Tutorial 1

17 February 2022 12:03

PROBABILITY THEORY – a crash course

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

S – set

finite countable → this course
uncountable ✗

Example: a set of all n-bit strings

2.) Probability events – $E \subseteq S$

Example: a string with exactly 3 symbols '1'.

3.) Probability function: – $p : S \to [0,1]$ ✓ $p(i) \geq 0$

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

Example: Uniform distribution of all n-bit strings

$$p(x) = \frac{1}{2^n}$$

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

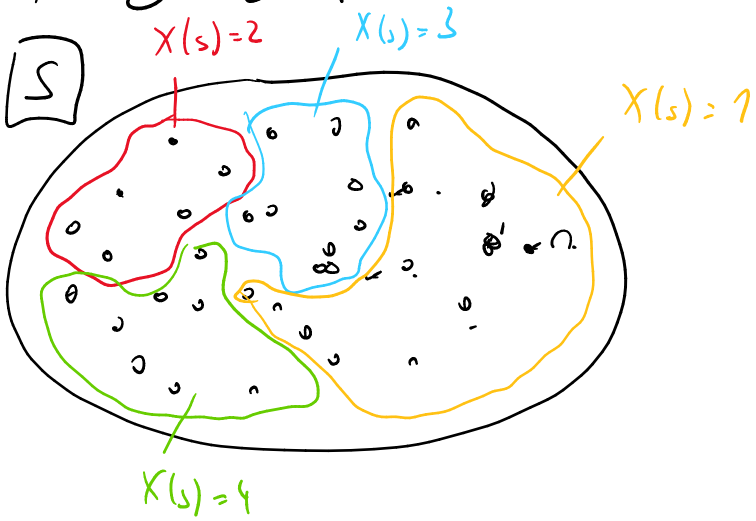
$$P(E) = \sum_{i \in E} p(i)$$

Example: What is probability to obtain a 5-bit string with exactly 3 symbols '1'. (Call this event E)

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

$\boxed{\text{RANDOM VARIABLES}}$ (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 a random n-bit string

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

$P_r(X=0) = 1/16$ $\quad$ $Pr\{X>1\} = 11/16$

$Pr(X=1) = 4/16$

$P_r(X=2) = 6/16$ $\quad$ $Pr[X \geq 2 \wedge X<4] = 10/16$

$Pr(X=2) = 6/16$ $\qquad Pr[X \geq 2 \wedge X < 4] = 10/16$

$Pr(X=3) = 4/16$

$Pr(X=4) = 1/16$

Expectation of random variables

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

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

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

$$= 2$$

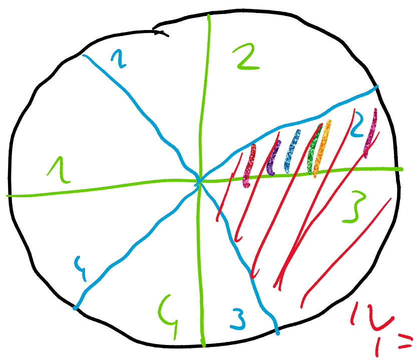
Conditional probabilities

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

$$Pr(X=i \mid Y=j) = Pr(X=i \wedge Y=j) / Pr(Y=j) \qquad [Pr(Y=j) \neq 0]$$

'X' $\quad$ Y $\qquad\qquad \emptyset$

S



$X=2, Y=2$

$Pr(X=2 \mid Y=2) \;/\; Pr(Y=2)$

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

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

X = number of symbols '1'

Y = parity of the string (even number of '1' $\Rightarrow Y=0$; odd number of "1" $\Rightarrow Y=1$)

