# Lecture 2

# Conditional Probability, Total Probability, Bayes' Theorem, and Independence

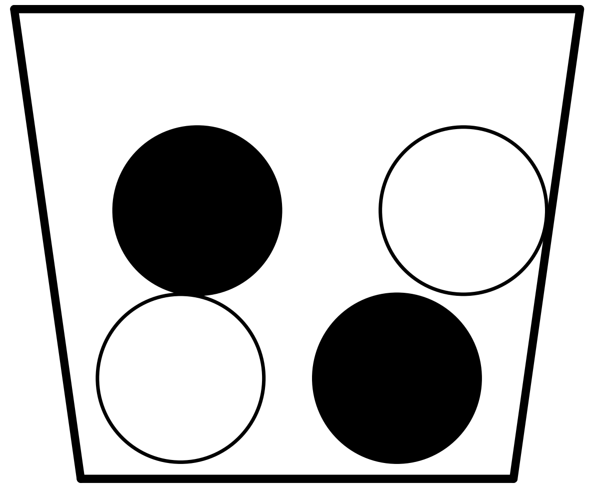
## Conditional Probability

Knowing that an event has occurred sometimes forces us to reassess the probability of another event; the new probability is the conditional probability.

Conditional probabilities correspond to updating one's beliefs when new information becomes available.

## Conditional Probability

An urn contains two white and two black marbles.



Events:

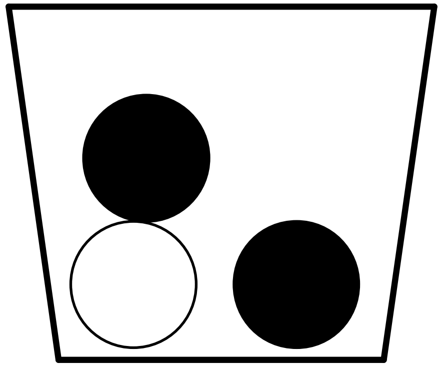
$A$ – 1st person removes a white marble;

$A^C$ – 1st person removes a black marble;

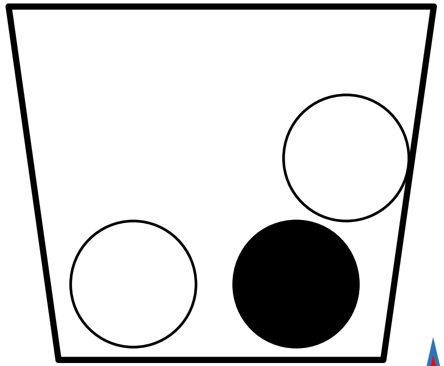
$B$ – 2nd person removes a white marble.

## Conditional Probability

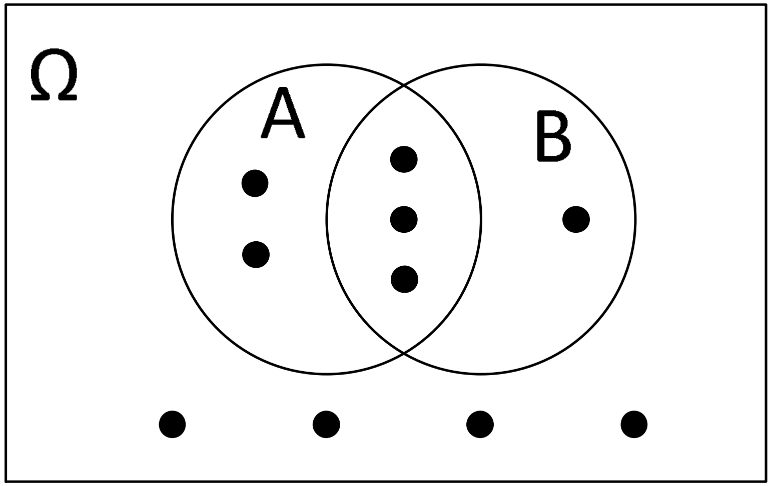
$$\Pr(B \text{ } given \text{ } A \text{ } has \text{ } occured) = \frac{1}{3}$$



$$\Pr(B \text{ } given \text{ } A^C \text{ } has \text{ } occured) = \frac{2}{3}$$



Assume that we pick a point at random from those shown in figure.

