A CRASH COURSE IN PROBABILITY THEORY

1.) **PROBABILITY SPACE** - all outcomes that can happen in a random experiment.

S - set
- finite
- countable
- uncountable - not in this course

EXAMPLE - a set of all n-bit strings

2.) **EVENTS** event $E \subseteq S$

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

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

space

n-bit strings with

$i \in E$

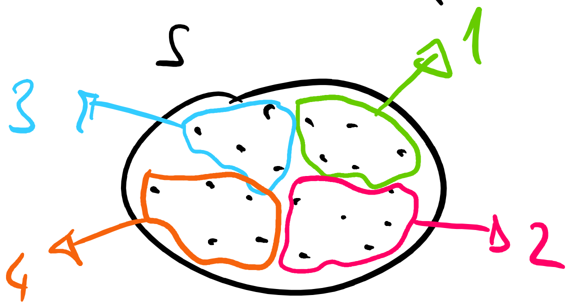
What is the probability to obtain a 5-bit string with exactly 3 symbols '1'? If probability function is uniform?

$$\binom{5}{3} = \frac{5!}{3!\,2!} = \frac{5 \cdot 4}{2} = 10$$

$$P(E) = 10 \cdot \frac{1}{32} = \frac{5}{16}$$

RANDOM VARIABLES

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



ESSENTIALLY X is a division of probability space S into mutually exclusive and collectively exhaustive set of events.

EXAMPLE X - is the number of '1' in n-bit string

Q For $n=4$. What is the distribution of X?

$$\Pr\{X=0\} = \frac{1}{16}$$
$$\Pr\{X=1\} = 4/16$$
$$\Pr\{X=2\} = 6/16$$
$$\Pr\{X=3\} = 4/16$$
$$\Pr\{X=4\} = 1/16$$

$$\Pr\{X=i\} = \binom{n}{i}/16$$

$$\Pr\{X=7.7, 9\} = 0$$

EXPECTATION OF A RANDOM VARIABLE

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

EXAMPE $E(X) =$ (X from previous example)

$$\sum_{i=0}^{4} i \cdot Pr\{X=i\} = 0 \cdot Pr[0] + 1 \cdot Pr[1] + 2 \cdot Pr[2] + 3 \cdot Pr\{3\} + 4 \cdot Pr[4]$$

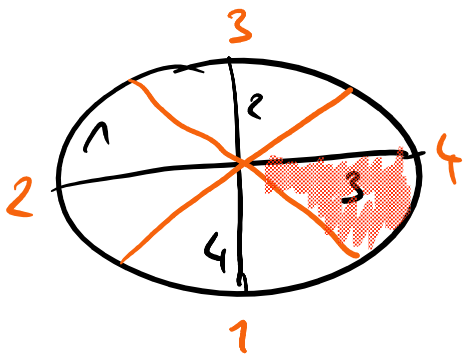
$$= 0 + 1/4 + 2 \cdot \frac{6}{16} + 3 \cdot \frac{1}{4} + 4 \cdot \frac{1}{16}$$

$$= 0 + 1/4 + 3/4 + 3/4 + 1/4 = \underline{\underline{2}}$$

CONDITIONAL PROBABILITIES

GIVEN 2 r.v. X and Y (over the same random experiment) define conditional probability of X given Y=j

$$Pr(X=i \mid Y=j) = Pr(X=i, Y=j) / Pr(Y=j) \quad (Pr(Y=j) \neq 0)$$



Intuitively we are creating a new Probability space equal to $(Y=j)$.

$\frac{1}{Pr(Y=j)}$ renormalizes the original probabilities.

