Random Variables

Ex. Flip a coin 3 times.

X = # heads among 3 flips

$s \in S$	$X(s)$
HHH	3
HHT	2
HTH	2
HTT	1 *
THH	2
THT	1 *
TTH	1 *
TTT	0

← a function

Defn: Random Variables

A random variable (RV) X is a function

$$X : S \to \mathbb{R}$$

also called a <u>random variate</u>

or <u>real-valued random variable</u>

or a <u>univariate random variable</u>

($\mathbb{R}$ not $\mathbb{R}^n$)

Ex.

1. toss two dice

   X = sum of two dice

2. toss a coin 25 times,

   X = length of the longest chain of consecutive Hs

3. observe rainfall

   X = crop yield

We'd like to say, e.g.

$$\mathbb{P}(\underbrace{X=1}_{\text{abuse of notation}})$$

recall: $\mathbb{P} : 2^S \to \mathbb{R}$

what we really mean

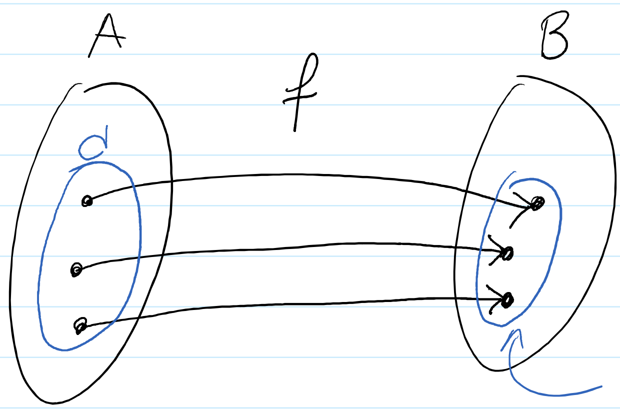
X = # heads in 3 flips

$$\mathbb{P}(X=1) = \mathbb{P}(\{HHT, THT, TTH\}) = 3/8$$

"$X = 1$" short-hand for

$$\underbrace{\{s \in S \mid X(s) = 1\}}_{\text{inverse image of } \{1\} \text{ under } X} \subset S$$

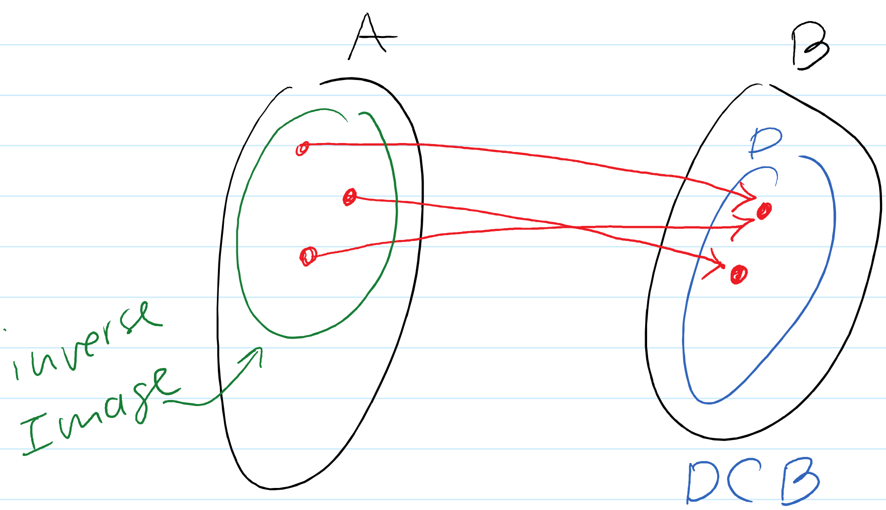
Review:



$Im(C) = f(C) \subset B$

$C \subset A$

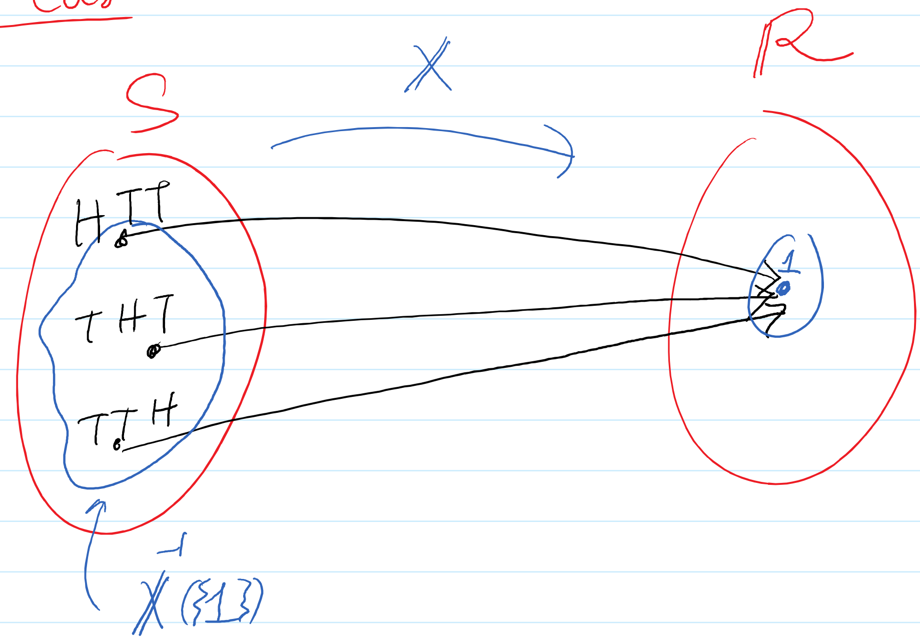
Image: $f(C) = \{f(x) \mid x \in C\}$



$D \subset B$

Inverse Image: $f^{-1}(D) = \{a \in A \mid f(a) \in D\}$

Prob. case:



Carefully:

$$\mathbb{P}(X = 1) = \mathbb{P}(\underbrace{X^{-1}(\{1\})}_{\subset S})$$

Notation: If X is a RV and $A \subset \mathbb{R}$

we write

$$\mathbb{P}(X \in A)$$

which means

$$\mathbb{P}(X^{-1}(A)).$$

Ex. X = # heads in 3 coin tosses

$P(X = 1 \text{ or } 2)$

$= P(X \in \{1,2\})$

$= P(X^{-1}(\{1,2\}))$

$= P(\{HTT, THT, TTH, HHT, HTH, THH\})$

$= 6/8$

Defn: Support of a RV

If X is a RV its support is the set of possible values of X, the image of S under X.

Ex. Cont. prev. ex.

$$\text{Support}(X) = \{0, 1, 2, 3\}$$

Notice $P(X=5) = 0$

more generally, if $A \subset \mathbb{R}$, $\text{Support}(X) \cap A = \emptyset$

then $P(X \in A) = 0$.

Informal Defn Types of RVs.

1. discrete: support that is finite or countable

Ex X = sum of two dice

countable

ex. X = sum of two dice

ex. X = # customers visiting shop

(2) continuous: support is uncountably infinite

ex. X = waiting time for a bus

support $= [0, \infty)$

Defn: Cumulative Distribution Function (CDF)

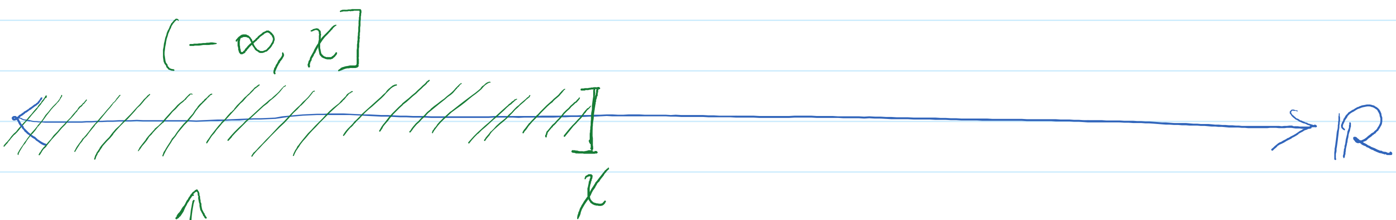
If X is a RV then its CDF is a function $F : \mathbb{R} \to \mathbb{R}$

