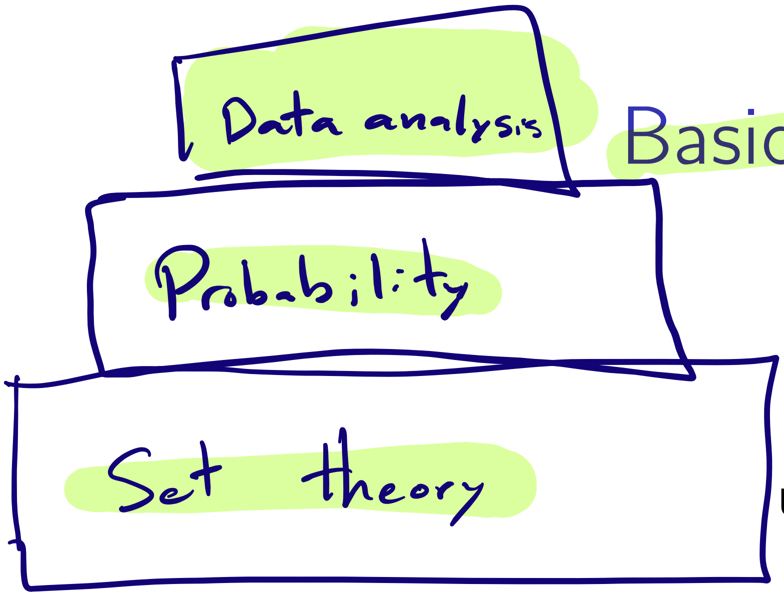
STAT 515 Lec 01 slides

Basics of sets, probability



Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Experiment

An *experiment* is a process which generates an outcome such that there is

(i) more than one possible outcome
(ii) the set of possible outcomes is known
(iii) the outcome is not known in advance

Sample space and sample points

- The *sample space* S of an experiment is the set of possible outcomes.
- The outcomes in a sample space S are called *sample points*.

pop of SC

Exercise: Give the sample space S for the following experiments:

1. Roll of a 6-sided die. $S = \{1, 2, 3, 4, 5, 6\}$
2. Number of COVID-19 cases confirmed in SC tomorrow. $S = \{0, 1, \ldots, N\}$ or $S = \{0, 1, \ldots\}$
3. Number of stars counted in a patch of space.
4. Blood type of a randomly selected student. $S = \{O+, O-, A+, A-, \ldots\}$
5. Time until you drop your new phone and crack the screen. $S = [0, \infty)$
6. Proportion of people in a sample with antibodies to a virus. $S = [0, 1]$
7. Deviation of today's temperature from the historical average.

$S = (-\infty, \infty)$

Event

An *event* is a collection of possible outcomes of an experiment, that is any subset of S (including S itself).

- Usually represent events with capital letters A, B, C, . . .
- Say an event A occurs if the outcome is in the set A.
- So events are equivalent to sets. Can refer to events as sets, to sets as events.
- Often refer to members of sets as *elements* of the set.

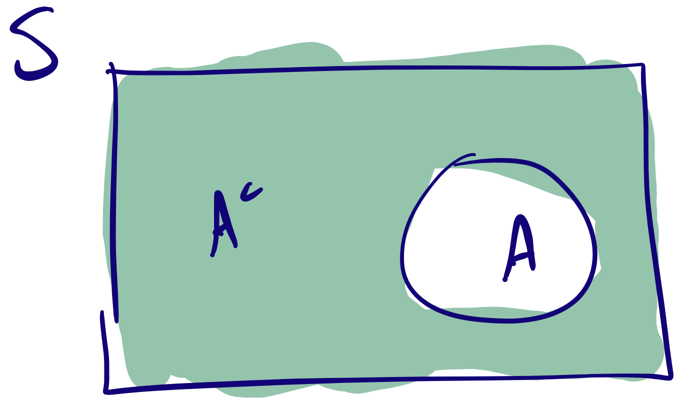
Exercise: Express the following events as subsets of the sample space.

