Probability introduction

Randomness

Probability motivation

Outcomes

Sample spaces

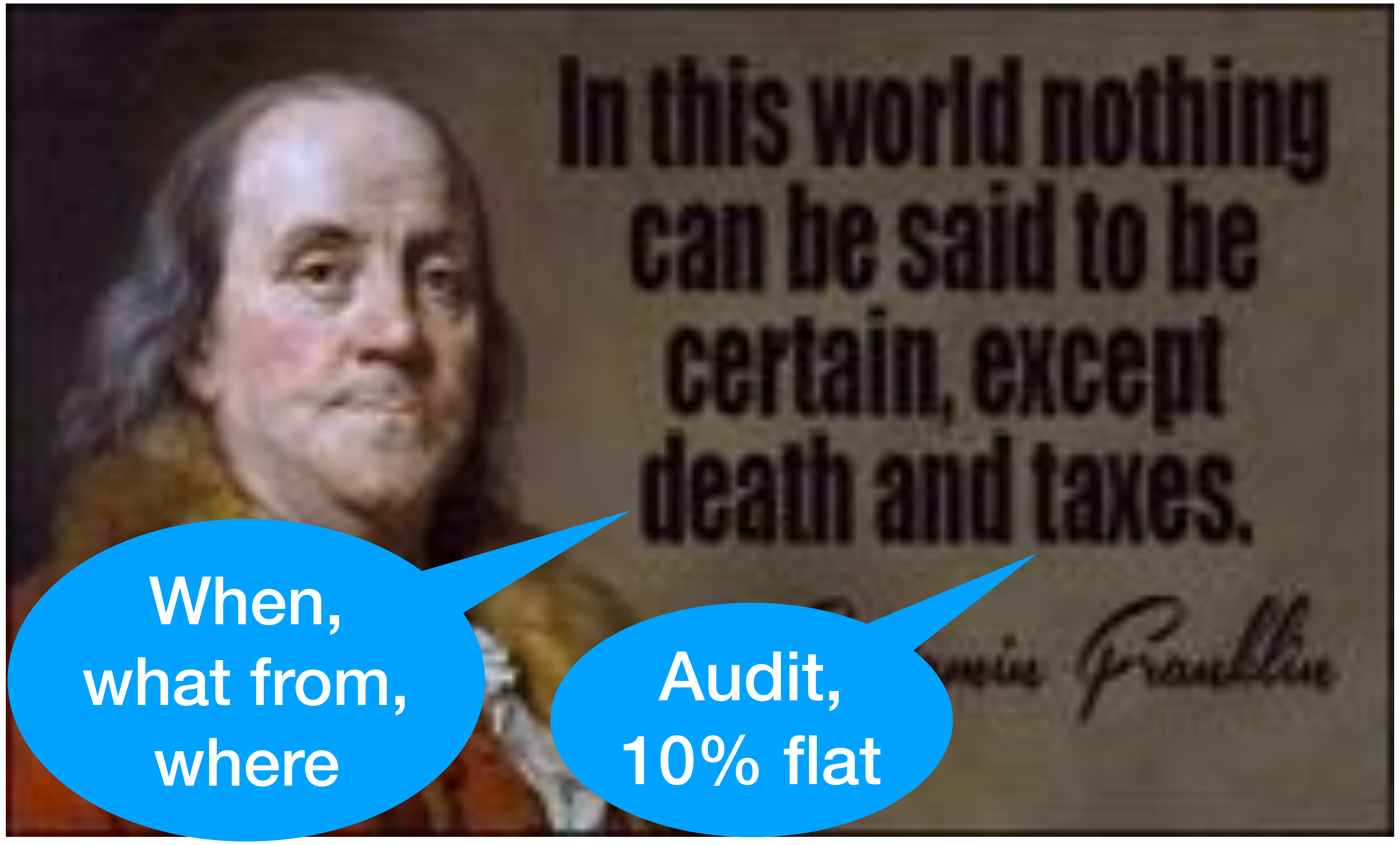
Distributions

Why Probability?

Some things in life are certain

Most are a little less predictable

Physicians	illness, medication
Farmers	rain, diet trends
Investors	stock price, economy
Advertisers	views, competition
Consumers	availability, sale
Students	food line, grade, parents, job, date, game



Random Phenomena

Give up? Reason intelligently?

Learn	Range	Average	Variability
Infer	Structure	Change	Relations
Predict	Future	Likelihood	Guarantees
Benefit	Compete	Plan	Build

Coming to Terms

As with sets

Need terminology

Discuss

Concisely

Precisely

Process of generating and observing data

Individual and collection of observations

Meaning of probability

Several approaches

Intuitive

Axioms

Data

Experiments

Probability developed in part to aid science

Process

Generating random data

Observing outcome

Experiment

Biology

Engineering

Business

...

Sociology

Unified approach

Applies generally

Understand

Analyze

generalize

Our experiments

Start simple

Get very complex

Outcomes and Sample Space

Possible experiment results are called outcomes, or observations

Set of possible outcomes is the sample space, denoted Ω

Notetation

Some use S or U

Experiment	Ω
Coin	{ h, t }
Die	{ 1, 2, ..., 6 }
Gender	{ m, f }
Age	$\mathbb{N}$
Temperature	$\mathbb{R}$

typically lower-case h, t, x

Elements, sets

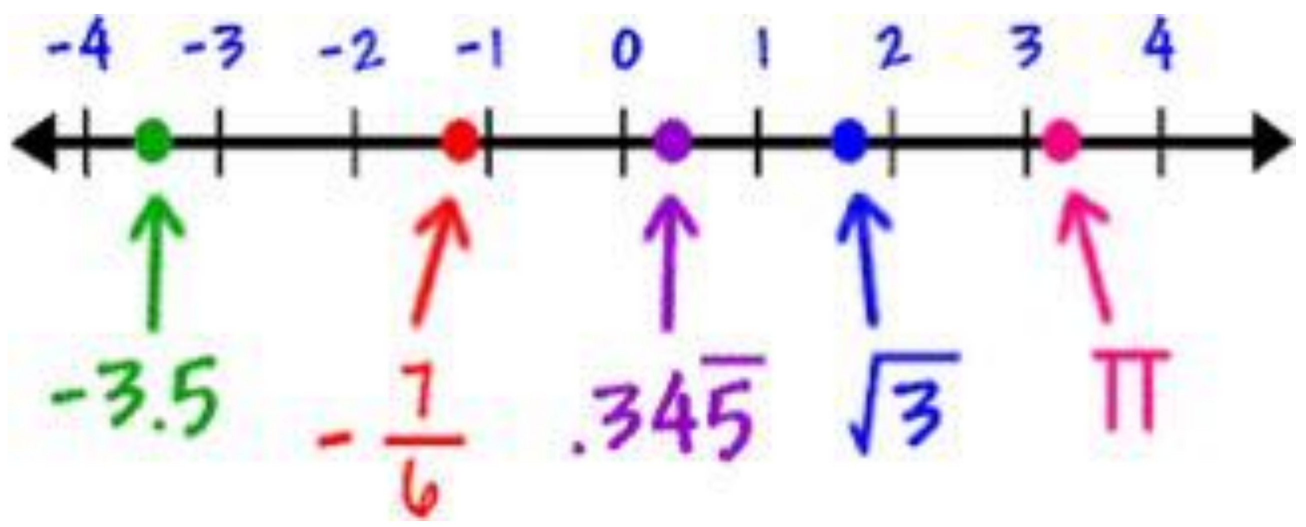
Next lecture: subsets

Two Sample-Space Types

Finite or countably infinite sample space is discrete

{ h, t } { 1, 2, …, 6 } $\mathbb{N}$ $\mathbb{Z}$ { words } { cites } { people }

Uncountably infinite sample space - continuous



$\mathbb{R}$ { temperatures } $\underbrace{\text{\{ salaries \}} \quad \text{\{ prices \}}}_{\text{upgraded}}$

Discrete spaces Easier to understand, visualize, analyze

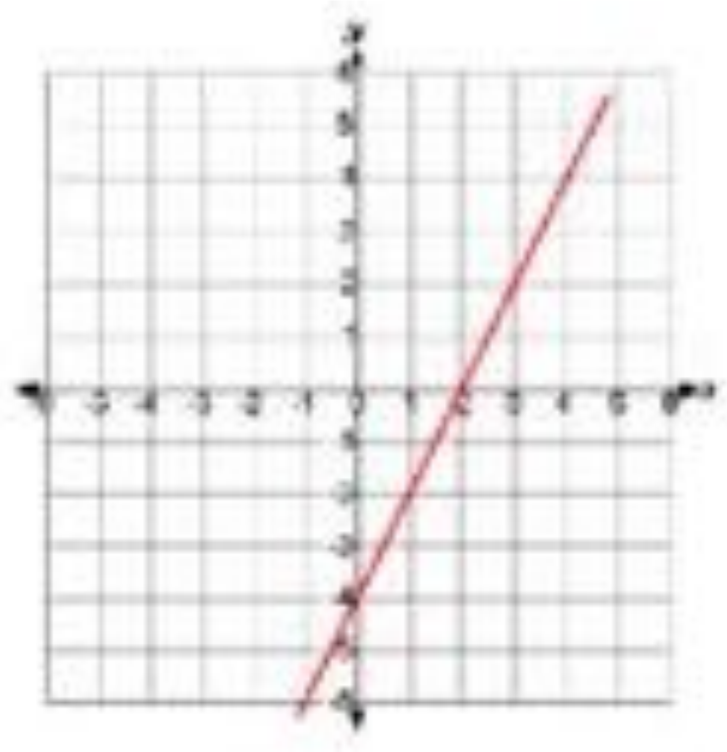
Important First

Next few topics: Discrete

Continuous Important Conceptually harder Later

Random Outcomes

Algebra Unknown value denoted by x $2x - 4 = 0$



Solve Before After

$x \in \mathbb{R}$ $x = 2$

Notetation

If needed use $y, z, \ldots$

Probability Random value of outcome denoted by X X - coin flip outcome

Experiment Before After

$X \in \Omega$ get h $X = h$

get t $X = t$

Notetation

If needed use $Y, Z, \ldots$

Probability of Outcome

The **probability**, or **likelihood**, of an outcome $x \in \Omega$, denoted $P(x)$, or $P(X=x)$, is the fraction of times x will occur when experiment is repeated many times

