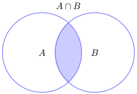
Chapter 9 (Part 1) Notes

Essential Probability Rules
STP-231

Arizona State University

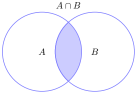


Probability Theory

- **Probability theory**: the science of uncertainty or chance

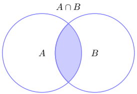
- **Randomness**: individual outcomes are uncertain, but there is nonetheless a regular distribution of outcomes with repetition

- **Probability of an outcome**: for random phenomenon, the proportion of times the outcome occurs over many repetitions



Probability and the Life Sciences

- Probability can determine the likelihood of a certain event
- In life sciences: example is is the likelihood of twins being born
- Do outside factors influence this occurence? Do twins run in mother;s line of the family? Has mother taken fertility drugs? Is the mother a hollywood actress?
- Conclusions in statistical analysis are reported in probabilities:
- What is the likelihood that a particular sample is different from a population or from another sample?
- How likely or unlikely is an experimental result



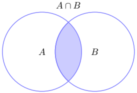
Frequentist vs Personal Probabilities

Frequentist approach

- A very large random sample can be used to approximate probabilities of random phenomenon
- Long run probability

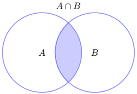
Personal

- Subjective probability based on one's own experience and judgement
- A probability nonetheless



Probability Models

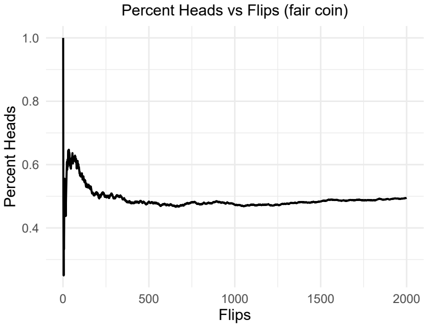
- **Sample space** ($\mathcal{S}$): Includes all possible outcomes of a random phenomenon
- **Event**: An outcome or set of outcomes of a random phenomenon, and is a subset of the sample space
- A probability model of a random phenomenon mathematically defines a sample space $\mathcal{S}$.
- How do we assign probabilities of events within $\mathcal{S}$?

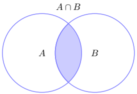


Example: Flipping a coin

Example: flipping a coin. $\mathcal{S}$ is the possible outcomes, i.e. some enumeration of H and T (heads/tails). Denote as $\mathcal{S} = \{H, T\}$. The event E in this example is how many heads we get,

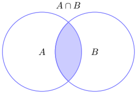
$$\Pr(E) \leftrightarrow \frac{\text{Number of times we flip head}}{\text{Number of times we flip the coin}}$$





Probability Rules

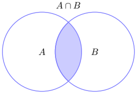
- Probability of an event, $\Pr(E)$, is a value that is always between 0 and 1, inclusive
- Probability of an impossible event is 0
- Probability of a certain event is 1 (event will always happen)
- The sum of probabilities of all events equals 1
- When the two events have no outcomes in common, they can never happen together, i.e. their joint probability is zero
- The probability that one or the other occurs is the sum of their individual probabilities minus probability both happen, $\Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A \text{ and } B)$
- The probability an event does not occur is 1 minus the probability event does occur, i.e. $\Pr(E^c) = 1 - \Pr(E)$



Discrete Probability Models

- A probability model with a sample space made up of alist of every individual outcome

- The probability of any event is the sum of the probabilities of the outcomes making up the event



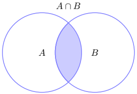
Equal Likelihood Model

- Every possible sample from the total number of possible outcomes is equally likely to be chosen
- The probability of an event E is defined as

$$\Pr(E) = \frac{\text{\# of outcomes in event}}{\text{Total outcomes in sample space}}$$

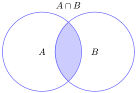
The procedure:

- Identify all possible outcomes in the sample space
- Identify the event
- Identify the outcomes that belong in the event
- Calculate ratio



Example

A test consists of two true or false problems. Suppose the answers to both questions are true. A student guesses both questions by flipping a coin. If it lands on heads, the student chooses true as the final answer and if lands on tails, the student chooses false as the final answer. What is the probability that a student gets exactly one out of two questions correct?

