This week: 2.1 - 2.5 Discrete and continuous r.v.

2.1 Random variables (r.v.)

A universal function containing all info about probabilities related to a general r.v. X is distribution function (df or cdf) of X defined as

$$F(x) = P(X \le x), \quad -\infty < x < \infty.$$

Some basic properties:

① $F(a) \le F(b)$, if $a < b$, $\quad 0 \le F(x) \le 1$

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

Why? if $a < b$, then $\{X \le a\} \subset \{X \le b\}$, and $\{a < X \le b\} = \{X \le b\} \setminus \{X \le a\}$. Hence

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

② $P(X = x) = F(x) - F(x-)$, where $F(x-)$ is the left-hand limit of F at x.

because $\{X = x\} = \bigcap_{n=1}^{\infty} \{x - \frac{1}{n} < X \le x\}$ ← decreasing sequence,

$$P(X = x) = \lim_{n\to\infty} P(x - \tfrac{1}{n} < X \le x) = \lim_{n\to\infty} \left[F(x) - F(x - \tfrac{1}{n})\right] =$$

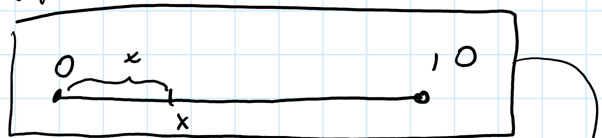
$$= F(x) - \lim_{n\to\infty} F(x - \tfrac{1}{n}) = F(x) - F(x-).$$

Ex 1. Let X be uniform in $(0, 10)$:

for $0 < a < b < 10$ $\quad P(a < X < b) = \frac{b-a}{10}$.

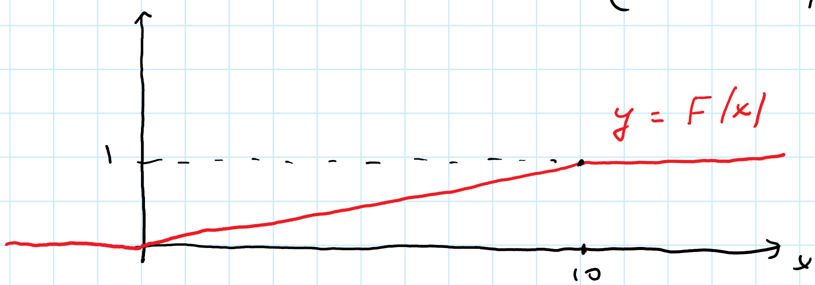
for $0 < a < b < 10$, $P(a < X \le b) = \frac{b-a}{10}$.

Find df of X and sketch it.



Answer

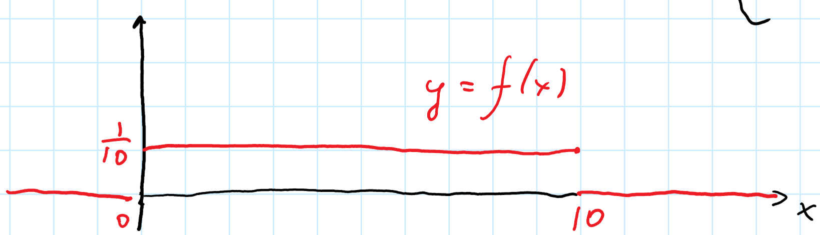
$$F(x) = P(X \le x) = \begin{cases} 0, & x \le 0 \\ x/10, & 0 < x \le 10 \\ 1, & x > 10 \end{cases}$$



$F(x)$ is continuous:

$F(x) - F(x-) = P(X = x) = 0$.

Comment: $f(x) = F'(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{10}, & 0 < x < 10 \\ 0, & x > 10 \end{cases}$ is pdf of X



all values in $(0,10)$ "equally likely"

By FTC, $F(x) = \int_{-\infty}^{x} f(t)\,dt$, $-\infty < x < \infty$.

Ex2. Let X be binomial$(n=2, p=\frac{1}{2})$.

Find df of X and sketch it.

