Florida State University Bayesian Workshop

Applied Bayesian Analysis for the Social Sciences

Day 1: Probability Review

Ryan Bakker
University of Georgia

Random Variables

- Flip a coin. Will it be heads or tails?
- The outcome of a single event is **random**, or unpredictable
- What if we flip a coin 10 times? How many will be heads and how many will be tails?
- Over time, patterns emerge from seemingly random events. These allow us to make probability statements.

Heads or Tails?

A coin toss is a random event [H or T] unpredictable on each toss but a stable pattern emerges of 50:50 after many repetitions.

- The French naturalist, Buffon (1707-1788) tossed a coin 4040 times; resulting in 2048 heads for a **relative frequency** of

  2048 /4040 = .5069

Heads or Tails?

- The English mathematician John Kerrich, while imprisoned by Germans in WWII, tossed a coin 5,000 times, with result 2534 heads . What is the Relative Frequency?
- 2,534 / 5,000 = .5068

Kerrich. That's it.

Heads or tails?

- A computer simulation of 10,000 coin flips yields 5040 heads. What is the relative frequency of heads?
- 5040 / 10,000 = .5040

Each of the tests is the result of a **sample** of fair coin tosses.

Sample outcomes vary.

- Different samples produce different results. True, but the **law of large numbers** tells us that the greater the number of repetitions the closer the outcomes come to the true probability, here .5.

A single event may be **un**predictable but the **relative frequency** of these events is **lawful** over an infinite number of trials\repetitions.

Random Variables

- "**X**" denotes a random variable. It is the outcome of a **sample of trials.**
- **"X**," some event, is **un**predictable in the short run but lawful over the long run.
- This "Randomness" is <u>not necessarily unpredictable</u>. Over the long run X becomes probabilistically predictable.
- We can never observe the "real" probability, since the "true" probability is a concept based on an infinite number of repetitions/trials. It is an "idealized" version of events

To figure the odds of some event occurring you need 2 pieces of information:

1. A list of all the possibilities – all the possible outcomes (sample space)
2. The number of ways to get the outcome of interest (relative to the number of possible outcomes).

Take a single Dice Roll

- Assuming an evenly-weighted 6-sided dice, what are the odds of rolling a 3?
- How do you know?
  - 6 possible outcomes (equally likely)
  - 1 way to get a 3
  - p(Roll=3) = 1 / 6

- What are the chances of rolling numbers that add up to “4” when rolling two six-sided dice?

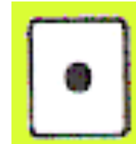
- What do we need to know?
  – All Possible Outcomes from rolling two dice
  – Outcomes that would add up to 4

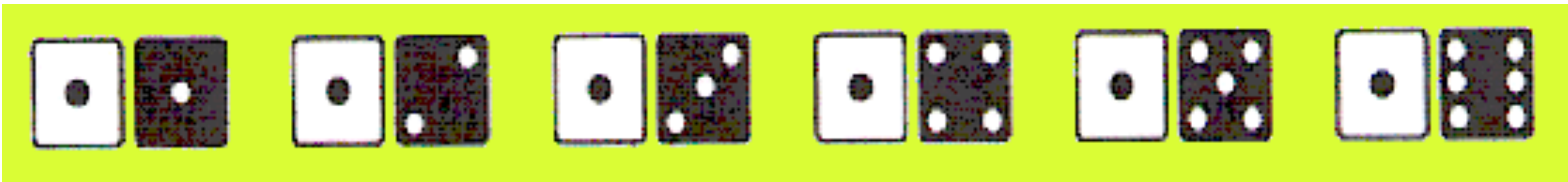
How Many Ways can the Two Dice Fall?

Let’s say the dice are different colors (helps us keep track.

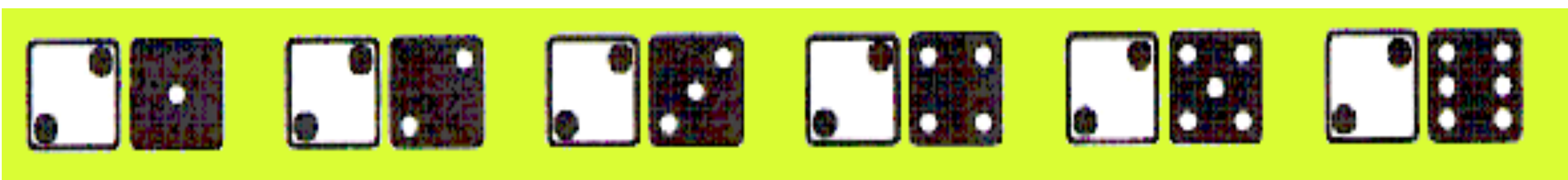
The White Dice could come out as:

We know how to figure out probabilities here, but What about the other dice?

