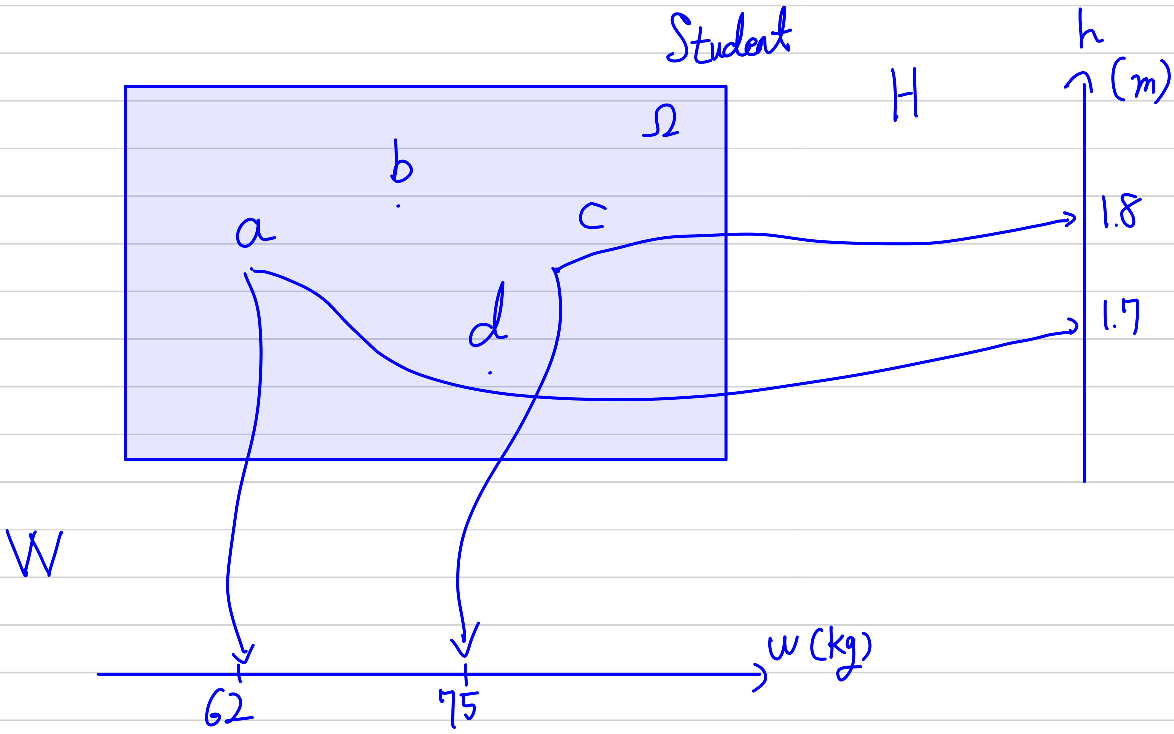
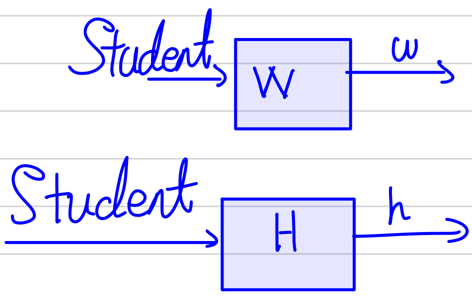
강의 5

Random variable

Random variable : 확률적 실험에 기반하여 얻을 수 있는 수
Discrete : 유한하거나 셀 수 있는 Set

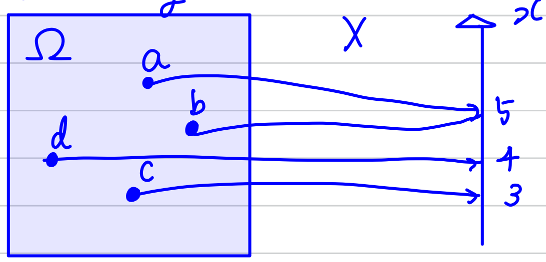




Notation | random variable X : 표본공간 Ω에 존재하는 어떤 한 함수
| numerical value x : A real number

X와 Y가 표본공간 Ω의 random variables일 때,
X와 Y로 이뤄진, X+Y, X/Y, 것들 또한 random variable이다.