defined for $x \in \mathbb{R}$

a number

$$F(x) = \mathbb{P}(X \leq x)$$

a RV



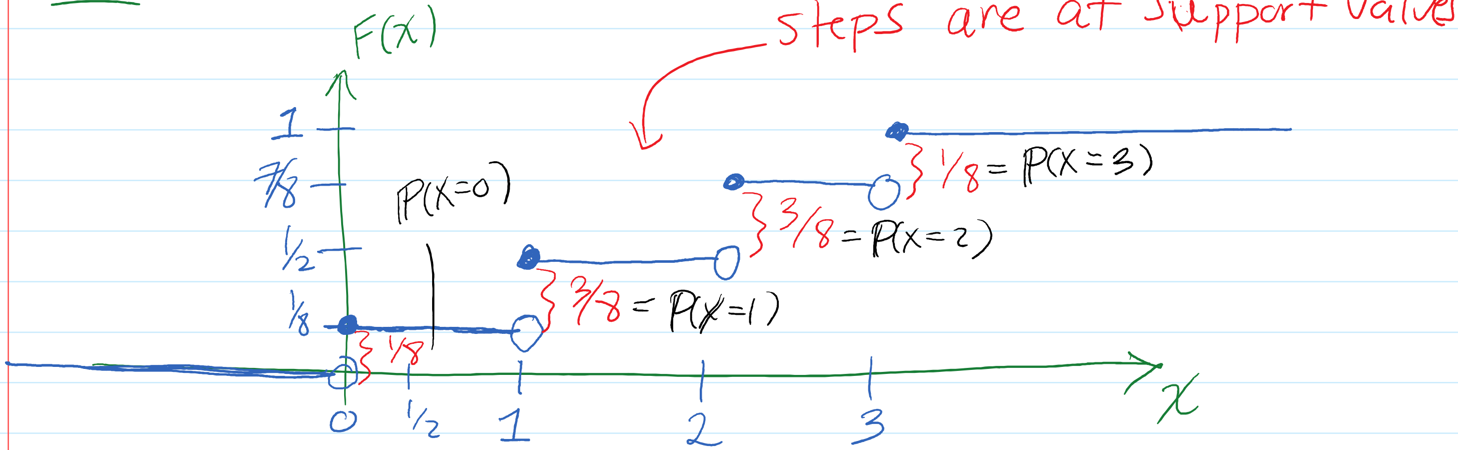
F(x) is prob of being here

I wrote $\mathbb{P}(X \leq x)$

$$= \mathbb{P}(X \in (-\infty, x])$$

$$= \mathbb{P}(X^{-1}((-\infty, x]))$$

Ex. Toss a coin 3 times.



$$F(0) = \mathbb{P}(X \le 0) = \mathbb{P}(X = 0) = 1/8$$

$$F(1/2) = \mathbb{P}(X \le 1/2) = \mathbb{P}(X = 0) = 1/8$$

$$F(.9) = \mathbb{P}(X \le .9) = \mathbb{P}(X = 0) = 1/8$$

$$F(1) = \mathbb{P}(X \le 1) = 4/8 = 1/2$$

$$F(1.5) = \mathbb{P}(X \le 1.5) = \mathbb{P}(X \le 1) = 1/2$$

$$F(2) = \mathbb{P}(X \le 2) = 7/8$$

$$F(3) = \mathbb{P}(X \le 3) = 1$$

$$F(4) = \mathbb{P}(X \le 4) = 1$$

$$F(-1) = \mathbb{P}(X \le -1) = 0$$

Facts:

① $\boxed{0 \le F(x) \le 1}$

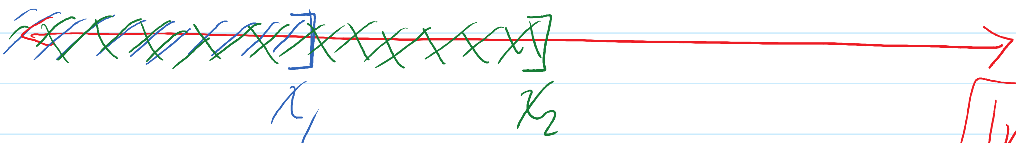
Pf. $F(x) = P(\;\sim\;) \in [0,1]$

(2) $\boxed{\lim_{x\to-\infty} F(x) = 0, \quad \lim_{x\to\infty} F(x) = 1}$

(3) $\boxed{F \text{ is non-decreasing}}$

Pf. $x_1 < x_2$ then $F(x_1) \le F(x_2)$

$(-\infty, x_1] \subset (-\infty, x_2]$



$$\begin{array}{ccc}
F(x_1) & & F(x_2) \\
= & & = \\
P(X \le x_1) & & P(X \le x_2) \\
= & & = \\
P(X \in \underbrace{(-\infty, x_1]}) & \subset & P(X \in \underbrace{(-\infty, x_2]}) \\
= & & = \\
P(\underbrace{X^{-1}((-\infty, x_1])}) & \le & P(\underbrace{X^{-1}((-\infty, x_2])}) \\
 & \subset &
\end{array}$$

