Primer on probability and statistics

Notation :

- Lower case : e.g. x, y, w, z $\begin{cases} \text{event realizations, data points,} \\ \text{column vectors (row vectors e.g. } w^T) \\ \text{functions, probability measures, etc.} \end{cases}$

- Upper case : e.g X, Y, A, B $\begin{cases} \text{random variables} \\ \text{matrices} \\ \text{some functions (e.g. cdf } F) \end{cases}$

- Caligraphics : $\{$ sets, operators , e.g. $\mathcal{F}$, $\mathbb{E}$ $\}$

Probability space :

$(\Omega, \mathcal{F}, p)$

Ω : sample space

$\mathcal{F}$: σ-algebra set of all events

p : probability measure

$p : \mathcal{F} \rightarrow [0, 1]$,

$X : \Omega \rightarrow \mathcal{X}$, e.g. $\mathbb{R}^n$

X : random variable

Discrete random variables :

A discrete r.v. X is an event taking values in a finite or countably infinite discrete space $\mathcal{X}$.

We will denote the probability of $X = x$ as $p(X = x)$ or $\boxed{p(x)}$ (more formally : $p(\{w \in \Omega : X(w) = x\})$)

$\begin{cases} \text{i) } 0 \leq p(x) \leq 1 \\ \text{ii) } \sum_x p(x) = 1 \end{cases}$, here $p(x)$ is called the probability mass function of X.

$x \in X$

iii) $p(\{\emptyset\}) = 0$ probability mass ... (pmf)

Continuous random variables : (events that take continuous values)

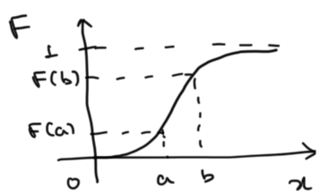
Define the events : $A = (X \leq a)$, $B = (X \leq b)$, $W = (a <$

Observe , $\underbrace{B = A \cup W}$, $A \cap W = \{\emptyset\}$

$$p(B) = p(A) + p(W) - \cancel{p(A \cap W)}^{0}$$

$$\Rightarrow \quad p(W) = p(B) - p(A)$$

Define a function $F(x) := p(X \leq x) \rightarrow$ cummulative distribut function of (cdf)



$$p(W) = p(a < X \leq b) = F(b) - F(a)$$

properties of a cdf

i.) $0 \leq F(x) \leq 1$

ii) $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$

iii) $x \leq y \Rightarrow F(x) \leq F(y)$ monotonically non decreasing function

If the cdf is differentiable , define $p(x) := \frac{d}{dx} F(x)$

$\hookrightarrow$ probability density functor (pdf) of X

By definition :

$$p(W) = p(a < X \leq b) = \int_a^b p(x)\, dx$$

Properties of a pdf :

i.) $p(x) \geq 0$

ii) $\int_{-\infty}^{\infty} p(x)\, dx = 1$

iii) $\int p(x)\, dx = p(x \in A)$

⊛ We require that $p(x) \geq 0$, it is possi that $p(x) \geq 1 \ \forall x$, as long as long as : $\int_{-\infty}^{+\infty} p(x)\, dx = 1$

$x \in A$