# STAT 25100 Lecture 6
# 3.3 Bayes Rule & Law of Total Probability
# 3.4 Independence

Yibi Huang
Department of Statistics
University of Chicago

# Law of Total Probability and Bayes' Rule

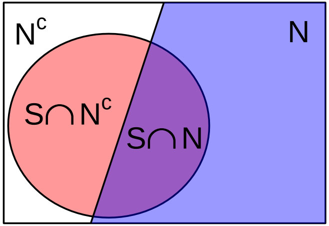
# Example – A Nervous Job Applicant

An job applicant has been invited for an interview.
Suppose The probability that

- he is nervous is $\mathrm{P}(N) = 0.7$,
- the interview is successful given he is nervous is $\mathrm{P}(S \mid N) = 0.2$,
- the interview is successful given he is not nervous is $\mathrm{P}(S \mid N^c) = 0.9$.

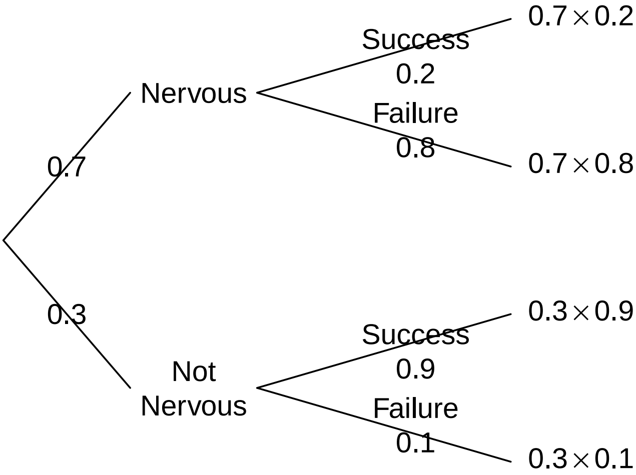
What is the probability that he succeeds in the interview?

$$\begin{aligned} \mathrm{P}(S) &= \mathrm{P}(S \cap N) + \mathrm{P}(S \cap N^c) \\ &= \mathrm{P}(N)\mathrm{P}(S \mid N) + \mathrm{P}(N^c)\mathrm{P}(S \mid N^c) \\ &= 0.7 \times 0.2 + 0.3 \times 0.9 = 0.41. \end{aligned}$$



# Tree Diagram for the Nervous Job Applicant Example

Another look of the computation:



# Nervous Job Applicant Example Continued

Conversely, given the interview is successful, what is the probability that the job applicant is nervous during the interview?

$$
\begin{aligned}
\mathrm{P}(N \mid S) &= \frac{\mathrm{P}(N \cap S)}{\mathrm{P}(S)} \\
&= \frac{\mathrm{P}(N \cap S)}{0.41} \qquad \mathrm{P}(S) = 0.41 \text{ from previous page} \\
&= \frac{\mathrm{P}(N)\mathrm{P}(S \mid N)}{0.41} \qquad \text{since } \mathrm{P}(N \cap S) = \mathrm{P}(N)\mathrm{P}(S \mid N) \\
&= \frac{0.7 \times 0.2}{0.41} = \frac{14}{41} \approx 0.34.
\end{aligned}
$$

# Bayes' Rule = Bayes' Theorem = Bayes' Formula

The problem in the previous slide is an example of **Bayes' Rule**.

Knowing $\mathrm{P}(B \mid A), \mathrm{P}(B \mid A^c)$, and $\mathrm{P}(A)$, is there a way to know $\mathrm{P}(A \mid B)$?

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

# Bayes' Rule = Bayes' Theorem = Bayes' Formula

The problem in the previous slide is an example of **Bayes' Rule**.

Knowing $\mathrm{P}(B \mid A), \mathrm{P}(B \mid A^c)$, and $\mathrm{P}(A)$, is there a way to know $\mathrm{P}(A \mid B)$?

$$\begin{aligned} \mathrm{P}(A \mid B) &= \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \\ &= \frac{\mathrm{P}(A)\mathrm{P}(B \mid A)}{\mathrm{P}(B)} \quad \text{since } \mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B \mid A) \end{aligned}$$

# Bayes' Rule = Bayes' Theorem = Bayes' Formula

The problem in the previous slide is an example of **Bayes' Rule**.