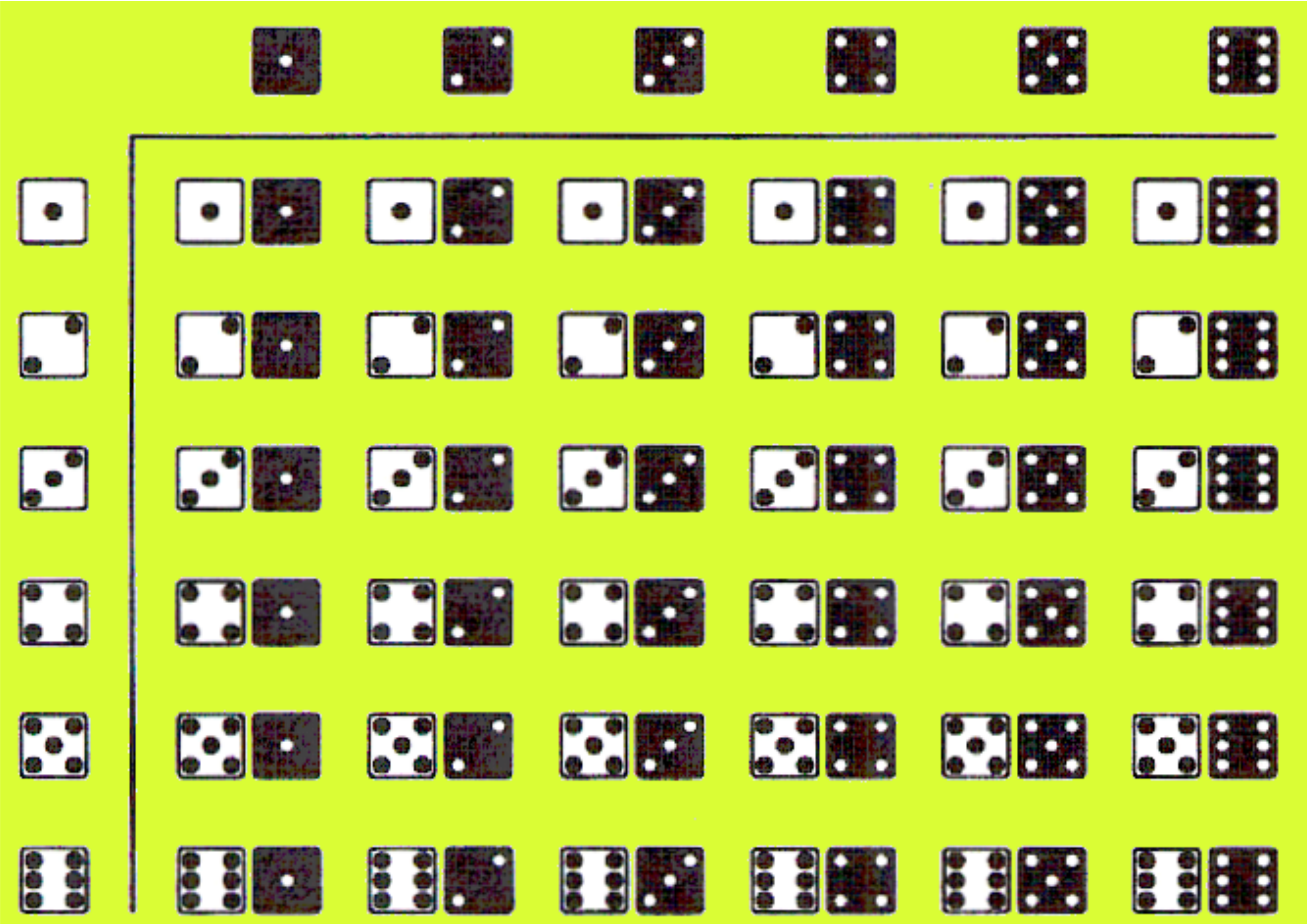
- When the white die shows  , there are six possible outcomes.



- When the white die shows  , there are six more possible outcomes.