EXAMPLE: $S = \{0,1\}^4$

X - number of '1'

Y - parity of the string (even (=0) or odd (=1) number of '1')

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

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

$$Pr(X=3 \mid Y=0) = 0$$

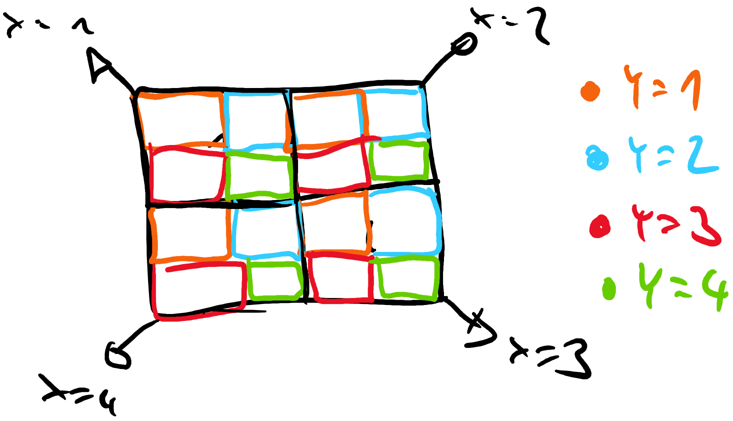
$$Pr(X=3 \mid Y=1) = 1/2 = Pr(X=3, Y=1) = 4/16 = 1/4$$

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

INDEPENDENCE of r.v.

R.v. X and Y are independent if for all i, j

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



EXAMPLE: ARE X and Y from the previous example independent?

$$Pr(X=3) = 1/4$$
$$Pr(X=3 \mid Y=0) = 0$$

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

is Y and Z independent?

0000	1100	1111
0001	0011	1110
0010	0110	1101
0100	1001	1011
1000	1010	0111
	0101	

$$Pr(Z=1) = 1/2$$
$$Pr(Z=0) = 1/2$$
$$Pr(Z=1 \mid Y=1) = \frac{4}{8} = 1/2$$
$$Pr(Z=0 \mid Y=1) = 1/2$$
$$Pr(Z=1 \mid Y=0) = \frac{4}{8} = 1/2$$
$$Pr(Z=0 \mid Y=0) = 1/2$$

✓ Independent

LINEARITY OF EXPECTATION

$W = X+Y+Z$ where X, Y, Z are r.v.?

W is a proper r.v.

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

$$E\left(\sum_i a_i X_i\right) = \sum_i a_i E(X_i) \qquad a_i \in \mathbb{R}$$

often much simpler to calculate

$$E\left(\sum_i X_i\right) = \boxed{\sum_i E(X_i)}$$

often much simpler to calculate

THIS HOLDS FOR DEPENDENT set of r.v. as well!

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

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

$$= 2 + 1/2 + 1/2 = 3$$

NOTE: $E(X_1 \cdot X_2) \neq E(X_1) \cdot E(X_2)$

$E(X_1 \cdot X_2) = E(X_1) \cdot E(X_2)$ ← For independent X_1 and X_2 only

---

RQUICKSORT

IN: Collection of numbers S

OUT: Ordered list of elements in S

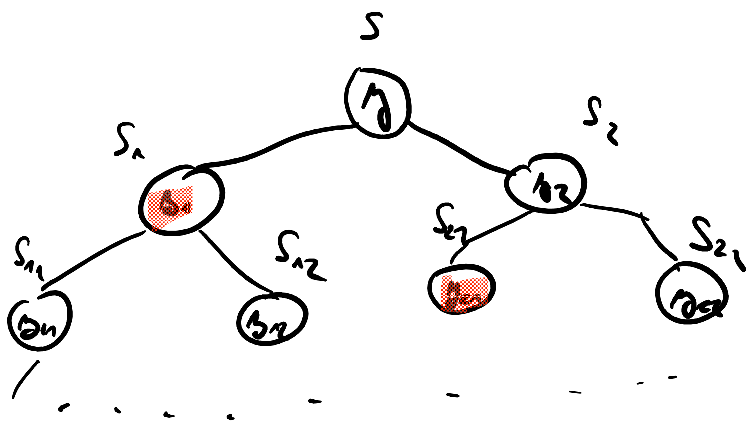
1.) Choose a pivot $y \in S$ uniformly at random

2.) Create S_1 which contains all $s < y$

Create S_2 which contains all $s > y$

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

What is the expected number of comparisons.



* Are these two elements compared in the run?