Fair coin

As # experiments $\to \infty$, fraction of heads (or tails) $\to \frac{1}{2}$

heads has probability ½ | $P(h) = \frac{1}{2}$ | $P(X=h) = \frac{1}{2}$

tails has probability ½ | $P(t) = \frac{1}{2}$ | $P(X=t) = \frac{1}{2}$

Fair die

As # experiments $\to \infty$, fraction of 1's (or 2,…,6) $\to \frac{1}{6}$

1 has probability ⅙ | $P(1) = \frac{1}{6}$ | $P(X=1) = \frac{1}{6}$

$P(x)$ | $P(X=x)$ | probability of x | fraction of times x will occur

Probability Distribution Function

P(x) is the fraction of times outcome x occurs

$P(h) = ½$

$P(1) = ⅙$

Viewed over the whole sample space, a pattern emerges

Coin	$P(h) = ½$	$P(t) = ½$	
Die	$P(1) = ⅙$	…	$P(6) = ⅙$
Rain	$P(\text{rain}) = 10\%$	$P(\text{no rain}) = 90\%$	

Sample space Ω and distribution P define the whole probability space

P maps outcomes in Ω to nonnegative values that sum to 1

$P: \Omega \rightarrow \mathbb{R}$

$P(x) \geq 0$

$\sum_{x \in \Omega} P(x) = 1$

Probability distribution function (PDF)

distribution

Probability introduction

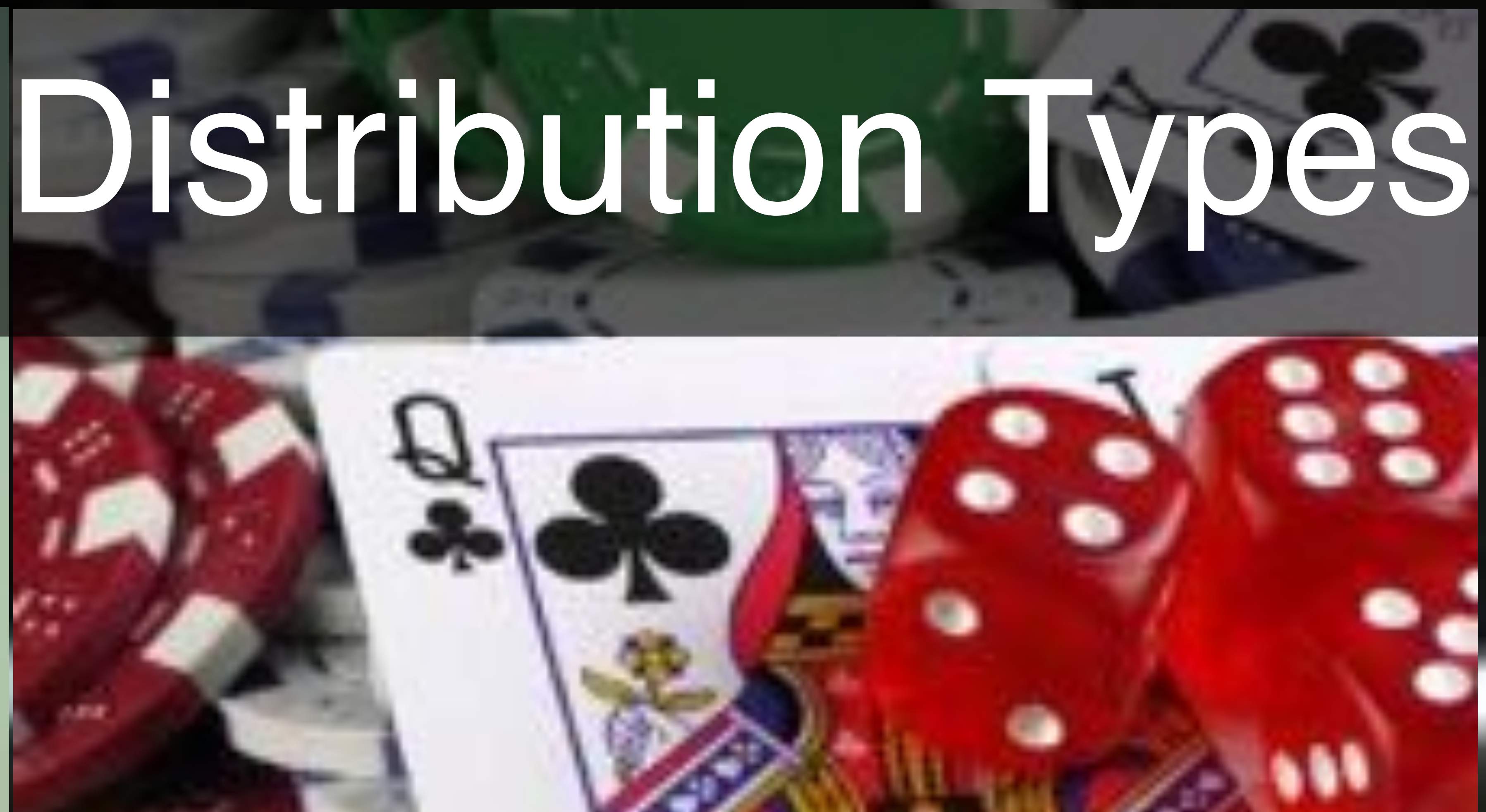
Randomness

Probability motivation

Outcomes

Sample spaces

Distributions



Distribution Types

Distribution Types

- Uniform sample spaces
- Coin, die, cards
- Non-uniform spaces
- Tetrahedral die

Uniform Probability Spaces

Generally, outcomes may have different probabilities

Rain P(rain) = 10% P(no rain) = 90%

Uniform (equiprobable) spaces

Uniform distribution

All outcomes are equally likely

Coin $P(h) = P(t) = ½$

Drastically simplifies probability specification

Uniform Probability Spaces

All outcomes are equally likely

$\forall\, x \in \Omega \quad P(x) = p$

$1 = \sum_{x \in \Omega} P(x) = \sum_{x \in \Omega} p = |\Omega| \cdot p$

$p = 1/\,|\Omega|$

Notetation

Draw Uniformly, Randomly

Coin $\quad P(h) = P(t) = p \quad 1 = P(h) + P(t) = 2 \cdot p \quad p = ½$

Uniform spaces $\quad$ Every outcome has probability $1/\,|\Omega|$

All you need to know is $|\Omega|$!

