Events vs Variables

- Assign outcomes of experiments to variables
- But why?
- Example: 6-side dice rolled

$\{6\} \rightarrow \mathbb{N}$

$\{1, 6\}$

$$N \in \mathbb{N}, \quad \overbrace{1 \leq N \leq 6}^{\text{constraint}}$$

type

7

RV

Maths	Constant	Variable	
Programming		Const	Variable

Never changes

Assigned once.
Never changes afterwards

$x = 3$

~~3~~5

Can always change

$x=3$
$x=5$
$x=7$
$x=$"string"

Random variables (RVs)

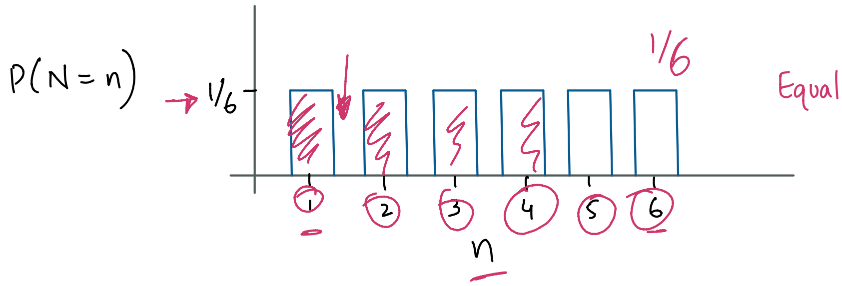
- A RV can take on a value from any set. Each value has a probability associated with it.

	n=1	n=2				n=6
N	1	2	3	4	5	6
$P(N=n)$	1/6	1/6	1/6	1/6	1/6	1/6

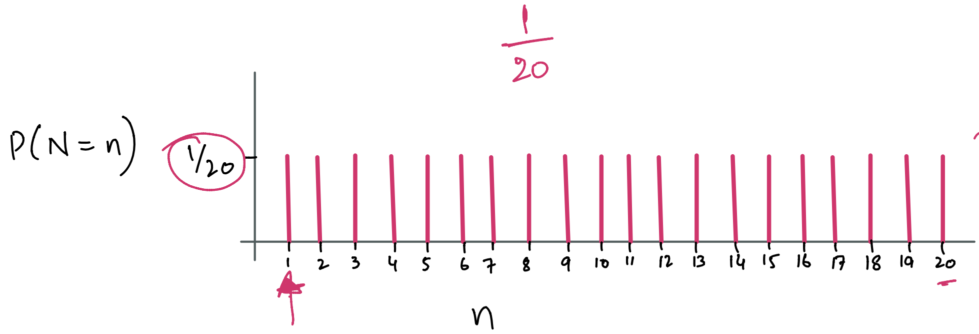
- Other examples:
  - Pick any person and measure their height
  - Pick any character and convert it to ASCII.

$E \rightarrow \square \; \mathbb{R}$

$\mathbb{N}$



— Similar to the histogram we saw earlier ...



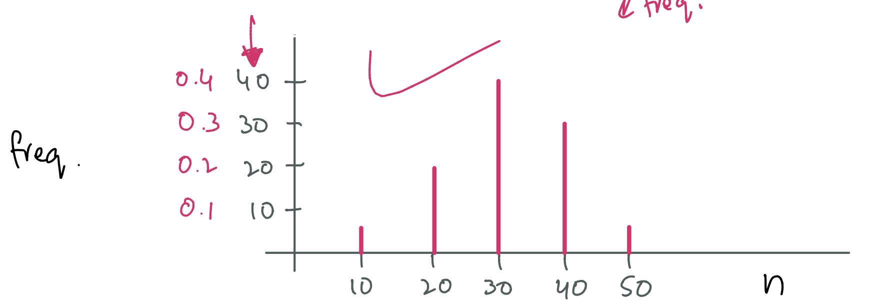
This is called a "distribution".

N is a 'discrete random variable'.

(Side note: Another example)

Measure ages of 100 people.

freq. distribution



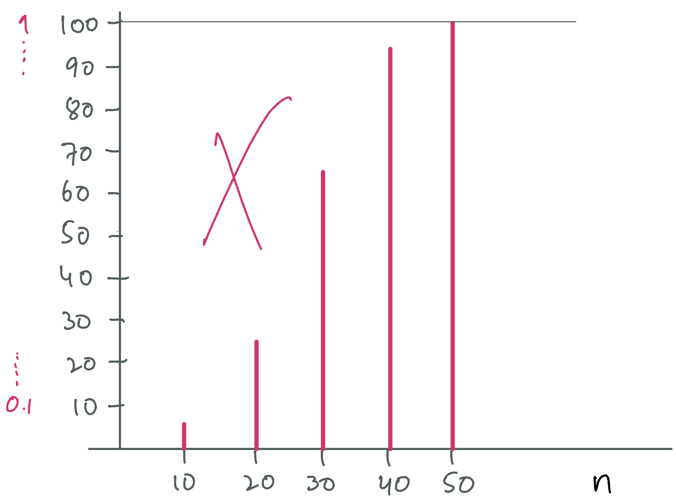
Freq. table

~~Weight~~ Age	Frequency
10	5/100
20	20/100
30	40/100
40	30/100
50	5/100
	100

This is a "probability frequency distribution". (PFD)