Answer Range of $X = \{0, 1, 2\}$, pmf of X:

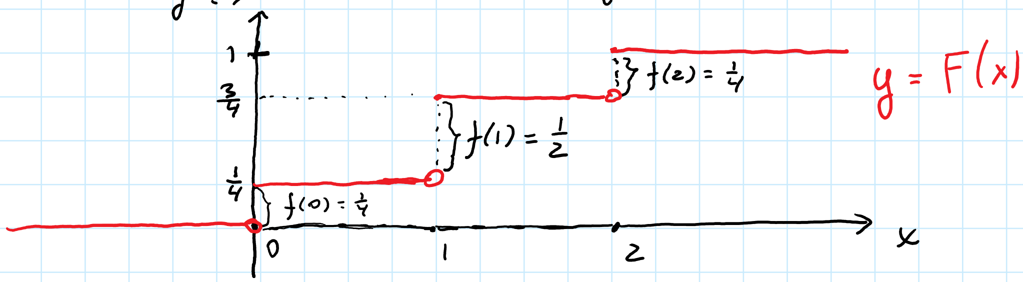
$f(0) = P(X=0) = (\frac{1}{2})^2 = \frac{1}{4} = f(2) = P(X=2)$

$f(1) = P(X=1) = 2 \cdot (\frac{1}{2})^2 = \frac{1}{2}$. Then

$F(x) = P(X \le x) = \sum P(X=y) = \sum f(y) =$

$$F(x) = P(X \le x) = \sum_{y \le x} P(X = y) = \sum_{y \le x} f(y) =$$

$$= \begin{cases} 0, & x < 0 \\ \frac{1}{4}, & 0 \le x < 1 \\ \frac{1}{4} + \frac{1}{2} = \frac{3}{4}, & 1 \le x < 2 \\ \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1, & 2 \le x \end{cases}$$



<u>Properties of general df $F(x) = P(X \le x) = P(X \in (-\infty, x])$</u>

① $0 \le F(x) \le 1$; $F(x) \le F(y)$ if $x \le y$

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

③ $F(x)$ is right-hand continuous and has left-hand limits.

<u>Why?</u> For instance, $\{X < \infty\} = \bigcup_{n=1}^{\infty} \{X \le n\}$ (↖ increasing)

$1 = P(X < \infty) = \lim_{n \to \infty} P(X \le n) = \lim_{n \to \infty} F(n)$.

$$\boxed{\text{Def. A function } F: \mathbb{R} \to [0,1] \text{ satisfying ①, ②, ③ is called df.}}$$

<u>More properties of F:</u>

4. $P(X = b) = F(b) - F(b-)$

5. $P(a < X \le b) = F(b) - F(a)$ | $P(X > a) = 1 - F(a)$

$P(a < X < b) = F(b-) - F(a)$ | $P(X \ge a) = 1 - F(a-)$.

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

2.2. Law of averages (LLN)

Let $A_1, A_2, \ldots$ be independent with $P(A_i) = p$,
$P(A_i^c) = q = 1-p$.

Comments 1. Think about an experiment performed repeatedly whose outcome is A or "not A" $= A^c$ (A = "success", A^c = "failure"):

A_i = "A in ith experiment" = "success in ith experiment".

For instance, a coin with $P(H) = p$ is tossed repeatedly
$A_i = H_i$ = "H in the ith toss".

Consider $\boxed{S_n = \sum_{i=1}^{n} 1_{A_i}}$: it is the number of times A happened in n trials.

Recall $1_A = \begin{cases} 1 & \text{if } A \text{ happens} \\ 0 & \text{otherwise.} \end{cases}$

2. S_n is binomial (n, p):
$$P(S_n = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \ldots, n.$$

3. the quantity $\frac{S_n}{n}$ is average (sample proportion, relative frequency) of "successes" in n trials;
A's

Thm 1. (Bernoulli, 1695)
$$\frac{S_n}{n} \longrightarrow p = P(A) \text{ as } n \to \infty \text{ with probability } 1.$$
or $\boxed{\frac{S_n}{n} \approx p \text{ for large } n.}$