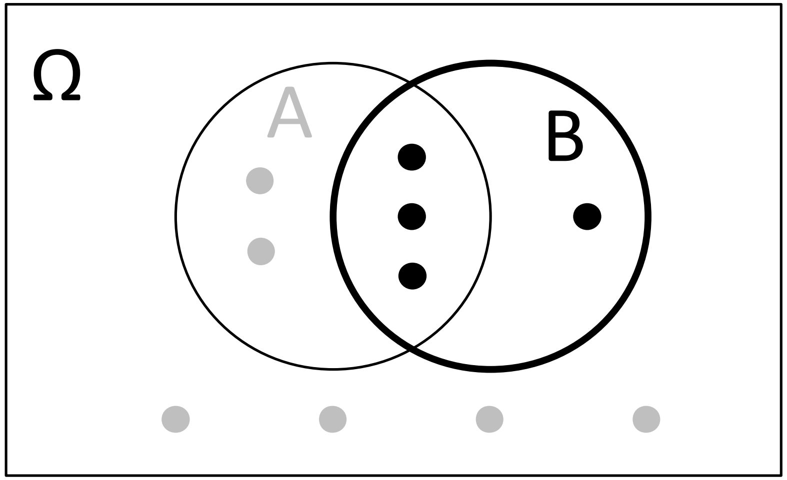



If $\Pr(A)$, $\Pr(B)$, and $\Pr(A \cap B)$ denote the probabilities of picking the point from $A$ , $B$, and $A \cap B$, respectively, then

$$\Pr(A) = \frac{5}{10} \qquad \Pr(B) = \frac{4}{10} \qquad \Pr(A \cap B) = \frac{3}{10}$$

If we restrict our attention to only those trials in which $B$ has occurred, then, obviously,

$$\Pr(A\ given\ B\ has\ occured) = \frac{3}{4}$$




We can deduce that

$$\Pr(A\ given\ B\ has\ occured) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

## Conditional Probability

Let $A$ and $B$ be arbitrary events in a given probability space $\Omega$, with $\Pr(B) \neq 0$.

Then we define the conditional probability of A, given B, as

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

hypothesis
premise

Notice that, actually, every probability may be regarded as a conditional probability, with the condition $\Omega$, since

$$\Pr(A|\Omega) = \frac{\Pr(A \cap \Omega)}{\Pr(\Omega)} = \frac{\Pr(A)}{\Pr(\Omega)} = \Pr(A)$$

Conversely, every conditional probability $\Pr(A|B)$ may be regarded as an unconditional probability in a new, reduced sample space $\Omega' = B$.

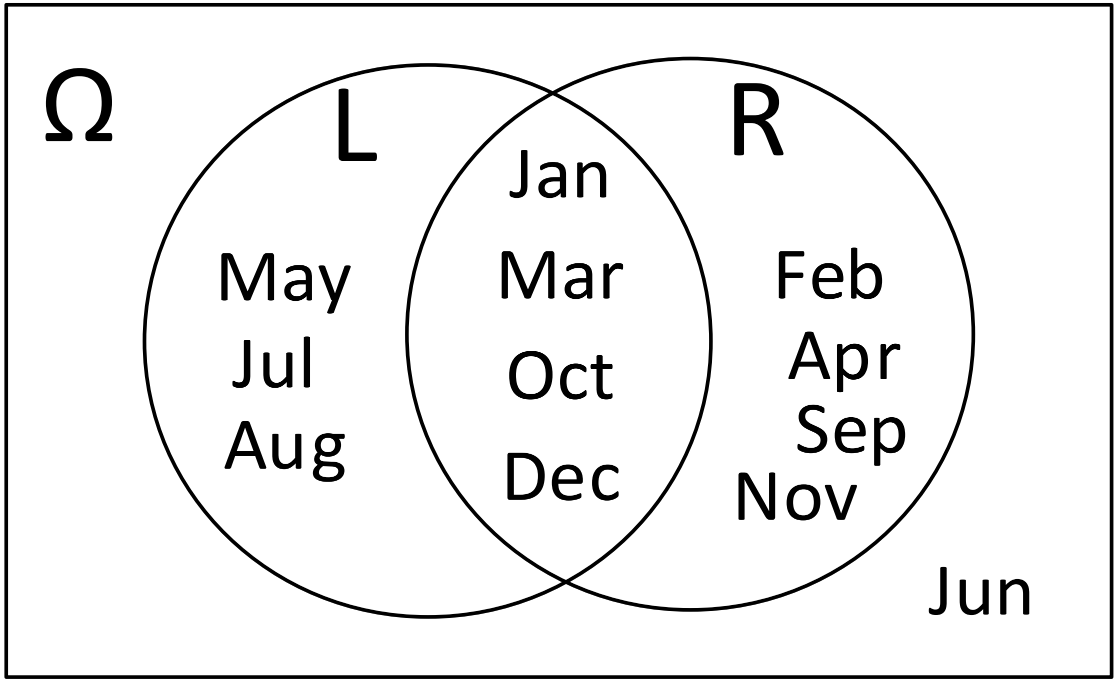
Assume we ask the next person we meet on the street in which month his/her birthday falls.

