Week	Topic	Reading
1	• Data, Models, and Information • Elementary statistics: Definitions • Overview of R	OIS 1 (ISL 1)
2	• Elementary statistics: Applications & Plots	OIS 1 (ISL 1)
3	• Introduction to data analysis with R • Review of tabular and graphical displays of data	ITR 1, 2, 5, 6, 7, 12
4	• Random variables: expectation and variance • Joint and conditional probability • Bayes rule	OIS 2
5	• Random variables: distributions (normal, binomial, poisson)	OIS 3

Definitions, basic concepts, R practice

Lots of definitions and pen-and-paper practice

Intro to Probability Theory: A bunch of definitions & problems

(lots of definitions & equations, followed by some playing of online games)

Intro to Probability Theory: A bunch of definitions & problems

Why bother?

- **Good background for all of the rest of this class**
- **Necessary for further statistics classes (Bayesian)**

Probability forms the foundation for *inference*

- A certain county has a population that is 50% women and 50% men. A jury is supposedly randomly selected. The jury ends up having a composition that is 40% women. Was there selection bias, or was this just due to random chance?

- In a randomized double-blind controlled experiment, a new surgery saved lives 60% of the time, while the old surgery saved lives only 55% of the time. Is this a big enough difference to replace the old surgery with the new one?

- A pack of potato chips is supposed to be manufactured to have an average weight of 10 ounces. 30 random bags of chips are weighed, and have an average weight of 9.6 ounces. Is the manufacturer cheating?
  - If the bags really have an average of 10 ounces, what is the probability we would get a sample average this low?

Basic probability

Roll a fair die once.

Notation: P(event)

P(roll a 1) =

P(roll at least a 4) =

P(roll at most 2) =

Basic probability

Roll a fair die once.

Notation: P(event)