1. Roll of a 6-sided die: $S = \{1,2,3,4,5,6\}$ ~~$S = [1,6]$~~ ?
   A = odd number rolled. $A = \{1, 3, 5\}$
2. Number of COVID-19 cases confirmed in SC tomorrow: $S = \{0, 1, \dots\}$
   A = no new cases. $A = \{0\}$
3. Number of stars counted in a patch of space:
   A = more than 1,000.
4. Blood type of a randomly selected student: $S = \{O+, O-, A+, \dots\}$
   A = has A antigen in RBCs. $A = \{A+, A-, AB+, AB-\}$
5. Time until you drop your new phone and crack the screen: $S = [0, \infty)$
   A = within the first month. $A = [0, 1)$
6. Proportion of people in a sample with antibodies to a virus:
   A = 60% or more with antibodies.
7. Deviation of today's temperature from the historical average:
   A = at least 10 degrees cooler.

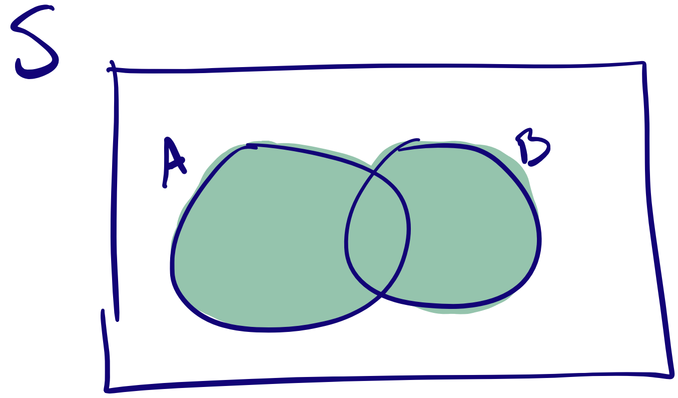
Elementary set operations

- The *complement set* of A is all the sample points in S which are not in A. Denote by A^c, "A complement"
- The *union set* of A and B is the set of sample points in A or in B or in both A and B. Write $A \cup B$. "A union B"
- The *intersection set* of A and B is the set of all point in both A and B. Write $A \cap B$, say "A intersect B".

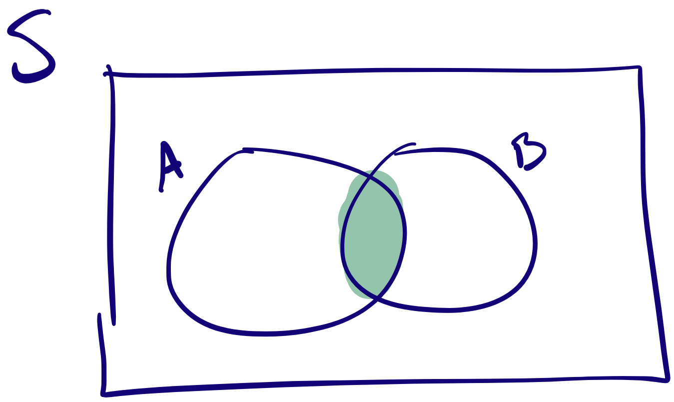
Draw pictures.



A^c



$A \cup B$



$A \cap B$

Exercise: Suppose we flip a coin three times and record the sequence of heads and tails.