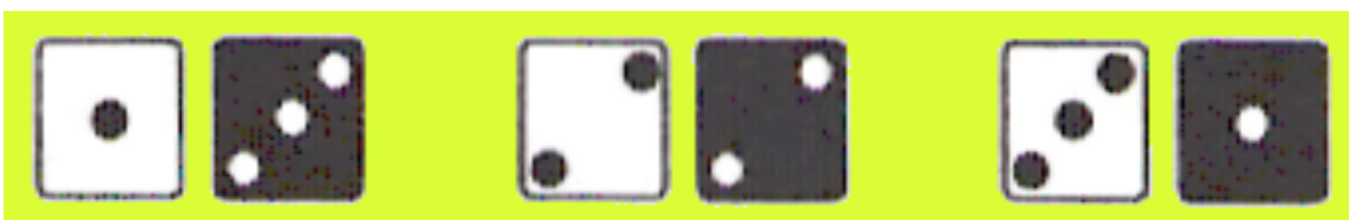


- We then just do that for all six possible outcomes on the white die



- Remember the Question: What is the probability of Rolling numbers that sum to 4?

- What do we need to know?
  - All Possible Outcomes from rolling two dice
    - (36--Check Previous Slide)
  - How many outcomes would add up to 4?



Our Probability is 3/36 = .08333

$$\text{Probability} = \frac{\text{Frequency of Occurrence}}{\text{Total \# outcomes}}$$

Frequency of occurrence **=** # of ways this one event could happen

Total # outcomes **=** # ways all the possible events could happen

Probability of a 7 is 6 ways out of 36 possibilities → p=.166

Frequency of Sum of 2 Dice

```
          6 -                             *  p = .167
            -
            -
          5 -              p = .139 *          *  p = .139
 F          -
 R          -
 E        4 -          p =.111 *                    *  p = .111
 Q          -
 U          -
 E          -
 N        3 -     p=.083  *                              * p=.083
 C          -
 Y          -
          2 -        * p=.056                          p=.056  *
            -
            -
          1 -   * p=.027                                   p=.027   *
              --+----+----+----+----+----+----+----+----+----+----+

              2.0       4.0       6.0       8.0      10.0      12.0
                                    SUM OF 2 DICE
```

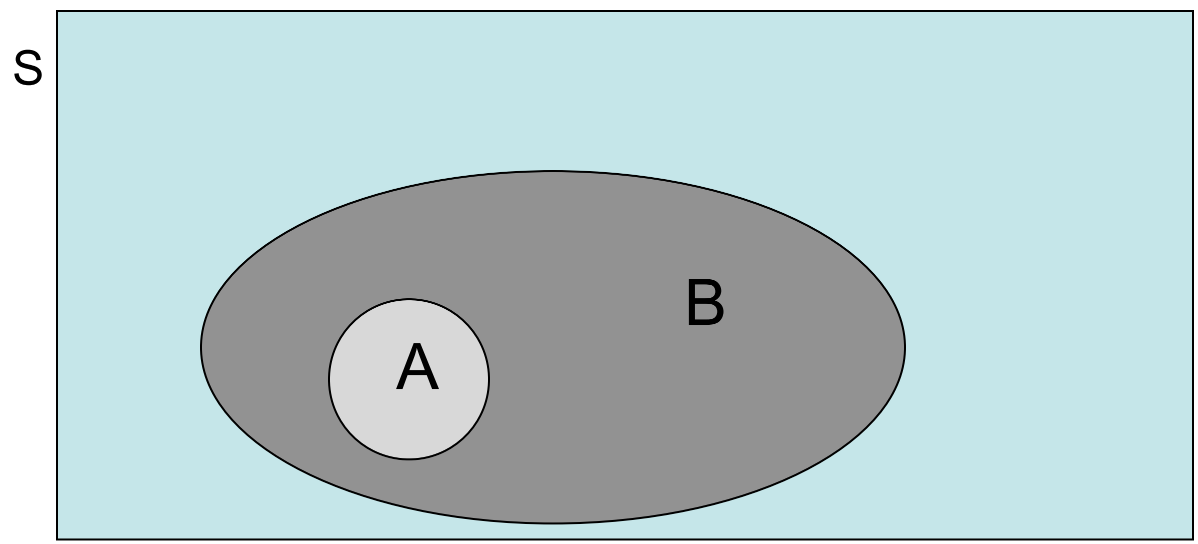
Review of Set Notation

- Capital Letters sets of points
- Lower case letters represent elements of the set
- For example:

