

➤ **Q:** What should a cdf look like?

- Theorem. If F_X is the cdf of a r.v. X , then it must satisfy the following properties:

(1) $0 \leq F_X(x) \leq 1$.

proof. $0 \leq F_X(x) = P(\{\omega \in \Omega : X(\omega) \in (-\infty, x]\}) \leq 1$.
Handwritten notes: $P_X((-\infty, x])$, a.p.m.

(2) $F_X(x)$ is nondecreasing, i.e., $F_X(a) \leq F_X(b)$ for $a < b$.

proof. For $a < b$, $(-\infty, a] \subset (-\infty, b]$,

$$F_X(a) = P_X((-\infty, a]) \leq P_X((-\infty, b]) = F_X(b).$$

Handwritten notes: by 2nd proposition in LNp.3-10

(3) For any $x \in \mathbb{R}$, $F_X(x)$ is continuous from the right, i.e.,

$$F_X(x) = F_X(x+) \equiv \lim_{t \downarrow x} F_X(t),$$

proof. Let x_n be a sequence s.t. $x_n \downarrow x$.

Let $E_n = (-\infty, x_n]$. Then, $E_n \downarrow (-\infty, x]$.

$$F_X(x) = P_X((-\infty, x]) = P_X\left(\lim_{n \rightarrow \infty} E_n\right)$$

Handwritten notes: $\bigcap_{n=1}^{\infty} E_n$ ($\because E_n$ decreasing)

Recall: $E_n \uparrow E$ or $E_n \downarrow E$
 $P(E) = P(\lim_{n \rightarrow \infty} E_n) = \lim_{n \rightarrow \infty} P(E_n)$
 (LNp.3-17)

$$\begin{aligned} &= \lim_{n \rightarrow \infty} P_X(E_n) = \lim_{n \rightarrow \infty} P_X((-\infty, x_n]) \\ &= \lim_{n \rightarrow \infty} F_X(x_n) \end{aligned}$$

Handwritten notes: $x_n \uparrow x, (-\infty, x_n] \uparrow (-\infty, x], P_X((-\infty, x]) = F_X(x-)$

(4) $\lim_{x \rightarrow \infty} F_X(x) = 1$ and $\lim_{x \rightarrow -\infty} F_X(x) = 0$,

proof. Let $x_n \downarrow -\infty$. Then, $E_n \equiv (-\infty, x_n] \downarrow \emptyset$.

$$\lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} P_X((-\infty, x_n])$$

$$= P_X\left(\lim_{n \rightarrow \infty} E_n\right) = P_X(\emptyset) = 0.$$

Similarly, if $x_n \uparrow \infty$, then $E_n \equiv (-\infty, x_n] \uparrow \mathbb{R}$, and

$$\lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} P_X((-\infty, x_n])$$

$$= P_X\left(\lim_{n \rightarrow \infty} E_n\right) = P_X(\mathbb{R}) = 1.$$

Recall. If $A \subset B$,
 $P(B \setminus A) = P(B) - P(A)$
 (LNp.3-10)

Handwritten notes: Why useful? Check () in LNp.5-8*

(5) $P_X(X > x) = 1 - F_X(x)$ and $P_X(a < X \leq b) = F_X(b) - F_X(a)$.

proof. $P_X(X > x) = 1 - P_X(\{X > x\}^c)$
 $= 1 - P_X(X \leq x) = 1 - F_X(x)$.

For $a < b$, $(-\infty, a] \subset (-\infty, b]$, and

$$\begin{aligned} P_X(a < X \leq b) &= P_X((-\infty, b] \setminus (-\infty, a]) \\ &= P_X((-\infty, b]) - P_X((-\infty, a]) = F_X(b) - F_X(a). \end{aligned}$$

transformation
pmf $\leftrightarrow$ cdf

(6) Moreover, if X is discrete with pmf f_X , then for $x \in \mathbb{R}$,

$$F_X(x) = \sum_{\substack{x_i \in \mathcal{X} \\ x_i \leq x}} f_X(x_i), \text{ and } f_X(x) = F_X(x) - F_X(x-).$$

proof.

$$F_X(x) = P_X(X \in (-\infty, x]) = \sum_{x_i \in (-\infty, x] \cap \mathcal{X}} f_X(x_i).$$

For $x_n \uparrow x$, $(-\infty, x_n] \uparrow (-\infty, x)$, and

$$F_X(x-) = \lim_{n \rightarrow \infty} F_X(x_n) = P_X((-\infty, x)).$$

$$\begin{aligned} \text{So, } f_X(x) &= P_X(\{x\}) = P_X((-\infty, x] \setminus (-\infty, x)) \\ &= P_X((-\infty, x]) - P_X((-\infty, x)) = F_X(x) - F_X(x-). \end{aligned}$$

(7) F_X has at most countably many discontinuity points.

proof. Let $\mathbb{D}$ be the collection of discontinuity points.