1. Give the sample space S.
2. Give the points in the event $A =$ at least two heads come up.
3. Give the points in the event $B =$ at least one tails comes up.
4. Give the points in the event $B \cap A$.
5. Give the points in the event $B^c \cup A^c$.
6. Give the points in the event $(B \cup A)^c$.
7. Give the points in the event $(B \cap A)^c$.
8. ~~Give the points in A^c.~~

① $S = \{HHH \quad HHT \quad HTH \quad THH \quad TTH \quad THT \quad HTT \quad TTT\}$

② $A = \{HHH \quad HHT \quad HTH \quad THH\}$ $\qquad A^c = \{TTH \quad THT \quad HTT \quad TTT\}$

③ $B = \{HHT \quad HTH \quad THH \quad TTH \quad THT \quad HTT \quad TTT\}$, $B^c = \{HHH\}$

④ $A \cap B = \{HHT \quad HTH \quad THH\}$

⑤ $B^c \cup A^c = \{HHH \quad TTH \quad THT \quad HTT \quad TTT\}$

⑥ $(B \cup A)^c = S^c = \emptyset$ ← The empty set

⑦ $(B \cap A)^c = \{HHT \quad HTH \quad THH\}^c = \{HHH \quad TTH \quad THT \quad HTT \quad TTT\}$

Exercise: Suppose we roll a 6-sided die twice. The sample space is

$$S = \left\{ \begin{array}{cccccc} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{array} \right\}.$$

Define the events

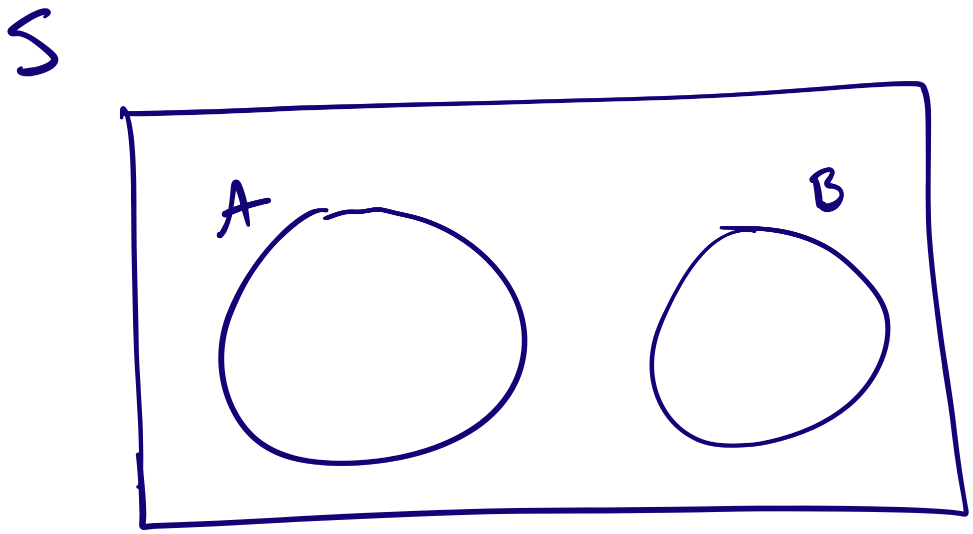
$$\begin{aligned} A &= \{\text{at least one of the rolls is odd}\} \\ B &= \{\text{sum of the rolls is a prime number}\} \\ C &= \{\text{sum of the rolls is an even number}\} \end{aligned}$$

1. Give $A \cup B$.
2. Give $A \cap B$.
3. Give $B \cap C$.

Mutual exclusivity/disjoint-ness

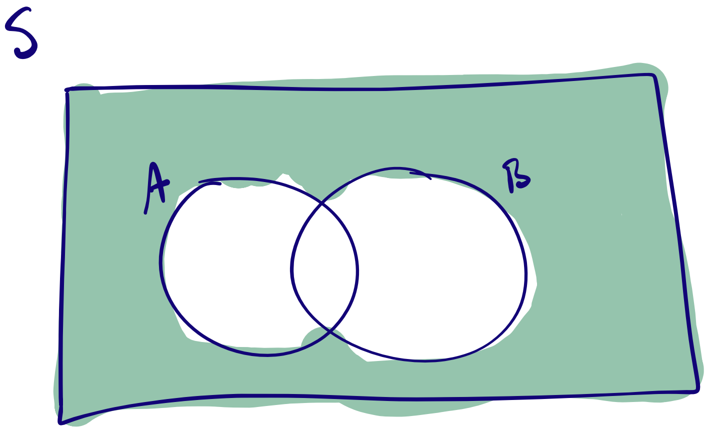
Two events A and B are called *mutually exclusive* or *disjoint* if $A \cap B = \emptyset$.