Knowing $\mathrm{P}(B \mid A), \mathrm{P}(B \mid A^c)$, and $\mathrm{P}(A)$, is there a way to know $\mathrm{P}(A \mid B)$?

$$\begin{aligned}
\mathrm{P}(A \mid B) &= \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \\
&= \frac{\mathrm{P}(A)\mathrm{P}(B \mid A)}{\mathrm{P}(B)} \quad \text{since } \mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B \mid A) \\
&= \frac{\mathrm{P}(A)\mathrm{P}(B \mid A)}{\mathrm{P}(B \cap A) + \mathrm{P}(B \cap A^c)} \quad \text{since } B = (B \cap A) \cup (B \cap A^c)
\end{aligned}$$

# Bayes' Rule = Bayes' Theorem = Bayes' Formula

The problem in the previous slide is an example of **Bayes' Rule**.

Knowing $\mathrm{P}(B \mid A), \mathrm{P}(B \mid A^c)$, and $\mathrm{P}(A)$, is there a way to know $\mathrm{P}(A \mid B)$?

$$\begin{aligned}
\mathrm{P}(A \mid B) &= \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \\
&= \frac{\mathrm{P}(A)\mathrm{P}(B \mid A)}{\mathrm{P}(B)} \quad \text{since } \mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B \mid A) \\
&= \frac{\mathrm{P}(A)\mathrm{P}(B \mid A)}{\mathrm{P}(B \cap A) + \mathrm{P}(B \cap A^c)} \quad \text{since } B = (B \cap A) \cup (B \cap A^c) \\
&= \frac{\mathrm{P}(A)\mathrm{P}(B \mid A)}{\mathrm{P}(A)\mathrm{P}(B \mid A) + \mathrm{P}(A^c)\mathrm{P}(B \mid A^c)} \quad \text{by Multiplication Rule}
\end{aligned}$$

# Medical Testing

A common application of Bayes' rule is in medical testing

- Let $D$ denote the event that an individual has the disease that we are testing for

# Medical Testing

A common application of Bayes' rule is in medical testing

- Let $D$ denote the event that an individual has the disease that we are testing for
- Let $T+$ denote the event that the test result is *positive*,
  and $T-$ denote the event that the test result is *negative*

# Medical Testing

A common application of Bayes' rule is in medical testing

- Let $D$ denote the event that an individual has the disease that we are testing for
- Let $T+$ denote the event that the test result is *positive*,
  and $T-$ denote the event that the test result is *negative*
- $\mathrm{P}(T+ \mid D)$ is called the *sensitivity* of the test

# Medical Testing

A common application of Bayes' rule is in medical testing

- Let $D$ denote the event that an individual has the disease that we are testing for
- Let $T+$ denote the event that the test result is *positive*,
  and $T-$ denote the event that the test result is *negative*
- $\mathrm{P}(T+ \mid D)$ is called the *sensitivity* of the test
- $\mathrm{P}(T- \mid D^c)$ is called the *specificity* of the test

# Medical Testing

A common application of Bayes' rule is in medical testing

- Let $D$ denote the event that an individual has the disease that we are testing for
- Let $T+$ denote the event that the test result is *positive*,
  and $T-$ denote the event that the test result is *negative*
- $\mathrm{P}(T+ \mid D)$ is called the *sensitivity* of the test
- $\mathrm{P}(T- \mid D^c)$ is called the *specificity* of the test
- Ideally, we hope $\mathrm{P}(T+ \mid D)$ and $\mathrm{P}(T- \mid D^c)$ both equal 1.
  However, medical tests are not perfect.
  They may give **false positives** and **false negatives**.

