Probability theory - a crash course

04 March 2021 15:55

1) PROBABILITY SPACE – a set of all possible outcomes of a random experiment

S – set

finite
countable → in this course

uncountable

EXAMPLE – a set of all n-bit strings

2.) EVENTS – $E \subseteq S$

Example – a string with exactly 3 symbols '1'

3.) PROBABILITY FUNCTION

$$p: S \to [0,1]$$

$$\sum_{i \in S} p(i) = 1$$

Example – Uniform distribution of all n-bit strings $p(x) = \frac{1}{2^n}$

$$P(E) = \sum_{i \in E} p(i)$$

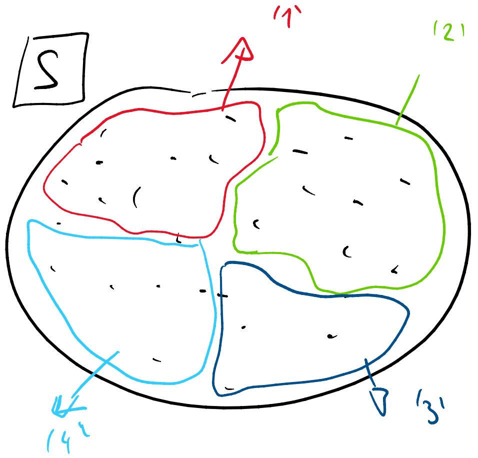
Example – What is the probability to obtain a 5-bit string with exactly three symbols '1' (call it event E)

$$P(E) = \sum_{i \in E} p(i) = \sum_{i \in E} \frac{1}{32} = \binom{5}{3} \cdot \frac{1}{32} = \frac{10}{32}$$

RANDOM VARIABLES

(X, Y, Z)

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



Essentially X is a division of probability space into mutually exclusive and collectively exhaustive set of events

Example: X - is the number of symbols '1' in n-bit string.

Example: For $n = 4$, what is the probability distribution of X.

$\Pr[X=0] = 1/16$

$\Pr[X=1] = 4/16$

$\Pr[X=2] = 6/16$

$\Pr[X=3] = 4/16$

$\Pr[X=4] = 1/16$

$\Pr[X > 1] = 11/16$

$\Pr[X \geq 2 \wedge X < 4] = 10/16$

EXPECTATION OF RANDOM VARIABLES

$$\mu(X) = E(X) = \sum_{\substack{i \in \mathbb{R} \\ \in \mathrm{Im}(X)}} i \cdot \Pr(X=i)$$

EXAMPLE: $E(X) = \sum_{i=0}^{4} i \Pr\{X=i\}$

$$= 0 \cdot \tfrac{1}{16} + 1 \cdot \tfrac{4}{16} + 2 \cdot \tfrac{6}{16} + 3 \cdot \tfrac{4}{16} + 4 \cdot \tfrac{1}{16}$$

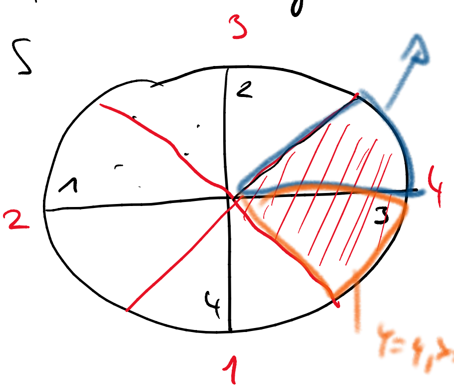
$$= 2$$

CONDITIONAL PROBABILITIES

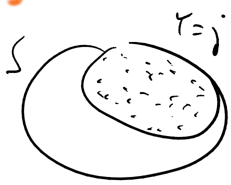
Given 2 random variables X and Y over the same random experiment define conditional probability of X given $Y=j$

$$\Pr(X=i \mid Y=j) = \Pr(X=i, Y=j) / \Pr(Y=j)$$

$[\Pr(Y=j) \neq 0]$



Intuitively we are creating a new probability space equal to event $Y=j$



$$\frac{1}{\Pr(Y=j)}$$ renormalizing original probability

<u>Examples:</u> $S = \{0,1\}^4$

X = number of '1'

Y = parity of the string (even number of '1' $Y = 0$, otherwise $Y = 1$)

$Pr\{Y = 0\} = 1/2$

$Pr\{Y = 1\} = 1/2$

$$Pr[X=3 \mid Y=0] = 0 = Pr\{X=3, Y=0\} / Pr\{Y=0\}$$

($Pr\{X=3, Y=0\} = 0$, $Pr\{Y=0\} = 1/2$)

$$Pr[X=3 \mid Y=1] = \frac{Pr\{X=3, Y=1\}}{Pr\{Y=1\}} = \frac{4/16}{1/2} = \frac{1}{2}$$

$Y=0$	$Y=1$
0000	1000
1100	0100
1010	0010
0110	0001
0011	0111
0101	1011
1001	1101
1111	1110

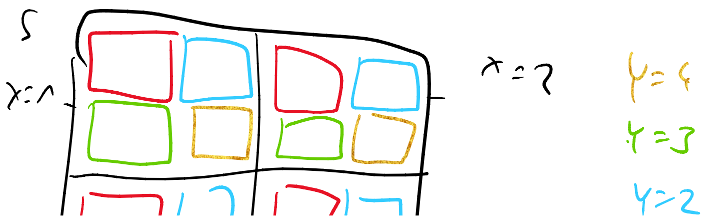
2 '1's

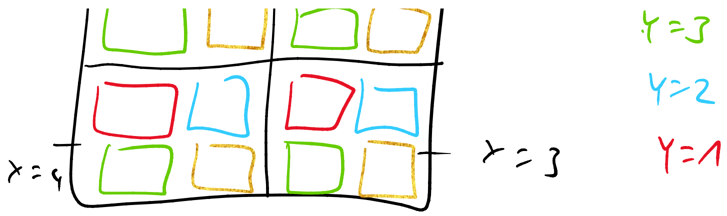
INDEPENDENCE

X and Y are independent for all i, j

$\Updownarrow$

$$Pr(X = i \mid Y = j) = Pr(X = i)$$





EXAMPLE Are X and Y from previous example independent?