What would be the probability that the answer is a “long” month (31 day) with a letter “r” in its name?

Let’s define events $L = \{Jan, Mar, May, Jul, Aug, Oct, Dec\}$ and $R = \{Jan, Feb, Mar, Apr, Sep, Oct, Nov, Dec\}$, and draw the Venn diagram.

The answer for the problem, obviously, is the probability of event $L \cap R$:

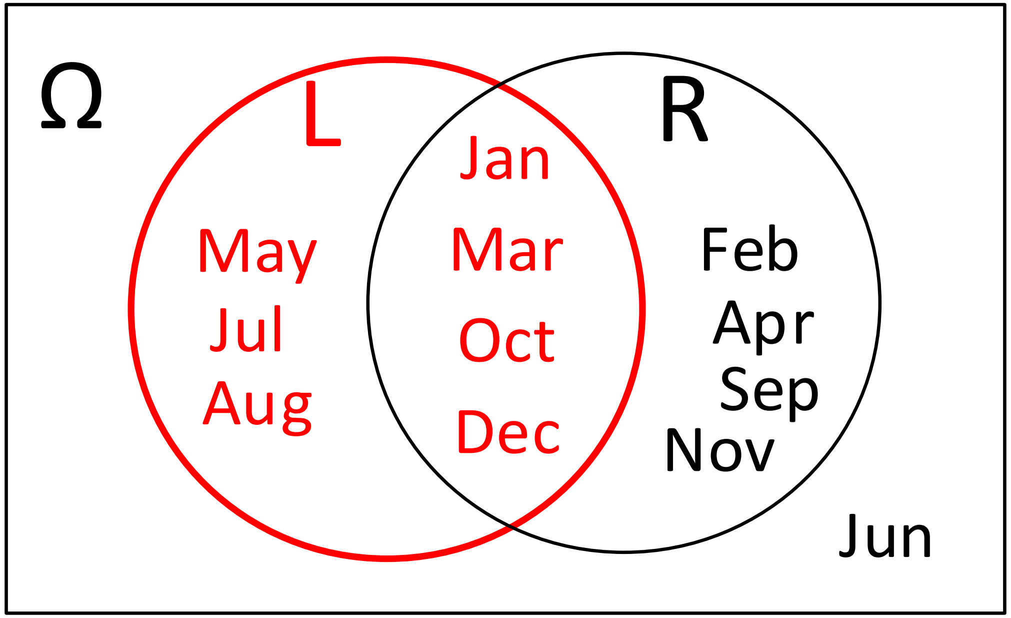
$$\Pr(L \cap R) = \frac{4}{12} = \frac{1}{3}$$