U

Fair Coin

Ω = { heads, tails } = { h, t }

$|\Omega| = 2$

Flip, or toss

Equally likely U $P(h) = P(t)$

$$P(h) = P(t) = \frac{1}{|\Omega|} = \frac{1}{2}$$

Fair Die

$\Omega = \{ 1, 2, 3, 4, 5, 6 \}$

$|\Omega| = 6$

Roll

Equally likely Ⓤ $P(1) = \ldots = P(6)$

$$P(1) = \ldots = P(6) = \frac{1}{|\Omega|} = \frac{1}{6}$$

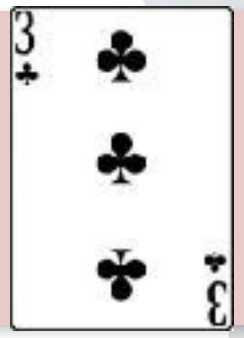
Deck of Cards

$\Omega = \{ \text{ cards } \}$

$|\Omega| = 52$

Draw a card

Equally likely Ⓤ

P() = … = P() $= \frac{1}{|\Omega|} = \frac{1}{52}$

Uniform → Non

Uniform, equiprobable, spaces

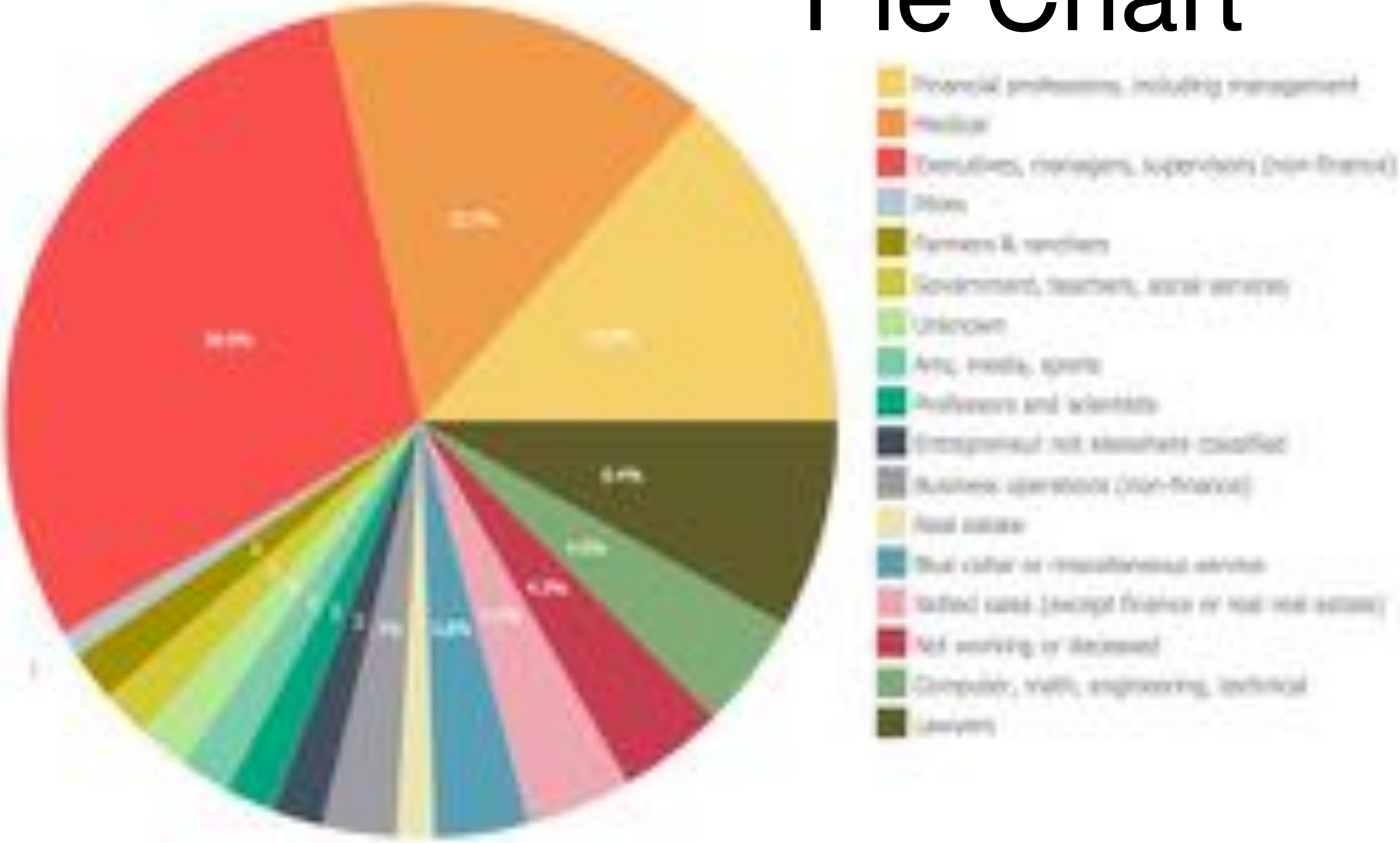
Coin Die Cards

In nature, nonuniform spaces abound

rain grades words illnesses

web pages people ….

Pie Chart



Challenge Non-uniform distribution we can remember

Tetrahedral Die

4-sided, pyramid die

Used in games, D&D

In games die is equiprobable

We will assume different probabilities

Easy to remember

Face	1	2	3	4
Probability	.1	.2	.3	.4

Conveniently, add to 1

Probability distribution

Do’s and Don’ts

Random notation may be confusing at first

Which expressions are valid?

P(X = 3) fair die: ⅙

P(3) $\stackrel{\text{def}}{=}$ P(X = 3)

P(x) specify x, e.g., for ∀x, P(x)=¼

P(1 = 3) 0

P(X) random value

Possible, but less common. Make sure this is what you mean.

P(x = 3)

Even less likely, probably wrong

Distribution Types

Uniform sample spaces

- Coin, die, cards

Non-uniform spaces

- Tetrahedral die

Events

Probability Events

Outcomes to subsets

Events

Occurrence

Probability

Events

Social event

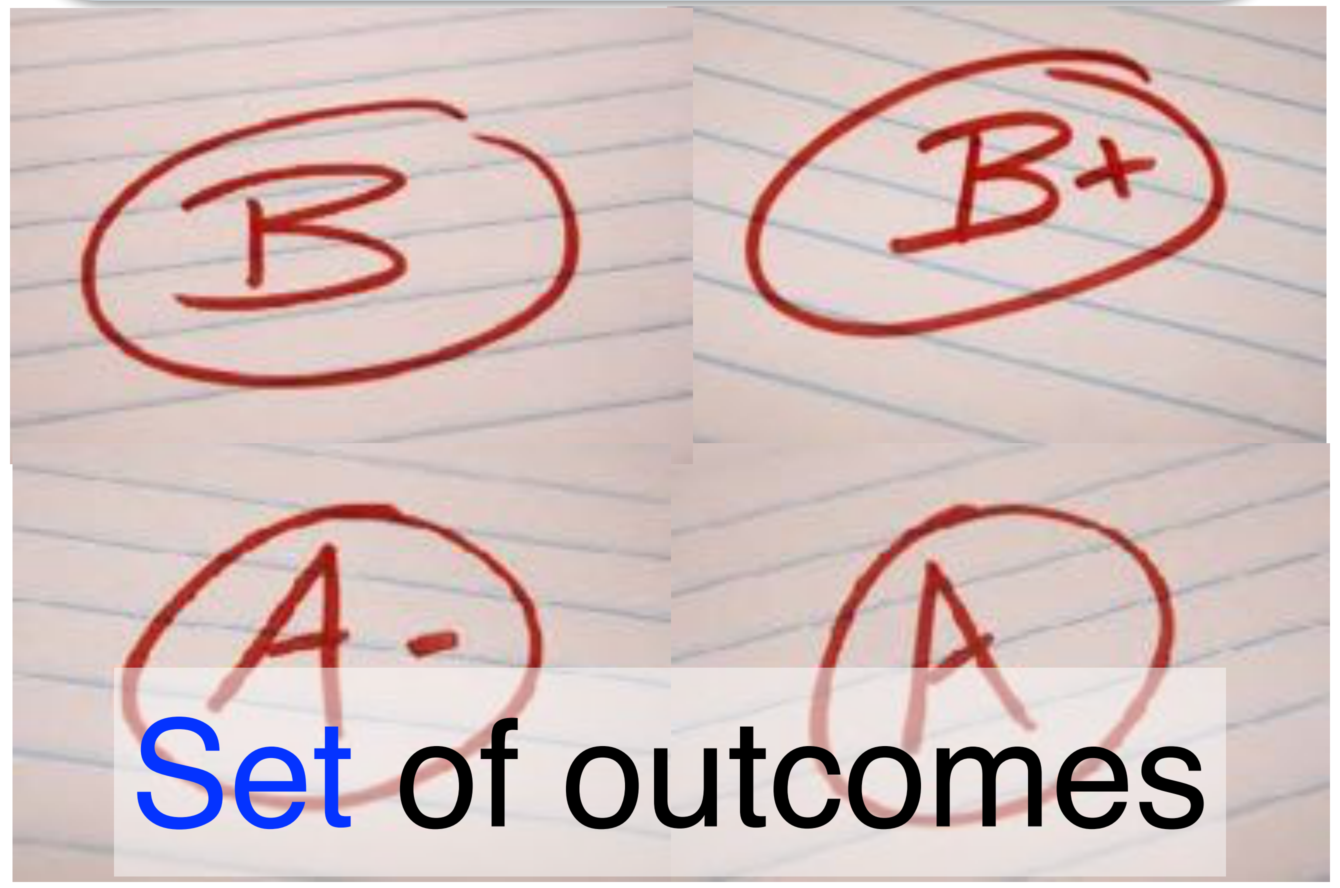
Sometimes — Care about one particular outcome

