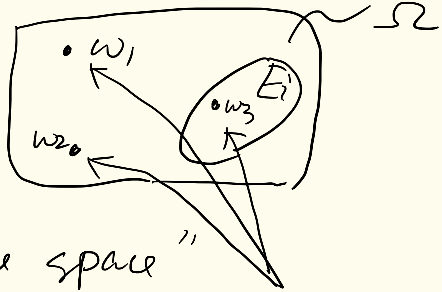
Review of Prob. Theory



Axioms of Prob:

1. $P(\Omega) = 1$
2. $P(E_i) \geq 0$
3. $P(E_1 \cup E_2 \cup \cdots) = \sum_{i=1}^{\infty} P(E_i)$,

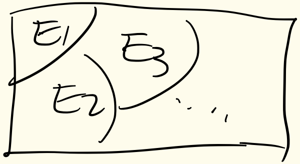
Ω: "sample space" (collection of all possible "outcomes")

E_i: "event" (a subset of Ω)
$E_i \subset \Omega$

ω: "outcome". $\omega \in \Omega$

E_i's are mutually exclusive,

i.e., $E_i \cap E_j = \phi$, $\forall i \neq j$



Propositions: $A' = \Omega \setminus A$

1. $P(A) + P(A') = 1$

2. $P(A) \leq 1$, 3. $P(A \cap B) + P(A \cup B) = P(A) + P(B)$

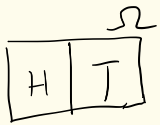
$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$

Ex: Consider games ("random experiments")

$\begin{cases} \underline{\text{finite}} \text{ outcomes} : (1), (2) \\ \underline{\text{countably infinite}} \text{ outcomes} \end{cases}$

(1) Flip a coin : $\Omega = \{H, T\}$

$\uparrow \frac{1}{2}$



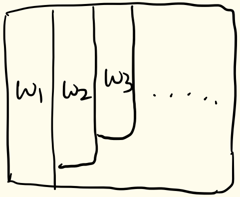
$P(\{H\}) = \frac{1}{2}$

(2) Roll a die : $\Omega = \{\text{"1"}, \text{"2"}, \text{"3"}, \ldots, \text{"6"}\}$ $P(\{\text{"6"}\}) = \frac{1}{6}$

(3) Roll a die until "6" appears.

$$\Omega = \{ \omega_1, \quad \leftarrow \text{ 6 on } 1^{st} \text{ roll} \qquad P(\{\omega_1\}) = \frac{1}{6}$$

$$\omega_2, \quad \leftarrow \text{ 6 on } 2^{nd} \text{ roll} \qquad P(\{\omega_2\}) = \frac{5}{6} \cdot \frac{1}{6} = \frac{5}{6^2}$$



$$\vdots$$

$$\omega_n, \quad \leftarrow \text{ 6 on } n^{th} \text{ roll} \qquad P(\{\omega_n\}) = \left(\frac{5}{6}\right)^{n-1} \cdot \frac{1}{6}$$

$$\vdots$$

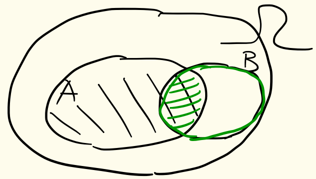
$$\}$$

$$\lim_{n\to\infty} \sum_{i=1}^{n} P(\{\omega_i\}) = \lim_{n\to\infty} \sum_{i=1}^{n} \left(\frac{5}{6}\right)^{i-1} \cdot \frac{1}{6} = \frac{1}{1-\frac{5}{6}} \cdot \frac{1}{6} = 1$$

$$\hookrightarrow = P\Big(\underbrace{\{\omega_1\} \cup \{\omega_2\} \cup \cdots}_{\Omega}\Big) = 1$$

Conditional Prob:

$$P(A|B) \triangleq \frac{P(A\cap B)}{P(B)} = \frac{P(A,B)}{P(B)}, \quad P(B)\neq 0$$

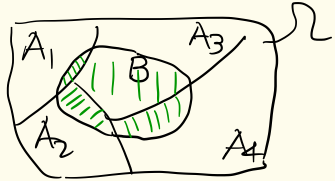


$$P(A) = P(A|\Omega)$$

Laws of total prob:

$$P(B) = \sum_{i=1}^{n} P(B, A_i), \quad \bigcup_{i=1}^{n} A_i = \Omega, \quad A_i \cap A_j = \emptyset, \forall i\neq j$$

$$= \sum_{i=1}^{n} P(B|A_i)\cdot P(A_i)$$



Bayes' Theorem:

$$P(A_j|B) = \frac{P(A_j, B)}{P(B)} = \frac{P(B|A_j)\cdot P(A_j)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}, \quad j=1,\dots,n$$