  $A = \{a_1, a_2, a_3\}$

Subsets

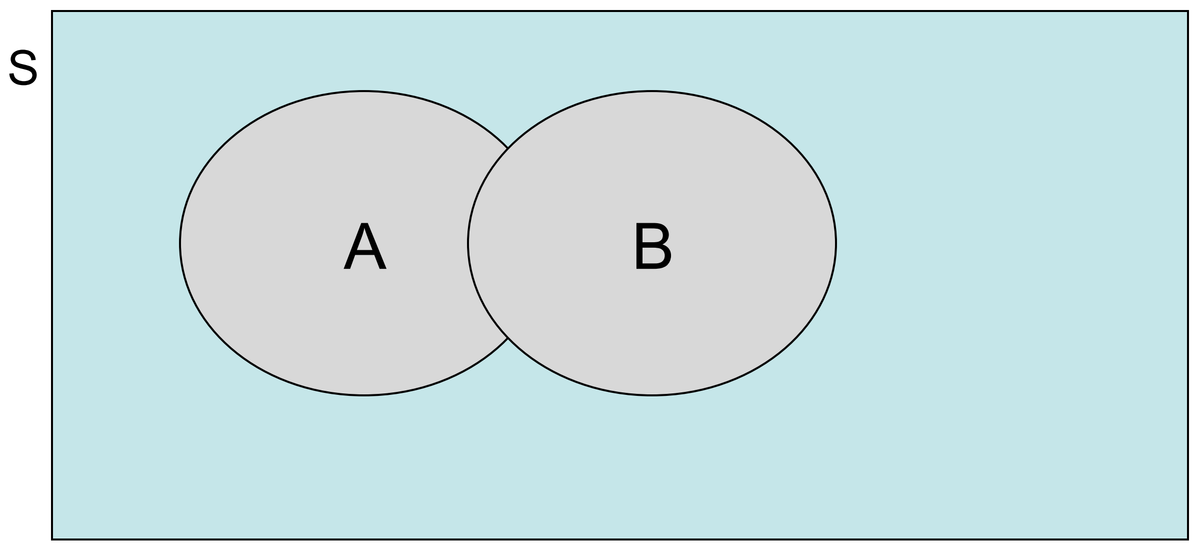
- Let S denote the full sample space (the set of all possible elements)
- For two sets A and B, if every element of A is also an element in B, we say that A is a *subset* of B $A \subset B$



Union

- The union of two arbitrary sets of points is the set of all points that are in at least one of the sets