Our horse will win race

Get B+ in class

Usually — Interested in a set of possible outcomes

Probabilistic event



Temperature > 10 degrees

Stock will close higher

Pass a course

Set of outcomes — Subset of the sample space Ω — Event

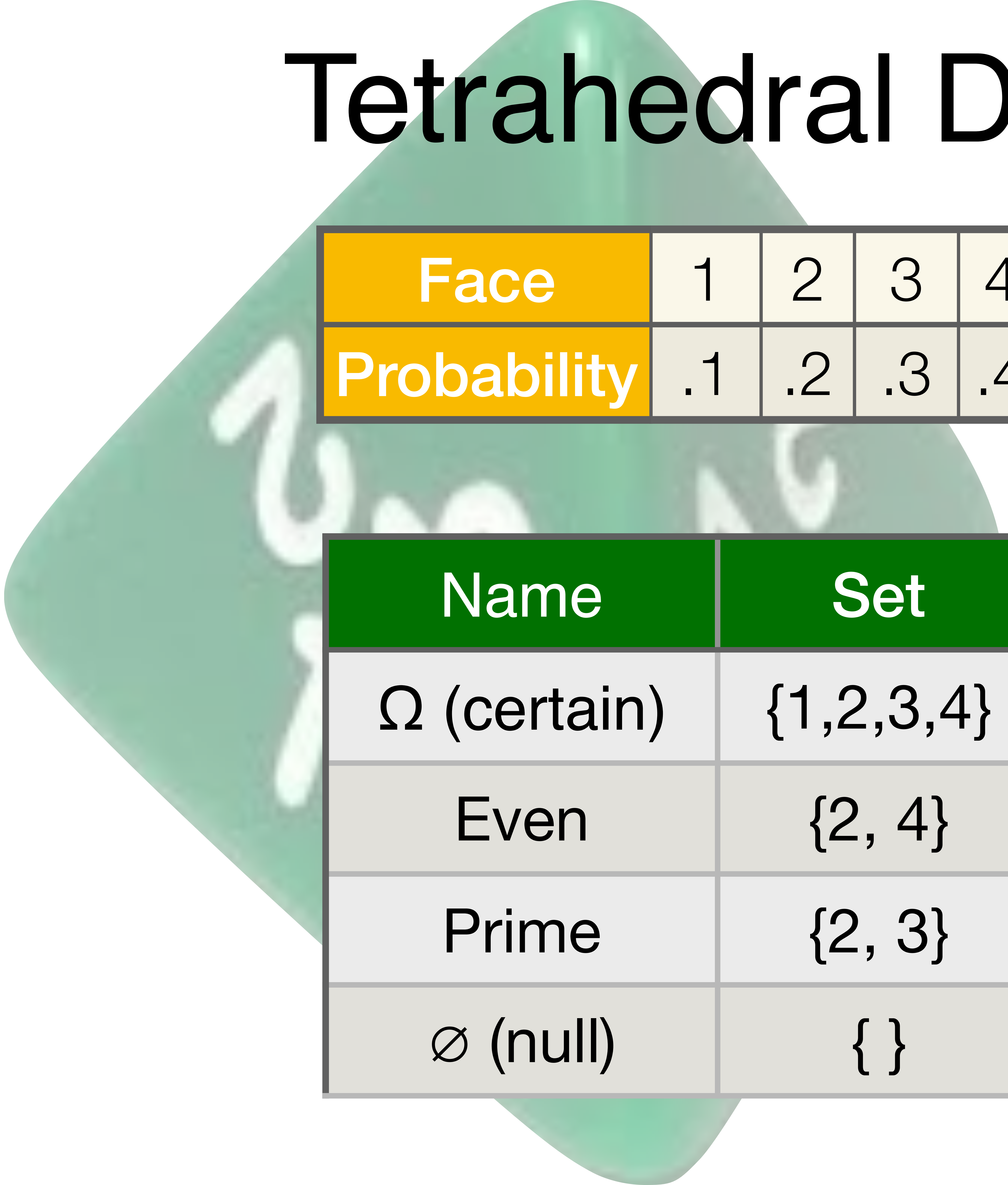
Die Events

Name	Set
Ω (certain)	{1, …, 6}
Even	{2, 4, 6}
Square	{1, 4}
> 4	{5, 6}

Complements

Name	Set
∅ (null)	{ }
Odd	{1, 3, 5}
Non square	{2, 3, 5, 6}
≤ 4	{1, 2, 3}

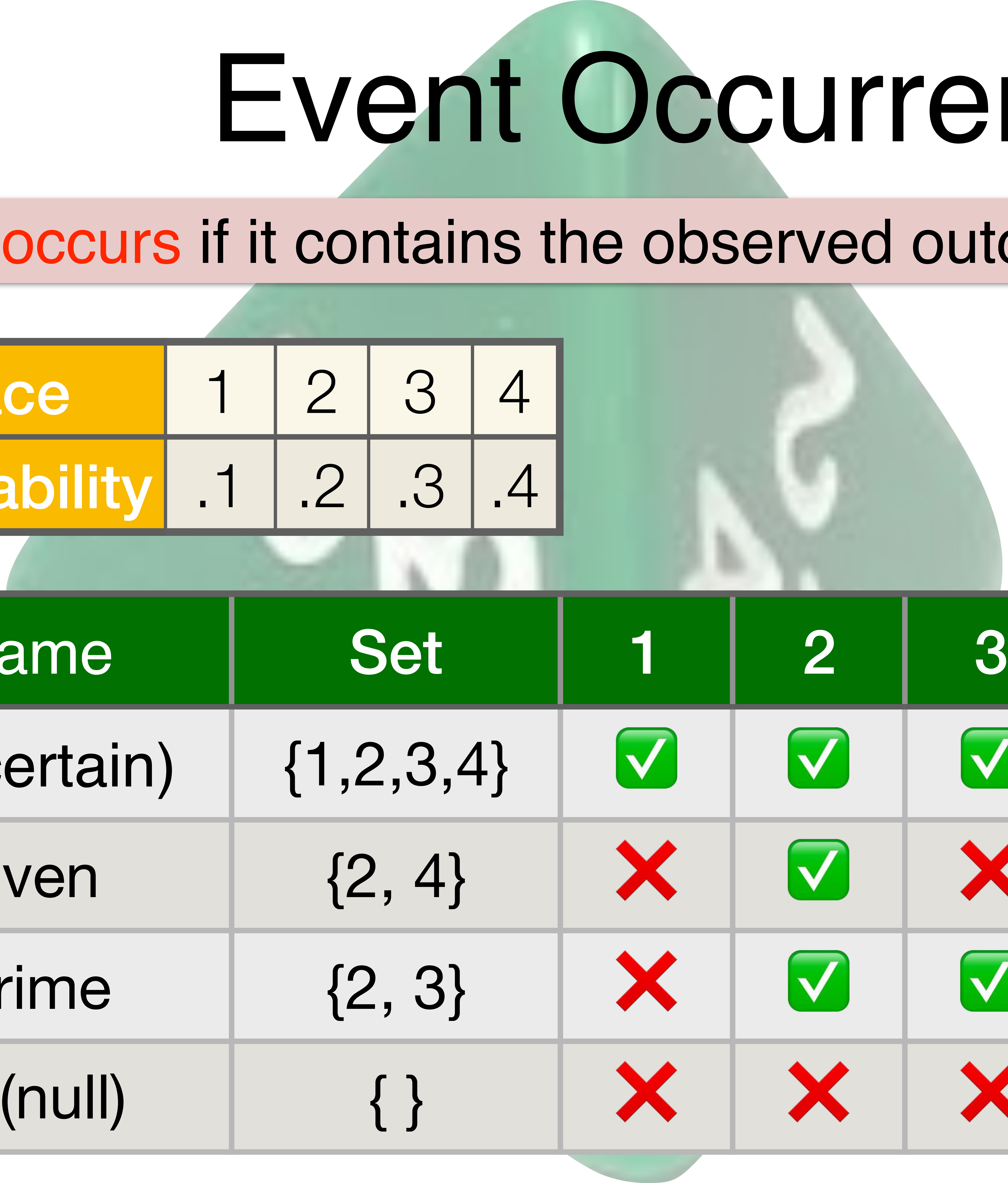
Tetrahedral Die



Face	1	2	3	4
Probability	.1	.2	.3	.4

Name	Set
Ω (certain)	{1,2,3,4}
Even	{2, 4}
Prime	{2, 3}
∅ (null)	{ }

Event Occurrence



An event occurs if it contains the observed outcome

E occurs if X∈ E

Face	1	2	3	4
Probability	.1	.2	.3	.4

Name	Set	1	2	3	4
Ω (certain)	{1,2,3,4}	✅	✅	✅	✅
Even	{2, 4}	❌	✅	❌	✅
Prime	{2, 3}	❌	✅	✅	❌
∅ (null)	{ }	❌	❌	❌	❌

Event Probability

Probability of event E

Fraction of experiments where E occurs, as # experiments grows

Even ④ ③ ② ① ④ ③ ③ ④ ② ④ ③

✅ ✅ ✅ ✅ ✅ ✅

times E occurs = Sum of # times 2 and 4 occur

Fraction of times E occurs = Sum of fraction of times 2 and 4 occur

$$P(E) = P(X \in E) = \sum_{x \in E} P(x)$$

Tetrahedral Die

Probability	.1	.2	.3	.4

Name	Set	1	2	3	4	P(E)
Ω (certain)	{1,2,3,4}	✅	✅	✅	✅	.1+.2+.3+.4 = 1
Even	{2, 4}	❌	✅	❌	✅	.2 + .4 = .6
Prime	{2, 3}	❌	✅	✅	❌	.2 + .3 = .5
∅ (null)	{ }	❌	❌	❌	❌	0

Uniform Spaces

$$P(x) = \frac{1}{|\Omega|}$$

$$P(E) = \sum_{x \in E} P(x) = \sum_{x \in E} \frac{1}{|\Omega|} = \frac{|E|}{|\Omega|}$$

Die Ω = {1,2,3,4,5,6} |Ω| = 6

Event	Set	$\lvert Event \rvert$	$P(Event) = \frac{\lvert Event \rvert}{6}$
Even	{2,4,6}	3	$\frac{3}{6} = \frac{1}{2}$
Square	{1,4}	2	$\frac{2}{6} = \frac{1}{3}$