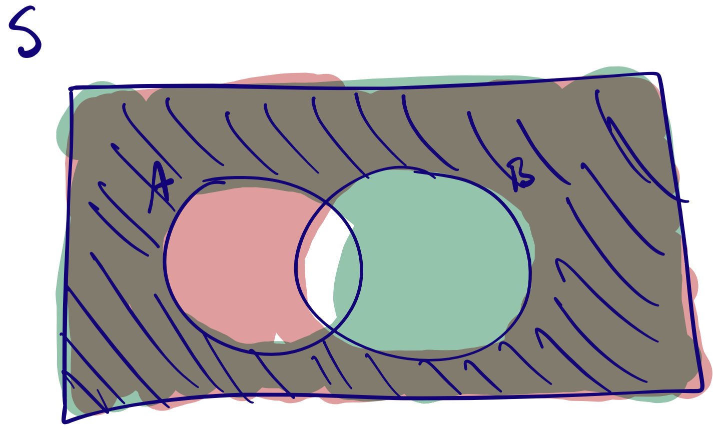
The set $\emptyset$ is the the *empty set*, which is the set containing no elements.



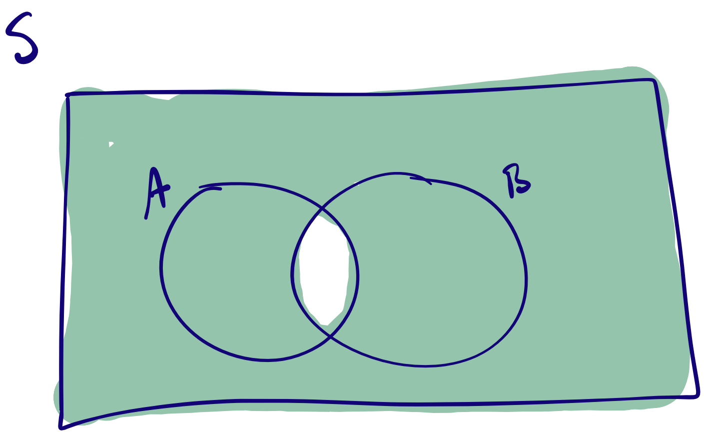
De Morgan's Laws

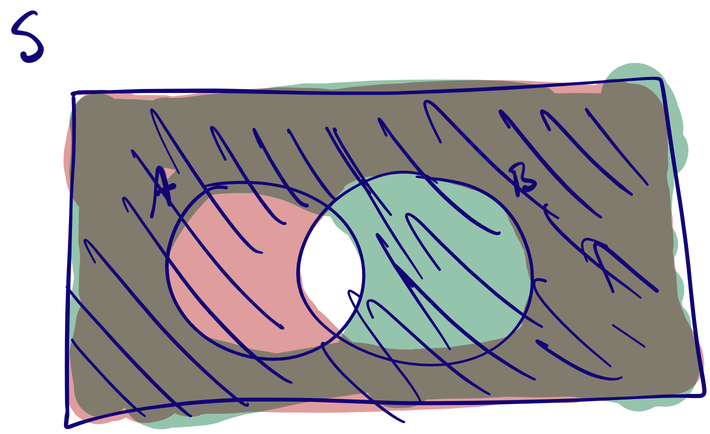
$(A \cup B)^c = A^c \cap B^c$





$(A \cap B)^c = A^c \cup B^c$





De Morgan's Laws

For any events A and B we have

$$(A \cup B)^c = A^c \cap B^c \quad \text{and} \quad (A \cap B)^c = A^c \cup B^c$$

Draw pictures.