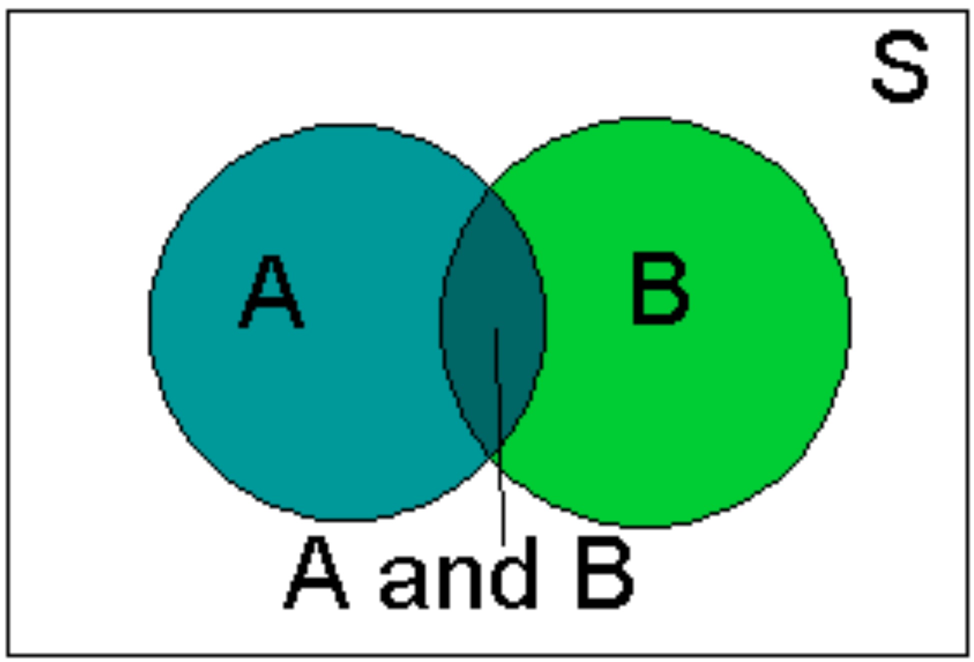
$$A \cup B$$



Intersection

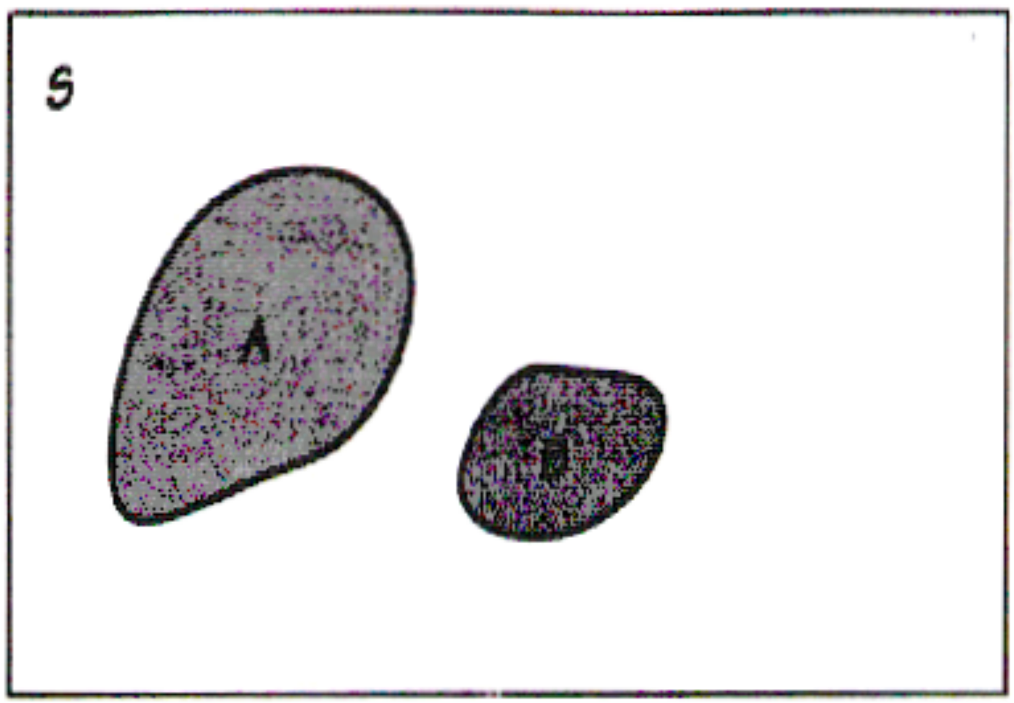
- The intersection of two arbitrary sets of points is the set of all points that are in both of the sets

$$A \cap B$$



Mutual Exclusivity

- Two events are said to be *disjoint* or *mutually exclusive* if none of the elements in set A appear in set B.



Independence

- We will give a more rigorous definition later, but…
- Two events are independent if the occurrence of A is unaffected by the occurrence or nonoccurrence of B.
- Example: You flip a coin—what is the probability of heads?
- You flip it 10 times, getting heads each time. What is the probability of getting heads again?