Probability that X=2, 4 ,or 6

Terminology

A, B Events

Intersect

Disjoint

Mutually exclusive

Do’s and Don’ts

$P(X \in \text{Even})$ die: $P(X \in \{2,4,6\}) = 3/6 = 1/2$

$P(\text{Even}) \stackrel{\text{def}}{=} P(X \in \text{Even})$

$P(1 \in \text{Even})$ 0

$P(2 \in \text{Even})$ 1

Possible, but less likely
Ensure you meant it

$P(x \in \text{Even})$ Even less likely, probably wrong

Probability **Events**

- Outcomes to subsets
- Events
- Occurrence
- Probability

Repeated Experiments

Repeated Experiments

- Compound experiments
- Repetition
- With and without replacement
- Order matters, or not

Composite Experiments

Experiments often consist of several parts

Student major year GPA

Ad product audience cost

Still can be viewed as a single experiment

Outcomes more complex

(CS, senior, 3.8)

(book, teenage, $9.99)

Outcome 3-tuple

Sample space Cartesian product

Combine smaller to analyze larger

Start simple

Independent Repetitions

Repetition — All experiments of same type

Daily stock prices | Daily temperatures

Coin flips | Die rolls | Card draws

Independecnt — Different components unrelated

First coin heads — Second coin 50% heads / tails ✅

Second coin more likely heads ❌

Two Coins

Independent experiments

Outcomes

Probabilities

		coin 2	½	½
			h	t
coin 1	½	h	¼	¼
	½	t	¼	¼

$\Omega = \{\ hh, ht, th, tt\ \} = \{h,t\}^2$

$|\Omega| = 4 = 2^2$

1 coin U → 2 coins U

$P(\ hh\ ) = P(\ ht\ ) = P(\ th\ ) = P(\ tt\ ) = 1/|\Omega| = 1/4$

Two Dice

Independent experiments

die 1 / die 2

	1	2	…	6
1	1/36	1/36	…	1/36
2	1/36	1/36	…	1/36
⋮	⋮	⋮		⋮
6	1/36	1/36	…	1/36

$\Omega = \{ 11, 12, \ldots, 66 \} = \{1, \ldots, 6\}^2$ $|\Omega| = 6^2 = 36$

1 die U → 2 dice U

$P(11) = P(12) = P(21) = \ldots = P(66) = 1/|\Omega| = 1/36$

Events

$$P(E) = P(X \in E) = \sum_{x \in E} P(x)$$

U → P(E) = |E| / |Ω|

2 coins U $|\Omega| = 2^2 = 4$

P(Different outcomes) = P({ht,th}) = 2 / |Ω| = 1/2

P(At least one h) = P({ ht,th,hh }) = 3 / |Ω| = 3/4

3 coins U $|\Omega| = 2^3 = 8$

P(Alternating) = P({ hth, tht }) = 2 / 8 = 1/4

P(odd # h's) = P({ htt, tht, tth, hhh}) = 4 / 8 = 1/2

Sampling

Many sources of randomness

Coin Die

Often sample (select) physical objects

Patients in a study Customers at a restaurant

Products for quality control Visitors to web pages

Cards from a deck Balls from an urn

Two sampling types With Without Replacement

Replacement

Sequentially select physical objects

With replacement

Replace (reuse) selected element

Experiments independent

Outcomes can repeat

Like coins dice

Without replacement

Do not replace selected element

Experiments dependent

Outcomes cannot repeat

Difference largest for small Ω

Replacement and Repetition

Sampling (selection) with without replacement

repeat as if from scratch

repeat with smaller set

Same element can cannot be selected again

✅  coin

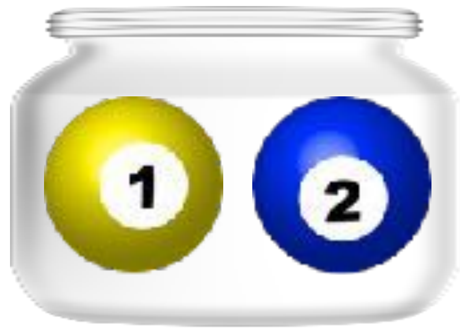
cards  ❌

✅  die

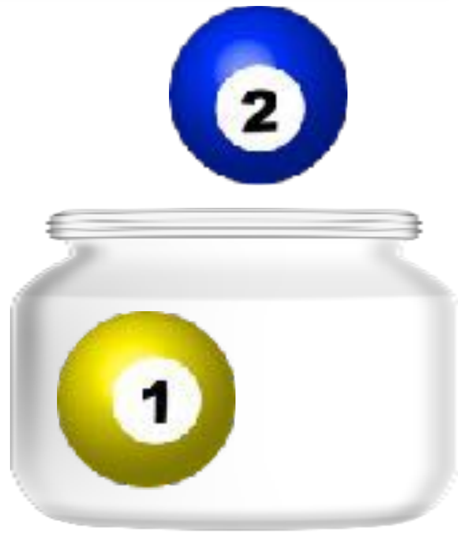
people  ❌

Balls in a Jar

1 and 2 ball in a jar



Pick one ball



Pick second ball

→ 2 1

Second selection - from a subset

Selection without replacement

Probabilities

		2nd ball: 1	2nd ball: 2
1st ball	½ 1	0	½
	½ 2	½	0

$\Omega = \{$ 1 2, 2 1 $\}$ $|\Omega| = 2$

U P(1 2) = P(2 1) = ½

Drawing Cards

Six cards

2-permutations of {1,..,6}

Draw one

Without replacement, draw a second

Outcomes

$\Omega = \{12, \ldots, 16, 21, \ldots, 26, \ldots, 65\} = [6]^{\underline{2}}$

$|\Omega| = 6^{\underline{2}} = 6 \cdot 5 = 30$

Probabilities

$i \neq j$ $P(i,j) = \frac{1}{6} \cdot \frac{1}{5} = 1/30$

U $P(i,j) = 1/|\Omega| = 1/30$

card 2

card 1	1	2	…	6
1	0	1/30	…	1/30
2	1/30	0		1/30
⋮	⋮			
6	1/30	1/30	…	0

Summary

2 experiments

	Original	With replacement	Without replacement
		Outcomes can repeat	Outcomes cannot repeat
Sample space	Ω_1	$\{ (x,y) : x,y \in \Omega_1 \} = \Omega_1^2$	$\{ (x,y): x,y \in \Omega_1,\ x \neq y \} = \Omega_1^{\underline{2}}$
| Sample space |	$\|\Omega_1\|$	$\|\Omega_1\|^2$	$\|\Omega_1\|^{\underline{2}} = \|\Omega_1\| \cdot (\|\Omega_1\|-1)$

If original sample space is uniform

Uniformity		Uniform	Uniform
P (element)	$1 / \|\Omega_1\|$	$1 / \|\Omega_1\|^2$	$1 / [\ \|\Omega_1\| \cdot (\|\Omega_1\|-1)\ ]$

Order Does Not Matter

Sometimes order matters　Stock　10 → 50

Sometimes does not matter　Elections　Dem　Rep

When order does not matter

Tuple of outcomes → Set of outcomes

(2,5)　(5,2) → {2,5}　Event { (2,5), (5,2) }

(4,4) → {4,4}

Probability　{1,…,6} twice　With and w/o replacement

With Replacement

2 cards ∈ {1,..,6} with replacement

P({1,2}) = P({ (1,2), (2,1) } = P((1,2)) + P((2,1)) = 2/36

P({1,1}) = P({ (1,1) } = 1/36

	1	2	…	6
1	1/36	1/36	…	1/36
2	1/36	1/36	…	1/36
⋮	⋮	⋮		⋮
6	1/36	1/36	…	1/36

Check

$$\binom{6}{2} \cdot \frac{2}{36} + \binom{6}{1} \cdot \frac{1}{36} = \frac{5}{6} + \frac{1}{6} = 1$$

Without Replacement

2 cards ∈ {1,..,6} without replacement

	1	2	…	6
1	0	1/30	…	1/30
2	1/30	0		1/30
⋮	⋮			
6	1/30	1/30	…	0

$i \neq j$ $P(i,j) = \frac{1}{6} \cdot \frac{1}{5} = 1/30$

$P(\{1,2\}) = P(\{(1,2), (2,1)\}) = P((1,2)) + P((2,1)) = 2/30$

$P(\{1,1\}) = 0$

Check

$$\binom{6}{2} \cdot \frac{2}{30} = 1$$

Alternative View

2 cards $\in$ {1,..,6} sequentially without replacement

$P(\{1,2\}) = P(1,2) + P(2,1) = 2/30$

Alternatively Select both cards simultaneously

$\Omega = \{ \{1,2\}, \{1,3\}, \ldots \{5,6\} \} = \binom{[6]}{2}$

$|\Omega| = \binom{6}{2}$ U

$P(\{1,2\}) = 1/15$

	1	2	…	6
1	0	1/30	…	1/30
2	1/30	0		1/30
⋮	⋮			
6	1/30	1/30	…	0

Poker Hands

Deck — 52 cards

Hand — 5 cards from deck

Card order does not matter

$\Omega = \{$  ,  ,  , $\dots\}$

$$|\Omega| = \binom{52}{5} = \frac{52!}{5!\cdot 47!} = \frac{52\cdot 51\cdot \overset{10}{\cancel{50}}\cdot 49\cdot \overset{2}{\cancel{48}}}{\cancel{5}\cdot\cancel{4}\cdot\cancel{3}\cdot\cancel{2}\cdot 1} = 2,598,960$$

