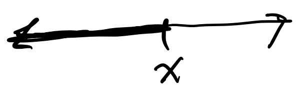
Have defined a random variable as a function of the outcome of a random experiment, denoted

$$X = X(\omega)$$

- One important type of event of interest is



$$\{X \leq x\} \quad \text{for } x \in \mathbb{R}$$

- Have defined the prob of this event to be the cdf, or distribution function of X:

$$F_X(x) = P(X \leq x), \forall x \in \mathbb{R}$$

- Have given 3 properties that a function $F: \mathbb{R} \to \mathbb{R}$ <u>must</u> satisfy to be a valid cdf $F_X(x)$ for some rv X. This means that the measure defined by $P(X \leq x) = F(x)$ will satisfy the axioms if F satisfies Properties 1 - 3.

Conversely, it can be shown that if a prob. measure

P satisfies the axioms, then $F_X(x)$ satisfies properties ①-③, plus:

④ $P(X > x) = 1 - F_X(x) \; \forall x \in \mathbb{R}$

This is because

$$\{X > x\} \cup \{X \leq x\} = S$$

and these events are disjoint, so

$$P(X > x) + P(X \leq x) = 1$$

(Axioms 2, 3)

so

$$P(X > x) = 1 - F_X(x)$$

From defn of F_X

(5) If $x_1 < x_2$ then

$$P\left(x_1 < X \leq x_2\right) = F_X\left(x_2\right) - F_X\left(x_1\right)$$

The proof will require that these are $<$ and $\leq$ respectively

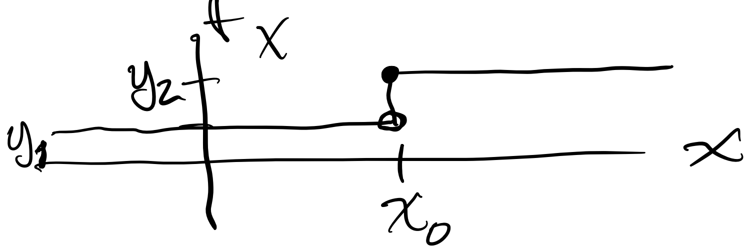
Proof left to the student

(6) For any $x \in \mathbb{R}$,

$$P(X = x) = F_X(x) - F_X\left(x^-\right)$$

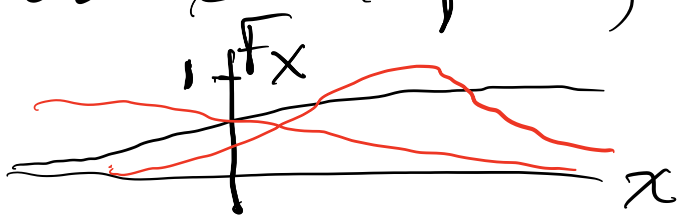
where $F_X\left(x^-\right) = \lim_{\varepsilon \to 0, \varepsilon > 0} F_X(x - \varepsilon)$

is the limit of F_X from the left at x.

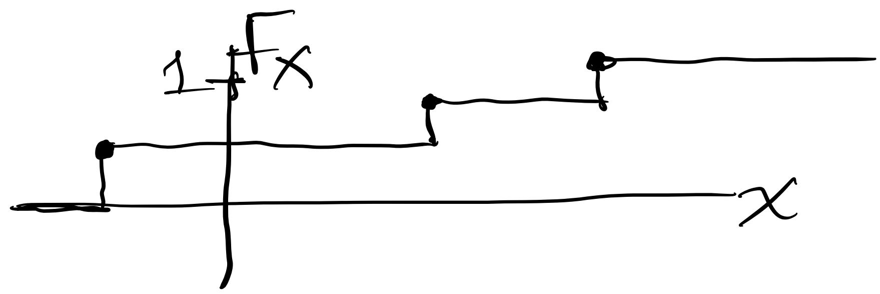


So $P(X = x_0) = y_2 - y_1$

So a valid cdf can look like these two examples,



or



but graphs in red are <u>not</u> valid cdfs

In practice, the cdf is used sometimes, but more often the "pdf" is used

The Probability Density Function of X

Defn. The probability density function (or pdf or the density function) of a rv X is

$$f_X(x) = \frac{dF_X(x)}{dx}$$

for all $x \in \mathbb{R}$ where F_X is differentiable.