Axioms for Probabilities

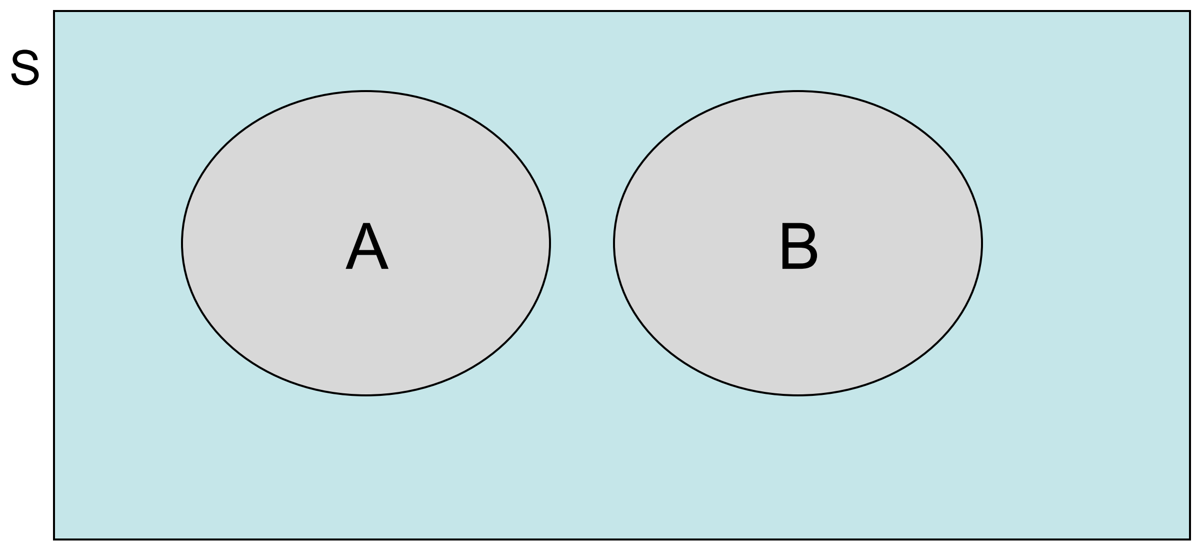
- The conventional rules for probabilities are named the Kolmogorov Axioms. They are:
  1. $P(A) \geq 0$
  2. $P(S) = 1$
  3. If A_1, A_2, A_3, … are pairwise mutually exclusive events in S, then:

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots) = \sum P(A_i)$$

Rules for Calculating Probabilities

- Simple Additive rule for disjoint events
  - a.k.a. the “or” rule

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



Example:

- One community is 75% white (non-hispanic), 10% black (non-hispanic), and 15% hispanic. They choose their mayor at random to maximize equality.
- What is the probability that the next mayor will be non-white?

$$P(Black \cup Hispanic) = P(Black) + P(Hispanic)$$

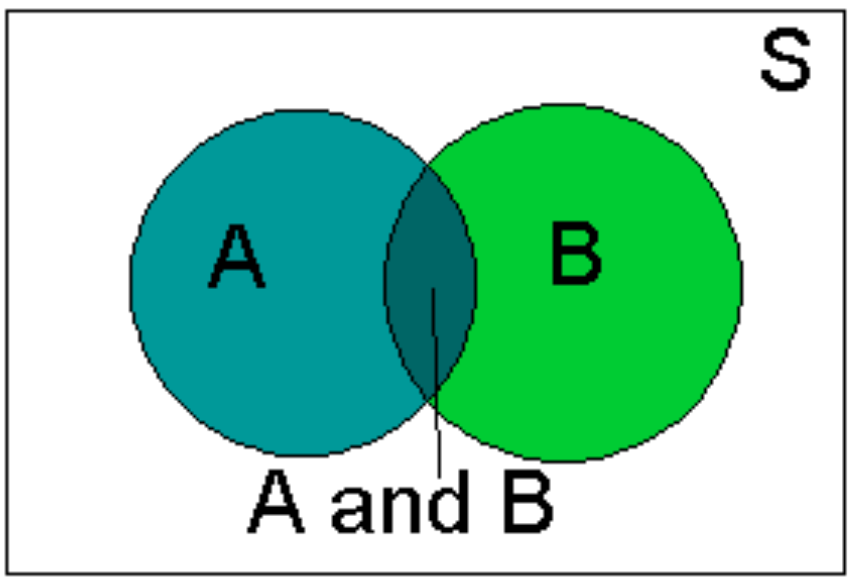
$$P(Black \cup Hispanic) = .1 + .15$$

$$P(Black \cup Hispanic) = .25$$

Rules for Calculating Probabilities

- Simple Multiplicative rule for independent events
  - a.k.a. the "and" rule

$$P(A \cap B) = P(A) * P(B)$$



Example:

- Suppose in that same mythical community (75% white, 10% black, 15% Hispanic) there was an even division of males and females. What is the probability of a white male mayor?

