

# CS70: On to probability.

## Modeling Uncertainty: Probability Space

1. Key Points
2. Random Experiments
3. Probability Space

# Key Points

- ▶ Uncertainty does not mean “nothing is known”
- ▶ How to best make decisions under uncertainty?
  - ▶ Buy stocks
  - ▶ Detect signals (transmitted bits, speech, images, radar, diseases, etc.)
  - ▶ Control systems (Internet, airplane, robots, self-driving cars, schedule surgeries in a hospital, etc.)
- ▶ How to best use ‘artificial’ uncertainty?
  - ▶ Play games of chance
  - ▶ Design randomized algorithms.
- ▶ Probability
  - ▶ Models knowledge about uncertainty
  - ▶ Optimizes use of knowledge to make decisions

# The Magic of Probability

**Uncertainty:** vague, fuzzy, confusing, scary, hard to think about.

**Probability:**

Precise, unambiguous, simple(!) way to reason about uncertainty.



Uncertainty = Fear



Probability = Serenity

Our mission: help you discover the serenity of Probability, i.e., enable you to think clearly about uncertainty.

Your cost: focused attention and practice on examples and problems.

# Random Experiment: Flip one Fair Coin

Flip a **fair** coin: (*One flips or tosses a coin*)



- ▶ Possible outcomes: Heads ( $H$ ) and Tails ( $T$ ) (*One flip yields either 'heads' or 'tails'.*)
- ▶ Likelihoods:  $H : 50\%$  and  $T : 50\%$

# Random Experiment: Flip one Fair Coin

Flip a **fair** coin:



What do we mean by **the likelihood of tails is 50%**?

Two interpretations:

- ▶ Single coin flip: 50% chance of 'tails' **[subjectivist]**

*Willingness to bet on the outcome of a single flip*

- ▶ Many coin flips: About half yield 'tails' **[frequentist]**

*Makes sense for many flips*

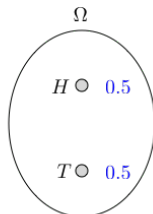
- ▶ Question: Why does the fraction of tails converge to the same value every time? **Statistical Regularity! Deep!**

# Random Experiment: Flip one Fair Coin

Flip a **fair** coin: model



Physical Experiment



Probability Model

- ▶ The physical experiment is complex. (Shape, density, initial momentum and position, ...)
- ▶ The Probability model is simple:
  - ▶ A set  $\Omega$  of **outcomes**:  $\Omega = \{H, T\}$ .
  - ▶ A **probability** assigned to each outcome:  
 $Pr[H] = 0.5, Pr[T] = 0.5$ .

# Random Experiment: Flip one Unfair Coin

Flip an **unfair** (biased, loaded) coin:



H: 45%

T: 55%

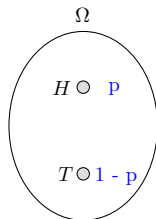
- ▶ Possible outcomes: Heads ( $H$ ) and Tails ( $T$ )
- ▶ Likelihoods:  $H : p \in (0, 1)$  and  $T : 1 - p$
- ▶ Frequentist Interpretation:
  - Flip many times  $\Rightarrow$  Fraction  $1 - p$  of tails
- ▶ Question: How can one figure out  $p$ ? Flip many times
- ▶ Tautology? No: **Statistical regularity!**

# Random Experiment: Flip one Unfair Coin

Flip an **unfair** (biased, loaded) coin: model



Physical Experiment



Probability Model



# Flip Two Fair Coins

- ▶ Possible outcomes:  $\{HH, HT, TH, TT\} \equiv \{H, T\}^2$ .
- ▶ Note:  $A \times B := \{(a, b) \mid a \in A, b \in B\}$  and  $A^2 := A \times A$ .
- ▶ Likelihoods:  $1/4$  each.



# Flip Glued Coins

Flips two coins glued together side by side:



Glued coins



50%



50%

- ▶ Possible outcomes:  $\{HT, TH\}$ .
- ▶ Likelihoods:  $HT : 0.5, TH : 0.5$ .
- ▶ Note: Coins are glued so that they show different faces.