From the Fundamental Theorem of Calculus, it can be shown that for $A \subset \mathbb{R}$,

$$\longrightarrow \boxed{P(X \in A) = \int_A f_X(x)\,dx}$$

For example,

- if $A = (-\infty, x]$, then

$$P(X \in A) = P(X \in (-\infty, x])$$

$$= P(X \leq x).$$ For this A,

$$\underbrace{P(X \in A)} = \int_{-\infty}^{x} f_X(\alpha)\, d\alpha$$

But this is $F_X(x)$, since for this A, $\{X \in A\} = \{X \leq x\}$

So we have shown that

$$F_X(x) = \int_{-\infty}^{x} f_X(\alpha)\, d\alpha$$

Common mistake:

$$\cancel{F_X(x) = \int_{-\infty}^{\infty} f_X(x)\, dx}$$

Not true. The integral must be from $-\infty$ to x, not to ∞

It can be shown that if a function $f(x)$ satisfies the properties

① $f(x) \geq 0, \forall x \in \mathbb{R}$

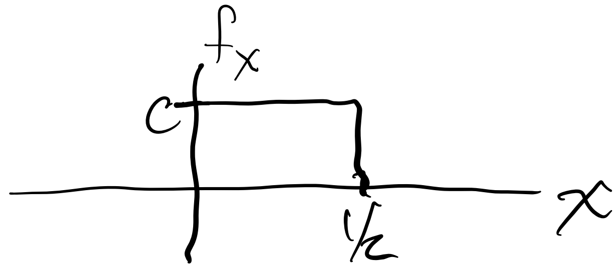
and ② $\displaystyle\int_{-\infty}^{\infty} f(x)\, dx = 1,$

then f is a valid pdf for some rv X.

Conversely, if P satisfies the axioms, or if the cdf satisfies its required 3 properties, then f_X will satisfy the 2 properties above.

Note that for a fixed $x_0 \in \mathbb{R}$, $f_X(x_0)$ is not a prob.

Example. Consider a rv X that is equally likely to take any value in $[0, \frac{1}{2}] \subset \mathbb{R}$. This means its pdf is:



for some constant c. Since

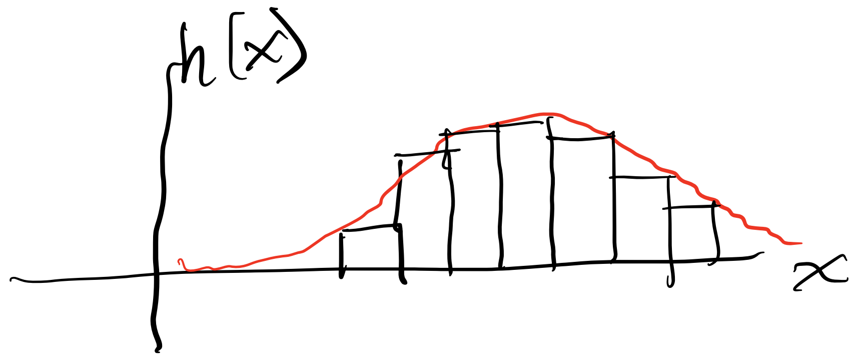
$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$, need

$\int_0^{1/2} c\,dx = 1$, so $c = 2$

So $f_X\left(\frac{1}{4}\right)$, for example,

is $2 \leftarrow$ not a prob., so it does not have to be ≤ 1.

To model rvs in practice, you might collect samples of the rv and plot a histogram :



Then fit a curve to estimate the pdf of X