$$P(White \cap Male) = P(White) * P(Male)$$

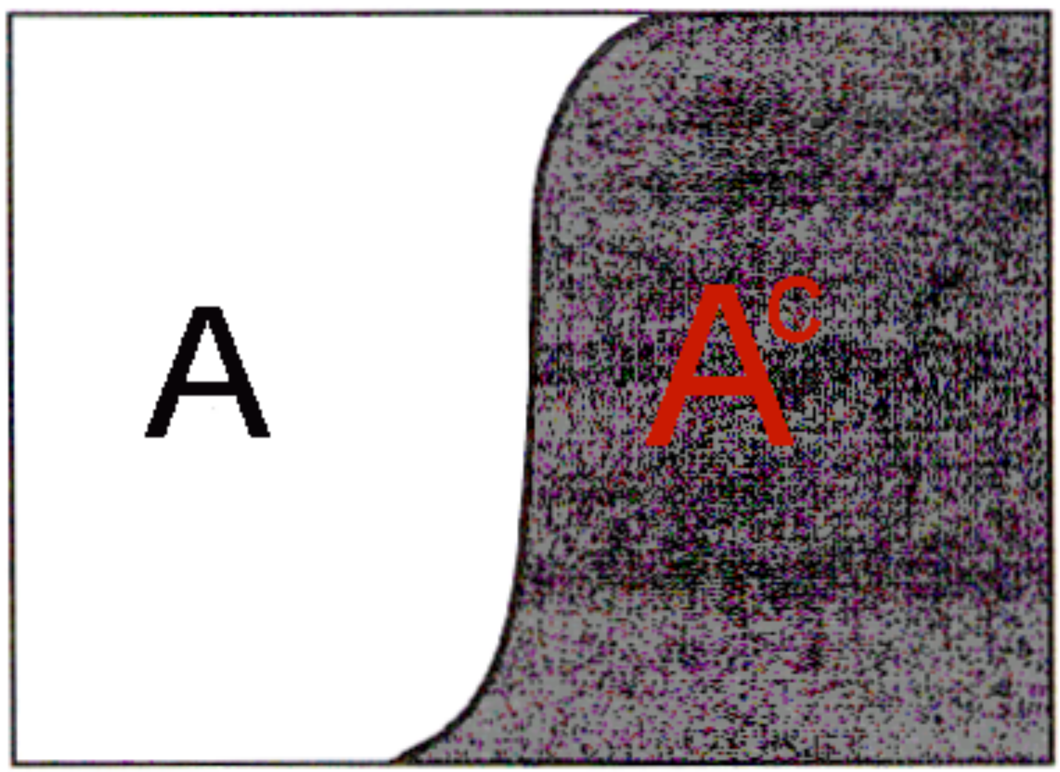
$$P(White \cap Male) = (.75) * (.5)$$

$$P(White \cap Male) = .375$$

Rules for Calculating Probabilities

- The Complement Rule

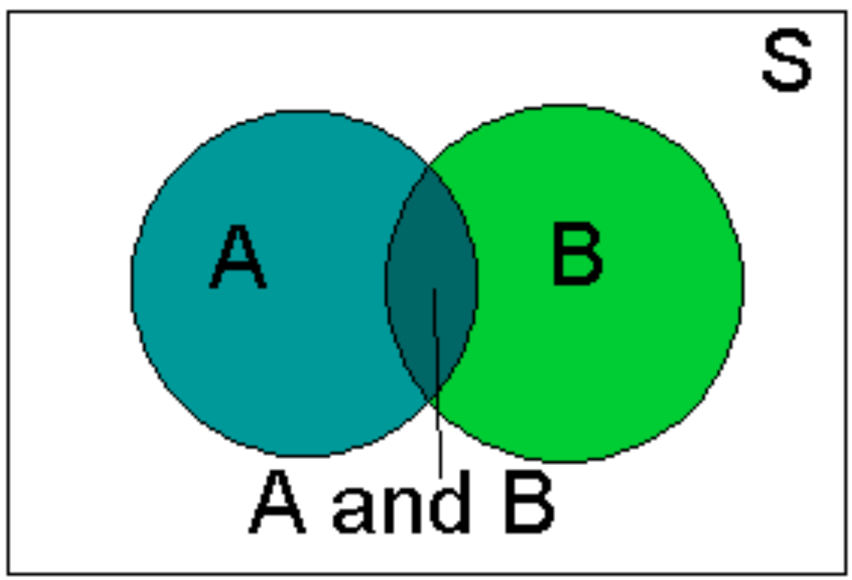
$$P(A^C) = 1 - P(A)$$



Rules for Calculating Probabilities

- Additive rule for events that are not mutually exclusive events

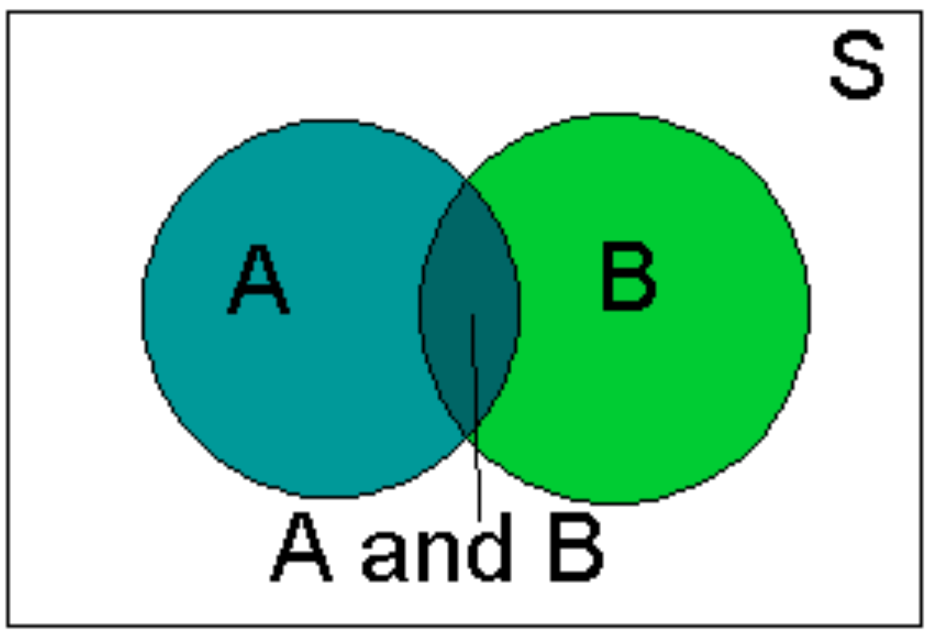
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Rules for Calculating Probabilities

- Multiplicative rule for conditional events

$$P(A \cap B) = P(A) \bullet P(B \mid A)$$



Conditional Probability

- Under some circumstances the probability of an event depends on another event.
- An unconditional probability asks what the chances are of rain tomorrow (event A).

$$P(A)$$

- A conditional probability says, “Given that rained today (event B), what are the chances of rain tomorrow? (event A)”
    - P(A|B)

Computing Conditional Probabilities

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)}$$

Independence

- Two events are said to be independent if

$$P(A \mid B) = P(A)$$

- Otherwise, the events are dependent

Bayes' Rule

- Suppose we knew P(B|A) but wanted to know P(A|B)?

$$P(B_j \mid A) = \frac{P(B_j)P(A \mid B_j)}{\sum_{i=1}^{k} P(B_i)P(A \mid B_i)}$$

Example

- Suppose you have been tested positive for a disease; what is the probability that you actually have the disease? Suppose the probability of having the disease is .01. The test is 95% accurate, and you tested positive. Do you have the disease?
