Background on Probability

"Experiment" $\longrightarrow$ Outcome $\rightarrow$ numbers

sample space Ω = {all outcomes}

Events : A, B, ... subset of Ω

A happens if $\overset{w}{\text{outcome}}$ $\in A$

$P(A)$, $P(B)$

outcome w/ $X(w)$, $Y(w)$ ... random variables

A canonical example (1)

$$\underset{\Omega}{\text{population}} \xrightarrow[\text{sampling}]{\text{random}} \underset{w}{\text{person}} \rightarrow \text{numbers}$$

$X(w)$: gender, discrete

$Y(w)$: height, continuous

A : the person is male (logical statement)

male sub-population (subset)

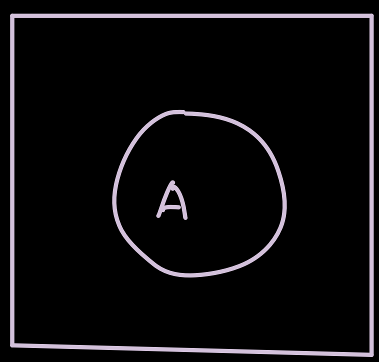
$\{w : X(w) = \text{male}\}$ $X=1$

$P(A) = P(X=1)$

Axiom 0: If all the outcomes are equally likely, then

$$P(A) = \frac{|A| \rightarrow \text{size of } A}{|\Omega| \rightarrow \text{size of } \Omega} \quad \text{(population proportion)}$$

Canonical example (2)



Ω

$\Omega = \{$ all points in $\square$ $\}$

uncountably infinitely many

Sampling a point

= randomly throw a point into Ω

$P($fall into $A) = P(A)$

$$= \frac{|A| \rightarrow \text{area of } A}{|\Omega| \rightarrow \text{area of } \Omega}$$

Probability is a **MEASURE**

Axiom 1 : $P(\Omega) = 1$

" " 2 : $P(A) \geq 0$

" " 3 : $A \cap B = \emptyset$, $P(A \cup B) = P(A) + P(B)$

" " 1D: length 2D: area 3D: volume

Monte Carlo:

population $\xrightarrow{\text{sampling}}$ random sample

$\downarrow$ N people $\qquad\qquad$ $\downarrow$ n people

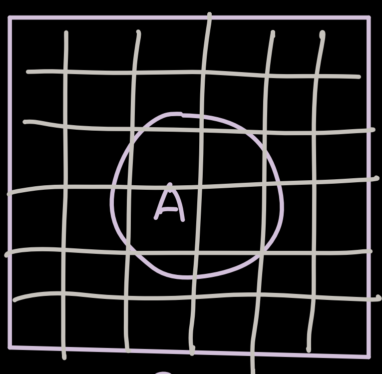
proportion of males $\xleftarrow{\text{estimate}}$ sample proportion

avg height of population $\longleftarrow$ sample average

If n is large enough $\rightarrow$ error of estimate is controllable

$\rightarrow$ variance can be reduced

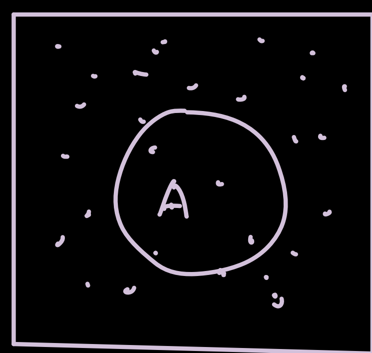
population



Ω

area $|\Omega| = 1$

sample space



n points

$|A| = P(A) \approx$ how often points fall into A

frequency

Infer population property from sample properties

(testing) $\qquad$ (training)

Random number generators

$$U \sim \text{Unif}[0, 1]$$

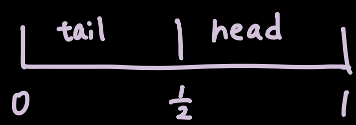
If $U \geq \frac{1}{2}$, $X =$ head (1)