For $x \in \mathbb{D}$, let $T_x = (F_X(x-), F_X(x))$.

Because $F_X(x-) \neq F_X(x) \Rightarrow F_X(x-) < F_X(x)$
 $\exists$ a rational number, denoted by r_x , in T_x .

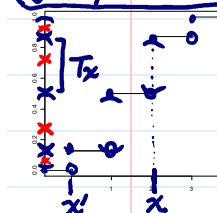
Because the set of rational numbers is a countable set,
 $\mathbb{D}$ is either finite or countably infinite.

Theorem. If a function F satisfies (2), (3), and (4), then F is a cumulative distribution function of some random variable.

proof. Skip. Out of the scope of the course.

discrete, continuous, or mixed

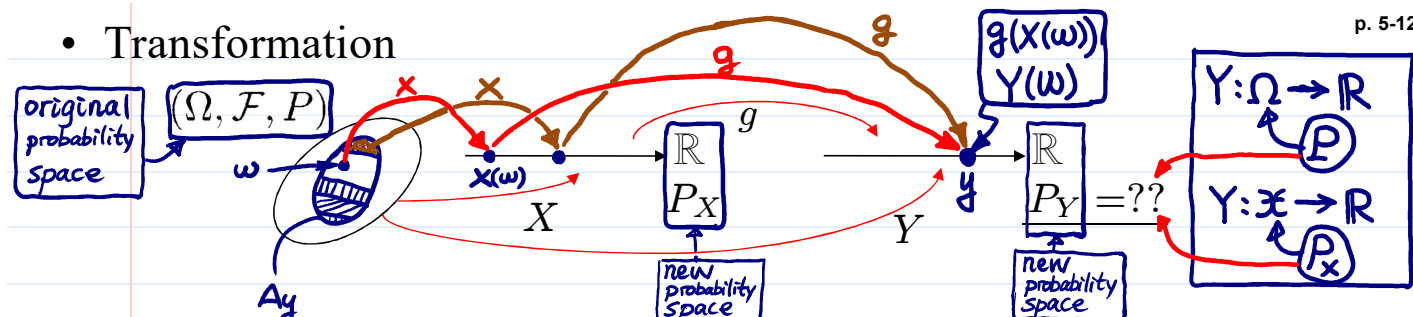
Q. F_X has at most how many jumps?



Thm in LNp.5-17 for pmf

(*) in LNp.5-8

Transformation



Theorem. Let X be a discrete r.v. with range $\mathcal{X}$ and pmf f_X ; let

$$Y = g(X) \quad (\omega)$$

then, the range of Y is $\mathcal{Y} = \{Y(\omega) : \omega \in \Omega\}$

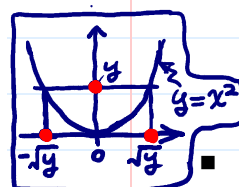
a countably many set? $\mathcal{Y} = \{g(x) : x \in \mathcal{X}\}$, $f_Y(y) = 0$ if $y \notin \mathcal{Y}$.

i.e., Y is a discrete r.v., and the pmf of Y is

$$f_Y(y) = \sum_{\substack{x \in \mathcal{X} \\ g(x) = y}} f_X(x).$$

proof. Since $\{\omega \in \Omega : Y(\omega) = y\} = \bigcup \{\omega \in \Omega : X(\omega) = x\}$,

$$f_Y(y) = \sum_{\substack{x \in \mathcal{X} \\ g(x) = y}} P(\{\omega \in \Omega : X(\omega) = x\}) = \sum_{\substack{x \in \mathcal{X} \\ g(x) = y}} f_X(x)$$



Example. If $Y = X^2$, then $f_Y(y) = f_X(\sqrt{y}) + f_X(-\sqrt{y})$.

Expectation (Mean) and Variance

- Q: We often characterize a person by his/her height, weight, hair color, How can we “roughly” characterize a distribution?