- We know:
  - The probability of anyone having the disease (.01)
  - The probability of testing positive for the disease conditional on having the disease (.95)
- We want to know the probability of having the disease if you tested positive for it…

Bayes' Rule

$$P(HaveIt \mid TestPos) = \frac{P(HaveIt)\,P(TestPos \mid HaveIt)}{P(HaveIt)\,P(TestPos \mid HaveIt) + P(NoHaveIt)\,P(TestPos \mid NoHaveIt)}$$

$$P(HaveIt \mid TestPos) = \frac{.01 \cdot .95}{.01 \cdot .95 + .99 \cdot .05}$$

$$P(HaveIt \mid TestPos) = \frac{.0095}{.0095 + .0495} = \frac{.0095}{.059} \approx .161$$

What? .161? Why so low?

- Out of 100 people who take this test, we expect only 1 would have the disease.
- However, 5 people would test positive even if they didn't have the disease.
- Out of those 6 people, only 1 actually has the disease…

Political Application

- In a certain population of voters, 40% are Republican and 60% are Democrats. It is reported that 30% of Republicans and 70% of Democrats support a particular issue. A randomly selected person is found to favor the issue—what is the conditional probability that they are a Democrat?

Work it out

- We want to know P(Dem | F_issue)

$$P(Dem \mid F\_issue) = \frac{P(Dem)P(F\_issue \mid Dem)}{P(Dem)P(F\_issue \mid Dem) + P(\text{Rep})P(F\_issue \mid \text{Rep})}$$

$$P(Dem \mid F\_issue) = \frac{.6 \cdot .7}{.6 \cdot .7 + .4 \cdot .3}$$

$$P(Dem \mid F\_issue) = \frac{.42}{.42 + .12} \approx .778$$