Discrete Random Variables (I)

Let

$X : S \longrightarrow \mathbb{R}$ $\rightsquigarrow$ random variable.

S = sample space of an experiment.

$\mathcal{X}$:= Values / image of X.

Often times when talking about X $\rightsquigarrow$ we will only specify $\mathcal{X}$ (not worry about S).

Suppose

$F_X \rightsquigarrow$ cdf of X

p_X or $f_X \rightsquigarrow$ pdf of X.

$\Big\}\rightsquigarrow$ Can use these to calculate

$\Downarrow$

$P(a \leq X \leq b)$

$\Updownarrow$

$$\boxed{P(X \text{ takes values in the interval } [a,b])}$$

X is said to be **Discrete** r.v $\xrightarrow{\text{if}}$ F_X is a step function

$\Updownarrow$

$\mathcal{X}$ will be a discrete subset of $\mathbb{R}$.

$\Updownarrow$

ie $\mathcal{X}$ is a

- finite set $\Updownarrow$ in bijection with $\{1,2,3,\ldots,N\}$ for some $N \in \mathbb{N}$
- Countably infinite set. $\Updownarrow$ in bijection with $\mathbb{N} = \{1,2,3,4,\ldots\}$

Goal for today $\rightsquigarrow$ Study r.v's with finite $\mathcal{X}$.

$\Updownarrow$

ie calculate pmf $p_X(x)$, μ_X and σ_X^2.

① Uniform Discrete Random Variable

Fix $N \in \mathbb{N}$.

> The r.v X has "uniform discrete distribution" with parameter N.

$\xrightarrow{\text{if}}$ $\mathcal{X} = \{1,2,3,\ldots,N\}$ and $p_X(x) = P(X=x) = \frac{1}{N}$ for $x = 1,2,3,\ldots,N$.

Note: ① $\sum_{x \in \mathcal{X}} p_X(x) = \underbrace{\frac{1}{N} + \frac{1}{N} + \cdots + \frac{1}{N}}_{N \text{ times}} = N \cdot \frac{1}{N} = 1$

and $p_X(x) = \frac{1}{N} > 0 \quad \forall x \in \mathcal{X}$

$\therefore$ p_X is a legitimate probability mass function!

② The distribution table for X is

x	1	2	3	4		$N-1$	N
$P(X=x)$	$\frac{1}{N}$	$\frac{1}{N}$	$\frac{1}{N}$	$\frac{1}{N}$		$\frac{1}{N}$	$\frac{1}{N}$

③

$$\mu_X = E(X) = 1 \cdot \frac{1}{N} + 2 \cdot \frac{1}{N} + 3 \cdot \frac{1}{N} + \cdots + N \cdot \frac{1}{N}$$
$$= \frac{1}{N}(1+2+3+\cdots+N)$$
$$= \frac{1}{N} \cdot \left(\frac{N \cdot (N+1)}{2}\right) = \frac{(N+1)}{2}$$

④ To calculate $V(X)$,

$$E(X^2) = 1 \cdot \frac{1}{N} + (2)^2 \cdot \frac{1}{N} + 3^2 \cdot \frac{1}{N} + \cdots + N^2 \cdot \frac{1}{N}$$
$$= \frac{1}{N}(1 + 2^2 + 3^2 + \cdots + N^2)$$
$$= \frac{1}{N}\left(\frac{N(N+1)(2N+1)}{6}\right) = \frac{(N+1)(2N+1)}{6}$$

$$\therefore V(X) = E(X^2) - E(X)^2 = \frac{(N+1)(2N+1)}{6} - \left(\frac{(N+1)}{2}\right)^2$$
$$= \frac{(N+1)(N-1)}{12}$$

⑤ Let $\{a_1, a_2, a_3, \ldots, a_N\} \subseteq \mathbb{R}$

$\Downarrow$

define $h : \mathcal{X} \longrightarrow \{a_1, a_2, \ldots, a_N\}$, $i \longmapsto a_i$

then if X has uniform distribution $\rightsquigarrow$ $E(h(X)) = \frac{a_1 + a_2 + \cdots + a_N}{N}$

$V(h(X))$ = variance of the numbers $\{a_1, a_2, a_3, \ldots, a_N\}$

Now fix $N, M, n \in \mathbb{N}$

Suppose a bag has N balls (identical except for color) M of which are red.

$\Downarrow$

Experiment: Choose n balls from the bag

$$\boxed{X = \# \text{ of red balls in the sample of } n \text{ balls.}}$$

$\Downarrow$

The calculation of $p_X(x) = P(X=x)$ depends on if

- Sampling without replacement $\Downarrow$ **Hypergeometric distribution** with parameters: N, M, n
- Sampling with replacement $\Downarrow$ **Binomial Distribution** with parameters: n, $p = \frac{N}{M}$

② Hyper geometric Distribution

> X is said to have hypergeometric distribution with parameters: N, M, n

$\xrightarrow{\text{if}}$ $\mathcal{X} = \{0,1,2,3,\ldots,n\}$

$$\boxed{p_X(x) = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}}$$

$P(X=x) = P\left(\text{getting } x \text{ successes in a sample of size } n \text{ sampled without replacement from a population of size } N \text{ with } M \text{ successes}\right)$

Note: x must satisfy:

$$\boxed{\max(0, n-N+M) \leq x \leq \min(n, M)}$$

> If X has hypergeometric distribution with parameters: N, M, n

$\rightsquigarrow$ $E(X) = n \cdot \frac{M}{N}$

$$V(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot \left(\frac{M}{N}\right) \cdot \left(1 - \frac{M}{N}\right)$$

$\Downarrow$

if we set $p = \frac{M}{N}$

$$\boxed{E(X) = n \cdot p \quad \text{and} \quad V(X) = \left(\frac{N-n}{N-1}\right) \cdot n \cdot p \cdot (1-p)}$$

③ Binomial Distribution

> X has Binomial Distribution with parameters: n, p

$\xrightarrow{\text{if}}$ $\mathcal{X} = \{0,1,2,\ldots,n\}$

for $x \in \mathcal{X}$

$$\boxed{p_X(x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{(n-x)}}$$

Note:

$$1 = 1^n = (p + (1-p))^n = \sum_{x=0}^{n} \binom{n}{x} p^x \cdot (1-p)^{(n-x)} = \sum_{x \in \mathcal{X}} p_X(x)$$

$\therefore$ p_X is a pmf.

> If X has Binomial Distribution with parameters: n, p

$\rightsquigarrow$

$$\boxed{E(X) = n \cdot p \qquad V(X) = n \cdot p \cdot q}$$

④ Bernoulli Trials

$\Downarrow$

Experiment with exactly two outcomes

- success
- Failure

Let $p \in (0,1)$ and let

$\boxed{P(\text{Success}) = p}$ $\boxed{P(\text{Failure}) = 1-p}$

> X is a Bernolli random variable with parameter: p

$\rightsquigarrow$ $\mathcal{X} = \{0,1\}$

and

$$\boxed{p_X(x) = \begin{cases} p & \text{if } x=1 \\ (1-p) & \text{if } x=0 \end{cases}}$$

In this case we have

$$\boxed{E(X) = p \quad \text{and} \quad V(X) = p \cdot (1-p)}$$