האוניברסיטה העברית בירושלים
המחלקה לסטטיסטיקה
הסתברות ותהליכים מקריים
מורה הקורס: עופר קלע

פונקציה $f:\mathbb{R}\to\mathbb{R}$ נקראת קמורה אם ורק אם לכל $0<\lambda<1$ ולכל $x,y\in\mathbb{R}$ מתקיים כי

$$f(\lambda x+(1-\lambda)y)\le\lambda f(x)+(1-\lambda)f(y)$$

טענה:
f קמורה אם ורק אם

$$g(x,y)=\frac{f(y)-f(x)}{y-x}$$

היא פונקציה לא יורדת ב-x,y על $\{(x,y)|x\neq y\}$.
הוכחה:
תחילה נשים לב כי $g(x,y)=g(y,x)$ לכל $x\neq y$ ולכן מספיק לבדוק מקרים בהם $x<y$.
נסתכל על $x<z<y$. אז מתקיים כי

$$\frac{f(z)-f(x)}{z-x}\le\frac{f(y)-f(x)}{y-x}$$

אם ורק אם

$$f(z)\le\frac{y-z}{y-x}f(x)+\frac{z-x}{y-x}f(z)$$

ומתקיים גם כי

$$z=\frac{y-z}{y-x}\,x+\frac{z-x}{y-x}\,y$$

מכיוון שלכל $x<z<y$ ניתן לבחור $\lambda=\frac{y-z}{y-x}$ ולכל $0<\lambda<1$ ולכל $x<y$ ניתן לבחור $z=\lambda x+(1-\lambda)y$ נובע כי

$$f(\lambda x+(1-\lambda)y)\le\lambda f(x)+(1-\lambda)f(y)$$

לכל $0<\lambda<1$ אם ורק אם לכל $x<z<y$ מתקיים כי

$$\frac{f(z)-f(x)}{z-x}\le\frac{f(y)-f(x)}{y-x}$$

באופן דומה, זה גם שקול ל-

$$\frac{f(y)-f(x)}{y-x}\le\frac{f(y)-f(z)}{y-z}$$

נניח כי $a<b,c<d$

$$\frac{f(b)-f(a)}{b-a}\le\frac{f(d)-f(a)}{d-a}\le\frac{f(d)-f(c)}{d-c}$$

ולכן

$$g(a,b) \leq g(c,d)$$

מכיוון ש-$g(x,y) = g(y,x)$ נובע כי אי השוויון האחרון מתקיים לכל בחירה של a,b,c,d עם $a \leq c$ ו- $b \leq d$ עם $a \neq b$ וגם $c \neq d$.

מסקנה

לכל $x \in \mathbb{R}$ קיימים הגבולות

$$\begin{aligned} D^+ f(x) &= \lim_{h \downarrow 0} \frac{f(x+h) - f(x)}{h} \\ D^- f(x) &= \lim_{h \downarrow 0} \frac{f(x-h) - f(x)}{-h} \end{aligned}$$

גבולות אילה סופיים ומתקיים כי

$$D^- f(x) \leq D^+ f(x)$$

כמו כן מתקיים לכל x,y ולכל $D^- f(x) \leq a \leq D^+ f(x)$ כי

$$f(y) \geq f(x) + a(y-x)$$

הוכחה: מכיוון ש-$x - 1 \leq x$ וכן $x \leq x + h$ נובע כי

$$\frac{f(x+h) - f(x)}{h} = \frac{f(x+h) - f(x)}{(x+h) - x} \geq \frac{f(x) - f(x-1)}{x - (x-1)} = f(x) - f(x-1) > -\infty$$

הפונצקיה מצד ימין לא עולה ב-h ולכן קיים הגבול כאשר $h \downarrow 0$ והוא סופי. עכשיו, אם ניקח $x < y$ ו-$0 < h < y - x$ נקבל כי

$$\frac{f(x+h) - f(x)}{h} = \frac{f(x+h) - f(x)}{(x+h) - x} \leq \frac{f(y) - f(x)}{y - x}$$

ואם נשאיף $h \downarrow 0$ נקבל כי

$$D^+ f(x) \leq \frac{f(y) - f(x)}{y - x}$$

באופן דומה נובע כי

$$\frac{f(x-h) - f(x)}{-h} = \frac{f(x-h) - h(x)}{(x-h) - x} = \frac{f(x) - f(x-h)}{h} \leq \frac{f(x+1) - f(x)}{(x+1) - x} = f(x+1) - f(x) < \infty$$