$U < \frac{1}{2}$, $X =$ tail (0)

transform U to X

Inversion Method

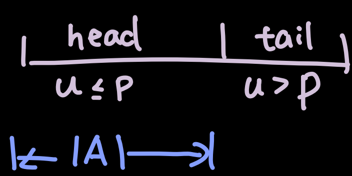
fair coin



biased coin

(customizable)

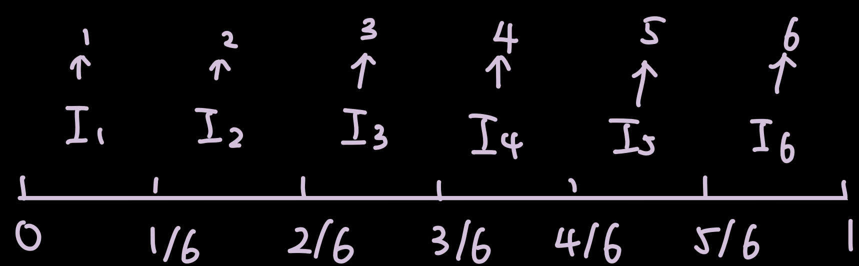
$$p(\text{head}) = p = \frac{|A|}{|\Omega|}$$ axiom 0



In 1D, $|A| =$ length

$|\Omega| = 1$

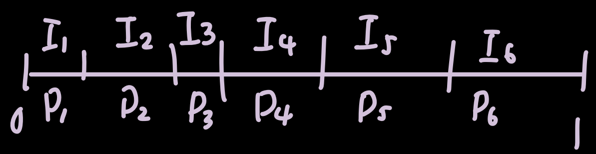
fair Die



If $U \in I_1$, $X = 1$

$U \in I_2$, $X = 2$

biased die



p_1 = length of I_1

p_2 = length of I_2

Probability Distribution

x	1	2	3	4	5	6
$p(x)$	p_1	p_2	p_3	p_4	p_5	p_6

$X \sim p(x)$ ← lower case

upper-case

Probability Mass Function

↓

random number that follows PMF.