$Pr(Y=0) = 1/2$

$Pr(Y=1) = 1/2$

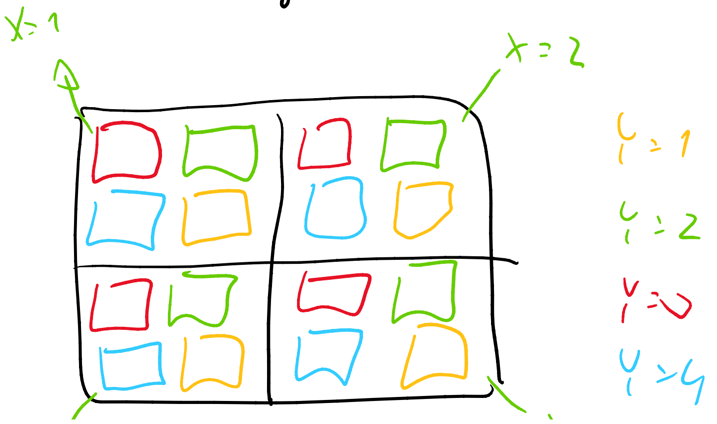
$Pr\{X=3 \mid Y=0\} = 0 = Pr\{X=3, Y=0\} / Pr(Y=0) =$

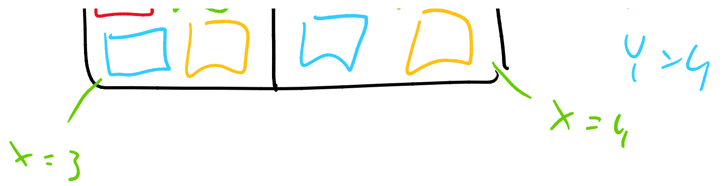
$$Pr\{X=3 \mid Y=1\} = 1/2 = \frac{Pr(X=3, Y=1)}{Pr\{Y=1\}} = \frac{4/16}{1/2} = \frac{8}{16} = \frac{1}{2}$$

Independence of random variables

X and Y are independent if for all $i, j \in \mathbb{R}$

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





Example: Are X and Y from previous example independent?

$\Pr(X=3) = 1/4$

$\Pr(X=3 / Y=0) = 0 \neq$

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

Are Z and Y independent?

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

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

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

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

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

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

1000	0111
1110	0100
1101	0010
1011	0001

1111	0101
1001	0110
1010	0011
1100	0000

LINEARITY OF EXPECTATION

$W = X + Y + Z$

there are typically easier to calculate

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

to calculate

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

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

$$= \underbrace{E(X)}_{2} + \underbrace{E(Y)}_{1/2} + \underbrace{E(Z)}_{1/2} = 3$$

$$E(X_1 \cdot X_2) \neq E(X_1) \cdot E(X_2) \quad \text{(only if idependent)}$$

The law of total probability

r.v. X and Y

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

$$= \sum_{j \in im(Y)} Pr(X=i \wedge Y=j)$$

RQuickSort

IN: A collection of numbers S

OUT: Ordered list of numbers in S

1.) if S contains a single element output S

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

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

Create S_2 which contains all

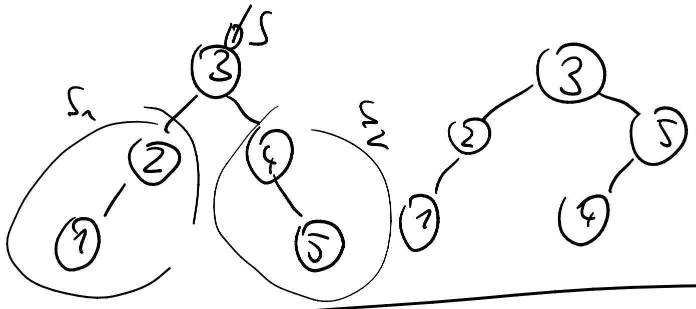
3.) Create S_1 which contains all $s \in S$, $s < y$