$\approx$ 2.6 million

P() ≈ 1/ 2.6 million

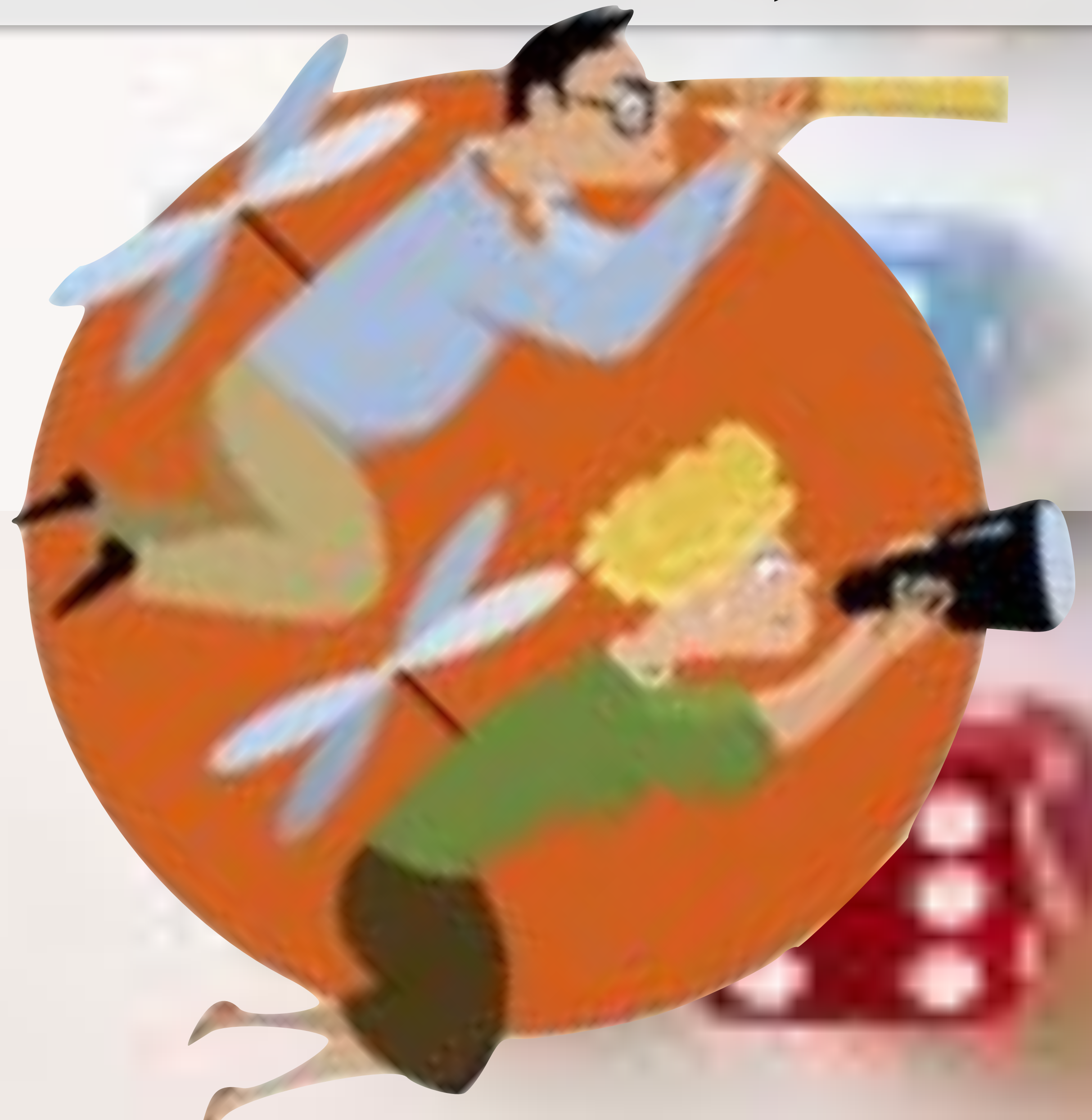
More soon

Compound experiments

Repetition

With and without replacement

Order matters, or not



Probability

Games of Chance

Uniform probability spaces

Probability of outcomes and events

Several games of chance

Dominoes

Tile | two parts | each with 0 to 6 dots

Only the two numbers matter - not their order

Ω { tiles } $\left\{ \begin{array}{l} \{0,0\}, \{1,1\}, \ldots, \{6,6\} \\ \{0,1\}, \{0,2\}, \ldots, \{5,4\}, \{5,6\} \end{array} \right\}$

Equiprobable

$$|\Omega| = \binom{7}{1} + \binom{7}{2} = 7 + \frac{7 \cdot 6}{2} = 7 + 21 = 28$$

Max value	0	1	2	…	6
# tiles	1	2	3		7
	{0,0}	{0,1} {1,1}	{0,2} {1,2} {2,2}		

$$1 + 2 + \ldots + 6 + 7 = \frac{7 \cdot 8}{2} = 28$$

Probabilities

Experiment Select one domino uniformly at random from the set

Outcomes $P(\{0,0\}) = P(\{1,2\}) = 1 / |\Omega| = 1/28$

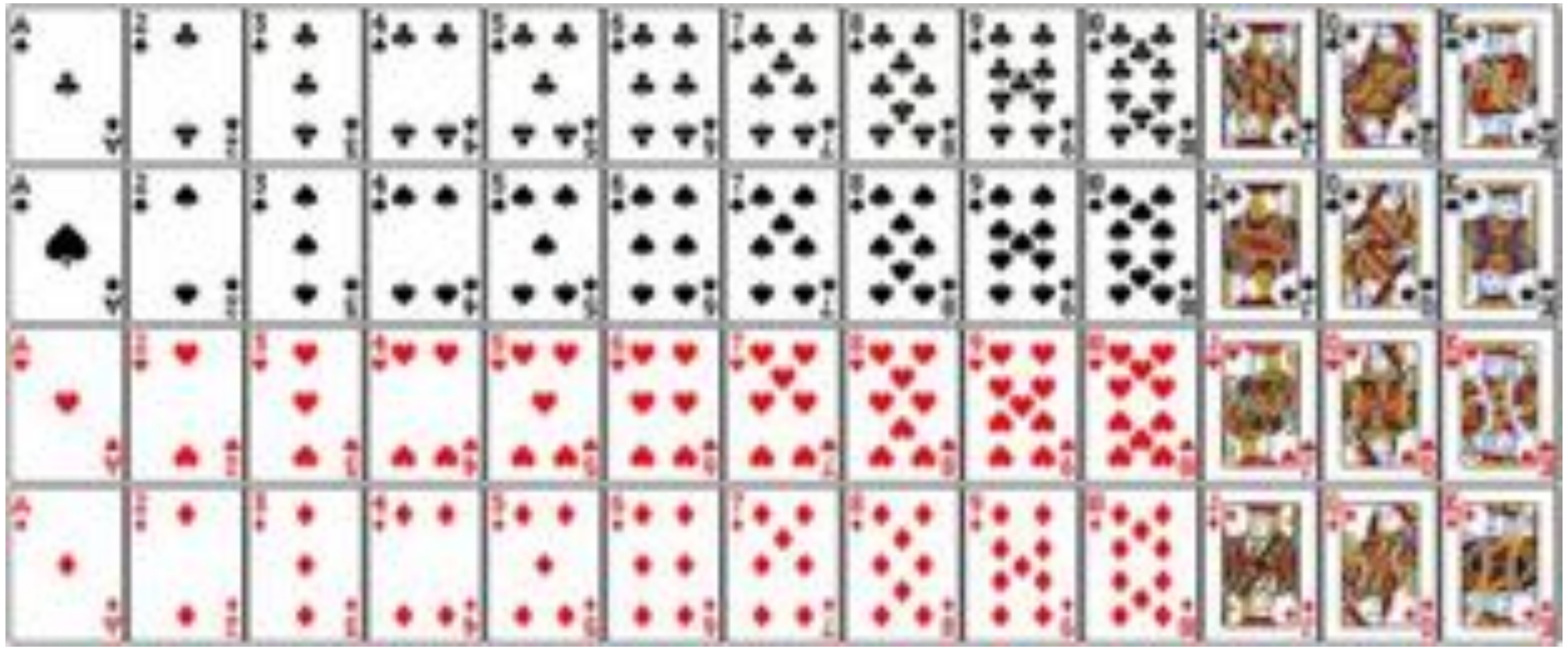
Events $P(\text{6 is one of the numbers}) = 7 / |\Omega| = 7/28 = 1/4$

$P(\text{Double}) = 7/28 = 1/4$

$P(\text{two different values}) = 21/28 = 3/4$

Cards

Standard deck: 52 cards



4 suits: clubs ($\clubsuit$), diamonds ($\diamondsuit$), hearts ($\heartsuit$), and spades ($\spadesuit$)

13 ranks: 2, 3, …, 10, Jack (J), Queen (Q), King(K), and Ace (A)

Single Card

$\Omega = \Big\{$  $\Big\}$

$|\Omega| = 52$

Equiprobable

$$P(\heartsuit) = \frac{|\heartsuit|}{52} = \frac{13}{52} = \frac{1}{4}$$

$$P(K) = \frac{|K|}{52} = \frac{4}{52} = \frac{1}{13}$$

$$P(Facecard) = P(J, Q, K) = \frac{3 \cdot 4}{52} = \frac{3}{13}$$

Poker Hands

Deck 52 cards

Hand 5 cards from deck

Card order does not matter

$\Omega = \{$  ,  ,  , $\ldots\}$

Equiprobable

$|\Omega| = \binom{52}{5} = 2,598,960$

Probability of several hands

Start with rarest

smaller sets are easier to count

Straight Flush

Straight — Consecutive rank sequence

Ace can precede 2 (A,2,3,4,5), or follow the K (10,J,Q,K,A)

