Sections 3.2 and 3.3

Timothy Hanson

Department of Statistics, University of South Carolina

Stat 205: Elementary Statistics for the Biological and Life Sciences

Section 3.3 Probability rules

1. Rule (1) $0 \leq \Pr\{E\} \leq 1$ for any event E.
2. Rule (2) If $E_1, E_2, \ldots, E_k$ are all possible experimental outcomes (smallest events possible), then

$$\sum_{i=1}^{k} \Pr\{E_i\} = \Pr\{E_1\} + \Pr\{E_2\} + \cdots + \Pr\{E_k\} = 1.$$

3. Rule (3) The probability that an event does not happen, E^C is $\Pr\{E^C\} = 1 - \Pr\{E\}$.

Probability of any event

- Let all experimental outcomes be listed as the smallest events $E_1, E_2, E_3, \ldots, E_k$.
- We can make new events from these, e.g. $A = \{E_2, E_4\}$.
- The probability of any event is the sum of the probabilities of the experimental outcomes in the event

$$\Pr\{A\} = \sum_{E_i \text{ in } A} \Pr\{E_i\}.$$

- Computing probabilities involves a lot of counting and summing.

Example 3.3.1 Blood type

- The smallest events possible are the individual experimental outcomes O, A, B, and AB. The proportions in the U.S. are $\Pr\{O\} = 0.44$, $\Pr\{A\} = 0.42$, $\Pr\{B\} = 0.10$, and $\Pr\{AB\} = 0.04$.
- All of these are between 0 and 1.
- $\Pr\{O\} + \Pr\{A\} + \Pr\{B\} + \Pr\{AB\} = 1$.
- The probability that a randomly selected individual *does not have* type AB is $\Pr\{AB^C\} = 1 - \Pr\{AB\} = 1 - 0.04 = 0.96$.
- The probability of either A or AB is

$$\Pr\{A, AB\} = \Pr\{A\} + \Pr\{AB\} = 0.42 + 0.04 = 0.46.$$

Union & intersection

- We can make a new event by taking the union or intersection of two events E_1 & E_2.
- The **union** of two events combines the experimental outcomes in both, written E_1 or E_2 (either can happen).
- The union happens if either E_1 or E_2 or both occur.
- The **intersection** of two events are the experimental outcomes *that are only in both at the same time*, written E_1 and E_2 (both must happen).
- The intersection of two events is the event that both occur.

Rolling a 6-sided die

- The experimental outcomes (smallest events possible) are 1, 2, 3, 4, 5, and 6. Each has probability $\frac{1}{6}$.
- Let $E_1 = \{1, 3, 5\}$ be the event that an odd number is rolled. Let $E_2 = \{1, 2, 3\}$ be the event that the roll is 3 or less.
- Union: E_1 or $E_2 = \{1, 2, 3, 5\}$.
- Intersection: E_1 and $E_2 = \{1, 3\}$.
- $\Pr\{E_1 \text{ or } E_2\} = \frac{4}{6}$.
- $\Pr\{E_1 \text{ and } E_2\} = \frac{2}{6}$.

Disjoint events

- If two events have no experimental outcomes in common, they are said to be disjoint.
- For the 6-sided die, $E_1 = \{1, 3, 5\}$ and $E_2 = \{2, 4, 6\}$ are disjoint.
- Question if E_1 and E_2 are disjoint, what is their intersection?
- Next slide is a Venn diagram for disjoint events...