# Enzyme-Linked Immunosorbent Assay (ELISA) Test for HIV

- $\mathrm{P}(T+ \mid D) = 0.98$ (sensitivity - positive for infected)
- $\mathrm{P}(T- \mid D^c) = 0.995$ (specificity - negative for not infected)
- $\mathrm{P}(D) = 1/300$ (prevalence of HIV in USA)

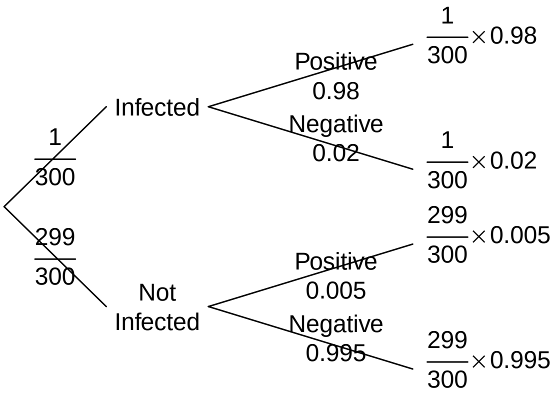
Suppose we apply ELISA test to the US population to screen for HIV.

a. What percentage of people will be tested positive?
b. What percentage of people that were tested positive actually are actually infected?

$$\mathrm{P}(T+) = \mathrm{P}(D)\mathrm{P}(T+ \mid D) + \mathrm{P}(D^c)\mathrm{P}(T+ \mid D^c)$$
$$= \frac{1}{300} \times 0.98 + \frac{299}{300} \times 0.005 = 0.00825$$
$$\mathrm{P}(D \mid T+) = \frac{\mathrm{P}(D)\mathrm{P}(T+ \mid D)}{\mathrm{P}(T+)} = \frac{1/300 \times 0.98}{0.00825} = 39.6\%$$

**This test is not confirmatory**. Need to confirm by a second test.

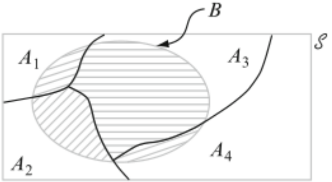
# Tree Diagram for the HIV Test




$$\mathrm{P}(D \mid T+) = \frac{(1/300) \times 0.98}{(1/300) \times 0.98 + (299/300) \times 0.005}$$

# Bayes' Rule for 3 or More Cases

- The 2 examples above both split the sample space into 2 parts $A$ or $A^c$ (nervous or not nervous, infected or not infected)
- In many cases, we need to calculate $\mathrm{P}(B)$ by splitting it into several parts, using the Law of Total Probability:




Suppose $A_1, A_2, \ldots, A_k$ are disjoint and $A_1 \cup A_2 \cup \cdots \cup A_k = S$ and $A_i \cap A_j = \emptyset$ for all $i \neq j$, then

$$\begin{aligned}\mathrm{P}(B) &= \mathrm{P}(B \cap A_1) + \mathrm{P}(B \cap A_2) + \cdots + \mathrm{P}(B \cap A_k) \\ &= \mathrm{P}(A_1)\mathrm{P}(B \mid A_1) + \mathrm{P}(A_2)\mathrm{P}(B \mid A_2) + \cdots + \mathrm{P}(A_k)\mathrm{P}(B \mid A_k).\end{aligned}$$

Using the Law of Total Probability, Bayes Rule becomes

$$\mathrm{P}(A_i \mid B) = \frac{\mathrm{P}(A_i)\mathrm{P}(B \mid A_i)}{\mathrm{P}(A_1)\mathrm{P}(B \mid A_1) + \mathrm{P}(A_2)\mathrm{P}(B \mid A_2) + \cdots + \mathrm{P}(A_k)\mathrm{P}(B \mid A_k)}$$

# Example (Bayes' Rule for 3 Cases)

At a gas station,

- 40% of the customers use regular gas ($A_1$),
- 35% use mid-grade gas ($A_2$), and
- 25% use premium gas ($A_3$).

Moreover,

- of those customers using regular gas, only 30% fill their tanks;
- of those using mid-grade, 60% fill their tanks;
- of those using premium, 50% fill their tanks.

