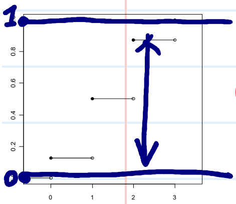
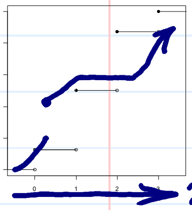
➢ **Q**: What should a cdf look like?

- Theorem. If F_X is the cdf of a r.v. X, then it must satisfy the following properties: (If not specified, the property holds for any r.v.'s (discrete, continuous, mixed))



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

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

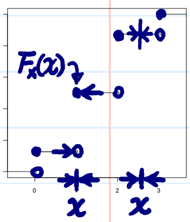


(2) $F_X(x)$ is nondecreasing, i.e., $F_X(a) \le F_X(b)$ for $a<b$.

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

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

(by 2nd proposition in LNp.3-10; $\because F_X(x) \equiv P_X((-\infty,x])$)



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

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

(cf. $F_X(x+) \leftrightarrow F_X(x-)$; for any such sequence)

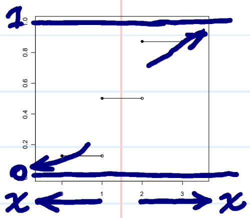
proof. Let x_n be a sequence s.t. $x_n \downarrow x$. (i.e. $x_1 \ge x_2 \ge x_3 \ge \cdots \ge x$ and $\lim_{n\to\infty} x_n = x$.)

Let $E_n = (-\infty, x_n]$. Then, $E_n \downarrow (-\infty, x]$. ($\bigcap_{n=1}^{\infty} E_n$ ($\because$ E_n decreasing))

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

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

($P_X((-\infty,x)) = F_X(x-)$; $x_n \uparrow x$, $(-\infty, x_n] \uparrow (-\infty, x)$)



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

proof. Let $x_n \downarrow -\infty$. Then, $E_n \equiv (-\infty, x_n] \downarrow \emptyset$. ($\bigcap_{n=1}^{\infty} E_n$)

$$\begin{aligned} \lim_{n\to\infty} F_X(x_n) &= \lim_{n\to\infty} P_X((-\infty, x_n]) \\ &= P_X\left(\lim_{n\to\infty} E_n\right) = P_X(\emptyset) = 0. \end{aligned}$$

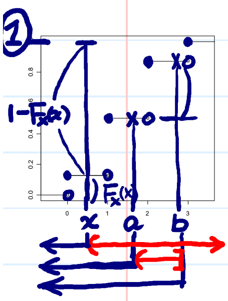
(for any such sequence)

Similarly, if $x_n \uparrow \infty$, then $E_n \equiv (-\infty, x_n] \uparrow \mathbb{R}$, and ($\bigcup_{n=1}^{\infty} E_n$)

$$\begin{aligned} \lim_{n\to\infty} F_X(x_n) &= \lim_{n\to\infty} P_X((-\infty, x_n]) \\ &= P_X\left(\lim_{n\to\infty} E_n\right) = P_X(\mathbb{R}) = 1. \end{aligned}$$

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

Why useful? Check (*) in LNp.5-8 (if $b \to \infty$)



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

proof.

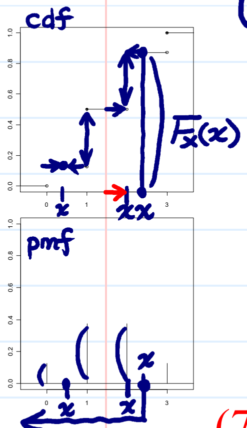
$$\begin{aligned} P_X(X > x) &= 1 - P_X(\{X > x\}^c) \\ &= 1 - P_X(X \le x) = 1 - F_X(x). \end{aligned}$$

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

$$P_X(a < X \le b) = P_X((-\infty, b] \setminus (-\infty, a])$$

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

transformation pmf↔cdf



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

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

proof. $F_X(x) = P_X(X \in (-\infty, x]) = \sum_{x_i \in (-\infty, x] \cap \mathcal{X}} f_X(x_i).$ (by (iv) in LNp.5-6)

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

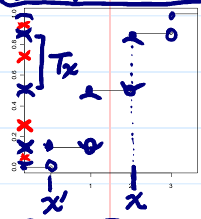
$$F_X(x-) = \lim_{n \to \infty} F_X(x_n) = P_X((-\infty, x)).$$

So, $f_X(x) = P_X(\{x\}) = P_X((-\infty, x] \setminus (-\infty, x))$

$$= P_X((-\infty, x]) - P_X((-\infty, x)) = F_X(x) - F_X(x-).$$

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

Q. F_X has at most how many jumps?



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 $F_X(x-) \neq F_X(x)$, ⇒ $F_X(x-) < F_X(x)$

$\exists$ a rational number, denoted by r_x, in T_x.

($x \in \mathbb{D}$ ← a one-to-one map → $r_x \in \mathbb{Q}$ ($\because$ if $x \neq x'$, then $r_x \neq r_{x'}$))

Because the set of rational numbers is a countable set,

$\mathbb{D}$ is either finite or countably infinite.