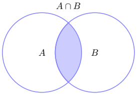


Example 2

Five students, Riley (R), Toro (T), Eddie (E), Joe (J), and Kelly (K), are all allergic to flaxseed. Two out of the five will be tested on a new genetically altered flaxseed to see if the two students are still allergic.

- What is the probability that Toro and Kelly (event A) are chosen for the study? 1/10 (see recording) The probability either or is in the study? 7/10
  $\Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A \text{ or } B)$

- What is the probability that Riley (Event B) is chosen for the study? 4/10. Write out all the combinations see Riley shows up 4 times.



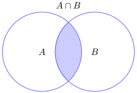
Random Sampling and Probabilities

When sampling from a population

- The probability for randomly choosing an individual with a certain characteristic is equivalent to the population proportion (relative frequency) of individuals with that charactertistic in the population

- EXAMPLE: A large population of the fruitflu Drosophila melanogaster is maintained in a lab. 30% of individuals are black, 70% are grey in the population. Suppose one fly is chosen at random from the population. Then the probability that a black fly is chosen is 0.3. More formally, define

$$E\text{: sampled fly is black, then } \Pr(E) = 0.3$$



Sampling from Populations
Continued

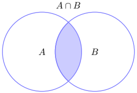
In a certain population of the freshwater sculpin, Cottus rotheus, the distribution of the number of tail vertebrae is shown in the table[1]

# of vertebrae	% of fish
20	3
21	51
22	40
23	6

Find the probability that the number of tail vertebrae in a fish randomly chosen from the population

- Equals 21
- Is less than or equal to 22
- is greater than 21
- is no more than 21

[1] 3.2.1 from Statistics for the Life Sciences by Samuels

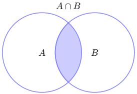


Sampling from Populations
Another Example

The effectives of a seed eating weevil on population control of a nonnative invasive species of a tree in South America called the Paraserianthes Iopantha was studied.

% Seed damage	# of trees
0-9	19
10-19	2
20-29	5
30-39	3
40-49	6
50-59	2
60-69	2

- Find probability of event that the tree has 20-29% seed damage
- At least 50% seed damage
- Less than 40% seed damage
- At least 30% but less than 59%
- Probability of all the previous 4 events occurring together



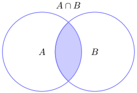
Continuous Probability Models

Density Curve

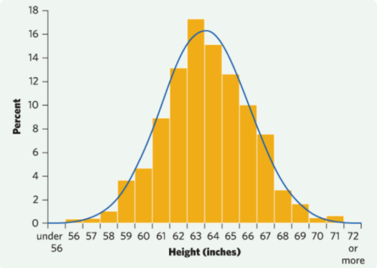
- Is always on or above the horizontal axis
- has area exactly 1 underneath
- Area under the curve and above any range of values on the horizontal axis is the proportion of all observations that fall within that range

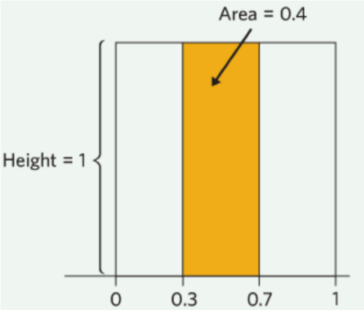
Continuous probability model:

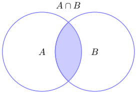
- Assigns probabilities as areas under a density curve



Pictoral Representation





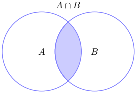


Example

Example 9.11 in the text, from previous slide.

