Expectation and Variance of Random Variables.

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.

We say X is

Discrete if

$F_X(x)$ is a step function

$\Updownarrow$

$\mathcal{X}$ is a discrete subset of $\mathbb{R}$

$\Updownarrow$

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

$\Updownarrow$

Calculations will usually involve discrete math and combinatorics.

$\Updownarrow$

$F_X(x) = P(X \leq x)$

$\Updownarrow$

$$\boxed{P(a \leq X \leq b) = F(b) - F(a^-)}$$

$a^- \rightsquigarrow$ 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

$\Updownarrow$

$F_X(x) = P_X(X \leq x)$

$\Updownarrow$

$$\boxed{P(a \leq X \leq b) = \int_a^b f_X(x)\,dx = F_X(b) - F_X(a)}$$

Note:

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

$P(X = c) = p_X(c)$

if X is continuous,

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

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

Expected Value 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

$\Updownarrow$

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

$\Updownarrow$

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

If X is Continuous

$\Updownarrow$

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

$\Updownarrow$

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

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

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 $\int_{-\infty}^{\infty} x \; f_X(x)\,dx$ 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

$\Updownarrow$

$$\boxed{\begin{aligned} V(X) &= E\left[(X-\mu_X)^2\right] \\ \sigma_X^2 &= \sum_{x \in \mathcal{X}} (x-\mu_X)^2 p_X(x) \end{aligned}}$$

If X is Continuous

$\Updownarrow$

$$\boxed{\begin{aligned} V(X) &= E\left[(X-\mu_X)^2\right] \\ \sigma_X^2 &= \int_{-\infty}^{\infty} (x-\mu_X)^2 f_X(x)\,dx \end{aligned}}$$

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

then

$$E\left[h(X)\right] = E\left[aX + b\right] = a\,E(X) + b$$

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

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

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

② Shortcut Formula for Variance

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

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

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

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

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

To calculate Variance

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

Note that,

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

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

$$\therefore \; V(X) = E(X^2) - (E(X))^2 = 5 - 4 = 1$$

$\Updownarrow$

Same as calculated above !!!