RANDOM VARIABLES

A random variable is a numeric encoding of the outcome of a statistical experiment. Precisely:

Defn: A random variable is a function from a sample space S to the real numbers.

random variable

$\mathbb{R} = (-\infty, \infty)$

I.e. a $\overbrace{\text{r.v.}}^{\text{random variable}}$ X is a function $\underbrace{X : S \to \mathbb{R}}$

"script" $\mathcal{X}$

X takes outcomes in S and returns real numbers

Denote by $\mathcal{X}$ the range of X, the specific set of values that X may take.

E.g. Flip a coin. $S = \{H, T\}$

Let X be 1 if H and 0 if T:

$$X(s) = \begin{cases} 1 & \text{if } s = H \\ 0 & \text{if } s = T \end{cases}$$

Range of X is $\mathcal{X} = \{0, 1\}$.

E.g. 3 coin flips.

$$S = \left\{ \begin{matrix} & HHT & TTH & \\ HHH & HTH & THT & TTT \\ & THH & HTT & \end{matrix} \right\}$$

Let X be # of heads:

$$X(s) = \begin{cases} 0 & \text{if } s = TTT \\ 1 & \text{if } s \in \{TTH, THT, HTT\} \\ 2 & \text{if } s \in \{HHT, HTH, THH\} \\ 3 & \text{if } s = HHH \end{cases}$$

Range of X is $\mathcal{X} = \{0, 1, 2, 3\}$.

Eg: chip animal, record dist. travelled.

$$S = [0, \infty)$$

Let X be dist. travelled:

$$X(s) = s$$

Range of X is $\mathcal{X} = [0, \infty)$

* We most of the time write X instead of $X(s)$.

* Range $\mathcal{X}$ also called the "support" of X

Expressing probabilities about a r.v. X:

$\mathcal{X}$ finite: e.g. $\mathcal{X} = \{0, 1\}$, $\mathcal{X} = \{1, 2, \dots, n\}$

Let X be a r.v. on the sample space $S = \{s_1, \dots, s_n\}$ which takes values in $\mathcal{X} = \{x_1, \dots, x_m\}$. Then for any $x \in \mathcal{X}$, we write

$$P_X(X = x) = P(\underbrace{\{s \in S : X(s) = x\}}_{\substack{\text{Event on which } X = x, \\ \text{i.e. set of outcomes } s \\ \text{for which } X(s) = x}})$$

little x represents any specific value
big X is the symbol for our r.v.

So $P_X(X = x)$ is the probability that X takes the value x.

Eg: 3 coin flips, X = # heads. So $\mathcal{X} = \{0, 1, 2, 3\}$.

Then
$$P_X(X=0) = P(TTT) = 1/8$$
$$P_X(X=1) = P(\{TTH, THT, HTT\}) = 3/8$$
$$P_X(X=2) = P(\{HHT, HTH, THH\}) = 3/8$$
$$P_X(X=3) = P(HHH) = 1/8$$

We often tabulate $P_X(X=x)$ like this:

x	0	1	2	3
$P_X(X=x)$	1/8	3/8	3/8	1/8

$\mathcal{X}$ countable: e.g. $\mathcal{X} = \mathbb{Z}^+$ $(\mathbb{Z}^+ = \{1, 2, \ldots\})$

* The same is done for $S = \{s_1, s_2, \ldots\}$, $\mathcal{X} = \{x_1, x_2, \ldots\}$, i.e. when S and $\mathcal{X}$ are countable, as in the $\mathcal{X}$-finite case

$\mathcal{X}$ finite, countable, or uncountable e.g. $\mathcal{X} = \{0,1\}$, $\mathcal{X} = \mathbb{Z}^+$, $\mathcal{X} = (-\infty, \infty)$.

Let X be a r.v. on any sample space S which takes values in $\mathcal{X}$. Then for any set $A \in \mathcal{E}_{\mathcal{X}}$, we write

special collection of sets of interest in $\mathcal{X}$

$$P_X(X \in A) = P(\{s \in S : X(s) \in A\})$$

set of values X may take

So $P_X(X \in A)$ is the probability that X takes a value in A.

E.g. Chip animal, record dist. travelled

$$S = [0, \infty), \quad \mathcal{X} = [0, \infty).$$

$$P_X(X \in (1,2)) = P(\{s \in [0,\infty) : X(s) \in (1,2)\}) = P((1,2))$$

How might we assign probabilities to intervals? (More on this soon)

* For a r.v. X, we call P_X the probability distribution of X.