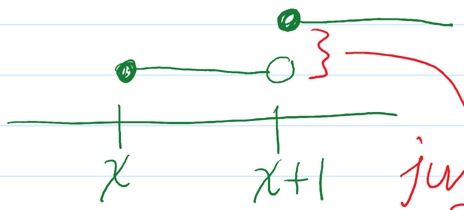
Inverse images preserve subset rel. $A \subset B$

$f^{-1}(A) \subset f^{-1}(B)$

$A \subset B$ then $P(A) \le P(B)$.

(4) $\boxed{P(a < X \le b) = F(b) - F(a)}$

Use case: discrete RVs

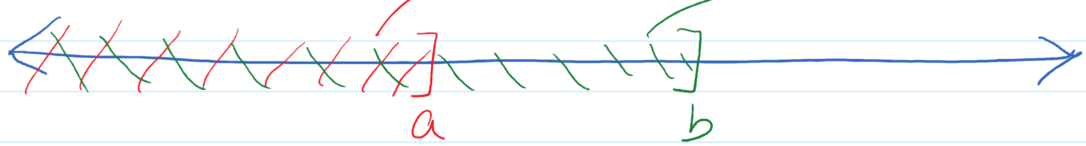


jump size

$$P(x < X \le x+1)$$
$$= P(X = x+1)$$
$$= F(x+1) - F(x)$$

justification:

$$P(X \in (a, b]) = P(a < X \le b) = F(b) - F(a)$$



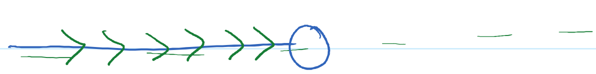
$$(-\infty, b] \setminus (-\infty, a] = (a, b]$$

5 $\boxed{F \text{ is right-continuous}}$

recall: cts fn g

$$\lim_{x \to a} g(x) = g(a)$$

right-cts: $\lim_{x \to a^+} F(x) = F(a)$



Note: if a fn is continuous then it is right continuous

Theorem: The function $F: \mathbb{R} \to \mathbb{R}$ is the CDF of some RV iff

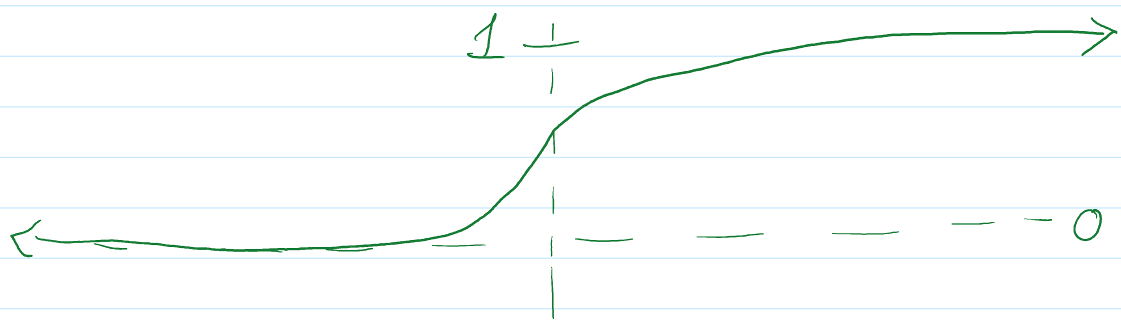
1. $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$

2. F is non-decreasing

3. F is right-continuous.

Ex. Let

$$F(x) = \frac{1}{1+e^{-x}}$$



Q: Is this a valid CDF?

Check 3 conditions:

1. $\lim F(x) = \frac{1}{1+e^{-(-\infty)}} = \frac{1}{1+e^{\infty}} = \frac{1}{\infty} = 0$

① $\lim_{x \to -\infty} F(x) = \frac{1}{1+e^{-(-\infty)}} = \frac{1}{1+e^{\infty}} = \frac{1}{\infty} = 0$

$\lim_{x \to \infty} F(x) = \frac{1}{1+e^{-\infty}} = \frac{1}{1+0} = 1$

② non-decreasing ?

✓ $\frac{dF}{dx} = \frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right) = \frac{e^{-x}}{(1+e^{-x})^2} > 0$

③ right continuous ? it

Yes, is continuos / differentiable.

So F is a valid CDF.