Let $B = \{$the next customer fills the tank.$\}$

# Example (Bayes' Rule for 3 Cases)

At a gas station,

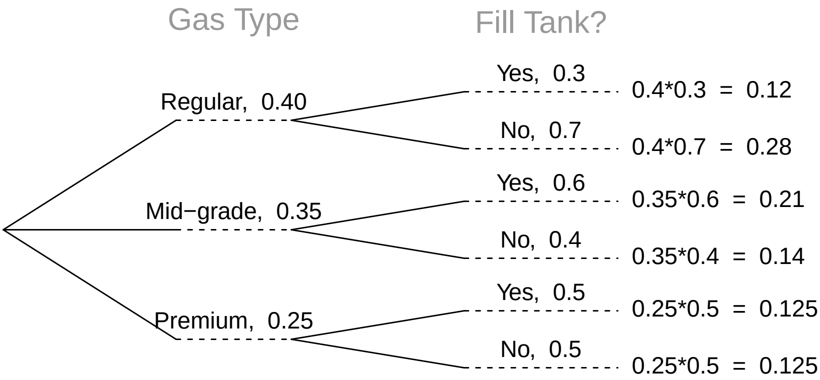
- 40% of the customers use regular gas ($A_1$),
- 35% use mid-grade gas ($A_2$), and
- 25% use premium gas ($A_3$).

Moreover,

- of those customers using regular gas, only 30% fill their tanks; $\mathrm{P}(B \mid A_1) = 0.3$
- of those using mid-grade, 60% fill their tanks; $\mathrm{P}(B \mid A_2) = 0.6$
- of those using premium, 50% fill their tanks. $\mathrm{P}(B \mid A_3) = 0.5$

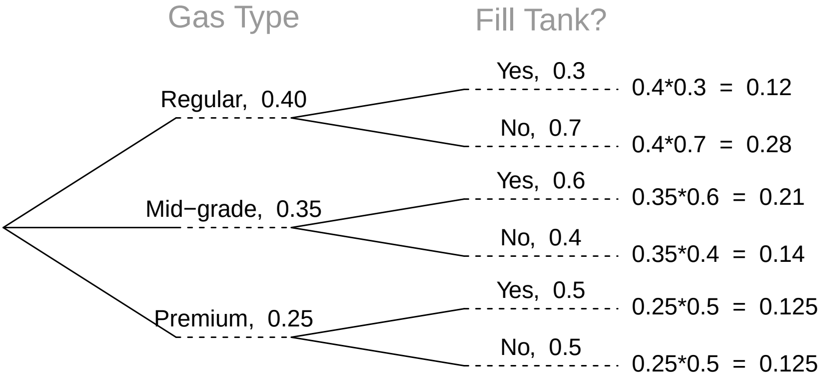
Let $B = \{$the next customer fills the tank.$\}$