לכל $h > 0$ ומכאן שגם בגבול. בפרט, מכיוון ש-

$$\frac{f(x-h) - f(x)}{-h} = \frac{f(x) - f(x-h)}{x - (x-h)} \leq \frac{f(x+h) - f(x)}{(x+h) - x} = \frac{f(x+h) - f(x)}{h}$$

נובע כי

$$D^- f(x) \leq D^+ f(x)$$

ולכן קבלנו כי לכל $x < y$ מתקיים כי

$$D^- f(x) \leq D^+ f(x) \leq \frac{f(y) - f(x)}{y - x}$$

וזה שקול ל-

$$f(y) \geq f(x) + D^- f(x)(y - x)$$

וגם ל-

$$f(y) \geq f(x) + D^+ f(x)(y - x)$$

כאשר $y < x$ אז

$$\frac{f(y) - f(x)}{y - x} = \frac{f(x) - f(y)}{x - y} \leq \frac{f(x) - f(x-h)}{x - (x-h)} = \frac{f(x-h) - f(x)}{-h}$$

עבור $0 < h < x - y$ ומכאן שבגבול נקבל כי

$$\frac{f(y) - f(x)}{y - x} \leq D^- f(x) \leq D^+ f(x)$$

ולכן מתקיים כי

$$f(y) \geq f(x) + D^- f(x)(y - x)$$

וגם

$$f(y) \geq f(x) + D^+ f(x)(y - x)$$

קבלנו איפה כי לכל $x \neq y$

$$f(y) \geq f(x) + D^- f(x)(y - x)$$

וגם

$$f(y) \geq f(x) + D^+ f(x)(y - x)$$

עבור $x = y$ אי השוויון הוא שוויון ולכן שני אי השוויונים הללו נכונים לכל x, y לבסוף אם נכפיל את אי השוויון הראשון ב-λ ואת השני ב-$1 - \lambda$ נקבל כי לכל x, y מתקיים כי

$$f(y) \geq f(x) + (\lambda D^- f(x) + (1 - \lambda) D^+ f(x))(y - x)$$

מכיוון שלכל $D^- f(x) \leq a \leq D^+ f(x)$ קיים $0 \leq \lambda \leq 1$ כך ש-

$$a = \lambda D^- f(x) + (1 - \lambda) D^+ f(x)$$

אז סיימנו.

מסקנה:
אם $f:\mathbb{R}\to\mathbb{R}$ קמורה אז היא גם רציפה.
הוכחה: אם f לא רציפה מימין אז זה סותר את קיום הגבול

$$\frac{f(x+h)-f(x)}{h}$$

כאשר $h\downarrow 0$. באותו אופן אם היא לא רציפה משמאל.
נניח עכשיו כי f היא פונקציה קמורה וכי X הוא משתנה מקרי עם $E|X|<\infty$. אז $EX\in\mathbb{R}$ וניקח

$$D^{-}f(EX)\le a\le D^{+}f(EX)$$

נקבל כי לכל $\omega\in\mathbb{R}$ עבורו $X(\omega)$ הוא מספר סופי מתקיים כי

$$f(X(\omega))\ge f(EX)+a(X(\omega)-EX)$$

מכאן נובע כי בהסתברות אחת מתקיים כי

$$f(X)\ge f(EX)+a(X-EX)$$

ולכן

$$f^{-}(X)\le (f(EX)+a(X-EX))^{-}\le |f(EX)|+|a|(|X|+|EX|)$$

ומכיוון ש-$E|X|<\infty$ נובע כי

$$Ef^{-}(X)<\infty$$

ולכן התוחלת של המשתנה המקרי $f(X)$ מוגדרת היטב. לכן מתקיים כי

$$Ef(X)\ge E\left(f(EX)+a(X-EX)\right)=f(EX)+a(EX-EX)=f(EX)$$

עכשיו, את כל מה שעשינו לעיל אפשר גם לעשות אם $f:[a,b]$ היא פונקציה קמורה ו-a או b (או שניהם) סופיים. התוצאה היא כי f רציפה על (a,b) (אך לא בהכרח ב-a או b) והנגזרות משמאל ומימין קיימות על (a,b) (אך לא בהכרח ב-a או b). בפרט נובע מכך כי f היא פונקצית בורל ולכן אם X הוא משתנה מקרי אז גם $f(X)$ הוא משתנה מקרי. עכשיו, כאשר a או b (או שניהם) סופיים, נצטרך לדרוש בפיתוח האחרון כי $EX\in(a,b)$. אך, מה היה קורה אם למשלa היה סופי ו-$EX=a$? מכיוון ש-$P(a\le X\le b)$ היה נובע מכך כי $P(X=a)=1$ ואז היינו מקבלים כי

