# Bayes' Theorem

**Example:** In an ML class, 40% of the students are from a science background and the remaining from engineering. Of the science students, 80% have no prior exposure to ML and this number is 70% for engineering students. Find the probability of a student being from a science background given ~~they~~ have some exposure to ML.

$$P(\underbrace{\text{science}}_{A_1} \mid \underbrace{\text{ML exposure}}_{B}) \equiv P(A_1 \mid B)$$

$$P(A_i \mid B) = \frac{P(A_i)\, P(B \mid A_i)}{P(B)} = \frac{P(A_i)\, P(B \mid A_i)}{\sum_j P(A_j) \cdot P(B \mid A_j)}$$

$$P(A_1|B) = \frac{P(A_1)\cdot P(B|A_1)}{P(A_1)\cdot P(B|A_1) + P(A_2)\cdot P(B|A_2)}$$

P(ML exposure | Science)

P(Science)

P(Engineering)

P(ML exposure | Engineering)

$$= \frac{0.4 \times 0.2}{0.4 \times 0.2 + 0.6 \times 0.3}$$

$$\boxed{P(A_1|B) = 0.3077}$$

<u>Random variable</u> is a variable that denotes the outcome of a stochastic experiment

Discrete

Continuous.

Two functions can be associated with a random variable that follows a given probability distribution.

Probability density function (pdf)

Cumulative distribution function (cdf).

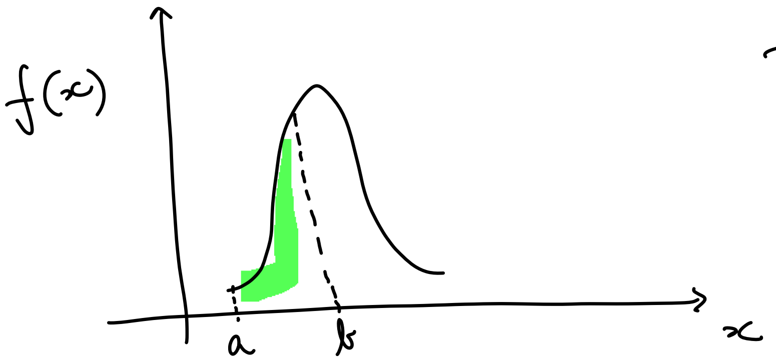
## Continuous random variables:

If a continuous random variable $X$ follows a probability distribution with the pdf $f(x)$, then the probability of $X$ lying in the interval $x$ and $(x+dx)$ is given as $\boxed{f(x)\, dx}$

$$P(x \leq X \leq x+dx) = f(x) \cdot dx$$

$$f(x) = \frac{P(x \leq X \leq x+dx)}{\phantom{dx}}$$

$dx \to \Delta x$




The probability of $X$ lying between two possible values $a$ and $b$ is:

$$\boxed{P(a \le X \le b) = \int_a^b f(x) \cdot dx.}$$

$$= F(b) - F(a)$$

Cumulative distribution function (cdf)

Considering $a \to -\infty$ and $b$ as $x$, we obtain the cdf:

$$\boxed{P(-\infty \le X \le x)} = \int_{-\infty}^{x} f(x') \cdot dx' = F(x) - \cancel{F(-\infty)} = \boxed{F(x)}$$

$F(-\infty) \to 0$

there is no possible value of $x$ below $-\infty$.

The cdf $\overline{F(x)}$ gives the probability that the random variable $X$ has a value less than or equal to the argument of the cdf, i.e. $\boxed{x}$

What will be $F(\infty)$:

$$F(\infty) = \int_{-\infty}^{\infty} f(x)\, dx = \boxed{1}$$

$\leftarrow$ the total probability of $X$ taking some value is 1.

$x \in \mathbb{R} \qquad x \in (-\infty, \infty)$

Often times, in ML, we deal with with many more variables. In such cases of multiple random variables, one can define the joint probability density function (jdf)

$$P(a \leq X \leq b,\ c \leq Y \leq d) = \int_a^b \int_c^d \underbrace{f(x,y)}_{\text{jdf}}\, dy\, dx$$

$$= \left[ \left[ F(x,y) \right]_{y=c}^{y=d} \right]_{x=a}^{x=b}$$

$$= \left[ F(x,d) - F(x,c) \right]_{x=a}^{x=b}$$

$$= F(b,d) - F(a,d) - F(b,c) + F(a,c)$$

Here, $F(x,y)$ is called the joint cumulative distribution function.

$$F(x,y) = P(-\infty < X \leq x, -\infty < Y \leq y)$$

Now, if $X$ and $Y$ are independent random variables,

$$f_{XY}(x,y) = f_X(x)\, f_Y(y)$$

Given the joint cdf, one can obtain the pdf by computing the derivative.

$$f_X(x) = \frac{dF_X(x)}{dx} \quad ; \quad f_X(x) = \frac{\partial F_{XY}(x,y)}{\partial x}$$