Flush — Same suit

Straight flush — Five consecutive same-suit cards

$$|SF| = \binom{10}{1} \cdot \binom{4}{1} = 40$$

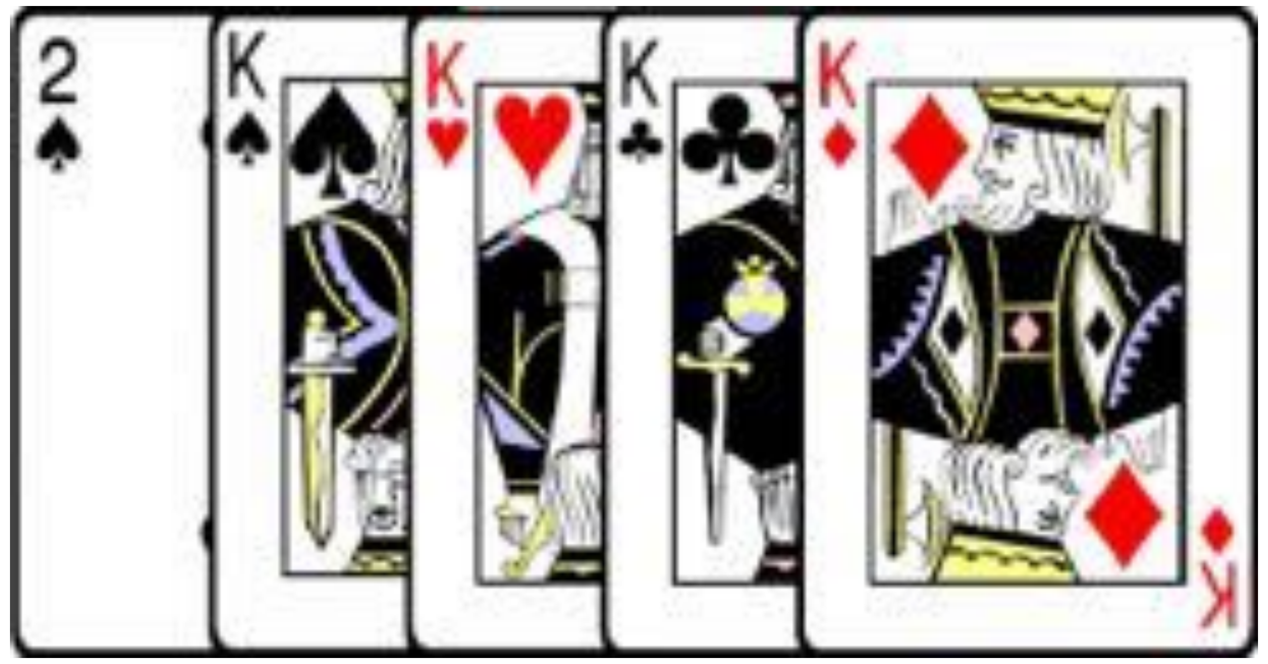
start location
A,2,…,10

suit

$$P(SF) = \frac{40}{\binom{52}{5}} = 0.0015\%$$

Four of a Kind

Four cards of the same rank, and another card



$$|FK| = \underbrace{\binom{13}{1}}_{\text{rank}} \cdot \underbrace{\binom{48}{1}}_{\text{another card}} = 624$$