Cumulative freq

CFD

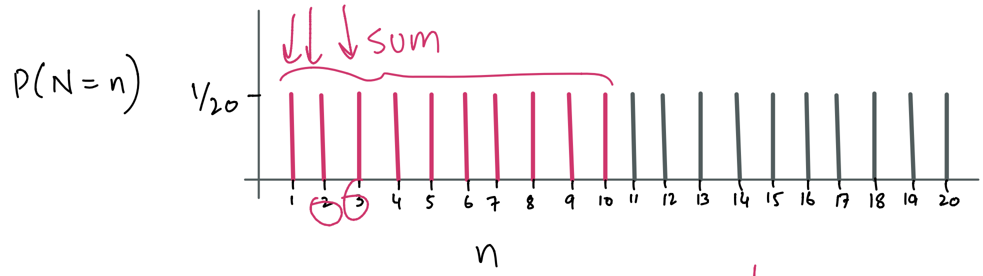


~~Weight~~ Age	Frequency
10	5/100
20	25/100
30	65/100
40	95/100
50	100/100 → 1

This is "cumulative freq distribution" (CFD)

Back to distributions:

$P(N <= 10)$



$$P(N \leq 10) = \sum_{i=1}^{10} P(N=i)$$ (mutually exclusive)

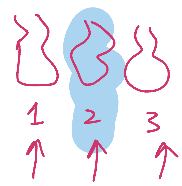
$P(N \leq 20)$ must equal 1

$P(N < 1)$ " " 0

How about 2 random variables?

" Bag i has i blue balls and 2 green balls".

Say $i = 6$



B=2

Random variables: B = bag picked.
C = color of ball picked

1 = blue 2 = green

4 6

P(4) + P(6)
→ OR

P(B=1, C=1)
+ → OR
P(B=1, C=2)

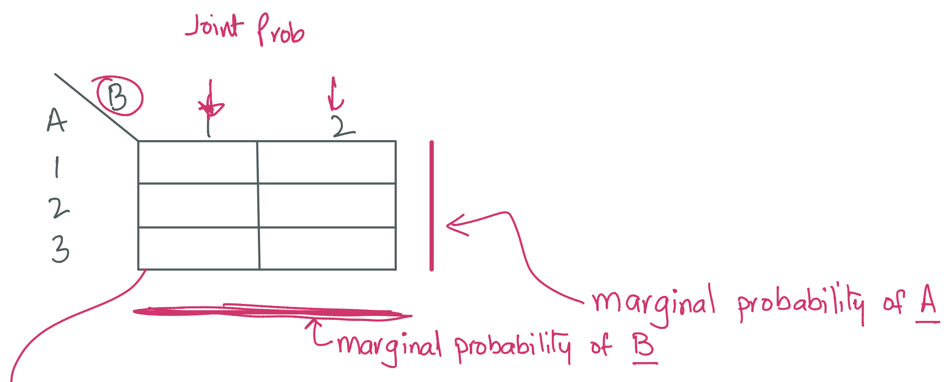
$P(B=b, C=c)$

$p(c=1)$

$P(\Omega)=1$

B \ C	C = 1	C = 2	P(B=b)
B = 1	1/18	= 2/18	1/6
B = 2	2/24	= 2/24	1/6
B = 3	= 3/30	= 2/30	1/6
B = 4	= 4/36	= 2/36	1/6
B = 5	= 5/42	= 2/42	1/6
B = 6	= 6/48	= 2/48	1/6
P(C=c)	p(c=1) = 0.594	p(c=2) = 0.405	1

1



Joint probability distribution of A and B

$$P(B=1) = \sum_i P(B=1, C=i)$$

$$P(B=1, C=1) + P(B=1, C=2)$$

(same rule of summation of mutually exclusive events)

$$P(C=1) = \sum_{i=1}^{6} P(B=i, C=1)$$

"Sum over all possible values of C."

But we already knew the 1/6 for $P(B=1)$!!

why we do we need to do this summation?

B \ C	C = 1	C = 2	P (B=b)
B = 1	1/18	2/18	
B = 2	2/24	2/24	
B = 3	3/30	2/30	
B = 4	4/36	2/36	
B = 5	5/42	2/42	
B = 6	6/48	2/48	

When we collect data from real world:

— Sensor (signal + noise)

We have important RVs and noise mixed together!

We are only able to measure joint probabilities.

But we are interested in marginals of one RV.

— Joint $\longrightarrow$ marginals