PMF (Probability Mass Function)



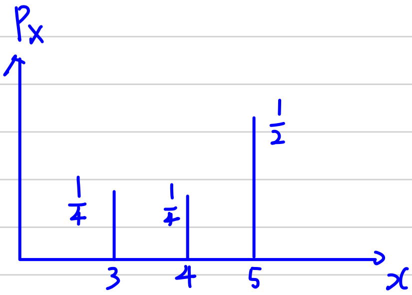
PMF는 random variable X의 "확률 법칙" 또는 "확률 분포"라고 할 수 있다.

특정 x값을 가졌을 때, "$X = x$"는 하나의 event라고 여길 수 있다.

Let us say that $x = 5$
Then event would be $X=5$ and the elements of X which satisfy $X=5$ can be written as

$\{w \mid X(w) = 5\}$ or $\{a, b\}$

In this case, the PMF of discrete r.v. X is P_X or $P(X=x)$
or $P(\{w \in \Omega \text{ and } X(w) = x\})$

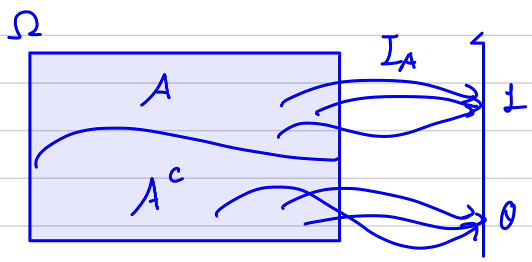


확률 법칙에 따라 $\sum_{x \in \Omega} P_X(x) = 1$, $P_X(x) \geq 0$

Bernoulli & Indicator

$$X = \begin{cases} 1 & \text{with parameter } p \\ 0 & \text{w.p. } 1-p \end{cases}$$

=> 1과 0으로 이루어진 r.v.를 Bernoulli r.v.라고 함

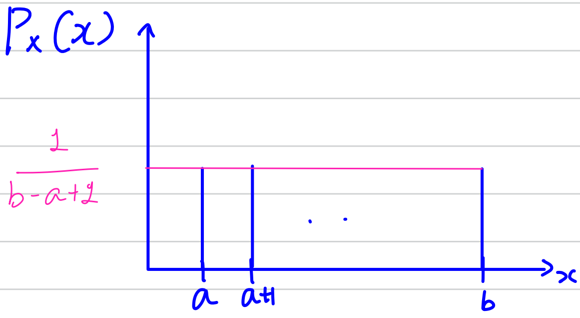


=> 이러한 r.v.를 A의 Indicator라고 하며, 다음 수식이 성립된다.

$$P_{I_A}(1) = P(I_A = 1) = P(A)$$

Discrete uniform r.v.

u.r.v. PMF



Parameter: a, b $\quad a \le b$, $a, b \in \mathbb{Z}$

Sample space : $\{a, a+1, \cdots, b\}$

r.v. : $X(w) = w$

Discrete Binomial r.v.

Experiment : n times independent tosses of a coin with $P(\text{Heads}) = p$

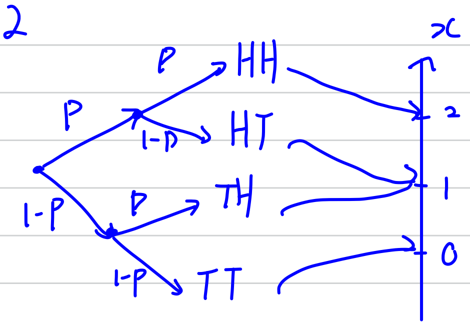
Sample space : A set of sequences of H and T, of length n

$$\Rightarrow \underbrace{HTTHHT \cdots HT}_{n}$$

r.v. X : The number of observed Heads

$n = 2$

ex)



Let us find a $P_X(1)$

$$P_X(1) = P(X = 1)$$

$$= P(HT) + P(TH)$$

$$= p(1-p) + p(1-p)$$

$$= 2p^1(1-p) = \binom{2}{1} p^1 (1-p)$$

(The expression $\binom{a}{b}$ equals ${}_aC_b$)

$$\therefore P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$= {}_nC_k \, p^k (1-p)^{n-k} \qquad k = (0 \sim n)$$

Discrete Geometric r.v.