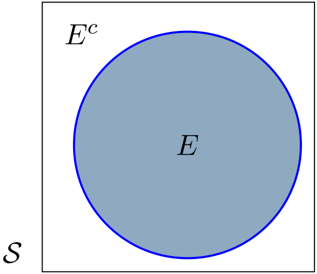
- $\mathcal{S} = \{\text{all numbers between 0 and 1}\}$
- Consider a random number generator that chooses numbers in $\mathcal{S}$. The random number generator will spread its output uniformly. The density curve is on the right.
- How do we assign probabilities to intervals?
- How do we assign probabilities individual values?

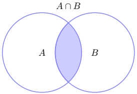
Venn Diagrams



Venn Diagrams

- Visuals to display events and relationships among events
- The sample space, $\mathcal{S}$, is the space that includes all possible outcomes in an experiment, and an event, E, is a subset of the sample space



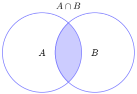


The Complement of E

- E^c : the event E does not occur
- For any event E:

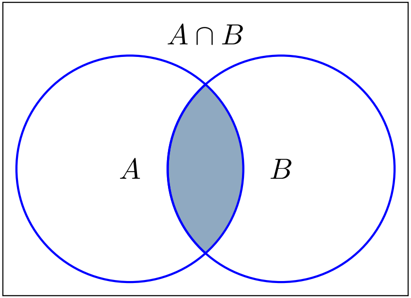
$$\Pr(E) = 1 - \Pr\left(E^c\right)$$
$$\Pr(E^c) = 1 - \Pr(E)$$

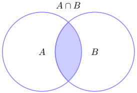
- For example, if event A is the suit hearts in a deck of cards, the complement is clubs, diamonds, and spades
- If event B is a worker making at least 40,000 dollars annually, the complement is workers making less than that



Intersection of Events

- A and B: The event both A and B occur, denoted $A \cap B$

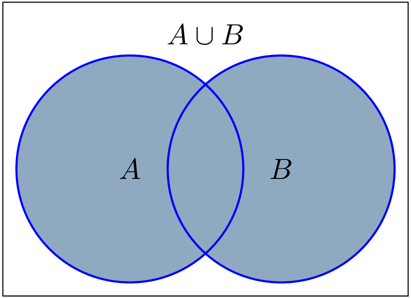


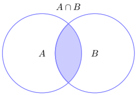


Union of Events

- A or B: The event either A or B occur, denoted $A \cup B$.

$$\Pr(A \text{ or } B) = \Pr(A) + \Pr(B) - \Pr(A \text{ and } B)$$



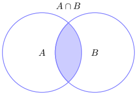


Example

When we roll a die, there are six possible outcomes. Define

- Event A: The die comes up even
- Event B: The die comes up odd
- Event C: The die comes up 6

List the outcomes in not A, $A\&C$, A or C, B & C

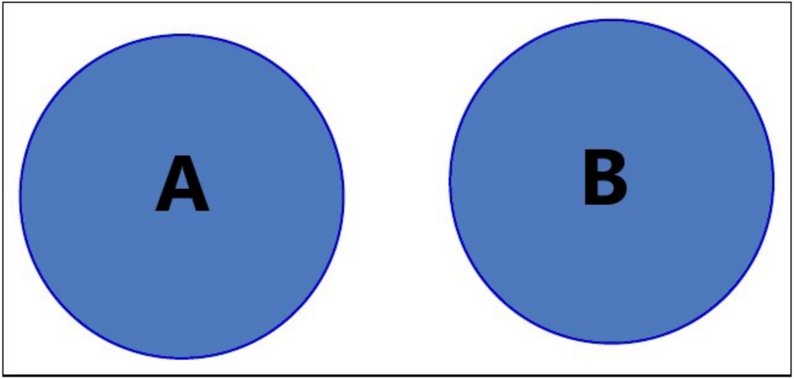


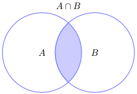
Mutually Exclusive

Mutually exclusive events: Two or more events that can not occur together. If A and B are mutually exclusive