$P(X = x) = p(x)$

Cumulative distribution function

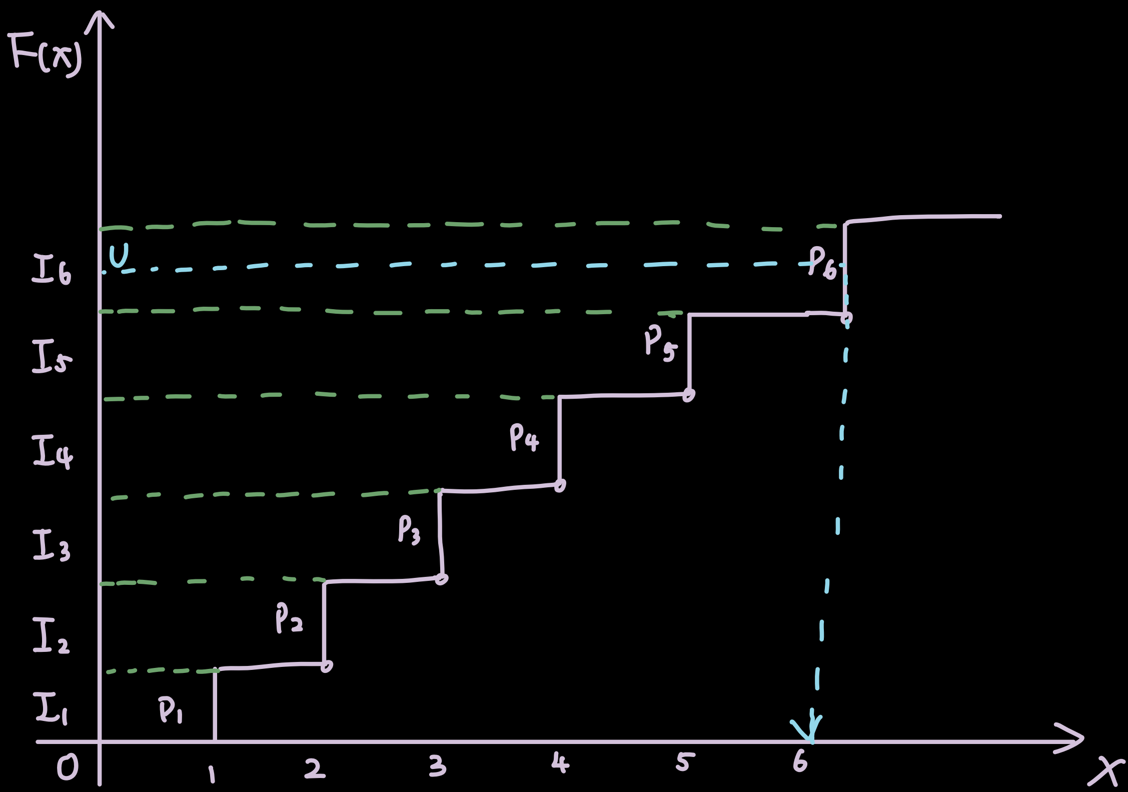
$F(x) = P(X \le x)$

↓

$x \in \mathbb{R}$

(all real #s)

$$F(x) = \begin{cases} 0 & x < 1 \\ p_1 & 1 \le x < 2 \\ p_1 + p_2 & 2 \le x < 3 \\ p_1 + p_2 + p_3 & 3 \le x < 4 \\ p_1 + p_2 + p_3 + p_4 & 4 \le x < 5 \\ p_1 + p_2 + p_3 + p_4 + p_5 & 5 \le x < 6 \\ 1 & x \ge 6 \end{cases}$$



$$X = F^{-1}(U)$$

Continuous Random Variables

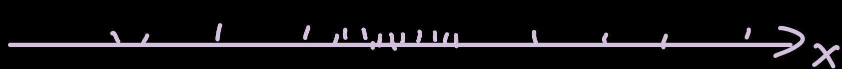
population $\xrightarrow{\text{sampling}}$ person $\longrightarrow$ height

Ω $\qquad$ ω $\qquad$ $X(\omega)$

X

e.g. 300 million people

scatter plot



Population density

$$\text{Density } f(x) = \frac{\#\text{ of points in } (x, x+\Delta x)/N}{\Delta x}$$

$$= \frac{N(x)/N}{\Delta x}$$

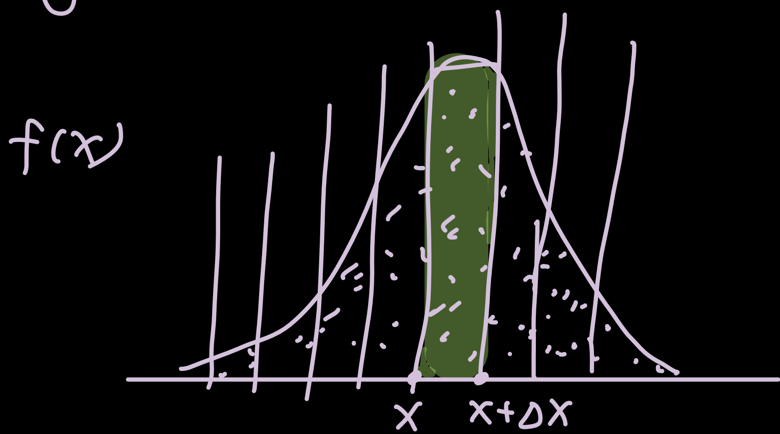
→ population proportion

→ area

$$= \frac{P(X \in (x, x+\Delta x))}{\Delta x} \quad \left(\lim_{\Delta x \to 0} \right)$$