# Flip two Attached Coins

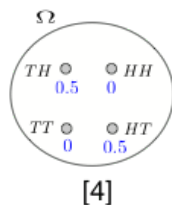
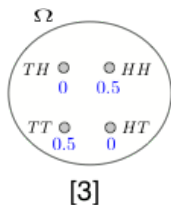
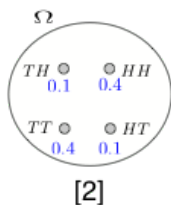
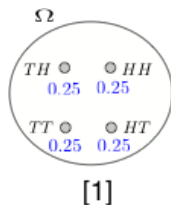
Flips two coins attached by a spring:



- ▶ Possible outcomes:  $\{HH, HT, TH, TT\}$ .
- ▶ Likelihoods:  $HH : 0.4, HT : 0.1, TH : 0.1, TT : 0.4$ .
- ▶ Note: Coins are attached so that they tend to show the same face, unless the spring twists enough.

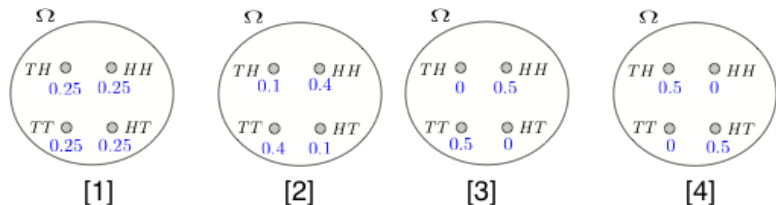
# Flipping Two Coins

Here is a way to summarize the four random experiments:



- ▶  $\Omega$  is the set of *possible* outcomes;
- ▶ Each outcome has a **probability** (likelihood);
- ▶ The probabilities are  $\geq 0$  and add up to 1;
- ▶ Fair coins: [1]; Glued coins: [3], [4];  
Spring-attached coins: [2];

# Flipping Two Coins



Important remarks:

- ▶ Each outcome describes the **two** coins.
- ▶ E.g.,  $HT$  is **one** outcome of each of the above experiments.
- ▶ **Wrong**: outcomes are  $\{H, T\}$  and one picks twice from that set.
- ▶ This viewpoint misses the relationship between the two flips.
- ▶ Each  $\omega \in \Omega$  describes one outcome of the **complete** experiment.
- ▶  $\Omega$  and the probabilities specify the random experiment.

# Poll

Flip two fair coins.

- (A) There are two outcomes. **No.**
- (B) The outcomes are H and T, twice. **No.**
- (C)  $\Omega = \{HH, HT, TH, TT\}$ . **Yes.**
- (D) A sample point  $\omega$  could be “H” **No.**
- (E) A sample point  $\omega$  could be “HT” **Yes.**

No to (A), (B), and (D).

# Flipping $n$ times

Flip a fair coin  $n$  times (some  $n \geq 1$ ):

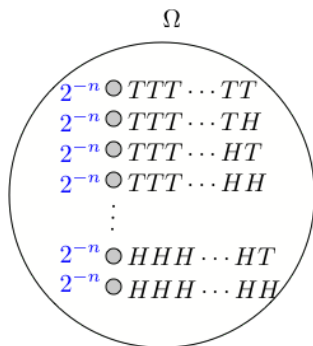
- ▶ Possible outcomes:  $\{TT \dots T, TT \dots H, \dots, HH \dots H\}$ .

Thus,  $2^n$  possible outcomes.

- ▶ Note:  $\{TT \dots T, TT \dots H, \dots, HH \dots H\} = \{H, T\}^n$ .

$A^n := \{(a_1, \dots, a_n) \mid a_1 \in A, \dots, a_n \in A\}$ .  $|A^n| = |A|^n$ .

- ▶ Likelihoods:  $1/2^n$  each.



# Poll

Flip coin  $n$  times.

(A)  $|\Omega| = 2^n$ .

(B) For  $n = 3$ ,  $HHH \in \Omega$ . Yes.

(C)  $Pr[HHH] = \frac{1}{2^3} = \frac{1}{8}$ .



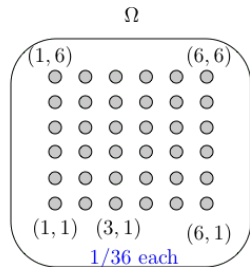
# Roll two Dice

Roll a **balanced** 6-sided die twice:

- ▶ Possible outcomes:  $\{1, 2, 3, 4, 5, 6\}^2 = \{(a, b) \mid 1 \leq a, b \leq 6\}$ .
- ▶ Likelihoods:  $1/36$  for each.



Physical Experiment



Probability Model

# Probability Space.

1. A “random experiment”:

- (a) Flip a biased coin;
- (b) Flip two fair coins;
- (c) Deal a poker hand.

