2.2. Law of averages (LLN)

Let $A_1, A_2, \dots$ 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):

$$P(S_n = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \dots, n.$$

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

Thm 1. (Bernoulli, 1695)

$\frac{S_n}{n} \longrightarrow p = P(A)$ as $n \to \infty$ with probability 1.

or $\boxed{\frac{S_n}{n} \approx p \text{ for large } n.}$

Steps of the proof.

1. It is shown that for each $\varepsilon > 0$ there is $\alpha \in (0, 1)$

so that

$$P\left(\left|\frac{S_n}{n}-p\right|>\varepsilon\right)\le 2a^n,\ n=1,2,\dots$$

2. We prove that only finite number of events $B_n=\left\{\left|\frac{S_n}{n}-p\right|>\varepsilon\right\}$ can happen, equivalently, equivalently, <u>for large n</u>

$$\left|\frac{S_n}{n}-p\right|\le\varepsilon: \quad \frac{S_n}{n}\approx p \text{ if } n \text{ is large.}$$

<u>Ex1</u>. Assume $P(B_n)\le C a^n$, $n\ge 1$, with $a\in(0,1)$. Find $P(\text{infinitely many } B_n)$.

<u>Answer</u>. $$P\Big(\bigcap_{n=1}^{\infty}\underbrace{\bigcup_{m\ge n}B_m}_{\uparrow\text{decreasing}}\Big)=\lim_{n\to\infty}P\Big(\bigcup_{m\ge n}B_m\Big)\le$$

$$\le\lim_n\sum_{m=n}^{\infty}P(B_m)\le\lim_n\sum_{m=n}^{\infty}Ca^m=\lim_n\frac{Ca^n}{1-a}=0$$

<u>Application</u>. Let $X_1, X_2, \dots$ be independent r.v. with the same df $F(x)$.

We say X_i are independent observations (copies) of X whose df is $F(x)$.

Consider

$$F_n(x)=\frac{\sum_{i=1}^{n}1_{\{X_i\le x\}}}{n},\quad -\infty<x<\infty.$$

$F_n(x)$ is relative frequency of $\{X\le x\}$ in n independent observations:

$A_i=1_{\{X_i\le x\}}$ are independent, $P(A_i)=P(X_i\le x)=$

$= F(x) = p$. By Bernoulli thm, $F_n(x) \approx F(x)$ for large n, $-\infty < x < \infty$.

$F_n(x)$ is called <u>empirical distribution function</u>.

<u>2.3</u>. Discrete and continuous r.v.

<u>Def</u>. X is <u>discrete r.v.</u> if X takes values in a finite or countable set.

<u>Claim</u>. All probability related to X can be found using the pmf of X

<u>Def</u>. The <u>pmf of X</u> is the function

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

<u>Note</u> 1. $P(X \in A) = \sum_{x \in A} P(X=x) = \sum_{x\in A} f(x)$.

2. df of X is the function

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

and $f(x) = P(X=x) = F(x) - F(x-)$

<u>Examples</u> 1. X is binomial (n, p),

$$f(x) = P(X=x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0, 1, \dots, n \\ 0, & \text{otherwise} \end{cases} \quad \Big| \quad q = 1-p$$

The most likely values of X are "at" $\boxed{n \cdot p}$

Recall for $n=2$, $p=\frac{1}{2}$, $P(X=0) = P(X=2) = \frac{1}{4}$, $P(X=1)=\frac{1}{2}$.

2. Coin with $P(H)=p$ is tossed repeatedly.

Let $X = \#$ of tosses needed for the 1st H to show up

X is waiting time for H. Then

Range of $X = \{1, 2, \quad \}$.

$$P(X=k) = P(T_1 \dots T_{k-1} H_k) = q^{k-1} p = \frac{q}{p} e^{k \overbrace{\ln q}^{<0}}, \; k=1,\dots$$

X is called geometric(p)

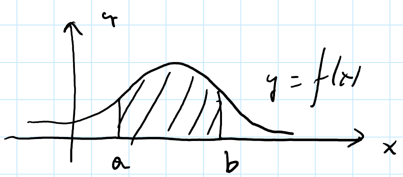
2. Continuous r.v.

Def. X is continuous r.v. if there is $f \geq 0$ on $\mathbb{R}$ so that df of X is given by

$F(X) = F_X(x) = \int_{-\infty}^{x} f(t)\,dt$, $-\infty < x < \infty$. | f is called pdf of X.

Note

1. $F(x)$ is continuous, $P(X=a) = F(a) - F(a-) = 0$,
2. $P(a < X \leq b) = F(b) - F(a) = \int_a^b f(x)\,dx$ = shadowed area



$P(X \in A) = \int_A f(x)\,dx$

3. $P(a < X \leq b) = \int_a^b f(x)\,dx \approx f(a)(b-a)$ if $a \approx b$
4. $f(x) = F'(x)$ for any x at which f is continuous.

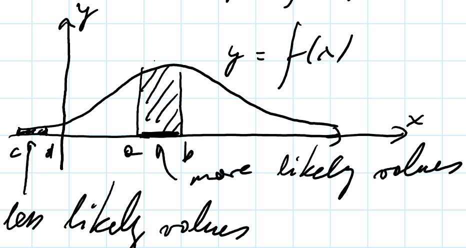
Recall $f(x) = F'(x) = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} \frac{P(x < X \leq x+h)}{h}$

, $\frac{P(x < X \leq x+h)}{}$ is average pdf in $(x, x+h]$.

Recall $f(x) = F'(x) = \lim_{h \to 0} \quad h \quad h \to 0 \quad h$

and $\dfrac{P(x < X \le x+h)}{h}$ is average pdf in $(x, x+h]$.

<u>Remark 1</u> pdf $f(x)$ is more "visual" than df $F(x)$



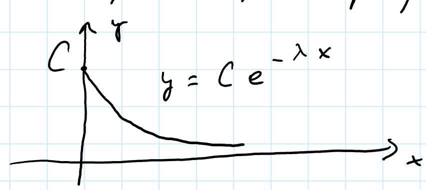
Values of X between a and b more likely than between c and d.

<u>Def.</u> $f \geq 0$ on $\mathbb{R} = (-\infty, \infty)$ is called pdf if

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

<u>Ex 1.</u> a) For what C, $f(x) = Ce^{-\lambda x}$, $x > 0$, is pdf?

Here $\lambda > 0$, $f(x) = 0$ if $x \leq 0$.



<u>Answer.</u>