Disjoint events

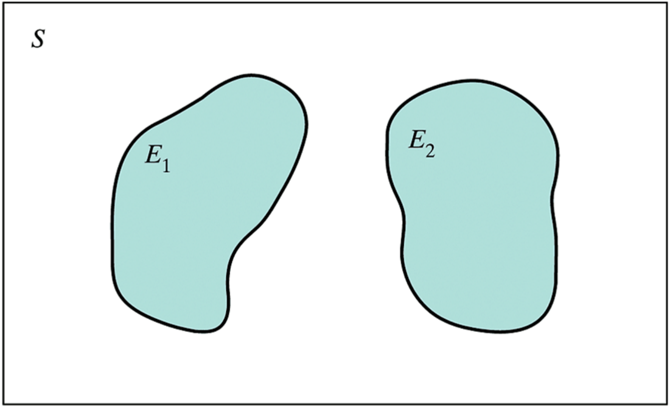


Figure 3.3.1 Venn diagram showing two disjoint events

Union and intersection

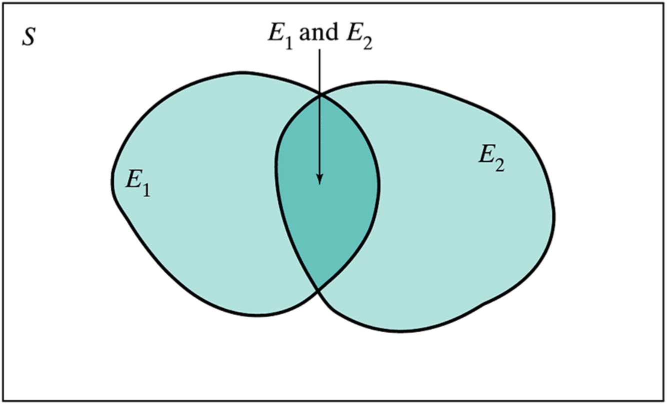


Figure 3.3.2 Venn diagram showing union (total shaded area) and intersection (middle area) of two events

Addition rules

- Rule (4) If E_1 and E_2 are disjoint then $\Pr\{E_1 \text{ or } E_2\} = \Pr\{E_1\} + \Pr\{E_2\}$.
- Rule (5) For any two events (disjoint or not) $\Pr\{E_1 \text{ or } E_2\} = \Pr\{E_1\} + \Pr\{E_2\} - \Pr\{E_1 \text{ and } E_2\}$.

Hair and eye color

Table 3.3.1 Hair color and eye color

		Hair color			
		Brown	Black	Red	Total
Eye color	Brown	400	300	20	720
	Blue	800	200	50	1,050
	Total	1,200	500	70	1,770

Relationship between hair and eye color for 1770 German men.

		Hair color			
		Brown	Black	Red	Total
Eye color	Brown	400	300	20	
	Blue	800	200	50	
	Total				1770

$$\Pr\{\text{Black hair}\} = \frac{300 + 200}{1770} = \frac{500}{1770} = 0.28.$$

		Hair color			
		Brown	Black	Red	Total
Eye color	Brown	400	300	20	
	Blue	800	200	50	
	Total				1770

$$\Pr\{\text{Blue eyes}\} = \frac{800 + 200 + 50}{1770} = \frac{1050}{1770} = 0.59.$$

		Hair color			
		Brown	Black	Red	Total
Eye color	Brown	400	300	20	
	Blue	800	200	50	
	Total				1770

$$\Pr\{\text{Black hair and blue eyes}\} = \frac{200}{1770} = 0.11.$$

		Hair color			
		Brown	Black	Red	Total
Eye color	Brown	400	300	20	
	Blue	800	200	50	
	Total				1770

$$\Pr\{\text{Black hair or blue eyes}\} = \frac{800 + 200 + 50 + 300}{1770} = \frac{1350}{1770} = 0.76.$$

Using Rule 5

Rule (5) gives

$$\begin{aligned}
\Pr\{\text{Black hair or blue eyes}\} &= \Pr\{\text{Black hair}\} + \Pr\{\text{blue eyes}\} \\
&\quad - \Pr\{\text{Black hair and blue eyes}\} \\
&= \frac{500}{1770} + \frac{1050}{1770} - \frac{200}{1770} \\
&= \frac{1350}{1770} \\
&= 0.76.
\end{aligned}$$

Conditional probability

- Sometimes we have information before we compute a probability.
- For example we might know an individual from the table has blue eyes.
- Knowing this fact *reduces the population* to only those who have blue eyes.
- This can change the probability of having black hair.
- The conditional probability of E_2 given E_1 is defined to be

$$\Pr\{E_2|E_1\} = \frac{\Pr\{E_2 \text{ and } E_1\}}{\Pr\{E_1\}}.$$

- In terms of computing conditional probabilities from a table, it's easiest just to ignore the part of the table that didn't happen!