2. A set of possible outcomes:  $\Omega$ .

- (a)  $\Omega = \{H, T\}$ ;
- (b)  $\Omega = \{HH, HT, TH, TT\}$ ;  $|\Omega| = 4$ ;
- (c)  $\Omega = \{ \underline{A\spadesuit A\diamondsuit A\clubsuit A\heartsuit K\spadesuit}, \underline{A\spadesuit A\diamondsuit A\clubsuit A\heartsuit Q\spadesuit}, \dots \}$   
 $|\Omega| = \binom{52}{5}$ .

3. Assign a probability to each outcome:  $Pr : \Omega \rightarrow [0, 1]$ .

- (a)  $Pr[H] = p, Pr[T] = 1 - p$  for some  $p \in [0, 1]$
- (b)  $Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4}$
- (c)  $Pr[ \underline{A\spadesuit A\diamondsuit A\clubsuit A\heartsuit K\spadesuit} ] = \dots = 1 / \binom{52}{5}$

# Probability Space: formalism.

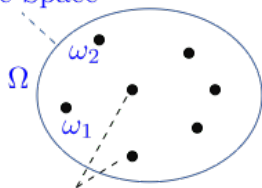
$\Omega$  is the **sample space**.

$\omega \in \Omega$  is a **sample point**. (Also called an **outcome**.)

Sample point  $\omega$  has a probability  $Pr[\omega]$  where

- ▶  $0 \leq Pr[\omega] \leq 1$ ;
- ▶  $\sum_{\omega \in \Omega} Pr[\omega] = 1$ .

Sample Space



Samples (Outcomes)

$$0 \leq Pr[\omega] \leq 1$$

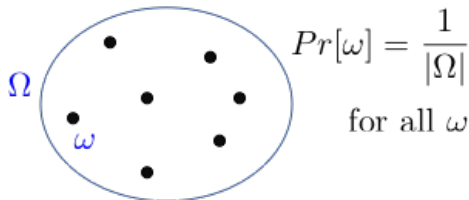
$$\sum_{\omega} Pr[\omega] = 1$$

# Probability Space: Formalism.

In a **uniform probability space** each outcome  $\omega$  is **equally probable**:

$$Pr[\omega] = \frac{1}{|\Omega|} \text{ for all } \omega \in \Omega.$$

## Uniform Probability Space



Examples:

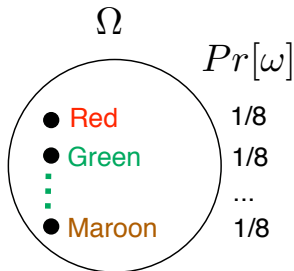
- ▶ Flipping two fair coins, dealing a poker hand are uniform probability spaces.
- ▶ Flipping a biased coin is not a uniform probability space.

# Probability Space: Formalism

Simplest physical model of a **uniform** probability space:



Physical experiment



Probability model

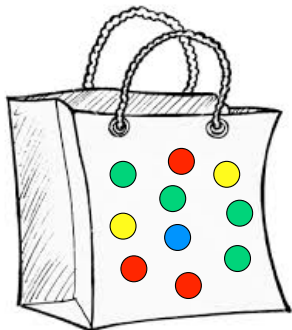
A bag of identical balls, except for their color (or a label). If the bag is well shaken, every ball is equally likely to be picked.

$$\Omega = \{\text{white, red, yellow, grey, purple, blue, maroon, green}\}$$

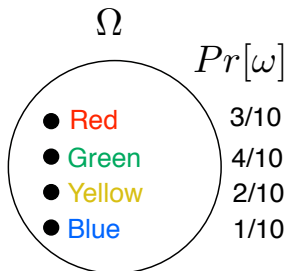
$$Pr[\text{blue}] = \frac{1}{8}.$$

# Probability Space: Formalism

Simplest physical model of a **non-uniform** probability space:



Physical experiment



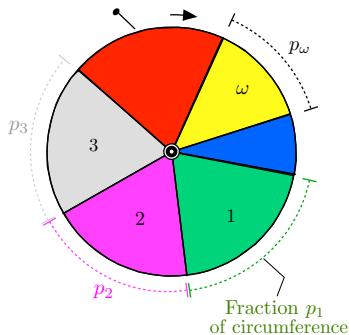
Probability model

$$\Omega = \{\text{Red, Green, Yellow, Blue}\}$$
$$Pr[\text{Red}] = \frac{3}{10}, Pr[\text{Green}] = \frac{4}{10}, \text{ etc.}$$

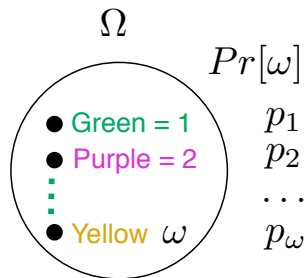
Note: Probabilities are restricted to rational numbers:  $\frac{N_k}{N}$ .