Exercise: A birthday party might have cake C and it might have ice cream I. Interpret in words the events

1. $(C \cap I)^c = C^c \cup I^c$ There will not be both cake and ice cream.
2. $(C \cup I)^c = C^c \cap I^c$ No cake and no ice cream.

Exercise: For two events A and B use our elementary set operations $\cup$, $\cap$, and the complement to give representations of the following:

- Both A and B occur. $A \cap B$
- Neither A nor B occurs; give two representations of this. $A^c \cap B^c = (A \cup B)^c$
- At least one of the events A and B occurs. $A \cup B$
- At least one of the events $\boxed{A \text{ and } B \text{ does not occur;}}$ give two representations. $A^c \cup B^c = (A \cap B)^c$
- One of the events A and B occurs but not the other.

$(A \cap B^c) \cup (B \cap A^c)$

Exercise: Suppose a wildebeest W, a crocodile C, and a giraffe G on a safari. Write down the following events using elementary set operations. You see...

1. a giraffe $\boxed{\text{but no}}$ wildebeest. (and not) $G \cap W^c$
2. all three types of animals. $W \cap C \cap G$
3. not all three types of animals. $(W \cap C \cap G)^c = W^c \cup C^c \cup G^c$
4. a giraffe and a wildebeest but no crocodile. $G \cap W \cap C^c$
5. not both a giraffe and a wildebeest. $(G \cap W)^c$
6. a giraffe and a wildebeest without seeing a crocodile or a crocodile and a wildebeest without seeing a giraffe. $((G \cap W) \cap C^c) \cup ((C \cap W) \cap G^c)$
7. exactly two of the three types of animal. $(W \cap C \cap G^c) \cup (W \cap C^c \cap G^c) \cup (W^c \cap C \cap G)$
8. exactly one of the three types of animal. $(W \cap C^c \cap G^c) \cup (W^c \cap C \cap G^c) \cup (W^c \cap C^c \cap G)$
9. at least one of the three types of animal. $W \cup C \cup G = (W^c \cap C^c \cap G^c)^c$
10. none of the animals. $W^c \cap C^c \cap G^c = (W \cup C \cup G)^c$

For an event A, we denote by $\boxed{P(A)}$ the probability that A occurs.

Examples:

1. Roll a die and let A be the event that ⚃ is rolled. Then $P(A) = 1/6$.
2. Roll two dice and let B be the event that ⚀, ⚀ is rolled. Then $P(B) = 1/36$.
3. Let C be the event that you get to park in your favorite spot. Then $P(C) =$ some number.

Now some axioms about probabilities...

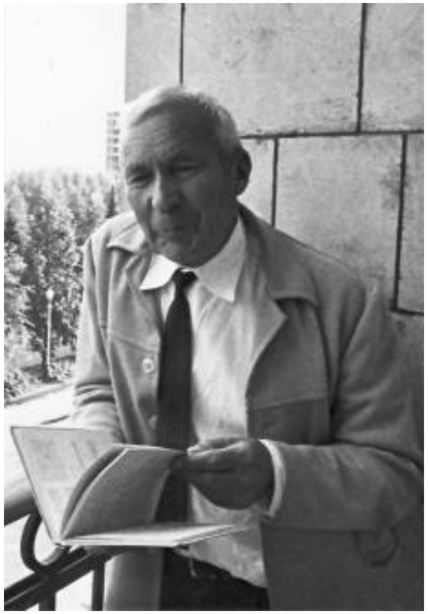
This is the first of the three Колмогоров $\boxed{\text{axioms}}$ of probability.

Kolmogorov

Axiom (Nonnegativity of probabilities)

For any event A we must have $P(A) \geq 0$.

Here is Колмогоров himself. Is he thinking about the axioms?