Independence:

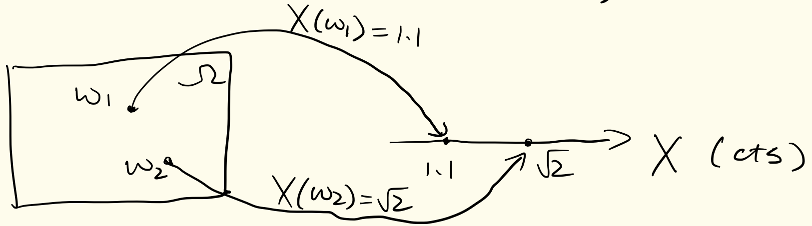
$$P(A \cap B) = P(A) \cdot P(B) \quad \text{or} \quad P(A|B) = P(A)$$

"condition-drop" test

Random Variables (r.v.s.): A map from sample space to $\mathbb{R}$

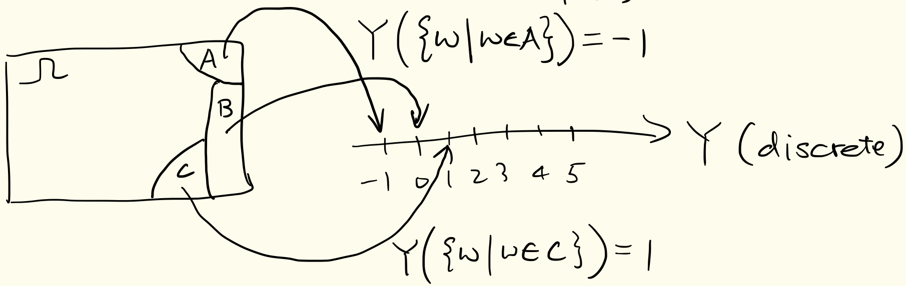
Ex: ① continuous r.v. $X: \Omega \longrightarrow \mathbb{R}$

$\omega \longmapsto X(\omega)$

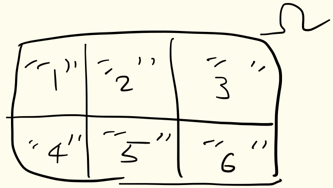


② discrete r.v. $Y: \Omega \longrightarrow \mathbb{Z}$

$\omega \longmapsto Y(\omega)$



Ex: Let X be the parity of roll of a die



$$X = \begin{cases} 0, & \text{even outcomes}, \\ 1, & \text{odd} \end{cases}$$

Event $\{X=0\} \triangleq \{\omega \mid X(\omega)=0\} = \{\text{"2"}, \text{"4"}, \text{"6"}\}$

$$\mathbb{P}[X=0] = \mathbb{P}[\{\omega \mid X(\omega)=0\}] = \mathbb{P}[\{\text{"2"}, \text{"4"}, \text{"6"}\}]$$

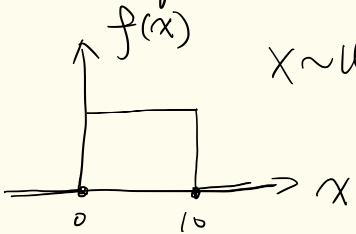
$$= \mathbb{P}[\{\text{"2"}\}] + \mathbb{P}[\{\text{"4"}\}] + \mathbb{P}[\{\text{"6"}\}] = \frac{1}{6} \times 3 = \frac{1}{2}$$

Cumulative Distribution/Density Function (CDF)

$$F(c) \triangleq \mathbb{P}[X \leq c] = \mathbb{P}[\{\omega \in \Omega \mid X(\omega) \leq c\}]$$

Percentile: $p \in [0,1]$. $(100 \cdot p)$th-percentile of the dist of r.v. X denoted by $\eta(p)$ is

$X \sim \text{Uniform}(0,10)$

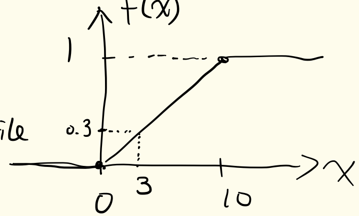


$$\eta(p) = F^{-1}(p),$$

$p = 0.3$

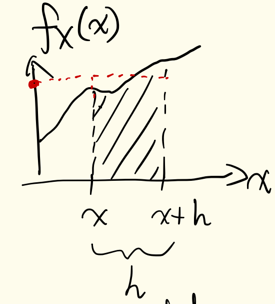
30th-percentile

$\eta(0.3) = 3$



Prob density function (pdf)