# Probability Space: Formalism

Physical model of a general **non-uniform** probability space:



Physical experiment



Probability model

The roulette wheel stops in sector  $\omega$  with probability  $p_\omega$ .

$$\Omega = \{1, 2, 3, \dots, N\}, Pr[\omega] = p_\omega.$$

## An important remark

- ▶ The random experiment selects **one and only one** outcome in  $\Omega$ .
- ▶ For instance, when we flip a fair coin **twice**
  - ▶  $\Omega = \{HH, TH, HT, TT\}$
  - ▶ The experiment selects *one* of the elements of  $\Omega$ .
- ▶ In this case, it's **wrong** to think that  $\Omega = \{H, T\}$  and that the experiment selects two outcomes.
- ▶ Why? Because this would not describe how the two coin flips are related to each other.
- ▶ For instance, say we glue the coins side-by-side so that they face up the same way. Then one gets *HH* or *TT* with probability 50% each. This is not captured by 'picking two outcomes.'



# Summary of Probability Basics

## Modeling Uncertainty: Probability Space

1. Random Experiment
2. Probability Space:  $\Omega$ ;  $Pr[\omega] \in [0, 1]$ ;  $\sum_{\omega} Pr[\omega] = 1$ .
3. Uniform Probability Space:  $Pr[\omega] = 1/|\Omega|$  for all  $\omega \in \Omega$ .

Bag of marbles.

With possibly different probabilities for each marble..

## Onwards in Probability.

Events, Conditional Probability, Independence, Bayes' Rule

## CS70: On to Events.

Events, Conditional Probability, Independence, Bayes' Rule

# Probability Basics Review

## Setup:

- ▶ Random Experiment.  
Flip a fair coin twice.
- ▶ Probability Space.
  - ▶ **Sample Space:** Set of outcomes,  $\Omega$ .  
 $\Omega = \{HH, HT, TH, TT\}$   
(Note: **Not**  $\Omega = \{H, T\}$  with two picks!)
  - ▶ **Probability:**  $Pr[\omega]$  for all  $\omega \in \Omega$ .  
 $Pr[HH] = \dots = Pr[TT] = 1/4$ 
    1.  $0 \leq Pr[\omega] \leq 1$ .
    2.  $\sum_{\omega \in \Omega} Pr[\omega] = 1$ .

## Set notation review

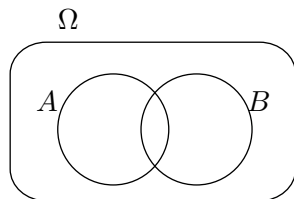


Figure: Two events

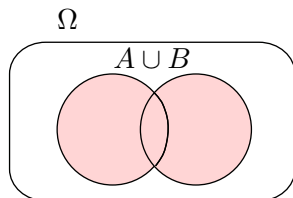


Figure: Union (or)

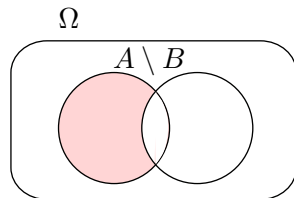


Figure: Difference ( $A$ ,  
not  $B$ )  
 $\Omega$

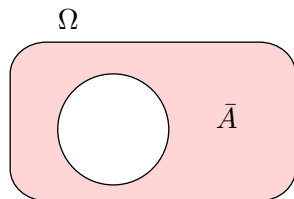


Figure: Complement  
(not)

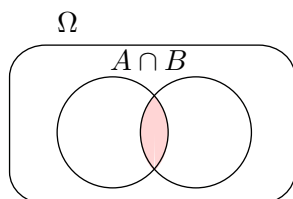


Figure: Intersection  
(and)

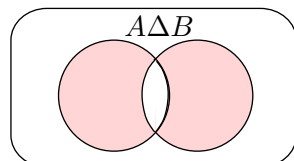


Figure: Symmetric  
difference (only one)

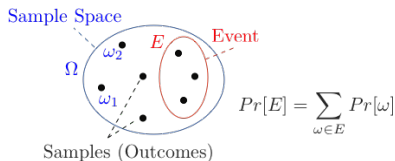
# Probability of exactly one 'heads' in two coin flips?

Idea: Sum the probabilities of all the different outcomes that have exactly one 'heads':  $HT, TH$ .