# Gas Station Example — Tree Diagram




**Q1**: What is the probability that the next customer requests premium gas and fills the tank?

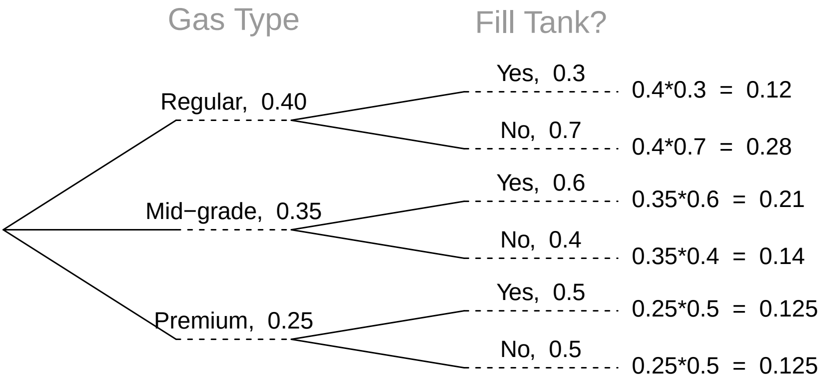
# Gas Station Example — Tree Diagram




**Q1**: What is the probability that the next customer requests premium gas and fills the tank?

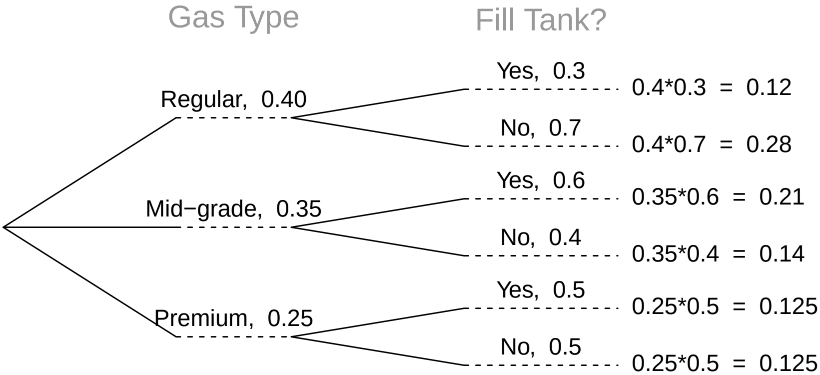
$$\mathrm{P}(A_3 \cap B) = \mathrm{P}(A_3)\mathrm{P}(B \mid A_3) = 0.25 \times 0.5 = 0.125.$$

# Gas Station Example — Tree Diagram




**Q2**: What is the probability that the next customer fills the tank.

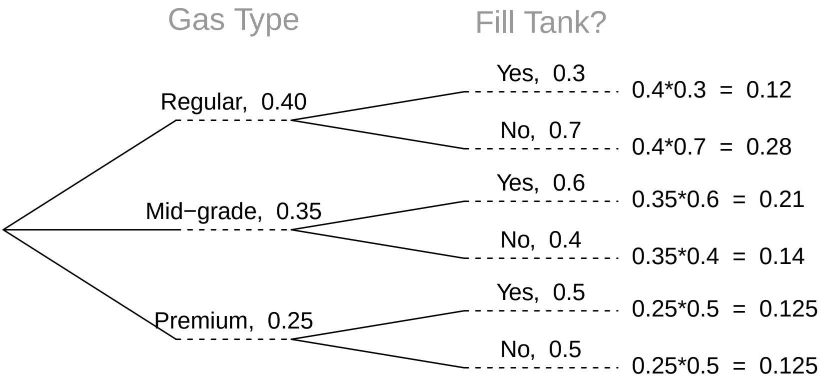
# Gas Station Example — Tree Diagram




**Q2**: What is the probability that the next customer fills the tank.

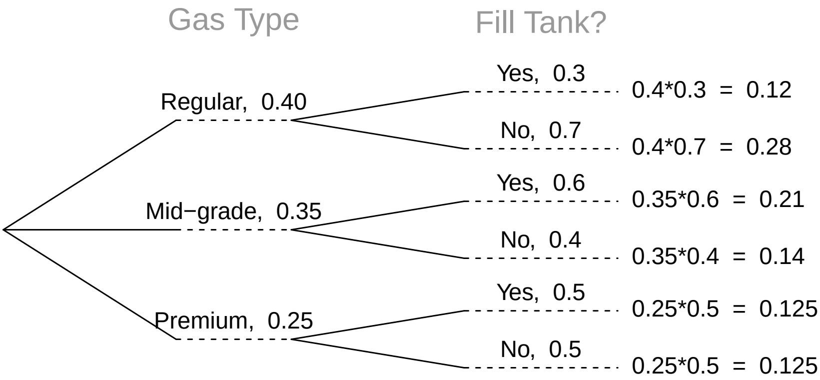
$$\begin{aligned} \mathrm{P}(B) &= \mathrm{P}(A_1)\mathrm{P}(B \mid A_1) + \mathrm{P}(A_2)\mathrm{P}(B \mid A_2) + \mathrm{P}(A_3)\mathrm{P}(B \mid A_3) \\ &= 0.4 \times 0.3 + 0.35 \times 0.6 + 0.25 \times 0.5 = 0.455 \end{aligned}$$

# Gas Station Example — Tree Diagram




**Q3**: If the next customer fills the tank, what is the probability that premium gas is requested?

# Gas Station Example — Tree Diagram




**Q3**: If the next customer fills the tank, what is the probability that premium gas is requested?

$$\mathrm{P}(A_3 \mid B) = \frac{\mathrm{P}(A_3 \cap B)}{\mathrm{P}(B)} = \frac{0.125}{0.455} \approx 0.275.$$

# Example — Law of Total Probability

If one rolls a pair of 6-face fair dice continuously until the sum of the two dice is a 5 or a 7, what is the probability that a sum of 5 appears before a sum of 7?