Axiom (Unity of the sample space probability)

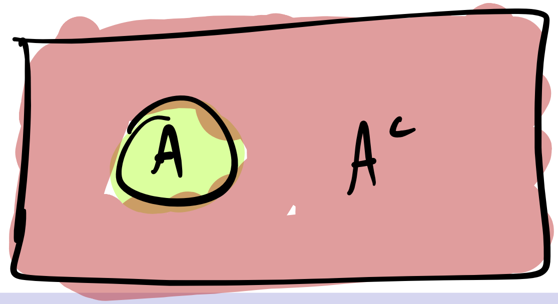
For a statistical experiment with sample sample S, $P(S) = 1$.

Examples:

- Rolling a die: $S = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$. We have $P(S) = 1$.
- MPG on next tank of gas: $S = [0, \infty)$. We have $P(S) = 1$.
- # students in class of 40 with cloth masks: $S = \{0, 1, \dots, 40\}$. $P(S) = 1$.

$P(S) = 1$

S



Axiom (Probability of a complement)

For an event A with complement A^c, $P(A^c) = 1 - P(A)$.

Examples:

- Roll a die and let $A = \{⚀, ⚁\}$. Then $P(A) = 1/3$. Give $P(A^c)$. $= 1 - P(A)$
  $= 2/3$
- We have $P(\emptyset) = 0$. Why?

$$P(\emptyset) = P(S^c) = 1 - P(S) = 1 - 1 = 0.$$

Result (Probability of a union)

For any two events A and B, we have $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Draw a picture...

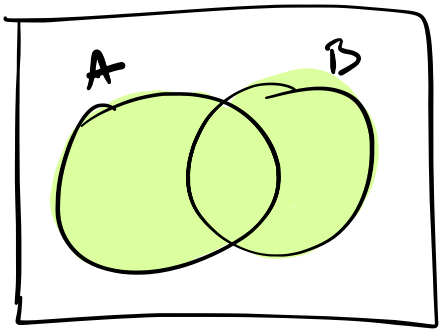
Exercise: Let F_1 and F_2 be the events that you meet friend 1 and friend 2 at a party, respectively. Suppose $P(F_1) = 0.4$, $P(F_2) = 0.5$ and $P(F_1 \cap F_2) = 0.01$.

Find the probabilities of the following:

1. You meet at least one of the two friends at a party. $P(F_1 \cup F_2)$
2. You do not meet either of your friends at a party. $P((F_1 \cup F_2)^c)$
3. You meet friend 1 but not friend 2 at a party. $P(F_1 \cap F_2^c)$

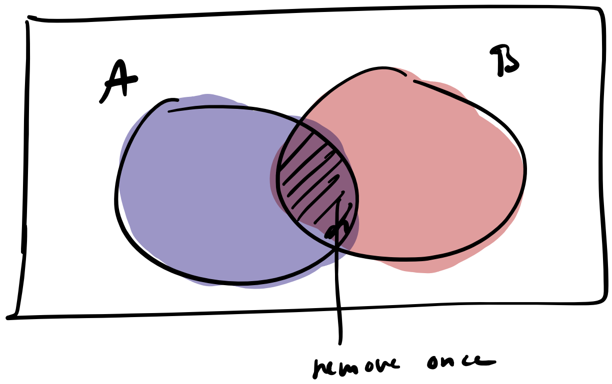
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

S



=

S



① $P(F_1 \cup F_2) = P(F_1) + P(F_2) - P(F_1 \cap F_2)$

$= 0.4 + 0.5 - 0.01$

$= 0.89$

② $P((F_1 \cup F_2)^c) = 1 - P(F_1 \cup F_2)$

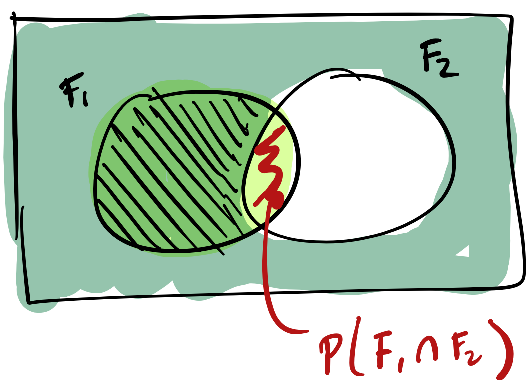
$= 1 - 0.89$

$= 0.11$

③ $P(F_1 \cap F_2^c) = P(F_1) - P(F_1 \cap F_2) = 0.4 - 0.01$

$= 0.39.$

S



Result (Nullity of empty set probability)