- $\Pr(A \& B) = 0$
- $\Pr(A \text{ or } B) = \Pr(A) + \Pr(B)$
- In general, if $A, B, C, \ldots$ are all mutually exclusive:
- $\Pr(A \& B \& C \& \ldots) = 0$
- $\Pr(A \text{ or} B \text{ or } C \text{ or } \ldots) = \Pr(A) + \Pr(B) + \Pr(C) + \ldots$

Mutually Exclusive
Picture





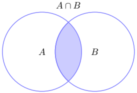
Example

Roll two dice, 36 equally likely outcomes

- A: Event that sum of the dice is 7
- B: Event that sum of the dice is 11
- C: Event that roll of the dice is odd
- D: Event that sum of the dice is 8
- E Event that roll of the dice is doubles

	⚀	⚁	⚂	⚃	⚄	⚅
⚀	2	3	4	5	6	7
⚁	3	4	5	6	7	8
⚂	4	5	6	7	8	9
⚃	5	6	7	8	9	10
⚄	6	7	8	9	10	11
⚅	7	8	9	10	11	12

Find the probabilities $\Pr(A)$, $\Pr(A \text{ or } B)$, $\Pr(C \text{ and } D)$, $\Pr(D \text{ and } E)$, $\Pr(D \text{ or } E)$



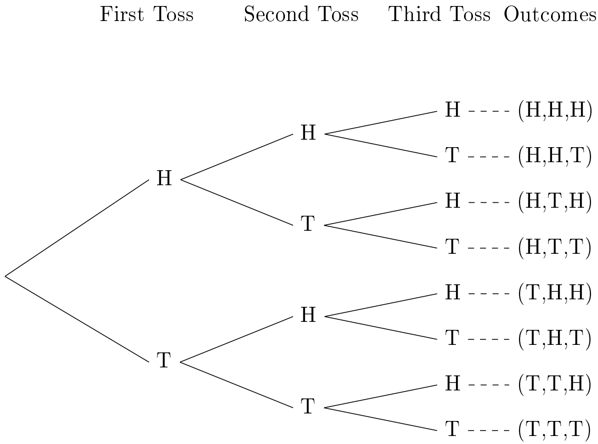
Example 2

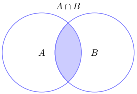
There are five associates on duty in a Staples office supply store: three women (Maggie, Linda, Vanessa) and two men (Gary, Pablo). An experiment consists of classifying the next customer's action. They will make a purchase from exactly one of the sales associates or buy nothing. What is the probability the next customer buys from Maggie or Pablo?

Action	Probability
Buy from Maggie	0.08
Buy from Linda	0.12
Buy from Vanessa	0.10
Buy from Gary	0.25
Buy from Pablo	0.15
Buy nothing	0.30

Probability Trees

Probability Tree





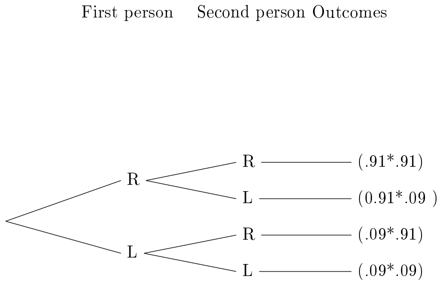
Using Probability Trees

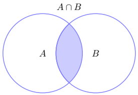
- Outcome probabilities:
- Travel ALONG branches
- When you are traveling along branches you multiply the probabilities
- Event probabilities
- Choose outcomes that satisfy event
- Sum those probabilities

Example

About 9% of people are left-handed. You randomly select two people

- What is the probability they are both right-handed?
- What is the probability that at least one right-handed person is selected?





Another Example

From 3.2.4 in Samuels Statistics for the Life Sciences

- Suppose that a disease is inherited via a sex-linked mode of inheritance so that a male offspring has a 50% chance of inheriting the disease. Further suppose that 51.3% of births are male. What is the probability that a randomly chosen child will be affected by the disease?