P(roll a 1) = 1/6

P(roll $<$ 4) = 3/6

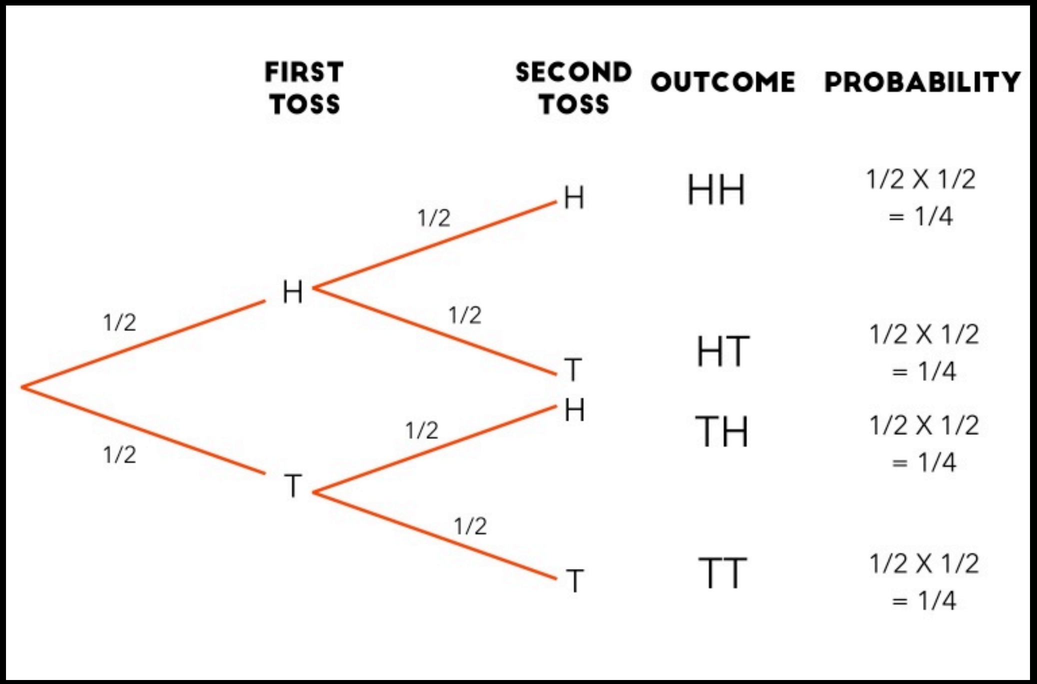
P(roll $<=$ 2) = 2/6

The classical theory of probability

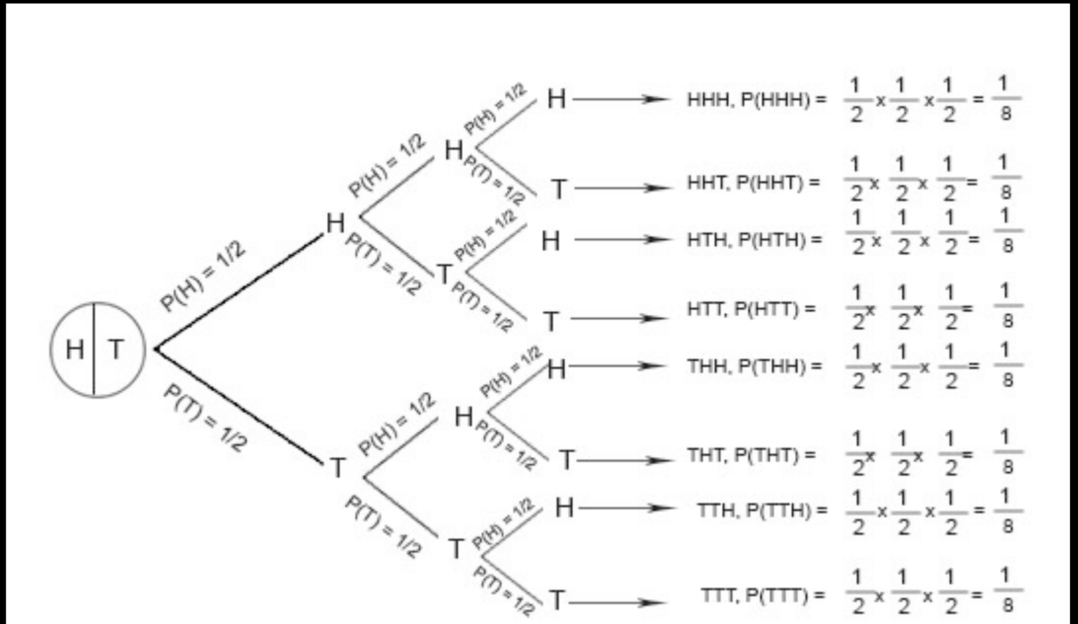
P(event) = $\frac{\text{\# of ways event can happen}}{\text{\# of total possible outcomes}}$

P(you get HH in two tosses of a fair coin) = ?