Bayes' Theorem

- Named after the Reverend Thomas Bayes. Published after his death.
- The conditional probability of E_2 given E_1 is defined to be

$$\Pr\{E_2|E_1\} = \frac{\Pr\{E_1|E_2\}\Pr\{E_2\}}{\Pr\{E_1\}}.$$

- It is the cornerstone of the statistical methods known as 'Bayesian Statistics'.

		Hair color			
		Brown	Black	Red	Total
Eye color	Brown				
	Blue	800	200	50	1050
	Total				

$$\Pr\{\text{Black hair}|\text{blue eyes}\} = \frac{200}{1050} = 0.19.$$

		Hair color			
		Brown	Black	Red	Total
Eye color	Brown			20	
	Blue			50	
	Total			70	

$$\Pr\{\text{Blue eyes}|\text{red hair}\} = \frac{50}{70} = 0.76.$$

		Hair color			
		Brown	Black	Red	Total
Eye color	Brown	400	300	20	720
	Blue				
	Total				

$$\Pr\{\text{Not red hair}|\text{brown eyes}\} = \frac{400 + 300}{720} = \frac{700}{720} = 0.97.$$

In biological sciences

- The hair/eyes example is not that interesting (to us at least).
- More relavent examples might be
  - $\Pr\{\text{heart attack this year}|\text{cholesterol} > 190\}$
  - $\Pr\{\text{cavities}|\text{don't floss}\}$
  - $\Pr\{\text{low stress}|\text{exercise}\}$
  - $\Pr\{\text{pertussis}|\text{vaccinated}\}$

Independent events

Two events E_1 and E_2 are **independent** if knowing one has occured does not change the probability of the other,

$$\Pr\{E_1|E_2\} = \Pr\{E_1\}.$$

Recall that $\Pr\{\text{Black hair}\} = 0.28$ and $\Pr\{\text{Black hair}|\text{blue eyes}\} = 0.19$. Are "black hair" and "blue eyes" independent events?

Section 3.2 Probability trees

- The probabilities from probability trees are given by

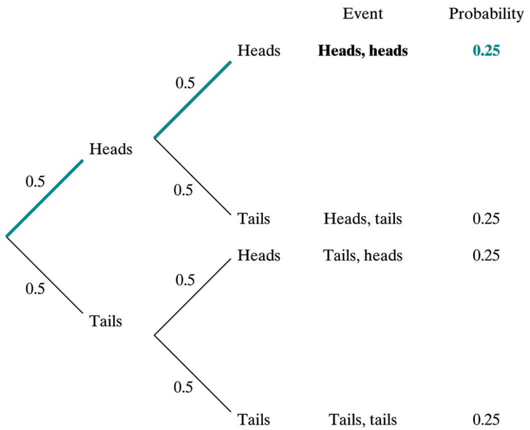
$$\Pr\{E_1 \text{ and } E_2\} = \Pr\{E_1\}\Pr\{E_2|E_1\}.$$

- This is the definition of conditional probability with both sides multiplied by $\Pr\{E_1\}$.
- Probability trees are used in the previous Section 3.2...

Probability trees

- If our event E is the result of two (or more) smaller experiments, probability trees provide a convenient way to find the probability of all possible outcomes.
- The first experiment results in either A or B.
- After the first experiment, a second experiment is carried out resulting in either C or D.
- The four possible events are AC, AD, BC, or BD and their probabilities are obtained from a probability tree by multiplying numbers going down the appropriate branch.

