightarrow \mathbb{R}$ פונקצייה מדידה בורל. האם גם $f(X), f(Y)$ ש"ה?
לפי ההגדרה זה לא ברור: $\mathbb{P}(f(X) \in (a, b)) \stackrel{?}{=} \mathbb{P}(f(Y) \in (a, b))$
אם $f^{-1}((a, b)) = (c, d)$ אין בעיה, כי אז $\mathbb{P}(X \in (c, d)) = \mathbb{P}(Y \in (c, d))$. אבל זה לא הולך. א, אחד לא הבטיח $f^{-1}((a, b)) = (c, d)$. אבל יש הגדרה שקולה:

$$\mathbb{P}(f(x) \in B) \stackrel{?}{=} \mathbb{P}(f(Y) \in B)$$

$$f^{-1}(B) = C \in \mathcal{B} \text{ וכאן } B \in \mathcal{B}$$

ואז

$$\mathbb{P}(X \in C) = \mathbb{P}(Y \in C)$$

אותו דבר נכון גם $X, Y : \Omega \rightarrow \mathbb{R}$ מ"מ, $U, V : \tilde{\Omega} \rightarrow \mathbb{R}$ מ"מ
אפשר לכתוב $(X, Y) : \Omega \rightarrow \mathbb{R}^2$ מ"מ דו ממדי. האם מהעובדה ש- X, Y מדידים בורל
נובע שגם המ"מ הדו ממדי הוא מדיד בורל? כן.
 $X : \Omega \rightarrow \mathbb{R}, Y : \Omega \rightarrow \mathbb{R}, X^{-1}((a, b)) \in \mathcal{F}, Y^{-1}((a, b)) \in \mathcal{F}$
ועכשיו ברור:
 $(X, Y) : \Omega \rightarrow \mathbb{R}^2, (X, Y)^{-1}((a, b) \times (c, d)) \in \mathcal{F}$ כעת
גם $(U, V) : \tilde{\Omega} \rightarrow \mathbb{R}^2, (U, V) : \tilde{\Omega} \rightarrow \mathbb{R}^2$ ש"ה אז $f(X, Y), f(U, V)$ הם ש"ה. כל עוד
 $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ פונקציה בורל. כמובן גם ליותר מממדים.
מה לגבי אי תלות?
 $X, Y : \Omega \rightarrow \mathbb{R}$ מ"מ ב"ת אם $\mathbb{P}(X \in (a, b), Y \in (c, d)) = \mathbb{P}(X \in (a, b))\mathbb{P}(Y \in (c, d))$
 $\forall a, b, c, d$
בדו מימד:

$$\mathbb{P}_{X,Y}((a, b) \times (c, d)) = \mathbb{P}_X((a, b)) \cdot \mathbb{P}_Y((c, d))$$

וראינו שככה מגדירים מידה דו ממדית משתי מידות: $\mathbb{P}_{X,Y} = \mathbb{P}_X \times \mathbb{P}_Y$ וקוראדורי דו
ממדי מבטיח שיש רק אחת כזאת.
אם X, Y ב"ת ו- $f, g : \mathbb{R} \rightarrow \mathbb{R}$ פונקציות בורל. האם $f(X), g(Y)$ ב"ת? אינטואיציה
הסתברותית אומרת שכן (מה שאני חושב ומה שהוא חושב) ננסה להוכיח:

$$\mathbb{P}(f(X) \in (a, b), g(Y) \in (c, d))$$

ורוצים לכתוב את זה איכשהו $X \in (\alpha, \beta), Y \in (\gamma, \delta)$. לפונקציות ממונוטוניות זה נכון.
לפונקציות אחרות תמונה הפוכה היא קבוצת בורל. ככה לא יעבוד. אבל יש לנו מידת
מכפלה! $P_{X,Y}(A \times B) = P_X(A) \cdot P_Y(B)$ וזה לכל קבוצות לא רק לקטעים. זה מעבר
לא טריוויאלי מתורת המידה. ועכשיו:

$$\begin{aligned} \mathbb{P}(f(X) \in A, g(Y) \in B) &= \mathbb{P}(x \in f^{-1}(A), y \in f^{-1}(B)) = \\ &= \mathbb{P}(x \in f^{-1}(A)) \cdot \mathbb{P}(y \in f^{-1}(B)) = \mathbb{P}(f(X) \in A) \cdot \mathbb{P}(g(Y) \in B) \end{aligned}$$

ואפשר באופן דומה להרחיב ל- n משתנים ב"ת.
עוד לא דיברנו על תוחלת.

$$\mathbb{E}X = \int_{\Omega} X dP = \int_0^{\infty} \mathbb{P}(X > x) dx - \int_0^{\infty} \mathbb{P}(X < -x) dx$$

זה נכון כש X אינטגרבילית. אבל אם לא אינטגרלית אז אם זה אינסוף מינוס מספר אז
התוחלת אינסוף. מספר מינוס אינסוף אז התוחלת מינוס אינסוף. עבור אינסוף מינוס
אינסוף התוחלת לא מוגדרת.
מה שחשוב שאם X, Y ש"ה אז $\mathbb{E}X = \mathbb{E}Y$.
אם $X : \Omega \rightarrow \mathbb{R}$ מ"מ אפשר להתבונן $X_*(\mathbb{R}, \mathcal{B}, P_X) \rightarrow \mathbb{R}$ כך ש- $X_*(x) = x$. כלומר
הופכים את R למרחב הסתברות. גם זה משתנה מקרי. וזה מצחיק. עכשיו נוכל לטעון ש-
 X ו- X^* הם ש"ה. בפרט, $\mathbb{E}X = \mathbb{E}X^*$. $\int_{\Omega} X dP = \int_{\mathbb{R}} x P_X(dx)$. וזאת הגדרה שקולה
של תוחלת. במונחי התפלגות. אפילו בלי מרחב הסתברות.