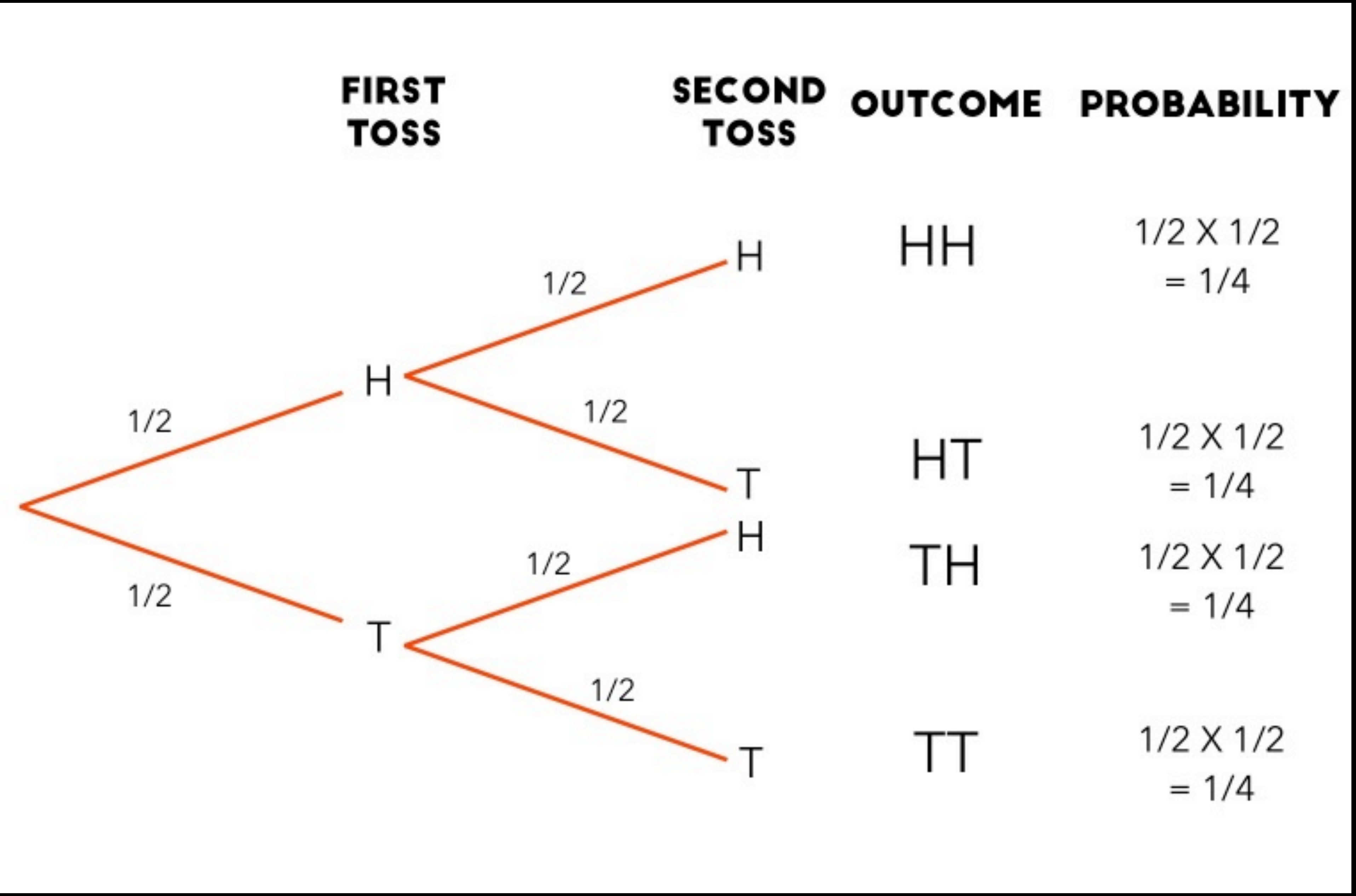
Make a tree diagram



P(you get HHH in three tosses of a fair coin) = ?



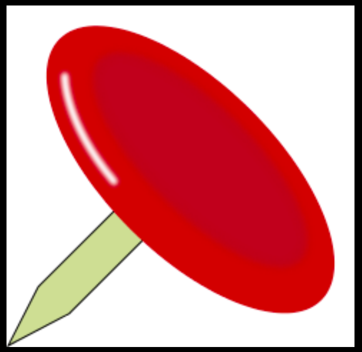
The classical theory of probability



But wait...

P(tack lands face down) = ?

P(tack lands face down) = ½ ?



P(comet hits earth tomorrow) = ?

P(comet hits earth tomorrow) = ½ ?

The relative frequency theory

P(event occurs) = the proportion of times the outcome would occur if we observed the random process an infinite number of times

$$P(\text{tack lands face down}) = \frac{\text{\# of times it lands face down}}{\text{\# of times we toss it}}$$

as # of times we tosses gets very large

Let n be the number of times we *repeat* the random process

As n gets larger, the approximation tends to get better,

More technically, the **Law of Large Numbers** says that as more observations are collected, the observed proportion of occurrences with a particular outcome after n trials converges to the true probability p of that outcome. i.e., The more times you repeat experiment, the better estimate of probability tends to get.

But wait...

P(comet hits earth tomorrow) = ?

We just repeat this experiment over and over, right?

Subjective theory

P(comet hits earth tomorrow) = ?

The subjective theory of probability says that probability is a statement about our knowledge or lack of knowledge. Probability is therefore expressed as degrees of rational belief.

Probability is a property of us, not of the world.

Probability: Practically

What is the probability of event #1 **or** event #2 occurring?

- $P(E_1 \text{ or } E_2)$ - General Addition Rule

What is the probability of event #2 **and** event #2 occurring?

- $P(E_1 \text{ and } E_2)$ - General Multiplication Rule

What is the probability of event #1 **given** event #2 (if event #1 depends on event #2)?

- $P(E_1 \mid E_2)$ - Conditional probability - marginal & joint probabilities; tree diagrams; Bayes' Theorem

How do these relate $P(E_1)$ and $P(E_2)$?