THE CUMULATIVE DISTRIBUTION FUNCTION

Defn: The cumulative distribution function (cdf) F_X of a r.v. X is

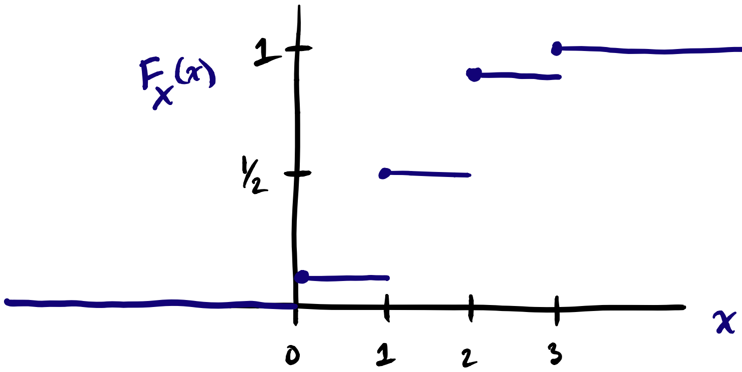
$$F_X(x) = P_X(X \le x) \quad \text{for all } x \in \mathbb{R}$$

could write $P_X(X \in (-\infty, x])$

E.g. 3 coin flips, X = # heads.

x	0	1	2	3
$P_X(X=x)$	1/8	3/8	3/8	1/8

$$\Rightarrow F_X(x) = \begin{cases} 0 & -\infty < x < 0 \\ 1/8 & 0 \le x < 1 \\ 4/8 & 1 \le x < 2 \\ 7/8 & 2 \le x < 3 \\ 1 & 3 \le x < \infty \end{cases}$$



Note:

* $F_X(x)$ defined for all $x \in \mathbb{R}$, not just $x \in \mathcal{X}$

e.g. $F_X(\tfrac{1}{2}) = P_X(X \le \tfrac{1}{2}) = P\left(\{s \in S : X(s) \le \tfrac{1}{2}\}\right) = P(\{TTT\}) = \tfrac{1}{8}$

* $F_X(x)$ has jumps at $x \in \mathcal{X}$ with jump sizes $P_X(X=x)$, $x \in \mathcal{X}$

* $F_X(x) = 0 \quad \forall\ x < 0$, $\quad F_X(x) = 1 \quad \forall\ x \ge 3$

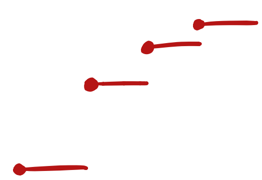
"if and only if"

Theorem: The function $F_X(x)$ is a cdf iff

(i) $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X(x) = 1$

(ii) $F_X(x)$ is nondecreasing in x

(iii) $F_X(x)$ is right-continuous; i.e. for every x_0, $\lim_{x \downarrow x_0} F(x) = F(x_0)$

(Right-continuity means jumps are like this: )

E.g. X = # coin flips to get a head, where on each flip, $P(\text{Head}) = p$.

$\mathcal{X} = \{1, 2, \dots\}$, and

x	1	2	3	...
$P_X(X=x)$	p	$\underbrace{(1-p)}_{1\text{ tail}}\underbrace{p}_{1\text{ head}}$	$\underbrace{(1-p)^2}_{2\text{ tails}}\underbrace{p}_{1\text{ head}}$	...

So we get $\underbrace{P_X(X=x) = (1-p)^{x-1}\,p}_{}$ for $x = 1, 2, \ldots$

called the geometric probability distribution

So for $x = 1, 2, \ldots$

$$P_X(X \le x) = P_X(X=1) + \cdots + P_X(X=x)$$

$$= \sum_{i=1}^{x} P_X(X=i)$$

$$= \sum_{i=1}^{x} (1-p)^{i-1}\, p$$

$$= p \sum_{i=1}^{x} (1-p)^{i-1}$$

$(1-p)^0 + (1-p)^1 + \ldots + (1-p)^{x-1}$

$$= p \sum_{i=0}^{x-1} (1-p)^i$$

$$= p \left[\frac{1-(1-p)^x}{1-(1-p)} \right]$$

$$= 1 - (1-p)^x$$

partial geometric series

Result:

For $t \neq 1$,

$$\sum_{i=0}^{n} t^i = \frac{1-t^{n+1}}{1-t} \quad \text{for } n = 1, 2, \ldots$$

Proof:

Let $S_n = \sum_{i=0}^{n} t^i = 1 + t + \ldots + t^n$

Then

$$S_n - t S_n = 1 + t + \ldots + t^n - (t + t^2 + \ldots + t^{n+1})$$

$$= 1 - t^{n+1}$$

So $S_n = \dfrac{1-t^{n+1}}{1-t}$

So for $x \in \mathbb{R}$,

$$F_X(x) = \begin{cases} 1-(1-p)^i & \text{for } x \in [i, i+1) \quad i = 0, 1, 2, \ldots \\ 0 & \text{for } x < 0 \end{cases},$$

Or we may equivalently write

$$F_X(x) = \left[1 - (1-p)^{\lfloor x \rfloor}\right] \mathbb{1}(x \ge 0)$$

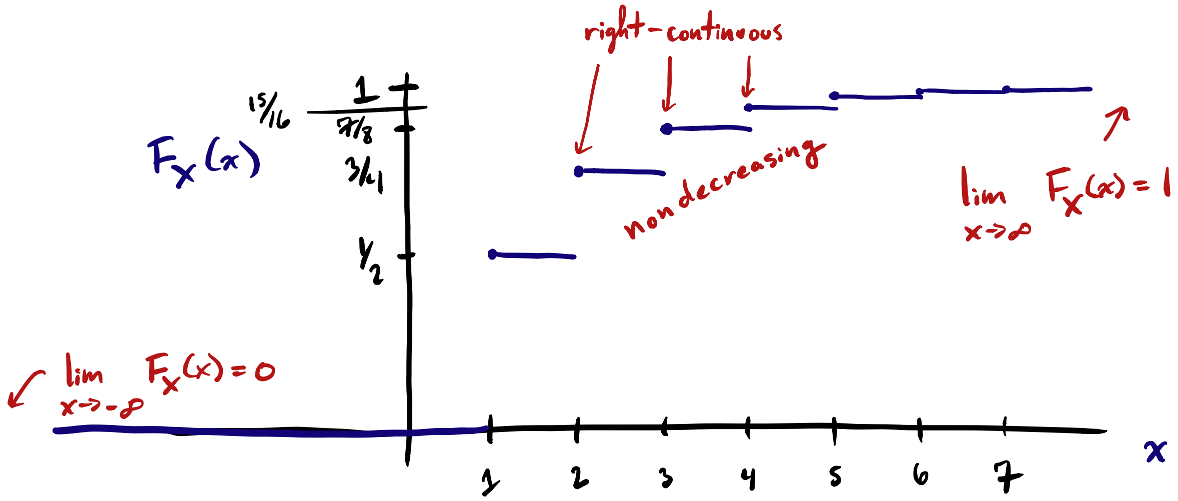
"floor" function

where $\lfloor x \rfloor$ is the greatest integer not exceeding x, e.g. $\lfloor 0.25 \rfloor = 0$, $\lfloor 1.87 \rfloor = 1$

and $\mathbb{1}(x \in A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$ is called the indicator function.

"Indicates" whether $x \in A$ by returning 1 or 0

For $p = 1/2$, $F_X(x)$ looks like



E.g. Chip animal, X = dist. travelled.

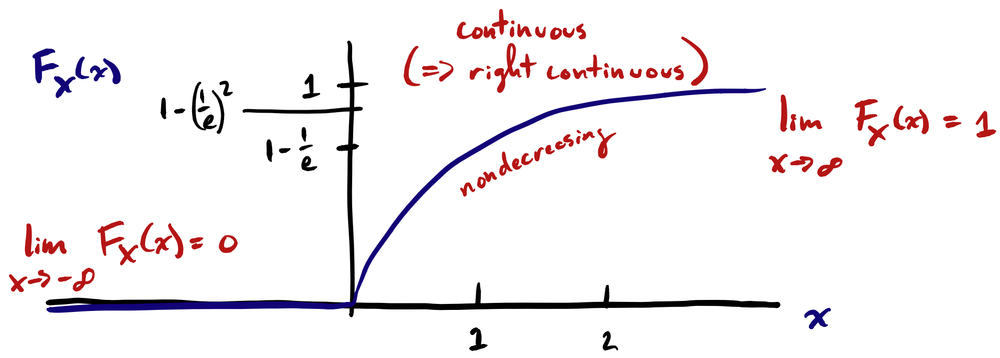
We might suppose that

This is the cdf of a probability distribution called the exponential distribution

$$F_X(x) = \begin{cases} 1 - e^{-x/\lambda} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$$

for some $\lambda > 0$.

For $\lambda = 1$, $F_X(x)$ looks like



* $F_X(x)$ is continuous (has no jumps); thus it is right-continuous

Defn: A r.v. X with cdf $F_X(x)$ is called a continuous r.v. if $F_X(x)$ is a continuous function of x and a discrete r.v. if $F_X(x)$ is a step function of x.