We have $P(\emptyset) = 0$.

Result (Probability of the union of mutually exclusive events)

If A *and* B *are mutually exclusive events, then* $P(A \cup B) = P(A) + P(B)$.

$P(A \cup B) = P(A) + P(B) - \underbrace{P(A \cap B)}_{=0 \text{ if } A \cap B = \emptyset} = P(A) + P(B)$ if $A \cap B = \emptyset$.

Exercise: You choose a ski slope at random from among 20 slopes, with 7 marked with a green circle, 5 with a blue square, 7 with a black diamond, and 1 with double black diamonds. Give the probability that you choose a slope marked with

1. a green circle. $P(G) = 7/20$
2. a green circle or a blue square. $P(G \cup B) = P(G) + P(B) = \frac{7}{20} + \frac{5}{20} = \frac{12}{20}$
3. anything but double black diamonds.

$P((\text{Double black diamond})^c) = 1 - P(\text{Double black diamond}) = 1 - \frac{1}{20} = \frac{19}{20}$

$A \overset{\text{subset}}{\subset} S$: A is a subset of S.

In some situations we can compute probabilities by counting sample points.

Computing probabilities when all sample points equally likely

If all outcomes in S are equally likely, for any event $A \subset S$, we have

$$P(A) = \frac{\#\{\text{sample points in } A\}}{\#\{\text{sample points in } S\}}.$$

Exercise: Suppose we roll two 6-sided dice. Find

1. $P(\text{sum of rolls equals 7})$
2. $P(\text{we roll doubles})$
3. $P(\text{sum of rolls at least 10})$
4. $P(\text{at least one roll is greater than 3})$

roll 1 roll 2

$$S = \left\{ \begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{matrix} \right\}.$$

1. $P(\text{Sum} = 7) = \dfrac{6}{36}$

2. $P(\text{Double}) = \dfrac{6}{36}$

3. $P(\text{Sum} \geq 10) = \dfrac{6}{36}$

4. $P(\text{At one roll} > 3) = \dfrac{27}{36}$

Exercise: A jukebox will play a song at random from a library of 50 songs, of which 20 are R&B, 10 are country, 12 are pop, and 8 are rock. Of the R&B songs, 15 are less than a decade old, and of each other genre, half of the songs are less than a decade old. Find the probability that the jukebox plays

1. A pop song. $P(\text{Pop}) = \frac{12}{50}$
2. An R&B song less than a decade old. $P(\text{R\&B} \cap < 10\text{ yrs}) = \frac{15}{50}$
3. A song less than a decade old. $P(< 10\text{ yrs}) = \frac{30}{50}$
4. A rock song more than a decade old. $P(R \cap \geq 10\text{ yrs}) = \frac{4}{50}$
5. A country song or any song that is more than a decade old.

	R&B	Country	pop	Rock	total
$\geq$ 10 yrs old	5	5	6	4	20
< 10 yrs old	15	5	6	4	30
total	20	10	12	8	50

5 $P(C \cup \geq 10\text{ yr}) = P(C) + P(\geq 10\text{ yr}) - P(C \cap \geq 10\text{ yr})$

$$= \frac{10}{50} + \frac{20}{50} - \frac{5}{50}$$

$$= \frac{25}{50}$$

$$= \frac{5 + 5 + 6 + 4 + 5}{50}$$

$$= \frac{25}{50}.$$