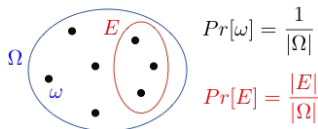
This leads to a definition!

## Definition:

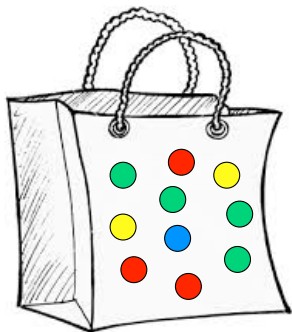
- ▶ An **event**,  $E$ , is a subset of outcomes:  $E \subset \Omega$ .
- ▶ The **probability of  $E$**  is defined as  $Pr[E] = \sum_{\omega \in E} Pr[\omega]$ .



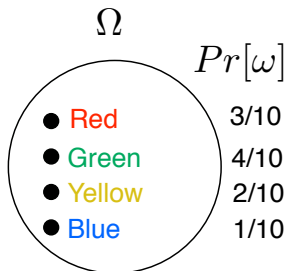
## Uniform Probability Space



## Event: Example



Physical experiment



Probability model

$$\Omega = \{\text{Red, Green, Yellow, Blue}\}$$
$$Pr[\text{Red}] = \frac{3}{10}, Pr[\text{Green}] = \frac{4}{10}, \text{ etc.}$$

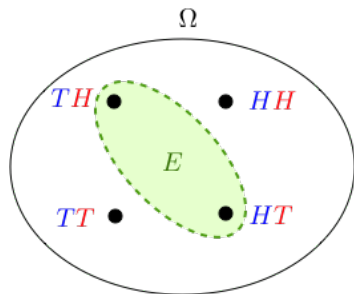
$$E = \{\text{Red, Green}\} \Rightarrow Pr[E] = \frac{3+4}{10} = \frac{3}{10} + \frac{4}{10} = Pr[\text{Red}] + Pr[\text{Green}].$$

# Probability of exactly one heads in two coin flips?

Sample Space,  $\Omega = \{HH, HT, TH, TT\}$ .

Uniform probability space:  $Pr[HH] = Pr[HT] = Pr[TH] = Pr[TT] = \frac{1}{4}$ .

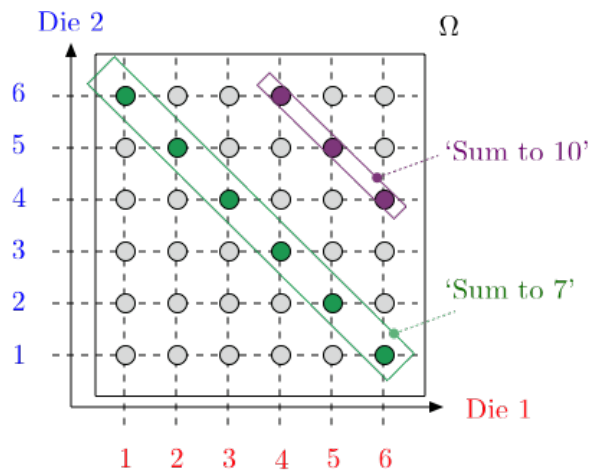
Event,  $E$ , "exactly one heads":  $\{TH, HT\}$ .



$$Pr[E] = \sum_{\omega \in E} Pr[\omega] = \frac{|E|}{|\Omega|} = \frac{2}{4} = \frac{1}{2}.$$



Roll a red and a blue die.



$$Pr[\text{Sum to } 7] = \frac{6}{36}$$

$$Pr[\text{Sum to } 10] = \frac{3}{36}$$

# Example and Polls: 20 coin tosses.

## 20 coin tosses

Sample space:  $\Omega =$  set of 20 fair coin tosses.

$$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}.$$

What is more likely?

(A)  $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$ , or

(B)  $\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0)$ ?

What is more likely?

(A) ( $E_1$ ) Twenty Hs out of twenty, or

(B) ( $E_2$ ) Ten Hs out of twenty?

# Example: 20 coin tosses.

## 20 coin tosses

Sample space:  $\Omega =$  set of 20 fair coin tosses.

$$\Omega = \{T, H\}^{20} \equiv \{0, 1\}^{20}; \quad |\Omega| = 2^{20}.$$

► What is more likely?

►  $\omega_1 := (1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)$ , or

►  $\omega_2 := (1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0)$ ?