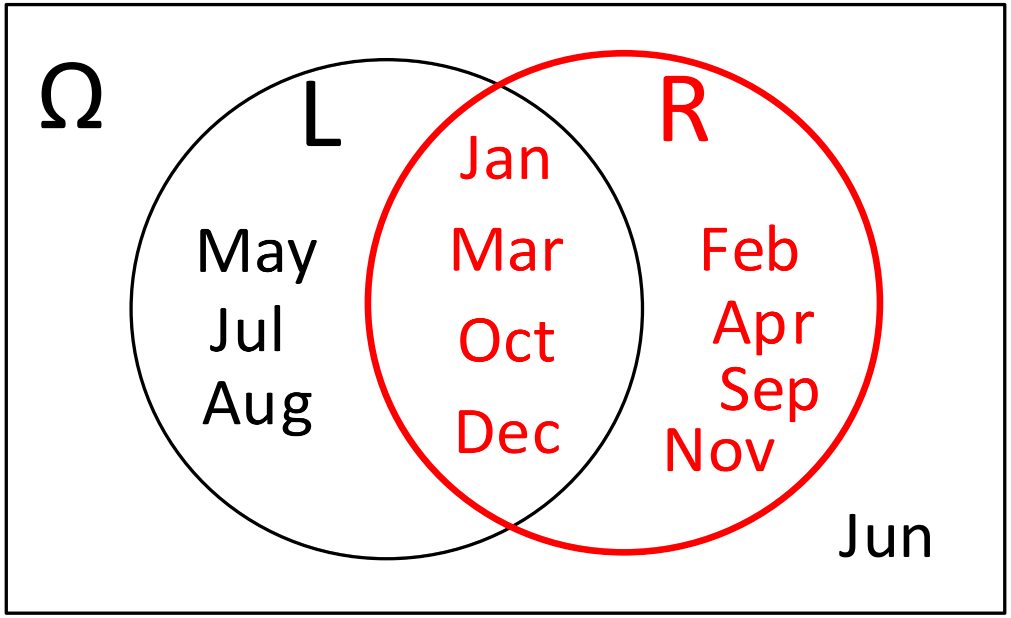


Suppose, it is known that the person we meet was born in a "long month," and we wonder whether he/she was born in a "month with the letter r."

Alternatively, if it is known that the person we meet was born in a "month with the letter r." What is the probability that he/she was born in a "long month."

# Conditional Probability




$$\Pr(R|L) = \frac{\Pr(R \cap L)}{\Pr(L)} = \frac{4}{12} \cdot \frac{12}{7} = \frac{4}{7}$$

$$\Pr(L|R) = \frac{\Pr(L \cap R)}{\Pr(R)} = \frac{4}{12} \cdot \frac{12}{8} = \frac{1}{2}$$

## The Multiplication Rule

From the definition of conditional probability we derive a useful rule by multiplying left and right by the denominator.

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} \Longrightarrow \Pr(A \cap B) = \Pr(A|B) \cdot \Pr(B)$$

Computing the probability of $A \cap B$ can hence be decomposed into two parts, computing $\Pr(B)$ and $\Pr(A|B)$ separately, which is sometimes easier than computing $\Pr(A \cap B)$directly

## The Multiplication Rule

Also,

$$\Pr(A \cap B) = \Pr(A|B) \cdot \Pr(B) = \Pr(B|A) \cdot \Pr(A)$$

Both ways are valid, but often one of $\Pr(A|B)$ and $\Pr(B|A)$ is easy and the other is not.

## The Multiplication Rule (*extended*)

Let $A_1, A_2, \ldots, A_k$ be $k$ events in a sample space $\Omega$. Then,

$$\Pr(A_1 \cap A_2 \cap \cdots \cap A_k) = \Pr(A_1) \cdot \Pr(A_2 | A_1) \cdot \Pr(A_3 | A_1 \cap A_2) \cdots \Pr(A_k | A_1 \cap A_2 \cap \cdots \cap A_{k-1})$$

Or, to put it shortly

$$\Pr\left(\bigcap_{i=1}^{k} A_i\right) = \Pr(A_1) \cdot \prod_{i=2}^{k} \Pr\left(A_i \middle| \bigcap_{j=1}^{i-1} A_j\right)$$

Conditional Probability

There are 4 red balls, 5 green balls and 2 black balls in an urn.

A ball is withdrawn from this urn three times *without replacement*.

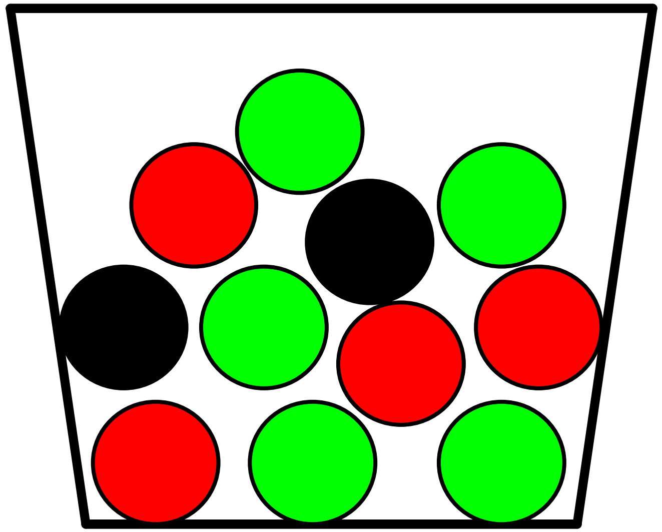
What is the probability that the first two balls are red, and the third ball is green?

