MATH 180A: INTRO TO PROBABILITY (FOR DATA SCIENCE)

Todd Kemp, APM 5202

www.math.ucsd.edu/~tkemp/180A

- ↳ Instructional team
- ↳ Course calendar (Google)
- ↳ Lecture schedule
- ↳ Homework — HW.0 due MONDAY, 09/30; HW.1 due FRIDAY, 10/04
- ↳ Datahub Labs → Lab.1 due MONDAY, 10/07
- ↳ Links to Piazza, Gradescope

Exams:

Wednesday, 10/23	8-10p	NO MAKEUP EXAMS
Wednesday, 11/20	8-10p	
Monday, 12/9	11:30a – 2:29p	MUST ATTEND

Think Pair Share

There are close to 180 students in this room. What are the odds that at least two share the same birthday?

(a) VERY unlikely

(b) $\frac{180}{365} \approx 50\%$

(c) $\binom{180}{2} / \binom{365}{2} \approx 25\%$

(d) VERY likely

The world around us is fundamentally random. 1.1

*
*
*
*
*

The modern rigorous foundation of probability theory goes back to

Ingredients

Sample Space Ω

Events $\mathcal{F}$

Probability Measure $\mathbb{P}$

Kolmogorov's Axioms

(i) For any event $A \subseteq \mathcal{F}$,

(ii) $\mathbb{P}(\Omega) =$ $\mathbb{P}(\emptyset) =$

(iii) If $A_1, A_2, A_3, \dots$ are events, then

$$\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \cdots) =$$

Eg. A fair die. $\Omega =$

$E =$ even, $O =$ odd

$T =$ divisible by 3

Eg. Two fair dice $\Omega =$

$D = \{$the sum of the two dice is 8$\}$

Uniform Probability Measure

Eg. A fair coin is tossed 3 times. $\Omega =$

$A = \{$ at least two tails$\}$

$B = \{$ exactly two tails$\}$

Random Sampling

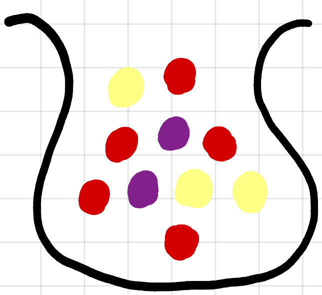
1.2

- choosing objects (perhaps repeatedly)
  (perhaps with replacement)

from a finite sample space –
with the uniform probability measure

$A \subseteq \Omega : \mathbb{P}(A) =$

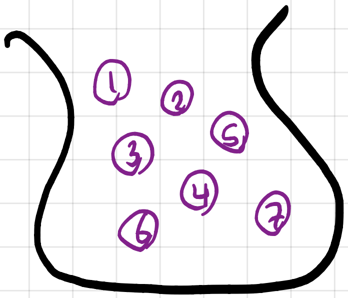
E.g.



An urn contains balls of various colors. We remove some number (either with or without replacement) and record some observation.

The sample space depends on the type of observation!

Sampling with replacement; order matters

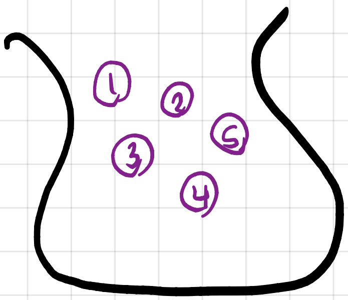


There are n balls, numbered $\{1, 2, \ldots, n\}$ in the urn. Choose one uniformly at random, record its #, then put it back. Do this k times.

$\Omega =$

* Repeated coin tosses are sampling with replacement ($n=$).
* Repeated die rolls are sampling with replacement ($n=$).

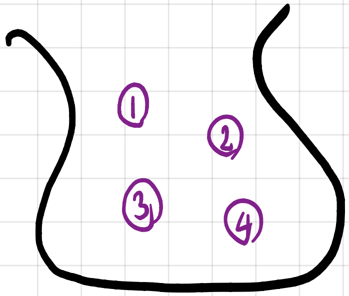
Sampling without replacement ; order matters



There are n balls, numbered $\{1, 2, \ldots, n\}$ in the urn. Choose one uniformly at random, record its #, then throw it away. Now there are $n-1$ balls in the urn. Sample another one, record it and throw it away. Repeat k times.

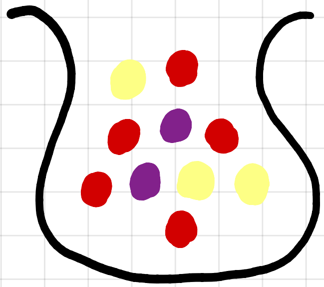
$$\Omega =$$

Sampling without replacement; order irrelevant



There are n balls, numbered $\{1, 2, \ldots, n\}$ in the urn. Select k of them uniformly at random. Record their labels, disregarding the order (e.g. always in increasing order).

$$\Omega =$$

E.g. 

An urn contains 10 balls:

2 blue

3 yellow

5 red

Problem: 3 balls are chosen without replacement.

$$P(2 \text{ yellow}, 1 \text{ red})$$