Answer: Both are equally likely:  $Pr[\omega_1] = Pr[\omega_2] = \frac{1}{|\Omega|}$ .

► What is more likely?

( $E_1$ ) Twenty Hs out of twenty, or

( $E_2$ ) Ten Hs out of twenty?

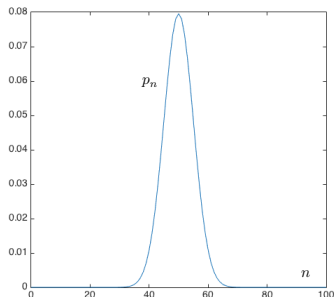
Answer: Ten Hs out of twenty.

Why? There are many sequences of 20 tosses with ten Hs; only one with twenty Hs.  $\Rightarrow Pr[E_1] = \frac{1}{|\Omega|} \ll Pr[E_2] = \frac{|E_2|}{|\Omega|}$ .

$$|E_2| = \binom{20}{10} = 184,756.$$

# Probability of $n$ heads in 100 coin tosses.

$$\Omega = \{H, T\}^{100}; |\Omega| = 2^{100}.$$



Event  $E_n = 'n \text{ heads}'; |E_n| = \binom{100}{n}$

$$p_n := Pr[E_n] = \frac{|E_n|}{|\Omega|} = \frac{\binom{100}{n}}{2^{100}}$$

Observe:

- ▶ Concentration around mean: **Law of Large Numbers**;
- ▶ Bell-shape: **Central Limit Theorem**.

## Exactly 50 heads in 100 coin tosses.

Sample space:  $\Omega = \text{set of 100 coin tosses} = \{H, T\}^{100}$ .

$$|\Omega| = 2 \times 2 \times \cdots \times 2 = 2^{100}.$$

Uniform probability space:  $Pr[\omega] = \frac{1}{2^{100}}$ .

Event  $E = \text{"100 coin tosses with exactly 50 heads"}$

$|E|$ ?

Poll: (A)  $100^{50}$

(B)  $2^{50}$ ,

(C)  $\binom{50}{100}$

(D)  $\binom{100}{50}$

(E) IDK - TBH.

Choose 50 positions out of 100 to be heads.

$$|E| = \binom{100}{50}.$$

$$Pr[E] = \frac{\binom{100}{50}}{2^{100}}.$$

# Calculation.

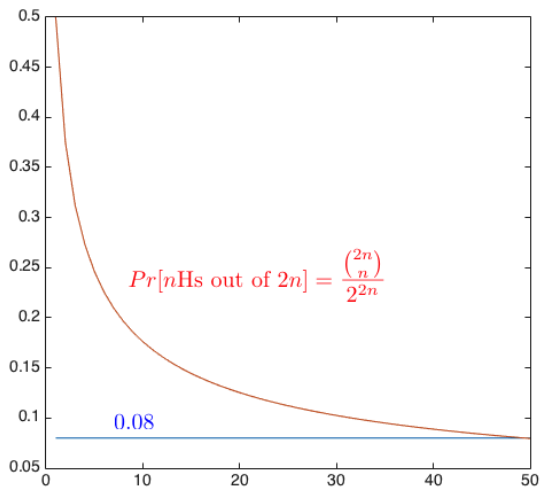
Stirling formula (for large  $n$ ):

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

$$\binom{2n}{n} \approx \frac{\sqrt{4\pi n}(2n/e)^{2n}}{[\sqrt{2\pi n}(n/e)^n]^2} \approx \frac{4^n}{\sqrt{\pi n}}.$$

$$Pr[E] = \frac{|E|}{|\Omega|} = \frac{|E|}{2^{2n}} = \frac{1}{\sqrt{\pi n}} = \frac{1}{\sqrt{50\pi}} \approx .08.$$

Exactly 50 heads in 100 coin tosses.



# Summary.

1. Random Experiment
2. Probability Space:  $\Omega$ ;  $Pr[\omega] \in [0, 1]$ ;  $\sum_{\omega} Pr[\omega] = 1$ .
3. Uniform Probability Space:  $Pr[\omega] = 1/|\Omega|$  for all  $\omega \in \Omega$ .
4. Event: "subset of outcomes."  $A \subseteq \Omega$ .  $Pr[A] = \sum_{\omega \in A} Pr[\omega]$
5. Some calculations.

Draw a marble from a bag.

(A) What is the set of marbles? Sample Space  $\Omega$ .

(B) What is a marble? Sample point:  $\omega \in \Omega$ .

(C) What is the set of blue marbles?  $A \subseteq \Omega$ . An event.