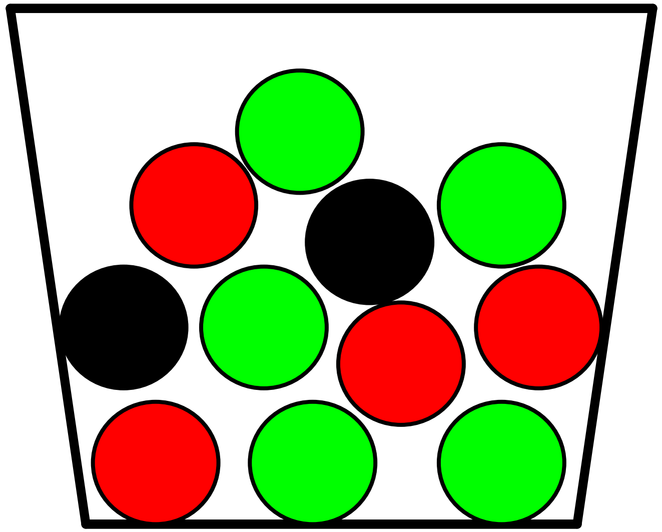
Let's denote with

$R_i$ - red ball is obtained at *i*th withdrawal;

$G_i$ - green ball is obtained at *i*th withdrawal;

$B_i$ - black ball is obtained at *i*th withdrawal.





Then, according to extended multiplication rule we need to find the following probabilities:

$\Pr(R_1)$ - red ball is obtained first;

$\Pr(R_2|R_1)$ - red ball is obtained second **given that** the first ball was also red;

$\Pr(G_3|R_1 \cap R_2)$ - green ball was obtained third, given that first two balls were red.

Since initially 4 out 11 balls were red:

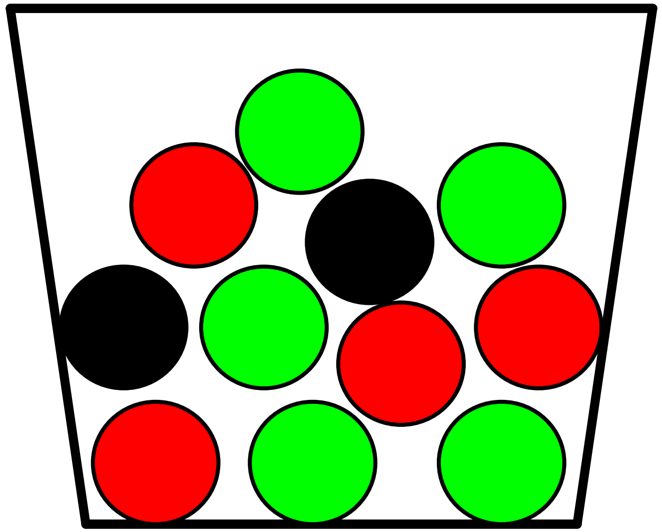
$$\Pr(R_1) = \frac{4}{11}.$$

After the 1st red ball is withdrawn

$$\Pr(R_2 | R_1) = \frac{3}{10}.$$

Finally, after two red balls were removed

$$\Pr(G_3 | R_1 \cap R_2) = \frac{5}{9}.$$

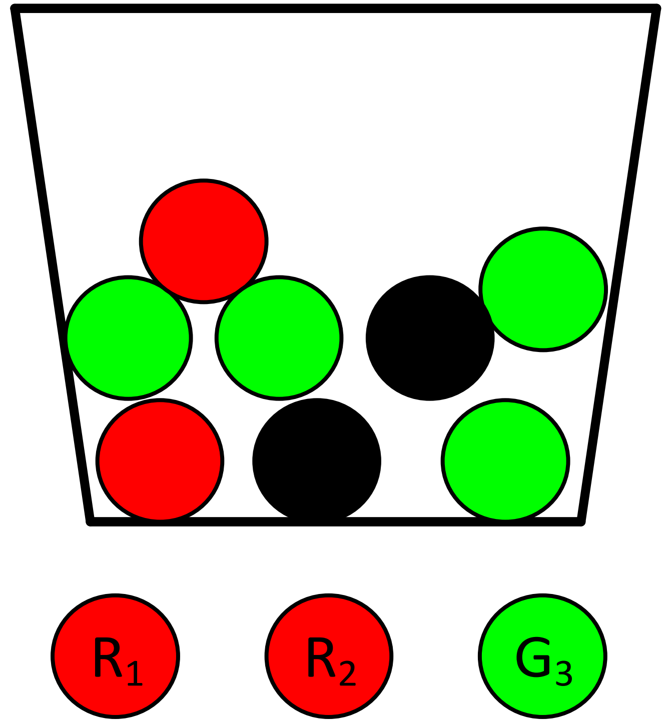


According to extended multiplication rule we have

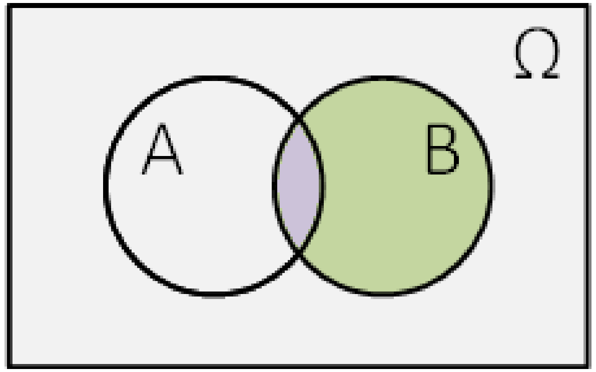
$$\Pr(R_1 \cap R_2 \cap G_3) =$$

$$= \Pr(R_1) \cdot \Pr(R_2 | R_1) \cdot \Pr(G_3 | R_1 \cap R_2) =$$

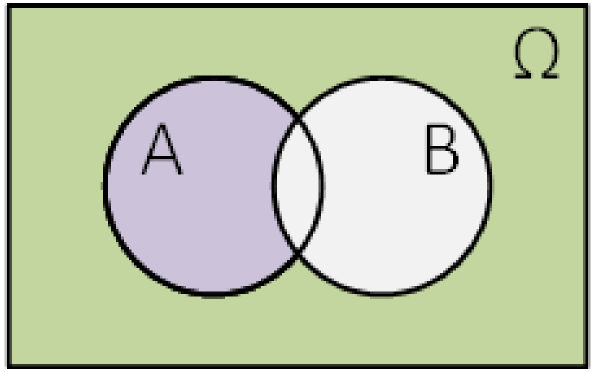
$$= \frac{4}{11} \cdot \frac{3}{10} \cdot \frac{5}{9} = \frac{2}{33}$$