Example 3.2.7 Two fair coin tosses

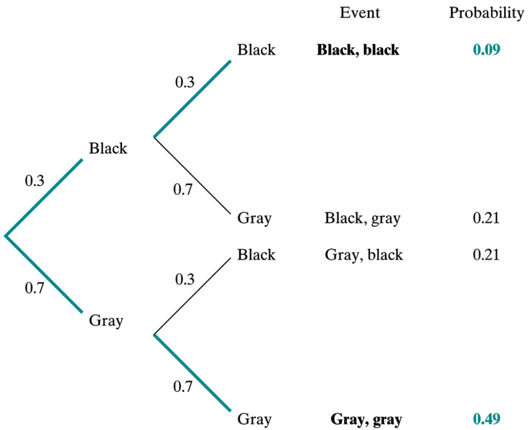


$$\Pr\{\text{Heads,Heads}\} = 0.5(0.5) = 0.25 \qquad \Pr\{\text{Heads,Tails}\} = 0.5(0.5) = 0.25$$

$$\Pr\{\text{Tails,Heads}\} = 0.5(0.5) = 0.25 \qquad \Pr\{\text{Tails,Tails}\} = 0.5(0.5) = 0.25$$

What is the probability of obtaining "heads" on both coin flips?

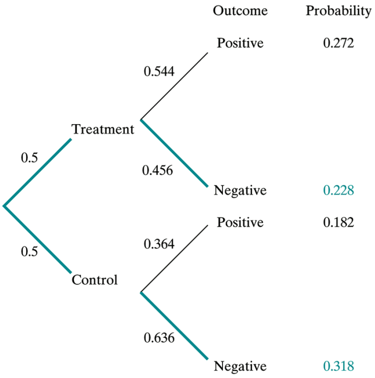
Example 3.2.8 sampling two flies



$$\Pr\{\text{Black,Black}\} = 0.3(0.3) = 0.09 \quad \Pr\{\text{Black,Gray}\} = 0.3(0.7) = 0.21$$
$$\Pr\{\text{Gray,Black}\} = 0.7(0.3) = 0.21 \quad \Pr\{\text{Gay,Gray}\} = 0.7(0.75) = 0.49$$

What is the probability of sampling the same color both times?

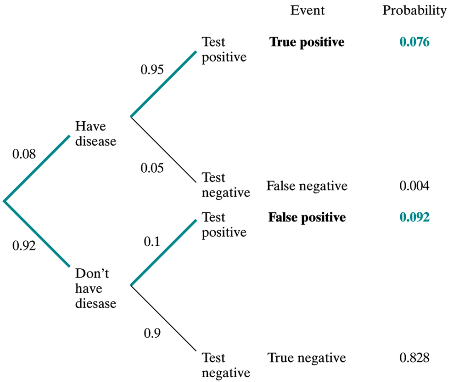
Example 3.2.9 nitric oxide example



$$\Pr\{\text{Treatment,Positive}\} = 0.5(0.544) = 0.272 \qquad \Pr\{\text{Treatment,Negative}\} = 0.5(0.456) = 0.228$$
$$\Pr\{\text{Control,Positive}\} = 0.5(0.364) = 0.182 \qquad \Pr\{\text{Control,Negative}\} = 0.5(0.636) = 0.318$$

What is the probability of a negative outcome?

Example 3.2.10 medical testing example



$$\Pr\{\text{Disease, Test positive}\} = 0.08(0.95) = 0.076 \qquad \Pr\{\text{Disease, Test negative}\} = 0.08(0.05) = 0.004$$

$$\Pr\{\text{No disease, Test positive}\} = 0.92(0.10) = 0.092 \qquad \Pr\{\text{No disease, Test negative}\} = 0.92(0.90) = 0.828$$

What is the probability of testing positive?

Example 3.2.11 A more relavent question...

- Given that my test comes up positive, what is the probability that I have the disease?
- The definition of conditional probability is

$$\begin{aligned}\Pr\{\text{disease}|\text{test positive}\} &= \frac{\Pr\{\text{disease and test positive}\}}{\Pr\{\text{test positive}\}}\\ &= \frac{0.076}{0.168}\\ &= 0.452.\end{aligned}$$