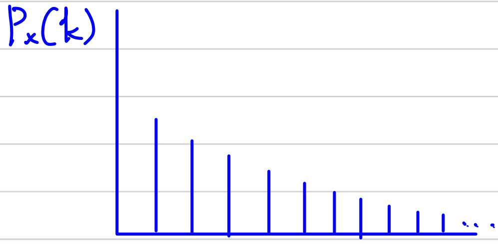
Experiment : Infinitely many independent tosses of a coin : $P(H) = p$

S.s. : A set of sequences of H and T

r.v. : The number of tosses until the first H

$$P_X(k) = P(X=k) = P(\underbrace{T\cdots T}_{k-1}\overset{\text{first Head}}{H}) = (1-p)^{k-1}p \qquad k = 1, 2, \cdots$$

Let $p = 1/3$



$$P(\text{no Heads ever}) \le P(T\cdots T) = \underbrace{(1-p)^k}_{\xrightarrow{k\to\infty}\, 0}$$

$$\Rightarrow P(X \to \infty) = 0$$

Expection / mean of a r.v.

Let $X = \begin{cases} 1, & \text{with parameter } 2/10 \\ 2, & \text{w.p. } 5/10 \\ 4, & \text{w.p. } 3/10 \end{cases}$

Point

$\to$ 1000번 중 200회

" 500

" 300

The average score : $\dfrac{1 \times 200 + 2 \times 500 + 4 \cdot 300}{1,000}$

Point Times

$$= 1 \times \boxed{\frac{200}{1,000}} + 2 \times \frac{500}{1,000} + 4 \times \frac{300}{1,000}$$

$\hookrightarrow$ Likelihood $= P(X=1)$

$$= 1 \cdot P(X=1) + 2 \times P(X=2) + 4 \times P(X=4)$$

$$\Rightarrow E[X] = \sum_{x \in \Omega} x P(X=x)$$

E of Bernoulli r.v.

$X = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{" } 1-p \end{cases}$

$E[X] = 1 \times p + 0 \cdot (1-p)$

$= \boxed{p}$

E of uniform r.v.

$X = 0 \sim n$ w.p. $\frac{1}{n+1}$

$E[X] = 0 \cdot \frac{1}{n+1} + 1 \cdot \frac{1}{n+1} + \cdots n \cdot \frac{1}{n+1}$

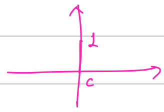
$= \frac{1}{n+1} \times (0 + 1 + \cdots n) = \boxed{\frac{n}{2}}$

Properties of E

1. If $X \geq 0$, then $E[X] \geq 0$

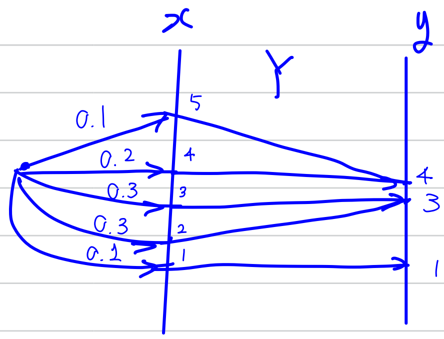
2. If $a \leq X \leq b$, then $a \leq E[X] \leq b$

3. If c is a constant, then $E[c] = c$



$E[c] = c \cdot p = c$

E for calculating $E[g(X)]$



Let X be a r.v. & $Y = g(X)$

Average $E[Y] = \sum_y y P_Y(y)$

$$= 1 \cdot 0.1 + 3 \cdot (0.3 + 0.3) + 4 \cdot (0.1 + 0.2)$$

$$= 1 \cdot 0.1 + 3 \cdot 0.3 + 3 \cdot 0.3 + 4 \cdot 0.1 + 4 \cdot 0.2$$

— Y를 한 번에 계산

— X를 고려한 후, 각각의 x의 기여도를 계산

$$E[Y] = E[g(X)] = \sum_x g(x) P_X(x)$$

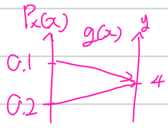
모든 x에 대해
특정 y에 대해

특정 y에 대해

Proof

$$\sum_y \sum_{x: g(x) = y} g(x) P_X(x) = \sum_x g(x) P_X(x)$$

ex) $y = 4$



$$\sum_y \sum_{x: g(x) = y} y P_X(x)$$

$$\sum_y y \boxed{\sum_{x: g(x) = y} P_X(x)}$$

→ A probability of specific value y

→ $P_Y(y)$

$$\sum_y y P_Y(y) = E[Y]$$

$$E[Y] = \sum_x Y P_X(x)$$

ex) $Y = x^2$

$$E[Y] = \sum_x x^2 P_X(x)$$

Linearity of E

$$E[aX+b] = aE[X]+b$$