X - number of '1'

Y - parity

$$Pr[X=3] = 1/4$$

$$Pr[X=3, Y=0] = 0 \quad \times$$

Z is the value of first bit of $\{0,1\}^4$

Are Z and Y independent?

$$Pr[Z=1] = 1/2$$

$$Pr[Z=0] = 1/2$$

$$Pr[Z=1 \mid Y=1] = 1/2$$

$$Pr[Z=0 \mid Y=1] = 1/2$$

$$Pr[Z=0 \mid Y=0] = 1/2$$

$$Pr[Z=1 \mid Y=0] = 1/2$$

LINEARITY OF EXPECTATION

$$W = X + Y + Z$$

w is a well defined random variable

$$E(W) = E(X+Y+Z) = E(X) + E(Y) + E(Z)$$

$$E\Big(\sum_i \underset{\text{scalars}}{a_i}\, \underset{\text{r.v.}}{X_i}\Big) = \sum_i a_i \underbrace{E(X_i)}_{\text{much easier to calculate}}$$

$$E(W) = \sum_{i,j,k} (i+j+k) \cdot Pr(X=i, Y=j, Z=k)$$

$$= \underset{2}{E(X)} + \boxed{\begin{matrix} E(Y) \\ \parallel \\ Pr(Y=0) \cdot 0 \\ + Pr(Y=1) \cdot 1 \end{matrix}} + \underset{1/2}{E(Z)} = \underset{?}{3}$$

$1/2$

$Pr(Y=1)$

$$E[X_1 \cdot X_2] \neq E(X_1) \cdot E(X_2)$$

equal if X_1 and X_2 are independent

THE LAW TOTAL PROBABILITY

r.v. X and Y

$$Pr(X=i) = \sum_{j \in im(Y)} \underbrace{Pr(X=i \mid Y=j) \cdot Pr(Y=j)}_{= Pr(X=i,\, Y=j)}$$

RQUICK SORT

IN: Collection of numbers S

OUT: Ordered list of elements in S

0.) if S contains a single element output it

1.) Choose a ==pivot== $y \in S$ ==uniformly at random==

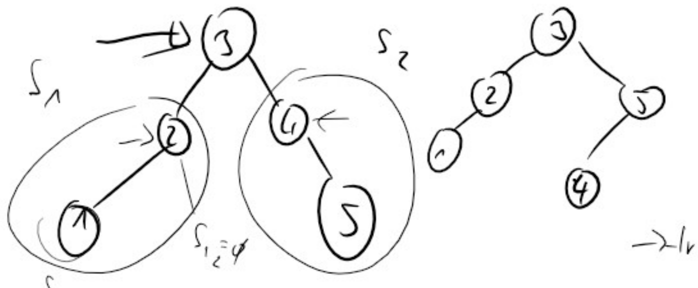
2.) Create S_1 which contains all $s \in S$, $s < y$

Create S_2 which contains all $s \in S$, $s > y$

3.) Output (rquicksort (S_1), y, rquicksort (S_2))

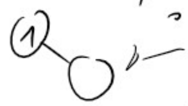
Probability space is a set of ordered trees with $|S| = n$ nodes

$S = \{1, 2, 3, 4, 5\}$



→ 4 trees with root 5

→ trees with root 1 ?

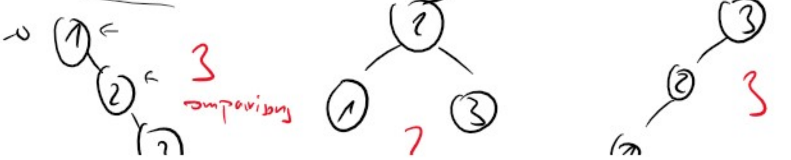


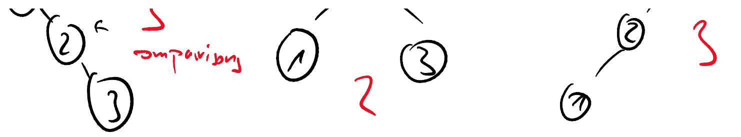
all ordered trees of size 4

$1/5 \cdot 1/2 \cdot 1/2$ $\qquad$ $1/5 \; 1/2 \; 1/2$

$n = 3$

$1/3 \cdot 1/2$ $\qquad$ $1/3$ $\qquad$ $1/3 \cdot 1/2$





How to calculate the number of comparisons

X_i - r.v. = number of comparisons in a given tree of size i

$$E(X_3) = \underbrace{4 \cdot \left(\tfrac{1}{3} \cdot \tfrac{1}{2}\right)}_{\text{4 trees with the same probability}} \cdot \underbrace{3}_{\text{num comparisons}} + \tfrac{1}{3} \cdot 2 = \boxed{\frac{8}{3}}$$

$$f(i) = E(X_i)$$

Q - how does this function scale with i?

$$O(i \cdot \log i)$$