$$f(EX)=f(a)=Ef(X)$$

ובאותו אופן עבור b.
מה לגבי תוצא דומה עבור תוחלת מותנה?
ובכן, במקרה זה היינו רושמים

$$f(X)\ge f(E[X|\mathcal{F}_0])+D^{+}f(E[X|\mathcal{F}_0])(X-E[X|\mathcal{F}_0])$$

כאשר $E[X|\mathcal{F}_0] \in (a,b)$ ולכן

$$E[f(X)|\mathcal{F}_0]1_{(a,b)}(E[X|\mathcal{F}_0]) \geq f(E[X|\mathcal{F}_0])1_{(a,b)}(E[X|\mathcal{F}_0])$$

כמו כן מכיוון ש-

$$(f(X) - f(a))1_{\{a\}}(E[X|\mathcal{F}_0]) \geq 0$$

בהסתברות אחת אז גם

$$E[(f(X) - f(a))1_{\{a\}}(E[X|\mathcal{F}_0])|\mathcal{F}_0] \geq 0$$

בהסתברות אחת. אך זה שקול ל-

$$E[f(X)|\mathcal{F}_0]1_{\{a\}}(E[X|\mathcal{F}_0]) \geq f(a))1_{\{a\}}(E[X|\mathcal{F}_0]) = f(E[X|\mathcal{F}_0]))1_{\{a\}}(E[X|\mathcal{F}_0])$$

ובאותו אופן עבור b. לכן מתקבל כי בהסתברות אחת

$$f(E[X|\mathcal{F}_0]) \leq E[f(X)|\mathcal{F}_0]$$

ברור כי כאשר $a = b$ (סופי) אז $X = a$ ו-$f(X) = f(a)$ בהסתברות אחת וכן $f(EX) = f(a)$.
לסיכום:

אי שוויון Jensen (מבטאים: יינסן ולא ג'נסן):
נניח כי X הוא משתנה מקרי עם $E|X| < \infty$ ונניח כי $P(a \leq X \leq b) = 1$ עבור $-\infty \leq a \leq b \leq \infty$. נניח כי $f : [a,b] \to \mathbb{R}$ קמורה. אז $Ef^-(X) < \infty$ ולכן $Ef(X)$ מוגדרת ומתקיים כי

$$f(EX) \leq Ef(X)$$

כמו כן לכל סיגמה שדה $\mathcal{F}_0 \subset \mathcal{F}$ מתקיים כי

$$f([EX|\mathcal{F}_0]) \leq E[f(X)|\mathcal{F}_0]$$

בהסתברות אחת.

הערה:
מתי מתקיים שוויון באי שוויון יינסן? נשים לב כי

$$f(X) - f(EX) - a(X - EX) \geq 0$$

בהסתברות אחת. שוויון באי שוויון יינסן מתקיים אם ורק אם התוחלת היא אפס ומכאן נובע כי בהסתברות אחת מתקיים כי

$$f(X) = aX + (f(EX) - aEX)$$

דהיינו, קיימים α, β כך שבהסתברות אחת

$$f(X) = \alpha X + \beta$$

האם זה בהכרח אומר ש-$Ef(X) = f(EX)$? לשם כך נסתכל על $g(x) = f(x) - \alpha x - \beta$. גם זו פונצקיה קמורה המקיימת בהסתברות אחת כי $g(X) = 0$ ולכן $Eg(X) = 0$. האם בכרח $g(EX) = 0$? התשובה היא שלא. למשל, אם $X \sim \text{Bern}(0.5)$ אז לכל פונצקיה f מתקיים כי

$$f(X) = f(0)(1-X) + f(1)X = (f(1) - f(0))X + f(0)$$

אך $f(0.5)$ יכול לקבל ערך שאינו בהכרח $\frac{f(0)+f(1)}{2}$.