Sol: Let $A = \{\text{a sum of 5 appears before a sum of 7}\}$ and

$$S_5 = \{\text{first roll} = 5\}, \quad S_7 = \{\text{first roll} = 7\}, \quad S_o = \{\text{first roll is not 5 or 7}\}.$$

Note $S_5$, $S_7$ and $S_n$ are disjoint and $S_5 \cup S_7 \cup S_o = S$ is the sample space. Moreover,

$$\mathrm{P}(S_5) = \frac{4}{36}, \quad \mathrm{P}(S_7) = \frac{6}{36}, \quad \mathrm{P}(S_o) = 1 - \mathrm{P}(S_5) - \mathrm{P}(S_7) = \frac{26}{36}.$$

By the Law of Total Probability,

$$\begin{aligned}\mathrm{P}(A) &= \mathrm{P}(S_5)\mathrm{P}(A \mid S_5) + \mathrm{P}(S_7)\mathrm{P}(A \mid S_7) + \mathrm{P}(S_o)\mathrm{P}(A \mid S_o) \\ &= \end{aligned}$$

# Independence

# Independence

Two events $A$ and $B$ are said to be independent if any of the following is true

- $\mathrm{P}(A \mid B) = \mathrm{P}(A)$ ................ $B$ happens doesn't affect how likely $A$ happens
- $\mathrm{P}(A \mid B) = \mathrm{P}(A \mid B^c)$ ..................................................... How likely $A$ happens is not affected by $B$ happens or not
- $\mathrm{P}(B \mid A) = \mathrm{P}(B)$ ................ $A$ happens doesn't affect how likely $B$ happens
- $\mathrm{P}(A \cap B) = \mathrm{P}(A) \times \mathrm{P}(B)$

If any of the identities above is true, then all remaining identities will also be true.

# Proof of $\mathrm{P}(A \mid B) = \mathrm{P}(A)$ implies $\mathrm{P}(B \mid A) = \mathrm{P}(B)$

$$
\begin{aligned}
\mathrm{P}(B \mid A) &= \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)} && \text{definition of conditional prob.} \\
&= \frac{\mathrm{P}(B)\mathrm{P}(A \mid B)}{\mathrm{P}(A)} && \text{Multiplication Rule} \\
&= \frac{\mathrm{P}(B)\mathrm{P}(A)}{\mathrm{P}(A)} && \text{since } \mathrm{P}(A \mid B) = \mathrm{P}(A) \\
&= \mathrm{P}(B)
\end{aligned}
$$

Thus, $\mathrm{P}(A \mid B) = \mathrm{P}(A)$ implies $\mathrm{P}(B \mid A) = \mathrm{P}(B)$.

# Proof of $\mathrm{P}(B \mid A) = \mathrm{P}(B)$ implies $\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B)$

$$
\begin{aligned}
\mathrm{P}(A \cap B) &= \mathrm{P}(A)\mathrm{P}(B \mid A) && \text{(by Multiplication Rule)} \\
&= \mathrm{P}(A)\mathrm{P}(B) && \text{(since } \mathrm{P}(B \mid A) = \mathrm{P}(B)\text{)}
\end{aligned}
$$

# Independent Events vs Disjoint Events

- If $A$ and $B$ are independent, $\mathrm{P}(A \cap B) = \mathrm{P}(A) \times \mathrm{P}(B)$.
- If $A$ and $B$ are disjoint: $A \cap B = \emptyset \Rightarrow \mathrm{P}(A \cap B) = 0$.
- If $\mathrm{P}(A) > 0$ and $\mathrm{P}(B) > 0$,
  - Independent events **cannot** be disjoint.
  - Disjoint events **cannot** be independent.
- Conceptually, $A$ and $B$ are disjoint means that one happens prevents the other from happening, so one's occurrence definitely affects the other's.

# Examples (Independent or not)

Rolling a black die and the white die, both 6-face and fair. Define

$$B_6 = \{\text{black die shows 6}\} \quad \rightarrow \quad \mathrm{P}(B_6) = 1/6$$
$$W_6 = \{\text{white die shows 6}\} \quad \rightarrow \quad \mathrm{P}(W_6) = 1/6$$
$$T_6 = \{\text{total is 6}\} \quad \rightarrow \quad \mathrm{P}(T_6) = 5/36$$
$$T_7 = \{\text{total is 7}\} \quad \rightarrow \quad \mathrm{P}(T_7) = 6/36 = 1/6$$

a. Are $B_6$ and $W_6$ independent?
b. Are $B_6$ and $T_6$ independent?

c. Are $B_6$ and $T_7$ independent?