$$f_X(x) \triangleq \lim_{h\to 0} \frac{\mathbb{P}[x \le X \le x+h]}{h}$$



$$= \lim_{h\to\infty} \frac{\mathbb{P}[\{\omega\in\Omega \mid X(\omega)\in[x, x+h]\}]}{h}$$

$$\geq 0$$

Trapezoidal rule

Expectation value (mean) $\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f_X(x)\,dx$

$$\mathbb{E}[h(X)] = \int_{-\infty}^{\infty} h(x) f_X(x)\,dx$$

Prob mass func (pmf)

$$p_Y(y) \triangleq \mathbb{P}[Y=y]$$

$$= \mathbb{P}[\{\omega\in\Omega \mid Y(\omega)=y\}]$$

$$\in [0,1]$$

$$\mathbb{E}[X] = \mu_X = \sum_x x\, p_X(x)$$

$$\mathbb{E}[X] = \sum_x h(x)\, p_X(x)$$

$$\mathbb{E}[aX+b] = a\cdot\mathbb{E}[X] + b$$

Variance

$$V(X) = \int_{-\infty}^{\infty} (x-\mu_x)^2 f_X(x)\,dx \quad ; \quad V(X) = \sum_x (x-\mu_x)^2 p_X(x)$$

or $Var(X)$

$$= \underbrace{\mathbb{E}[(X-\mu_x)^2]}_{\sigma^2} = \mathbb{E}[X^2] - \underbrace{(\mathbb{E}X)^2}_{\mu^2} \Longrightarrow \mathbb{E}[X^2] = \mu^2 + \sigma^2$$

$$V(aX+b) = a^2 V(X)$$

Moments: $m_k = \mathbb{E}[X^k]$: k^{th} moment

$$= \int_{-\infty}^{\infty} x^k f_X(x)\,dx \quad ; \quad = \sum_x x^k p_X(x)$$

Joint distribution: $\iint f_{XY}(x,y)\,dx\,dy = 1$ | $\sum_y \sum_x p_{XY}(x,y) = 1$

$$f_X(x) = \int_{y\in\mathbb{R}} f_{XY}(x,y)\,dy \quad \Big| \quad p_X(x) = \sum_y p_{XY}(x,y)$$

$p_Y(1)$

X \ Y	1	....	N
1	0.1		0.02
2	0.2		0.01

$p_X(1)$

Sum along the row/column to marginalize

(Joint) Independent pdf/cdf/pmf factors under $\mathbb{P}$, e.g.,
f, F, p

$$f(x_1, \dots, x_n) = f(x_1) \cdot \dots \cdot f(x_n) = \prod_{i=1}^{n} f(x_i) \quad \left(\coprod_i X_i\right)$$

independence

Conditional dist: $f_{Y|X}(y|x) = \frac{f(x,y)}{f(x)}$ | $\underbrace{p_{Y|X}(y|x)}_{\mathbb{P}[Y=y \mid X=x]} = \frac{p(x,y)}{p(x)}$

Covariance C_{XY} / $\mathrm{Cov}(X,Y) \triangleq \mathbb{E}[(X-\mathbb{E}X)(Y-\mathbb{E}Y)]$

$$= \mathbb{E}[XY] - (\mathbb{E}[X])(\mathbb{E}[Y])$$

Def: Uncorrelatedness: $\mathrm{Cov}(X,Y) = 0$ or $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$

① $X \perp\!\!\!\perp Y \Rightarrow \mathrm{Cov}(X,Y) = 0$

② $\left.\begin{array}{l} \mathrm{Cov}(X,Y) = 0 \\ X, Y \text{ are joint Gaussian/normal} \end{array}\right\} \Rightarrow X \perp\!\!\!\perp Y$

Proof for "$X \perp\!\!\!\perp Y \Rightarrow \mathrm{Cov}(X,Y)=0$":

$$\mathrm{Cov}(X,Y) = \underline{\mathbb{E}[XY]} - \mathbb{E}[X]\,\mathbb{E}[Y]$$

$$= \iint xy\, f(x,y)\, dx\, dy - \left(\int x f(x)\, dx\right) \cdot \left(\int y f(y)\, dy\right)$$

$$\overset{X \perp\!\!\!\perp Y}{=} \iint x y\, f(x) f(y)\, dx\, dy - (\qquad) \cdot (\qquad) = 0 \qquad \square$$

Correlation: $\quad r_{xy} \triangleq \mathbb{E}[XY]$

Def: Orthogonality (in prob/stat): $\mathbb{E}[XY] = 0$