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

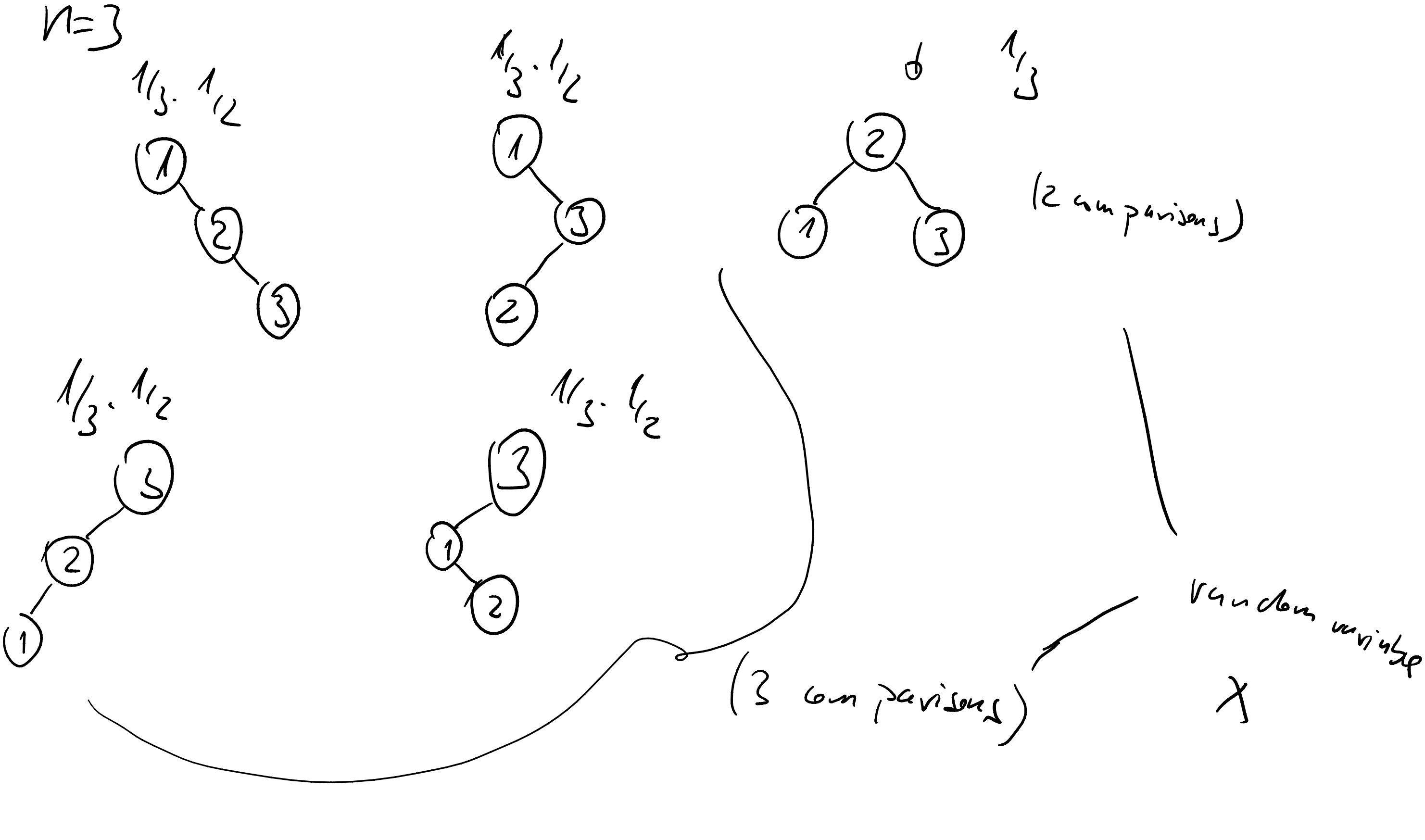
4.) Output $(\text{rquicksort}(S_1), y, \text{rquicksort}(S_2))$

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

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



$n = 3$



$$E(X) = 4 \cdot \frac{1}{6} \cdot 3 + \frac{1}{3} \cdot 2 = \frac{8}{3}$$

X_i - number of comparisons for input of size i

$f(i) \geq E(X_i)$

Can we find an upper bound $f(i)$ on $E(X_i)$ that shows $E(X_i)$ scales well?

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