Definition: If X is a discrete r.v. with pmf f_X and range $\mathcal{X}$, then the expectation (or called expected value) of X is

期望值

加權平均

$$E(X) = \sum_{x \in \mathcal{X}} x f_X(x), \quad \sum_{x \in \mathcal{X}} f_X(x) = 1$$

$$\mathcal{X} = \{x_1, \dots, x_n\}$$

$$\uparrow \quad \quad \quad \uparrow$$

$$f_1 \quad \dots \quad f_n$$

$$E(X) = \sum_{i=1}^n x_i \cdot \frac{1}{n} = \frac{x_1 + \dots + x_n}{n}$$

provided that the sum converges absolutely. i.e. $\sum_{x \in \mathcal{X}} |x| f_X(x) < \infty$

X: random value
cf.
E(X): fixed value
cf.
distribution of X: fixed

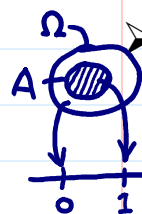
Example. If all value in $\mathcal{X}$ are equally likely, then $E(X)$ is simply the average of the possible values of X .

$\mathcal{X}$ is a finite set

Example (Committees, LNP.5-6). In the committees example,

$$E(X) = 0 \cdot \frac{5}{210} + 1 \cdot \frac{50}{210} + 2 \cdot \frac{100}{210} + 3 \cdot \frac{50}{210} + 4 \cdot \frac{5}{210} = \frac{2}{1}$$

on average, 2 women in the committees



Example (Indicator Function). a r.v.

For an event $A \subset \Omega$, the indicator function of A is the

$$1_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A, \\ 0, & \text{if } \omega \notin A. \end{cases} \quad \begin{aligned} \{1_A = 1\} &= \{A \text{ occurs}\} \\ \{1_A = 0\} &= \{A \text{ not occur}\} \end{aligned}$$

Its range $\mathcal{X}$ is $\{0, 1\}$ and its pmf is $f(x) = 0$, if $x \notin \mathcal{X}$

$$f(0) = P(A^c) = 1 - P(A) \quad \text{and} \quad f(1) = P(A),$$

for a p.m. P defined on Ω .

Note: Expectation (expected value) may not be a value that the r.v. can generate

So, $E(1_A) = 0 \cdot [1 - P(A)] + 1 \cdot P(A) = P(A)$.

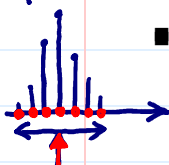
Intuitive Interpretation of Expectation

加權平均 ($\because \sum_{x \in \mathcal{X}} f_X(x) = 1$)

of all possible values (outcomes) of the r.v.

Expectation of a r.v. parallels the notion of a weighted average, where more likely values are weighted higher than less likely values.

pmf



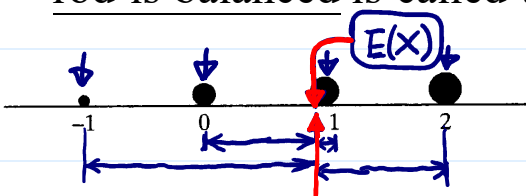
It is helpful to think of the expectation as the “center” of mass of the pmf.

$$0 = \sum_{x \in \mathcal{X}} x f_X(x) - E(X) \sum_{x \in \mathcal{X}} f_X(x) = \sum_{x \in \mathcal{X}} (x - E(X)) f_X(x) = 0$$

重心

center of gravity: If we have a rod with weights $f_X(x_i)$ at each possible points x_i 's then the point at which the rod is balanced is called the center of gravity.

槓桿原理:
力矩
= 距離 $\times$
作用力



$$p(-1) = .10, \quad p(0) = .25, \quad p(1) = .30, \quad p(2) = .35$$

$\wedge$ = center of gravity = .9

$$(-1) \times 0.1 + 0 \times 0.25 + 1 \times 0.3 + 2 \times 0.35 = 0.9$$

Recall.
Objective
probability
in LNp.3-19

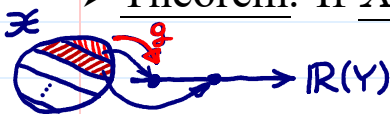
- Expectation can be interpreted as a long-run average ($\because$ Law of Large Number, Chapter 8)

e.g. repeat 10000 times,

0	2	2	-1	0	1	...	1
1st	2nd	3rd	4th	5th	6th		10000th
							their average ≈ 0.9
							deterministic

- Expectation of Transformation LNp.5-12

➤ Theorem. If X is a discrete r.v. with range $\mathcal{X}$ and pmf f_X ; let



$Y = g(X)$, discrete r.v.

and $\mathcal{Y}$ be the range of Y , f_Y be the pmf of Y , then

$$E(Y) \equiv \sum_{y \in \mathcal{Y}} y f_Y(y) \stackrel{\text{cf.}}{=} \sum_{x \in \mathcal{X}} g(x) f_X(x),$$

provided that the sum converges absolutely.