No "point probabilities" for continuous r.v.s

Let X be a continuous r.v.; consider event $\{X=x\}$:

$$\{X=x\} \subset \{x-\varepsilon < X \leq x\} \quad \text{for every } \varepsilon > 0$$

gives

$$0 \leq P_X(X=x) \leq P_X(x-\varepsilon < X \leq x) = F_X(x) - F_X(x-\varepsilon)$$

for every $\varepsilon > 0$, and

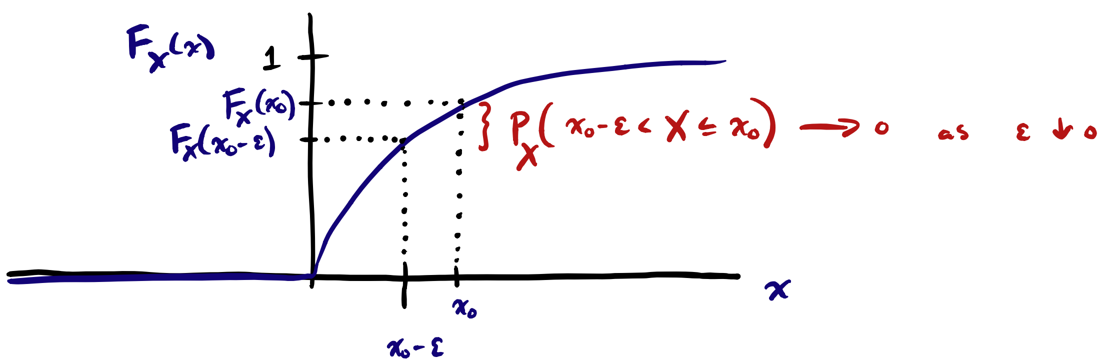
continuity of F_X => can pass limit inside

$$\lim_{\varepsilon \downarrow 0} F_X(x) - F_X(x-\varepsilon) = F_X(x) - F_X\left(\lim_{\varepsilon \downarrow 0} x-\varepsilon\right) = F_X(x) - F_X(x) = 0.$$

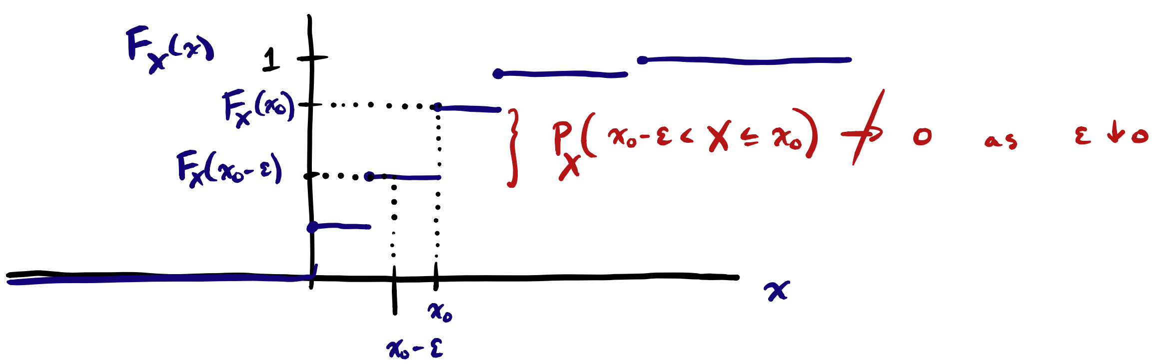
This gives $P_X(X=x) = 0$.

Illustration:

F_X continuous:



F_X a step function:



Thus for a continuous r.v. X, for any $a, b \in \mathbb{R}$,

$$P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = F_X(b) - F_X(a).$$

We are very often interested in more than one r.v. at a time.

Defn: Two r.v.s X and Y on the same sample space S with the same range $\mathcal{X}$ are identically distributed if for every $A \in \xi_{\mathcal{X}}$

$$P_X(X \in A) = P_Y(Y \in A)$$

All events of interest in the range $\mathcal{X}$

Theorem: The following two statements are equivalent

(a) The r.v.s X and Y are identically distributed

(b) $F_X(x) = F_Y(x)$ for every x

* Equivalence means each statement implies the other.

* From now on, say that two r.v.s are identically distributed if they have the same cdf.

E.g. X = # coin flips to get a head on Monday

Y = # coin flips to get a head on Tuesday

Then $F_X(x) = F_Y(x)$ for every x.

Because the universe is like that.