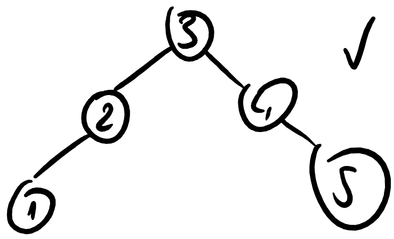
two elements s_i and s_j are compared iff one is in a subtree with the other as it's root

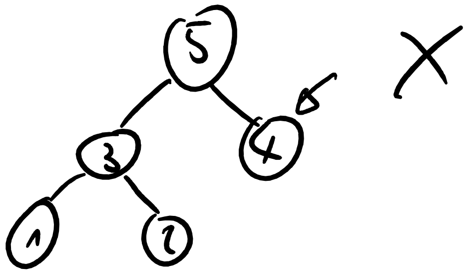
b

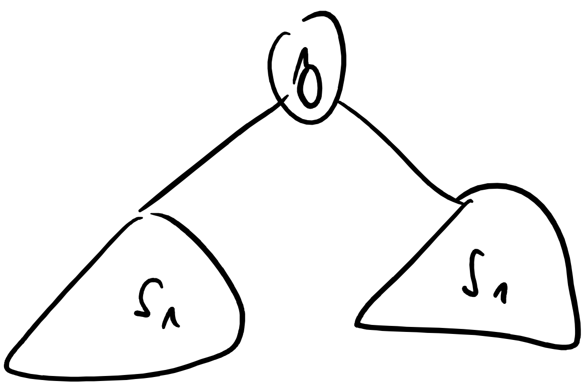
they are "above" each other

Probability space is the set of all "ordered" trees of elements S

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

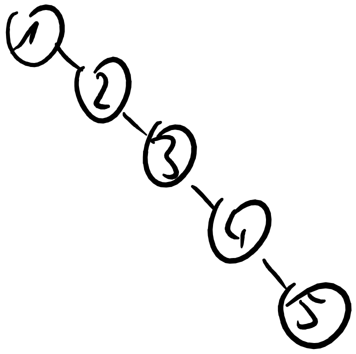




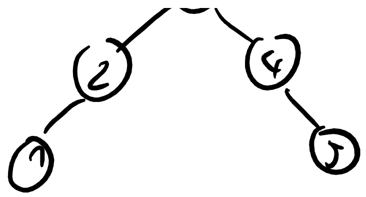


$\forall s \in S_1 \quad s < y$

$\forall s \in S_2 \quad s > y$



10 comparisons



6 comparisons

$\forall i,j$ define a r.v. S_{ij}

$S_{ij} = 1$ if i and j are compared in the run

$S_{ij} = 0$ otherwise

$S = \sum_{ij} S_{ij}$ -> total number of comparisons

Lin. of Exp

$$E(S) = E\left(\sum_{ij} S_{ij}\right) = \sum_{ij} E(S_{ij})$$

$$E(S_{ij}) = 0 \cdot P_r(S_{ij}=0) + 1 \cdot P_r(S_{ij}=1) = \underline{P_r(S_{ij}=1)} < P_{ij}^*$$

Step1: P_r that i and j get compared in the first recursion? $\frac{2}{|S|}$

Step2: S_1, γ, S_2 for i,j to get

compared in step?
depends on y

both are in S_1 (W.L.O.G)

Step k / condition on the fact that i, j were not compared yet

what is the probability they will get compared?

$$\frac{2}{|S_1|} \qquad |S_1| \geq |i-j|+1 \qquad s_i \in S$$

$$s_i < s_j \in S \quad i<j$$

in step k, conditioned on being in the same set

$$P_r(S_{ij}=1) = \frac{2}{|S_1|} \leq \frac{2}{|i-j|+1}$$

$$\underbrace{P_r(S_{ij}=1)}: \quad \underbrace{P_r(s_{ij}=1 \mid \text{they are not in the same})}_{=0} + P_r(s_{ij}=1 \mid \text{they are in the same set}) \mapsto \text{the law of total probability (see tutorial II)}$$

$$= P(S_{ij}=1 \mid \text{they are in the same set}) = \frac{2}{|S_1|} \leq \frac{2}{|i-j|+1}$$

$$E(S) \leq \sum_{i,j} \frac{2}{|i-j|+1} = \underbrace{\ldots}_{\text{some steps (see slides)}} = n \cdot \underbrace{\sum_{i=1}^{n} \frac{1}{i}}_{\Theta(\log n)} = \Theta(n.\log n)$$