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

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

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

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

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

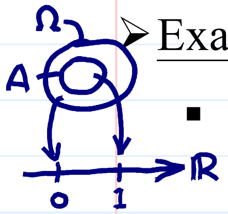
$\mathcal{X}=\{x_1,\ldots,x_n\}$, each with probability $1/n$

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

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

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

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 r.v.:

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

$\{\mathbb{1}_A=1\}=\{A \text{ occurs}\}$
$\{\mathbb{1}_A=0\}=\{A \text{ not occur}\}$

- 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

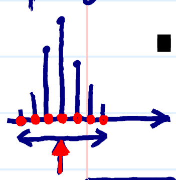
- So, $E(\mathbf{1}_A) = 0\cdot[1-P(A)]+1\cdot P(A) = P(A).$

➢ Intuitive Interpretation of Expectation

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

of all possible 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.



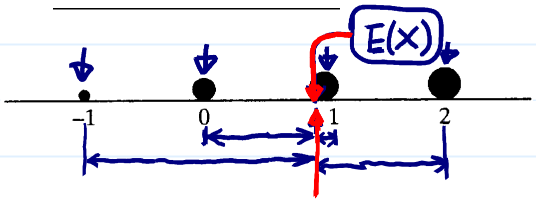
- It is helpful to think of the expectation as the "center" of mass of the pmf.

$0=\sum_{x\in\mathcal{X}} x f_X(x) - E(X)\underbrace{\sum_{x\in\mathcal{X}} f_X(x)}_{1}$
$=\sum_{x\in\mathcal{X}}(x-E(X))f_X(x)$

  - 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.

重心

槓桿原理: 力矩 = 距離 × 作用力



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

∧ = center of gravity = .9

$(-1)\times0.1+0\times0.25+1\times0.3+2\times0.35=$

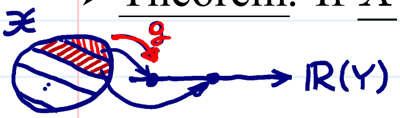
Recall Objective probability in LNp.3-19

- 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
Then, random
−1 → about 1000 times
0 → : 2500 :
1 → : 3000 :
2 → : 3500 :
→ 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.

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

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

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

provided that the sum converges absolutely. $\sum_{y\in\mathcal{Y}} |y| f_Y(y) < \infty$

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

by Thm in LNp.5-12

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

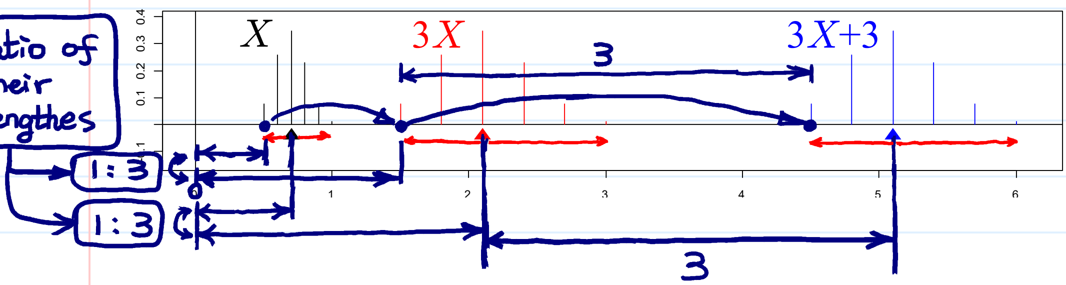
- Example. $Y = X^2$, $E(Y) \overset{\text{cf}}{=} \sum_{x\in\mathcal{X}} x^2 f_X(x) \equiv E(X^2)$. Y

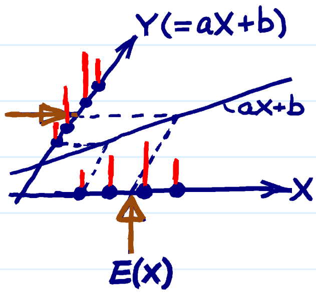
➢ Theorem. For $a, b \in \mathbb{R}$, $E(aX+b) = a\cdot E(X)+b$.

proof. fixed constants; a transformation of X: $Y = aX+b$

$$E(Y) = E(aX+b) = \sum_{x\in\mathcal{X}} (ax+b)f_X(x) = a\Big[\sum_{x\in\mathcal{X}} x f_X(x)\Big] + b\Big[\underbrace{\sum_{x\in\mathcal{X}} f_X(x)}_{=1}\Big]$$

ratio of their lengthes





- 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

a fixed constant

deviation of X from μ_X

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

E(Y) = Var(X)

provided that the sum converges. a tansformation of X, $Y = (X-\mu_X)^2$

same unit

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

having same unit as X

$\sigma_X = \sqrt{\sigma_X^2}$ is called the *standard deviation* of X. 標準差