# Examples (Independent or not)

Rolling a black die and the white die, both 6-face and fair. Define

$$B_6 = \{\text{black die shows 6}\} \quad \to \quad \mathrm{P}(B_6) = 1/6$$
$$W_6 = \{\text{white die shows 6}\} \quad \to \quad \mathrm{P}(W_6) = 1/6$$
$$T_6 = \{\text{total is 6}\} \quad \to \quad \mathrm{P}(T_6) = 5/36$$
$$T_7 = \{\text{total is 7}\} \quad \to \quad \mathrm{P}(T_7) = 6/36 = 1/6$$

a. Are $B_6$ and $W_6$ independent? Yes, as they involve difference dice.
b. Are $B_6$ and $T_6$ independent?

c. Are $B_6$ and $T_7$ independent?

# Examples (Independent or not)

Rolling a black die and the white die, both 6-face and fair. Define

$$
\begin{aligned}
B_6 &= \{\text{black die shows 6}\} \quad \rightarrow \quad \mathrm{P}(B_6) = 1/6 \\
W_6 &= \{\text{white die shows 6}\} \quad \rightarrow \quad \mathrm{P}(W_6) = 1/6 \\
T_6 &= \{\text{total is 6}\} \quad \rightarrow \quad \mathrm{P}(T_6) = 5/36 \\
T_7 &= \{\text{total is 7}\} \quad \rightarrow \quad \mathrm{P}(T_7) = 6/36 = 1/6
\end{aligned}
$$

a. Are $B_6$ and $W_6$ independent? Yes, as they involve difference dice.

b. Are $B_6$ and $T_6$ independent?
   No. As $B_6 \cap T_6 = \emptyset$, $\mathrm{P}(B_6 \cap T_6) = 0 \neq \mathrm{P}(B_6)\mathrm{P}(T_6)$

c. Are $B_6$ and $T_7$ independent?

# Examples (Independent or not)

Rolling a black die and the white die, both 6-face and fair. Define

$$B_6 = \{\text{black die shows 6}\} \quad \rightarrow \quad \mathrm{P}(B_6) = 1/6$$
$$W_6 = \{\text{white die shows 6}\} \quad \rightarrow \quad \mathrm{P}(W_6) = 1/6$$
$$T_6 = \{\text{total is 6}\} \quad \rightarrow \quad \mathrm{P}(T_6) = 5/36$$
$$T_7 = \{\text{total is 7}\} \quad \rightarrow \quad \mathrm{P}(T_7) = 6/36 = 1/6$$

a. Are $B_6$ and $W_6$ independent? Yes, as they involve difference dice.

b. Are $B_6$ and $T_6$ independent?
No. As $B_6 \cap T_6 = \emptyset$, $\mathrm{P}(B_6 \cap T_6) = 0 \neq \mathrm{P}(B_6)\mathrm{P}(T_6)$

c. Are $B_6$ and $T_7$ independent?
Yes, since

$$\mathrm{P}(B_6 \cap T_7) = \mathrm{P}(\text{black die} = 6, \text{white die} = 1) = \frac{1}{36} = \mathrm{P}(B_6)\mathrm{P}(T_7) = \frac{1}{6} \cdot \frac{1}{6}.$$

# Independence of 3 Events

Three events $E$, $F$, and $G$ are said to be *independent* if ALL of the following are true

$$\begin{aligned} \mathrm{P}(EFG) &= \mathrm{P}(E)\mathrm{P}(F)\mathrm{P}(G) \\ \mathrm{P}(EF) &= \mathrm{P}(E)\mathrm{P}(F) \\ \mathrm{P}(EG) &= \mathrm{P}(E)\mathrm{P}(G) \\ \mathrm{P}(FG) &= \mathrm{P}(F)\mathrm{P}(G) \end{aligned}$$

