Expectation and Variance of Random Variables

X is a random variable
$X : S \longrightarrow \mathbb{R}$

$\mathcal{X}$ = values of X

a typical value of x.

$c \rightsquigarrow$ particular value.

toss a coin five times
$X = \#$ tails in five tosses
$\mathcal{X} = \{0,1,2,3,4,5\}$
$x \in \mathcal{X}$, for $c \in \mathcal{X}$, $\{X = c\}$

$P(X=4) = P(\text{there were 4 tails in 5 tosses})$
$= 5 \cdot \frac{1}{32}$

Example 1)

x	0	1	2	3	4
$P(X=x)$	1/16	4/16	6/16	4/16	1/16

$$F_X(x) = \begin{cases} 0 & -\infty \le x < 0 \\ 1/16 & 0 \le x < 1 \\ 5/16 & 1 \le x < 2 \\ 11/16 & 2 \le x < 3 \\ 15/16 & 3 \le x < 4 \\ 1 & x \ge 4 \end{cases}$$

$P(\text{the number of heads in 4 tosses lie between 1 and 3 (inclusive)})$
$= P(1 \le X \le 3) = p_X(1) + p_X(2) + p_X(3)$
$= F_X(3) - F_X(1^-) = \frac{4}{16} + \frac{6}{16} + \frac{4}{16}$
$= \frac{15}{16} - \frac{1}{16} = \frac{14}{16}$ $= \frac{14}{16}$

1^- = largest value of X strictly less than 1 $= 0$

$F_X(\tfrac{7}{2}) = \sum_{\substack{y \in \mathcal{X} \\ y \le \frac{7}{2} = 3.5}} p_X(x)$
$= p_X(0) + p_X(1) + p_X(2) + p_X(3)$

$P(\tfrac{1}{2} \le X \le \tfrac{3}{2}) = F_X(\tfrac{3}{2}) - F_X(\tfrac{1}{2}^-)$
$= F_X(1) - F_X(0)$

$E(X) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{4}{16} + 2 \cdot \frac{6}{16} + 3 \cdot \frac{4}{16} + 4 \cdot \frac{1}{16}$
$= \underline{\underline{2}}$

Example 2) X is a r.v
$\mathcal{X}$ = values of $X = \{a_1, a_2, \dots, a_n\} \subseteq \mathbb{R}$

x	a_1	a_2	a_3	....	a_n
$P(X=x)$	$1/n$	$1/n$	$1/n$		$1/n$

$E(X) = a_1 \cdot \frac{1}{n} + a_2 \cdot \frac{1}{n} + \dots + a_n \cdot \frac{1}{n}$
$= \frac{1}{n} \cdot (a_1 + a_2 + \dots + a_n)$
$=$ mean $(a_1, a_2, \dots, a_n)$

General: $E(X) = p_1 a_1 + p_2 a_2 + \dots + p_n a_n$

$h(x) = -2x + 2$, $Y = h(X)$
$\mathcal{X} = \{0,1,2,3,4\}$
$h(\mathcal{X}) : \{2, 0, -2, -4, -6\} = \mathcal{Y}$

$h(x)$	2	0	-2	-4	-6
$P(Y=h(x))$	1/16	4/16	6/16	4/16	1/16

$E(h(X)) = 2 \cdot \frac{1}{16} + 0 \cdot \frac{4}{16} - 2 \cdot \frac{6}{16} - 4 \cdot \frac{4}{16} - 6 \cdot \frac{1}{16}$
$= \frac{2}{16} + 0 - \frac{12}{16} - \frac{16}{16} - \frac{6}{16}$
$= \frac{-32}{16} = -2$

$\mu_{-2X+2} = -2$ Recall $\mu_X = 2$
$-2 \cdot \mu_X + 2 = -2 \cdot 2 + 2$
$= -4 + 2 = -2 //$

$V(Y) = V(h(X)) = V(-2X+2) = (-2)^2 \cdot V(X)$
$= 4 \cdot 1$
$= 4 \cdot 1 = \underline{\underline{4}}$

$X \rightsquigarrow$ random variable
$Y = X + b$ $Var(X) = Var(Y)$

$V(h(X)) = (2-(-2))^2 \cdot \frac{1}{16} + \dots + (-6+2)^2 \cdot \frac{1}{16}$
$= \underline{4}$

Let $S \rightsquigarrow$ sample space

$X : S \longrightarrow \mathbb{R}$ be a random variable

$\mathcal{X} :=$ values of X.

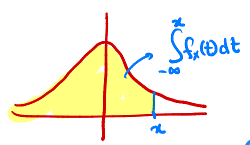
$F_X \rightsquigarrow$ "Cummulative distribution fuction" cdf of X.

$p_X(x)$ or $f_X(x) \rightsquigarrow$ "probability mass/density function" pmf/pdf of X.

$P(X=x) = p_X(x)$

We say X is

Discrete if $F_X(x)$ is a step function
$\Updownarrow$
$\mathcal{X}$ is a discrete subset of $\mathbb{R}$
$\Updownarrow$



$$\boxed{F_X(x) = \sum_{\substack{y \in \mathcal{X} \\ y \le x}} p_X(y)}$$

$\lim_{x \to \infty} F_X(x) = 1$

$\Updownarrow$
Calculations will usually involve discrete math and combinatorics.
$\Updownarrow$
$F_X(x) = P(X \le x)$
$\Updownarrow$
$$\boxed{P(a \le X \le b) = F(b) - F(a^-)}$$

$a^- \rightsquigarrow$ largest possible value of X strictly less than a.