$$P(FK) = \frac{624}{\binom{52}{5}} = 0.0240\%$$

Full House

Three cards of one rank and two cards of another rank

$$|FH| = 13^{\underline{2}} \cdot \binom{4}{3} \cdot \binom{4}{2} = 3744$$

rank of triple, rank of pair

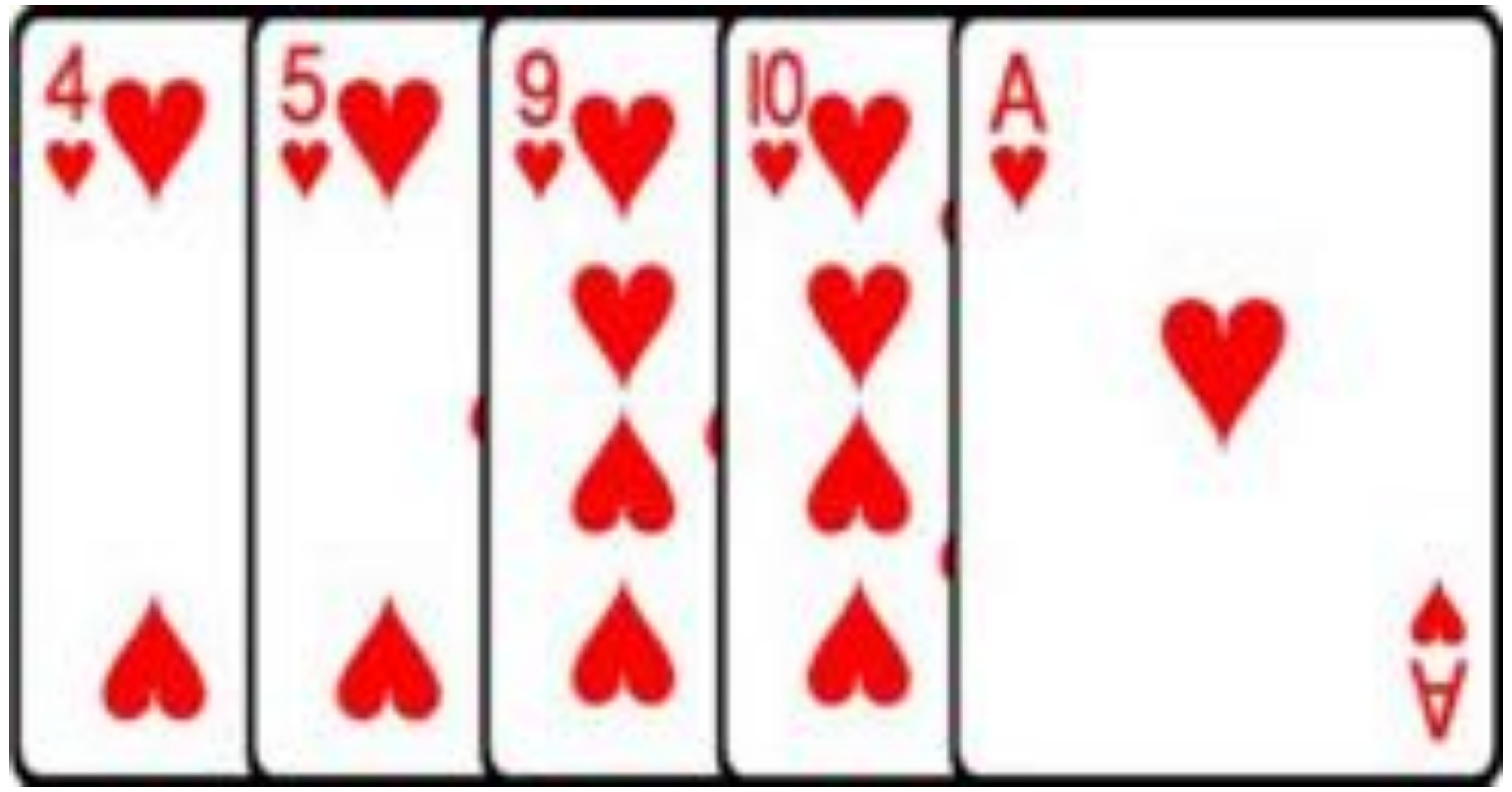
3 cards from rank of triple

2 cards from rank of pair

Falling power 13·12

$$P(FH) = \frac{3744}{\binom{52}{5}} = 0.1441\%$$

Flush



Five nonconsecutive same-suit cards

Nonconsecutive distinguishes from a straight flush

$$|F| = \binom{4}{1} \cdot \left(\binom{13}{5} - 10\right) = 5108$$

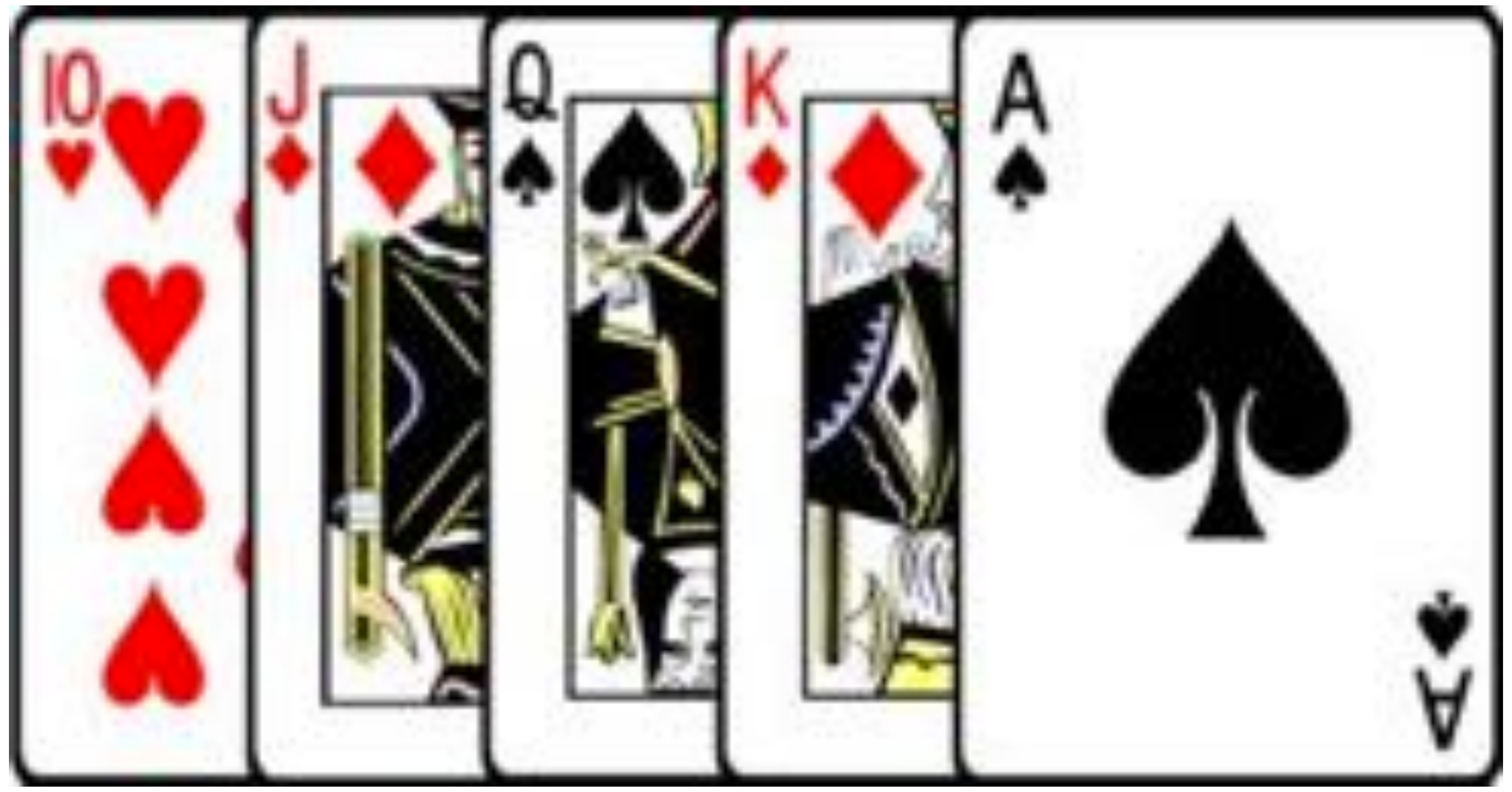
suit

5 cards from suit

5 consecutive same suit cards

$$P(F) = \frac{5108}{\binom{52}{5}} = 0.1965\%$$

Straight



Five consecutive cards, not all of the same suit

Ace can either precede 2 or succeed K

$$|S| = \binom{10}{1} \cdot \left(4^5 - \binom{4}{1}\right) = 10,200$$

Start location from ace to 10

Suit for each card

All same suit

$$P(S) = \frac{10200}{\binom{52}{5}} = 0.3925\%$$

Poker Probabilities

Notebook

Simulations for probabilities of several hands

Modify for other hands

Probability

Games of Chance

Uniform probability spaces

Probability of outcomes & events

Several games of chance

General

probability spaces

Probability

Non-Uniform Spaces

Tetrahedral Die

Side	1	2	3	4
Probability	.1	.2	.3	.4

Sum to 1

Even $\quad V = \{2,4\} \quad P(V) = .2 + .4 = .6$

Prime $\quad R = \{2,3\} \quad P(R) = .2 + .3 = .5$

General Subtraction

$P(R - V) = P(\{2,3\} - \{2,4\}) = P(\{3\}) = .3$

$P(R) - P(R \cap V) = P(\{2,3\}) - P(\{2,3\} \cap \{2,4\})$

$= P(\{2,3\}) - P(\{2\})$

$= .2 + .3 - .2 = .3$ ✅

Backgammon

Roll 2 dice

Moves determined by the two faces

order irrelevant

Outcomes

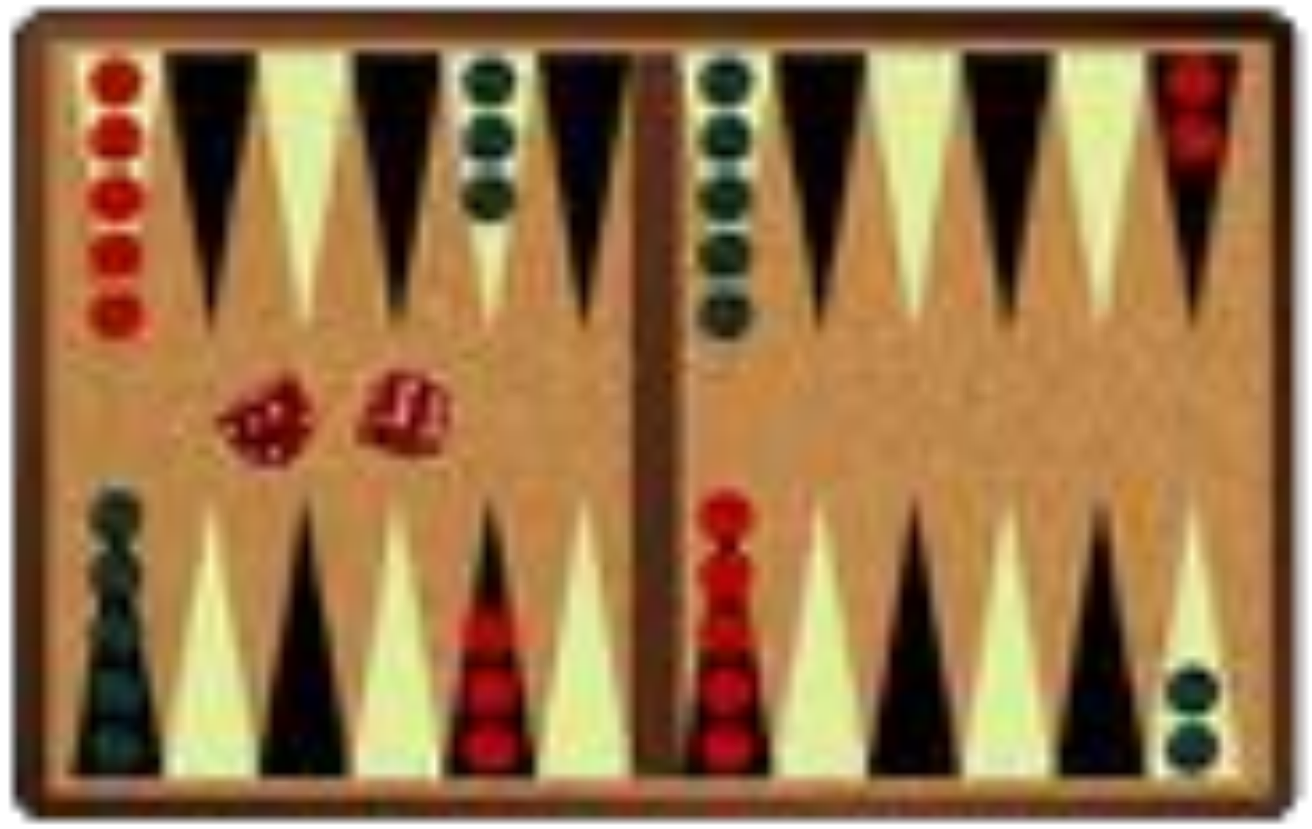
{1,1}, {1,2}, {1,3}, …, {5,6}, {6,6}

possible outcomes = ?

As in Dominos

$$\binom{6}{1} + \binom{6}{2} = 6 + \frac{6 \cdot 5}{2} = 21$$

Outcome Probabilities



$$P(1\,1) = P(1\,2) = \ldots = P(6\,6) = \frac{1}{36}$$

$$\left.\begin{aligned} P(\{1,1\}) &= p(1\,1) = \frac{1}{36} \\ P(\{1,2\}) &= p(1\,2, 2\,1) = \frac{1}{18} \end{aligned}\right\}$$

Note: not 1/21, not equiprobable

Check

$$\binom{6}{1} \cdot \frac{1}{36} + \binom{6}{2} \cdot \frac{1}{18} = 1$$

Event Probabilities

Both reasonable

$$P(\text{double}) = 6 \cdot P(\{1, 1\}) = \frac{1}{6}$$

= P(2nd equals 1st)

$$P(\text{not double}) = 1 - \frac{1}{6} = \frac{5}{6}$$

= P(2nd different from 1st)

Event Probabilities

$$P(\text{sum is } 2) = P(\{1,1\}) = \frac{1}{36}$$

$$P(\text{sum is } 3) = P(\{1,2\}) = P(1\,2, 2\,1) = \frac{2}{36}$$

$$P(\text{sum is } 4) = P(1\,3, 2\,2, 2\,1) = \frac{3}{36}$$

Hence sum 7 is the most likely case (6 ways to get it)

$$P(\max = m) = \frac{2m-1}{36} \qquad m = 1, ..., 6$$

$$P(\max \leq m) = \frac{m^2}{36}$$