↳ continuous

Histogram



each person is a small ball

small balls on top of each other

distribute balls into bins

probability of proportion $P(X \in (x, x+\Delta x))$

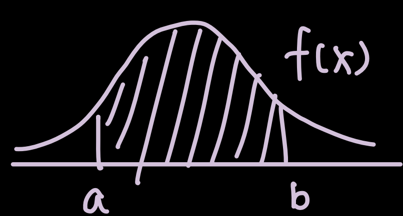
$$= f(x)\,\Delta x$$

$$= \text{Area of bin } (x, x+\Delta x)$$

$$P(X \in (a,b)) = \sum_{\text{bins} \in (a,b)} f(x)\,\Delta x$$

$$\xrightarrow[0]{\Delta x} \quad = \int_a^b f(x)\,dx$$

$=$



Another POV : Math definition of a histogram above

Ω = region under $f(x)$

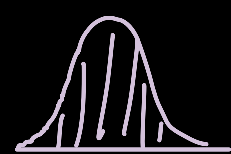
randomly sample a point ω

$\to (X(\omega), Y(\omega))$

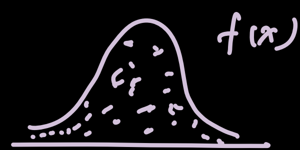
$X \sim f(x)$

$P(X \in (x, x+\Delta x)) =$

$= f(x)\Delta x$



$= 1$

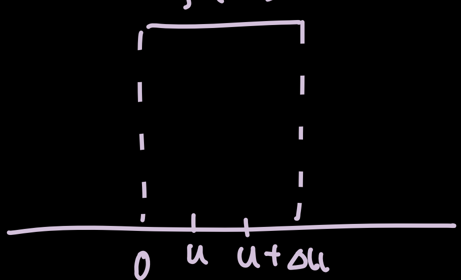


$\Downarrow$ project onto 1D

denser in the middle

$U \sim \text{Unif}[0, 1]$

$$f(u) = \begin{cases} 0 & u < 0 \\ 1 & 0 \le u \le 1 \\ 0 & u > 1 \end{cases}$$

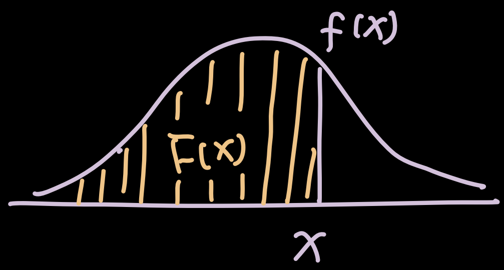


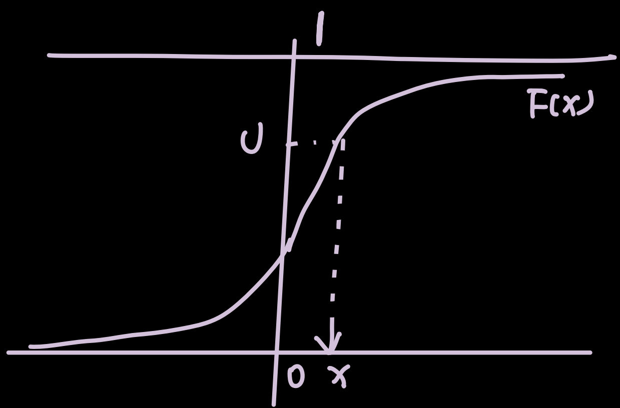
$P(U \in (u, u+\Delta u)) = \Delta u$

Want to generate $X \sim f(x)$ random #s

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)\,dx$$

cumulative density function





$U \sim Unif[0, 1]$

$X = F^{-1}(U)$ solve $F(x) = U$ for x.

Inversion method under continuous case is same as under discrete case.

e.g. $X \sim N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

e.g.

SAT	Percentile
1450	99%
x	F(x)
$F^{-1}(u)$	u