# Pairwise Independence $\neq$ Independence

Rolling a black die and the white die, both 6-face and fair. Define

$$
\begin{aligned}
B_6 &= \{\text{black die shows 6}\} \\
W_6 &= \{\text{white die shows 6}\} \\
T_7 &= \{\text{total is 7}\}
\end{aligned}
$$

Two pages ago, we just showed

- the independence of $B_6$ and $W_6$
- the independence of $B_6$ and $T_7$
- the independence of $W_6$ and $T_7$ can be shown similarly as $B_6$ and $T_7$

However, $B_6, W_6$ and $T_7$ are NOT independence since $B_6 \cap W_6 \cap T_7 = \emptyset$,

$$\mathrm{P}(B_6 \cap W_6 \cap T_7) = 0 \neq \mathrm{P}(B_6)\mathrm{P}(W_6)\mathrm{P}(T_7) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6}.$$

# Properties of Independence

If $E$, $F$, and $G$ are independent, then $E$ will be independent of any event formed by $F$ and $G$, like

$$F \cup G, \quad F^c \cap G, \quad F \cup G^c, \quad \text{etc.}$$

For example, let's prove the independence of $E$ and $F \cup G$.

$$\begin{aligned}
\mathrm{P}[E(F \cup G)] &= \mathrm{P}(EF \cup EG) && \text{(Distributive law)} \\
&= \mathrm{P}(EF) + \mathrm{P}(EG) - \mathrm{P}(EFG) && \text{(Inclusion-Exclusion formula)} \\
&= \mathrm{P}(E)\mathrm{P}(F) + \mathrm{P}(E)\mathrm{P}(G) - \mathrm{P}(E)\mathrm{P}(FG) && \text{(Definition of Independence)} \\
&= \mathrm{P}(E)[\mathrm{P}(F) + \mathrm{P}(G) - \mathrm{P}(FG)] \\
&= \mathrm{P}(E)\mathrm{P}(F \cup G) && \text{(Inclusion-Exclusion formula)}
\end{aligned}$$

# Independence of $n$ Events

The events $E_1, E_2, \ldots, E_n$ are said to be *independent* if for every subset $E_{i_1}, E_{i_2}, \ldots, E_{i_r}$, for $1 \le i_1 < i_2 < \ldots < i_r \le n$ of these events, we have

$$\mathrm{P}(E_{i_1} E_{i_2} \ldots E_{i_r}) = \mathrm{P}(E_{i_1})\mathrm{P}(E_{i_2}) \cdots \mathrm{P}(E_{i_r}).$$

Finally, we define an **infinite set** of events to be *independent* if every finite subset of those events is independent.

# Example — 5 or 7 First (Revisit)

If one rolls a pair of 6-face fair dice continuously until the sum of the two dice is a 5 or a 7, what is the probability that a sum of 5 appears before a sum of 7?

Sol. On p.13, we solved using the Law of Total Prob. Here is another approach. Let

$$E_n = \{\text{no 5 or 7 appears on the first } n-1 \text{ rolls, } n\text{th roll is 5}\}$$
$$A_k = \{k\text{th roll is neither 5 nor 7}\}$$
$$B_n = \{n\text{th roll is 5}\}$$

- $\{5 \text{ appears before } 7\} = \cup_{n=1}^{\infty} E_n$. Are $E_n$'s disjoint?
- $E_n = A_1 A_2 \ldots A_{n-1} B_n$. Are $A_1, A_2, \ldots, A_{n-1}, B_n$ independent?

$$\mathrm{P}(A_i) = \qquad\qquad , \mathrm{P}(B_j) =$$

$$\mathrm{P}(E_n) = \mathrm{P}(A_1)\mathrm{P}(A_2)\ldots\mathrm{P}(A_{n-1})\mathrm{P}(B_n) = \left(\frac{26}{36}\right)^{n-1} \frac{4}{36}$$

$$\mathrm{P}(5 \text{ before } 7) = \mathrm{P}(\cup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mathrm{P}(E_n) = \sum_{n=1}^{\infty} \left(\frac{26}{36}\right)^{n-1} \frac{4}{36} = \frac{4}{36} \frac{1}{1-\frac{26}{36}} = \frac{2}{5}.$$