To calculate $E(Y)$, not necessary to first obtain $f_Y(y)$ or to know the distribution of Y .

can be derived from f_X by the Thm in LNp.5-12

$$\sum_{y \in \mathcal{Y}} |y| f_Y(y) < \infty$$

proof. $\sum_{x \in \mathcal{X}} g(x) f_X(x) \stackrel{\text{cf.}}{=} \sum_{y \in \mathcal{Y}} \left\{ \sum_{\substack{x \in \mathcal{X} \\ g(x)=y}} g(x) f_X(x) \right\}$

by Thm in LNp.5-12

$\mathcal{X} = \bigcup_{y \in \mathcal{Y}} \{x \in \mathcal{X} \mid g(x)=y\}$
mutually exclusive & countably many
countably many

$$= \sum_{y \in \mathcal{Y}} y \sum_{\substack{x \in \mathcal{X} \\ g(x)=y}} f_X(x) = \sum_{y \in \mathcal{Y}} y f_Y(y)$$

- Example. $Y = X^2$, $E(Y) \stackrel{\text{cf.}}{=} \sum_{x \in \mathcal{X}} x^2 f_X(x) \equiv E(X^2)$.

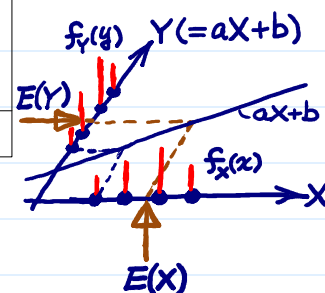
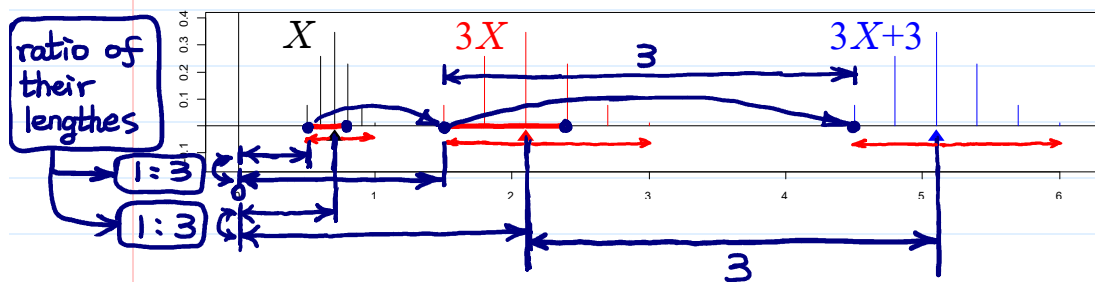
➤ Theorem. For $a, b \in \mathbb{R}$, $E(aX+b) = a \cdot E(X) + b$.

proof.

fixed constants

a transformation of X : $Y = aX+b$

$$E(Y) = E(aX+b) = \sum_{x \in \mathcal{X}} (ax+b) f_X(x) = a \left[\sum_{x \in \mathcal{X}} x f_X(x) \right] + b \left[\sum_{x \in \mathcal{X}} f_X(x) \right]$$



- Mean and Variance. 變異數

平均數

➤ Definition. The expectation of X is also called the mean of X and/or f_X . The variance of X (and/or f_X) is defined by

$$Var(X) \equiv E[(X - \mu_X)^2] = \sum_{x \in \mathcal{X}} (x - \mu_X)^2 f_X(x)$$

provided that the sum converges.

a transformation of X , $Y = (X - \mu_X)^2$

$$E(Y) = Var(X)$$

check graph in LNp.5-14

Why?

The $E(X)$ is often denoted by μ_X and $Var(X)$ by σ_X^2 . Also,

having same unit as X

$$\sigma_X = \sqrt{\sigma_X^2}$$

a fixed constant

is called the standard deviation of X .

標準差