בכיוון השני, אי שוויון יינסן יהיה אי שוויון חזק אם ורק אם X אינו קבוע בהסתברות אחת ומתקיים כי f קמורה ממש על $[a,b]$ במובן שלכל $x, y \in [a,b]$ ולכל $0 < \lambda < 1$ יתקיים כי

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)y$$

לדוגמה, e^x קמורה ממש על $\mathbb{R}$ ו-$-\log x$ קמורה ממש על $(0,\infty)$. לעומת זאת $|x|$, x^+ או x^- קמורות אך אינה קמורות ממש על $\mathbb{R}$.

לבסוף, די פשוט להראות כי אם $f : [a,b] \to \mathbb{R}$ אז $D^+ f(x)$ רציפה מימין ולא יורדת על (a,b) ואילו $D^- f(x)$ רציפה משמאל ולא יורדת על (a,b). בפרט מתקיים כי

$$\begin{aligned} D^+ f(x) &= D^- f(x+) \\ D^- f(x) &= D^+ f(x-) \end{aligned}$$

אוסף נקודות אי הרציפות של $D^+ f(x)$ ו-$D^- f(x)$ מתלכדות והוא לכל היותר בן מניה. בכל נקודות הרציפות של כל אחת מהן מתקיים כי $D^+ f(x) = D^- f(x)$. לבסוף, מכיוון ש-$D^\pm f(x)$ רציפה כמעט בכל מקום ביחס למידת לבג על $[a,b]$ נובע כי זו פונצקיה אינטגרבילית רימן על (a,b) ולא קשה לבדוק כי לכל $a < x < y < b$ מתקיים כי

$$f(y) - f(x) = \int_x^y D^+ f(s) ds = \int_x^y D^- f(x) dx$$

למעשה זה גם אומר שכל פונצקיה g המקיימת $D^- f(x) \le g(x) \le D^+ f(x)$ לכל $x \in (a,b)$ מתקיים כי g אינטגרבילית רימן על (a,b) ומתקיים כי לכל $x < y$

$$f(y) - f(x) = \int_x^y g(s) ds$$

ולכן כל פונצקיה כזאת היא צפיפות של f (פונצקיה קמורה, לא בהכרח פונצקית התפלגות ובפרט יכולה גם לקבל ערכים שליליים).

דבר אחרון, f נקראת קעורה אם $-f$ קמורה ולכן במקרה זה כל אי השוויונים (גם זה שבאי שוויון יינסן) מתהפכים. דוגמה: נניח כי $a_1, \ldots, a_n$ הם מספרים חיוביים וסופיים. אז אם נניח כי I מקבל את הערכים $1, \ldots, n$ בהסתברויות $1/n$ ו-$X = a_I$ אז

$$\log\left(\prod_{i=1}^n a_i\right)^{1/n} = \frac{1}{n}\sum_{i=1}^n \log a_i = E\log X \le \log EX \le \log\left(\frac{1}{n}\sum_{i=1}^n a_i\right)$$

ולכן

$$\left(\prod_{i=1}^n a_i\right)^{1/n} \le \frac{1}{n}\sum_{i=1}^n a_i$$

צד ימין נקרא הממוצע ההנדסי ואילו צד ימין הוא הממוצע החשבוני. מכיוון שמכך גם נובע כי

$$\frac{1}{\left(\prod_{i=1}^{n} a_i\right)^{1/n}} = \prod_{i=1}^{n} (1/a_i)^{1/n} \leq \frac{1}{n}\sum_{i=1}^{n}(1/a_i)$$

אז נובע גם כי

$$\frac{n}{\sum_{i=1}^{n} \frac{1}{a_i}} \leq \left(\prod_{i=1}^{n} a_i\right)^{1/n} \leq \frac{1}{n}\sum_{i=1}^{n} a_i$$

כאשר צד שמאל נקרא הממוצע ההרמוני. אי השוויון הזה מתקיים לכל אוסף של n ערכים חיוביים ומתקיים שוויון אם ורק אם $a_1 = a_2 = \ldots = a_n$. בדרך כלל מראים זאת באמצעות אינדוקציה. בעזרת אי שוויון יינסן התוצאה מידית.