Continuous if $F_X(x)$ is a continuous function
$\Updownarrow$
$\mathcal{X}$ is a union of intervals in $\mathbb{R}$.
$\Updownarrow$
$$\boxed{F_X(x) = \int_{-\infty}^{x} f_X(t)dt}$$
$\Updownarrow$
Calculations will typically involve calculus concepts
$\Updownarrow$
$F_X(x) = P_X(X \le x)$
$\Updownarrow$
$$\boxed{P(a \le X \le b) = \int_a^b f_X(x)\,dx = F_X(b) - F_X(a)}$$

Note:

① If X is discrete and $c \in \mathcal{X}$,

$P(X=c) = p_X(c) \neq 0$

if X is continuous,

$P(X=c) = P(c \le X \le c) = \int_c^c f_X(x)dx = 0$

for all $c \in \mathcal{X}$.

Expected Values of Random Variables

$\Updownarrow$
Average or mean value of the random variable X
$\Updownarrow$ denoted as
$\boxed{E(X) \text{ or } \mu_X}$

If X is Discrete
$\Updownarrow$
$$\boxed{E(X) = \sum_{x \in \mathcal{X}} x \cdot p_X(x)}$$

If X is Continuous
$\Updownarrow$
$$\boxed{E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\,dx}$$

Suppose $h(X)$ is a function of X.

$$\boxed{E(h(X)) = \sum_{x \in \mathcal{X}} h(x) \cdot p_X(x)}$$

$$\boxed{E(h(x)) = \int_{-\infty}^{\infty} h(x) f_X(x)\,dx}$$

Note: The definition of "expected value" make sense as long as

$\sum_{x \in \mathcal{X}} x\, p_X(x)$ or $\int_{-\infty}^{\infty} x f_X(x)\,dx$ exists.

This comment exists because it is possible that

$\sum_{x \in \mathcal{X}} x \cdot p_X(x)$ or $\int_{-\infty}^{\infty} x \cdot f_X(x)dx$ might not exist!

Variance of a Random Variable

$\Updownarrow$
Captures the variation from the expected value μ_X
$\Updownarrow$
Denoted by $\boxed{V(X) \text{ or } \sigma_X^2}$

If X is Discrete
$\Updownarrow$
$$\boxed{V(X) = E[(X-\mu_X)^2] \qquad \sigma_X^2 = \sum_{x \in \mathcal{X}} (x-\mu_X)^2 p_X(x)}$$

If X is Continuous
$\Updownarrow$
$$\boxed{V(X) = E[(X-\mu_X)^2] \qquad \sigma_X^2 = \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)\,dx}$$

The "standard deviation" is defined as $\boxed{\sigma_X = \sqrt{\sigma_X^2}}$

Properties of μ_X, σ_X^2, and σ_X

① Suppose $h(X) = aX + b \longrightarrow$ "linear function" of the r.v X.

then

$E[h(X)] = E[aX+b] = aE(X) + b$

ie $\boxed{\mu_{aX+b} = a\mu_X + b}$

$V(h(X)) = a^2 V(X) \longrightarrow \boxed{\sigma_{aX+b}^2 = a^2 \sigma_X^2}$

and $\boxed{\sigma_{aX+b} = |a| \sigma_X}$

② Shortcut Formula for Variance

$V(X) = E[(X-\mu_X)^2] = E[X^2 - 2X\mu_X + \mu_X^2]$
$= E(X^2) - 2\mu_X E(X) + E(\mu_X^2)$

$\therefore \boxed{V(X) = E(X^2) - (E(X))^2}$

Example:

Consider the experiment $\rightsquigarrow$ toss a fair coin four times.

$\boxed{X = \# \text{ of heads in four tosses.}}$

Then the probability distribution table of X is

x	0	1	2	3	4
$P(X=x)$	1/16	4/16	6/16	4/16	1/16

Then

$E(X) = 0 \cdot p_X(0) + 1 \cdot p_X(1) + 2 \cdot p_X(2) + 3 \cdot p_X(3) + 4 \cdot p_X(4)$
$= 0 \cdot \frac{1}{16} + 1 \cdot \frac{4}{16} + 2 \cdot \frac{6}{16} + 3 \cdot \frac{4}{16} + 4 \cdot \frac{1}{16}$
$= \frac{4+12+12+4}{16} = \frac{32}{16}$
$= 2$

$\therefore$ If we toss a fair coin 4 times repeatedly on average we would expect to get 2 heads in 4 tosses.

To calculate Variance

$V(X) = \sum_{x \in \mathcal{X}} (x-\mu_X)^2 \cdot p_X(x)$
$= (0-2)^2 \frac{1}{16} + (1-2)^2 \frac{4}{16} + (2-2)^2 \cdot \frac{6}{16} + (3-2)^2 \frac{4}{16} + (4-2)^2 \cdot \frac{1}{16}$
$= \frac{4}{16} + \frac{4}{16} + \frac{4}{16} + \frac{4}{16}$
$= \frac{16}{16} = 1$.

Note that,

$E(X^2) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{4}{16} + (2)^2 \cdot \frac{6}{16} + (3)^2 \cdot \frac{4}{16} + (4)^2 \cdot \frac{1}{16}$
$= \frac{4 + 4 \cdot 6 + 9 \cdot 4 + 16 \cdot 1}{16} = \frac{4+24+36+16}{16}$
$= \frac{80}{16} = 5$

Also $(E(X))^2 = (2)^2 = 4$.

$\therefore V(X) = E(X^2) - (E(X))^2 = 5 - 4 = 1$
$\Updownarrow$
Same as calculated above!!!