Continuous random variables

- Events that take continuous values.

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

By construction , observe : $B = A \cup W$, $A \cap W = \{\emptyset\}$

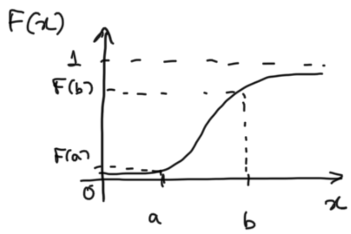
$$p(B) = p(A) + p(W) - \cancel{p(A \cap W)}^{0} \Rightarrow \boxed{p(W) = p(B) - p(A)} \quad (1)$$

Now let us define the function $F(x) := p(X \leq x)$

↳ cumulative distribution function of the continuous random variable X. (cdf)

By definition,

F is a monotonically non-decreasing function :



- $0 \leq F(x) \leq 1$
- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to +\infty} F(x) = 1$
- If $x \leq y \Rightarrow F(x) \leq F(y)$ (monotonically non-decreasing)

$$(1) \Rightarrow p(W) = p(a < X \leq b) = F(b) - F(a)$$

If the cdf $F(x)$ is also differentiable (not always the case!), then define $f(x) := \frac{d}{dx} F(x)$

↳ probability density function of the continuous random variable X. (pdf)

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

Notice that $f(x) > 0$, but it is possible that $f(x) > 1$, for any given continuous random variabe X, as long as the density integrates to 1 , i.e.:

Properties of a p.d.f : $\begin{cases} \int_{-\infty}^{+\infty} f(x)\,dx = 1 \\ f(x) \geq 0 \\ \int_{x \in \mathcal{X}} f(x)\,dx = p(x \in \mathcal{X}) \end{cases}$

Recap:

1. A random variable X is an event that can take possible differen values $x \in \mathcal{X}$
   $\hookrightarrow \mathcal{X}$: event space, i.e. the collection of all possible outcomes.
2. A discrete / continuous random variable can be characterized by its probability mass/density function.
3. Often we will encounter two scenarios:
   i.) Given some realization x, what $p(x)$?
   ii) Given some $p(x)$, generate events x that are distributed according to $p(x)$? [Sampling problem]

from now on
(*) $p(x)$ denotes a pdf.

- Let X be a continuous random variable with a pdf $p(x)$

– Mean / Expected value of X : $\mathbb{E}_{x \sim p(x)}[X] = \int_{x \in \mathcal{X}} x\, p(x)\, dx = \mu$
(First order moment / statistic of X)

– $\mathbb{V}ar[X] := \mathbb{E}_{x \sim p(x)}[(X-\mu)^2] = \int_{x \in \mathcal{X}} (x-\mu)^2 p(x)\, dx$
2nd-order moment
⋮

– n-th order moment / statistic : $\mathbb{E}_{x \sim p(x)}[(X-\mu)^n] = \int_{x \in \mathcal{X}} (x-\mu)^n p(x)\, dx$

Basic rules of probability

	Discrete r.vs	Continuous r.vs
$p(x)$	pmf	pdf

sum rule	$p(x) = \sum_{y \in \mathcal{Y}} p(x,y)$	$p(x) = \int_{y \in \mathcal{Y}} p(x,y)\,dy$
product rule	$p(x,y) = p(x\|y)\,p(y)$	$p(x,y) = p(x\|y)\,p(y)$
Bayes	$p(x\|y) = \dfrac{p(y\|x)\,p(x)}{\underbrace{\sum_{x \in \mathcal{X}} p(y\|x)\,p(x)}_{p(y)}}$	$p(x\|y) = \dfrac{p(y\|x)\,p(x)}{\underbrace{\int_{x \in \mathcal{X}} p(y\|x)\,p(x)\,dx}_{p(y)}}$

- Properties of the expectation of a random variable :

1. $\mathbb{E}[a] = a$, for any constant a.
2. $\mathbb{E}_{x \sim p(x)}[a f(x)] = a\, \mathbb{E}_{x \sim p(x)}[f(x)]$, $\int_{x \in \mathcal{X}} a f(x) p(x)\,dx = a \int_{x \in \mathcal{X}} f(x) p(x)\,dx$
3. $\mathbb{E}[a f(x) + b g(x)] = a\,\mathbb{E}[f(x)] + b\,\mathbb{E}[g(x)]$, Linearity

- Properties of variance :

$$\sigma^2 := \mathbb{V}\mathrm{ar}[X] = \mathbb{E}_{x \sim p(x)}\left[(X-\mu)^2\right] = \int_{x \in \mathcal{X}} (x-\mu)^2 p(x)\,dx \geq 0$$

$\mathbb{V}\mathrm{ar}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, this identity holds since :

$$\underbrace{\int_{x \in \mathcal{X}} (x-\mu)^2 p(x)\,dx}_{} = \underbrace{\int_{x \in \mathcal{X}} x^2 p(x)\,dx}_{\mathbb{E}[X^2]} + \underbrace{\mu^2 \cancelto{1}{\int_{x \in \mathcal{X}} p(x)\,dx}}_{\mu^2} - 2\mu \overbrace{\int_{x \in \mathcal{X}} x\, p(x)\,dx}^{\mathbb{E}[X] := \mu}$$

$$- 2\mu^2$$

$$\Rightarrow \quad \mathbb{V}\mathrm{ar}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

- Standard deviation : $\sigma := \mathrm{Std}[X] = \sqrt{\mathbb{V}\mathrm{ar}[X]} = \sqrt{\sigma^2}$

* Statistical independence :

Two random variables X, Y are statistically independent if :

$$p(x, y) = p(x)\,p(y)$$

Then:

- $p(y|x) = p(y)$
- $p(x|y) = p(x)$
- $\mathbb{V}ar[X+Y] = \mathbb{V}ar[X] + \mathbb{V}ar[Y]$ (this is <u>not</u> true if X, Y are not independent

In fact, if X, Y are not independent:

$$\mathbb{V}ar[X+Y] = \mathbb{V}ar[X] + \mathbb{V}ar[Y] + \text{cov}[X,Y] + \text{cov}[Y,X]$$

... next time we continue to define the concepts of covariance / correla