2 rules of probability

$0 \leq P(E) \leq 1$ (or $0\% \leq P(E) \leq 100\%$)

0 implies the event is impossible and
1 implies the event is certain

Also, $P(E) + P(\text{not } E) = 1$ and so

$P(\text{not } E) = 1 - P(E)$

e.g. P(not getting a Queen in a deck of 52 cards)
=1 – P(getting a Queen in a deck of 52 cards)
= 1 – 4/52 = 48/52

Disjoint and non-disjoint outcomes

Disjoint (mutually exclusive) outcomes: Cannot happen at the same time.

- The outcome of a single coin toss cannot be a head and a tail.
- A student both cannot fail and pass a class.
- A single card drawn from a deck cannot be an ace and a queen.

Non-disjoint outcomes: Can happen at the same time.

- A student can get an A in Stats and A in Programming in the same semester.

Addition Rule of disjoint outcomes

If A_1 and A_2 represent two <u>disjoint</u> outcomes, then the probability that one of them occurs, i.e. one or the other occurs, is:

$P(A_1 \text{ or } A_2) = P(A_1) + P(A_2)$

Addition Rule of disjoint outcomes

If A_1 and A_2 represent two disjoint outcomes, then the probability that one of them occurs, i.e. one or the other occurs, is:

$P(A_1 \text{ or } A_2) = P(A_1) + P(A_2)$

Draw 1 card
P(Jack or King) = 4/52 + 4/52

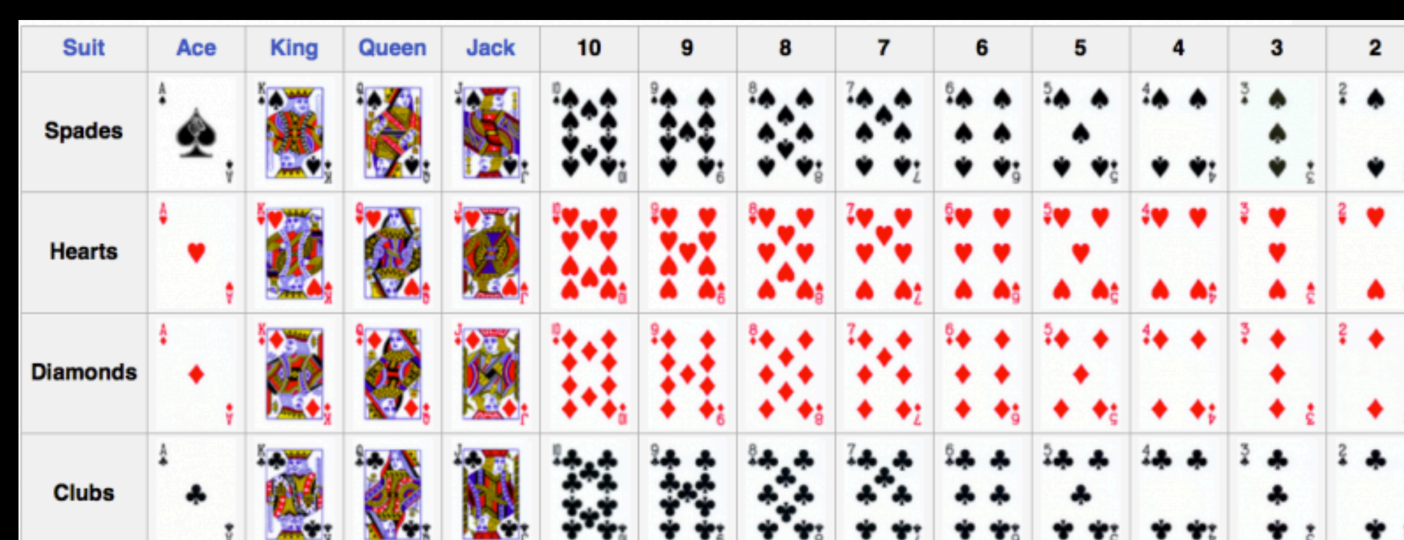
Addition Rule of disjoint outcomes

If A_1 and A_2 represent two <u>disjoint</u> outcomes, then the probability that one of them occurs, i.e. one or the other occurs, is:

$P(A_1 \text{ or } A_2) = P(A_1) + P(A_2) + \ldots$

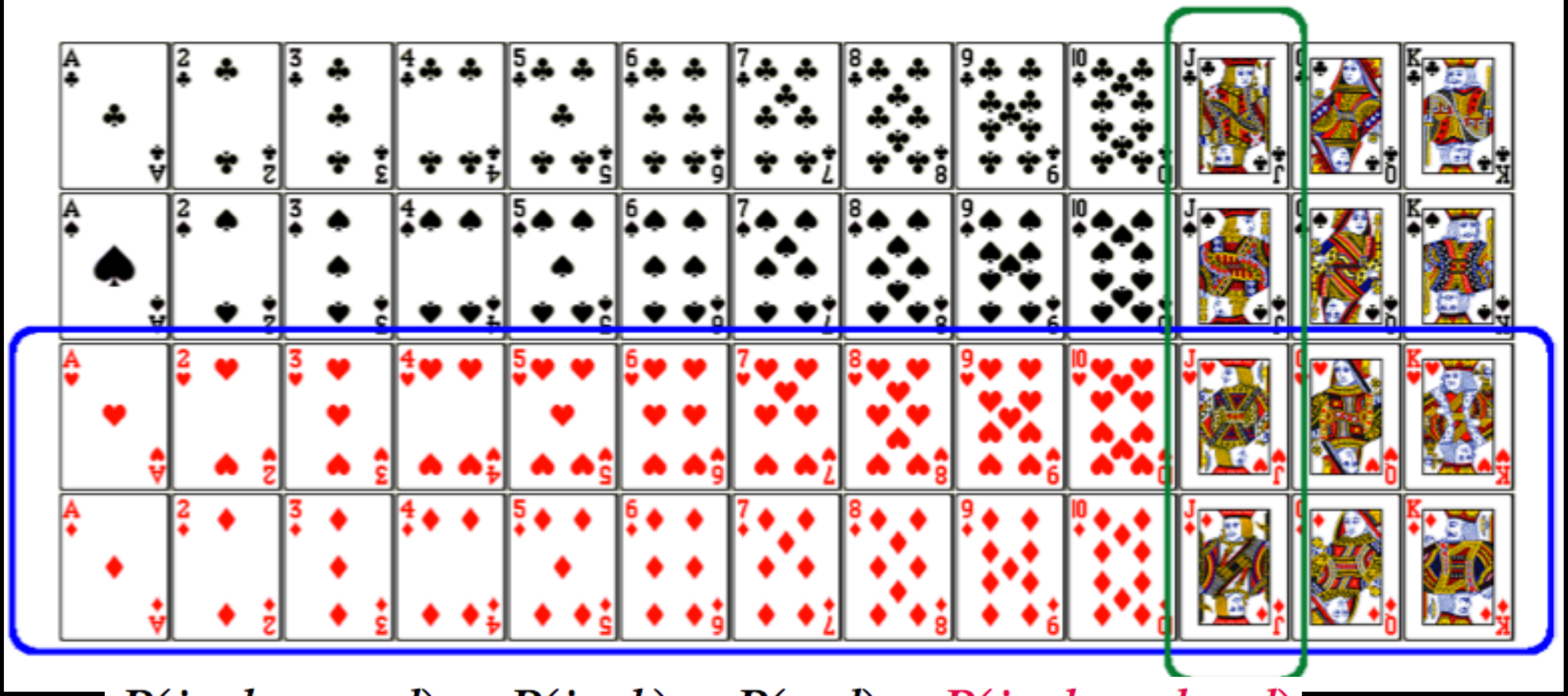
Draw 1 card
P(Jack or King) = 4/52 + 4/52

Draw 1 card
P(Jack or Red) = 4/52 + 13/52 ?



Union of non-disjoint events

What is the probability of drawing a jack or a red card from a well shuffled full deck?



$$P(\textit{jack or red}) = P(\textit{jack}) + P(\textit{red}) - P(\textit{jack and red})$$

$$= \frac{4}{52} + \frac{26}{52} - \frac{2}{52} = \frac{28}{52}$$

http://www.milefoot.com/math/discrete/counting/cardfreq.htm

Practice

What is the probability that a randomly sampled student thinks marijuana should be legalized or they agree with their parents' political views?

	Share Parents' Politics		
Legalize MJ	No	Yes	Total
No	11	40	51
Yes	36	78	114
Total	47	118	165

(1) (40 + 36 - 78) / 165
(2) (114 + 118 - 78) / 165
(3) 78 / 165
(4) 78 / 188
(5) 11 / 47

Recap

General addition rule

P(A or B) = P(A) + P(B) - P(A and B)

Note: For disjoint/mutually exclusive events, P(A and B) = 0, so the above formula simplifies to P(A or B) = P(A) + P(B)