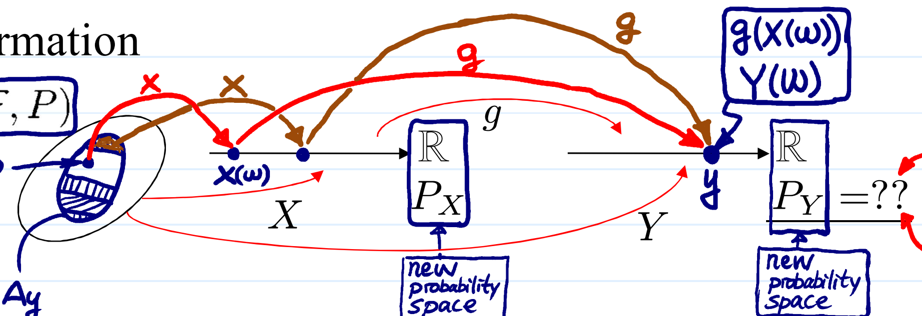
Thm in LNp.5-7 for pmf

■ Theorem. If a function F satisfies (2), (3), and (4), then F is a cumulative distribution function of some random variable. (in LNp.5-9~10)

proof. Skip. Out of the scope of the course. (discrete, continuous, or mixed)

$(*)$ 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 $= \{Y(\omega) : \omega \in \Omega\}$

$$\mathcal{Y} = \{g(x) : x \in \mathcal{X}\},$$

(a countably many set?) → $f_Y(y) = 0$, if $y \notin \mathcal{Y}$.

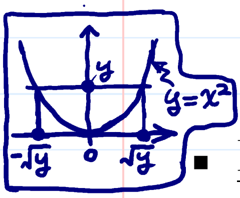
i.e., Y is a discrete r.v., and the pmf of Y is

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

(countably many)

proof. Since $\{\omega \in \Omega : Y(\omega) = y\} = \bigcup_{x \in \mathcal{X},\, g(x) = y} \{\omega \in \Omega : X(\omega) = x\}$, (countably many; mutually exclusive; $P(A_y)$)

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



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

❖ **Reading**: textbook, Sec 4.1, 4.2, 4.10

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

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

provided that the sum converges absolutely.

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

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

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

➢ Example (Indicator Function).

- For an event $A\subset\Omega$, the indicator function of A is the r.v.:

$$\mathbf{1}_A(\omega)=\begin{cases}1, & \text{if } \omega\in A,\\ 0, & \text{if } \omega\notin A.\end{cases}$$

- Its range $\mathcal{X}$ is $\{0, 1\}$ and its pmf is

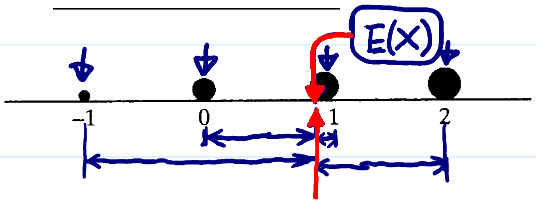
$$f(0)=P(A^c)=1-P(A) \quad \text{and} \quad f(1)=P(A),$$

for a p.m. P defined on Ω.

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

➢ Intuitive Interpretation of Expectation

- Expectation of a r.v. parallels the notion of a weighted average, where more likely values are weighted higher than less likely values.
- It is helpful to think of the expectation as the "center" of mass of the pmf.
  - 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.



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

$\wedge$ = center of gravity = .9

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

- Expectation of Transformation

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

$$Y = g(X),$$

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

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

provided that the sum converges absolutely.

proof. $\sum_{x\in\mathcal{X}} g(x)f_X(x) = \sum_{y\in\mathcal{Y}}\{\sum_{x\in\mathcal{X},\, g(x)=y} g(x)f_X(x)\}$

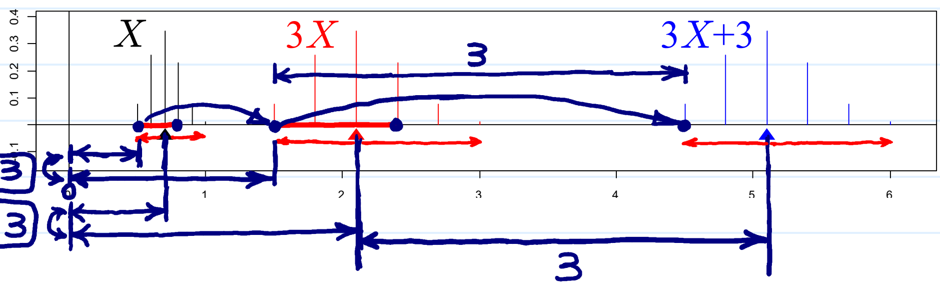
$= \sum_{y\in\mathcal{Y}} y \sum_{x\in\mathcal{X},\, g(x)=y} f_X(x) = \sum_{y\in\mathcal{Y}} y f_Y(y)$

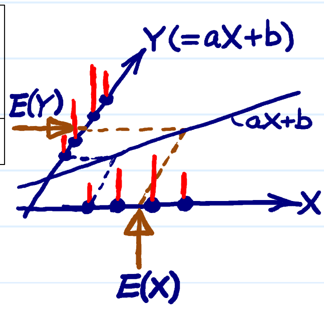
- Example. $Y = X^2$, $E(Y) = \sum_{x\in\mathcal{X}} x^2 f_X(x) \equiv E(X^2)$.

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

proof.

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





- Mean and Variance.

➤ 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.

- The $E(X)$ is often denoted by μ_X and $Var(X)$ by σ_X^2. Also, $\sigma_X = \sqrt{\sigma_X^2}$ is called the *standard deviation* of X.