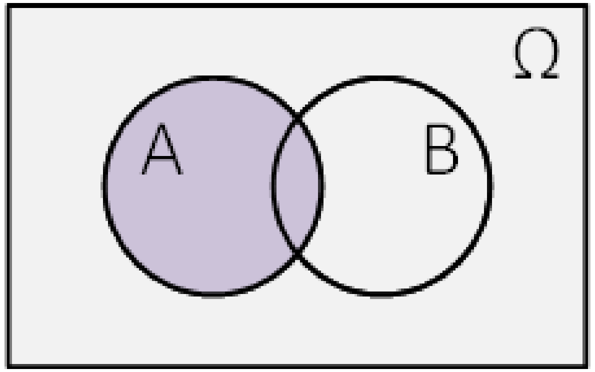
## Total Probability




$\Pr(A|B)$




$\Pr(A|B^C)$



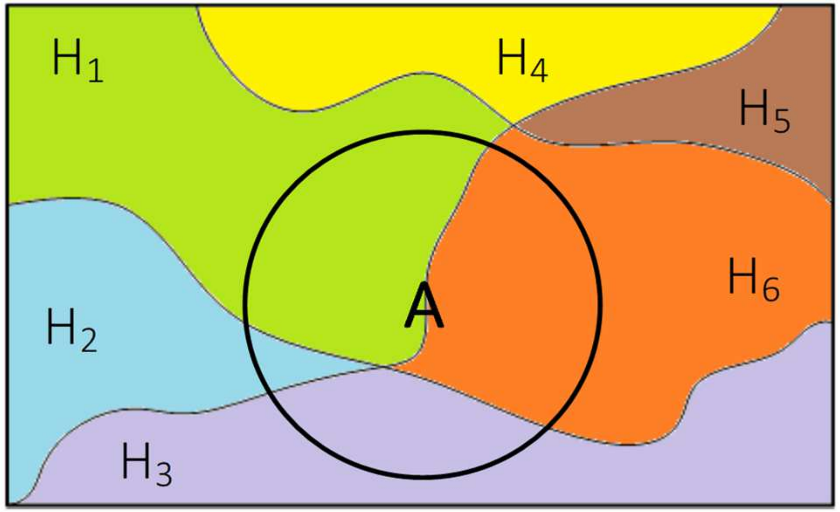

$\Pr(A)$

Basing on the definition of conditional probability, we can obtain

$$\Pr(A) = \Pr(A|B)\Pr(B) + \Pr(A|B^C)\Pr(B^C)$$

## Total Probability Formula

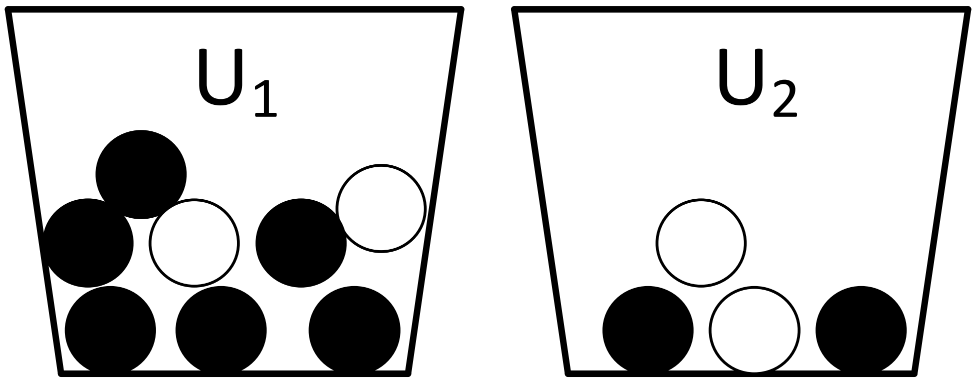
Furthermore, if $H_1, H_2, \ldots, H_k$ form a partition of the sample space $\Omega$, then




$$\Pr(A) = \sum_{i=1}^{k} \Pr(A|H_i) \cdot \Pr(H_i)$$

## Total Probability

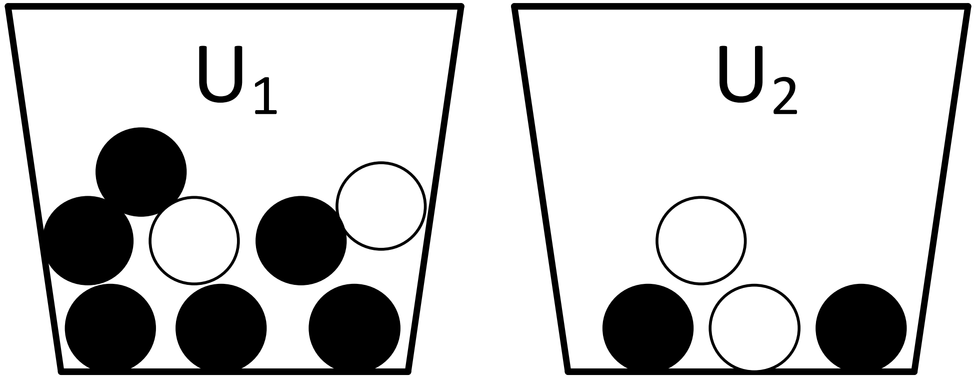
Suppose we have two urns, with the first one containing two white and six black balls and the second one containing two white and two black balls.




We pick an urn at random and then pick a ball from the chosen urn at random.

What is the probability of picking a white ball?

Let us denote the events that we choose urn 1 by $U_1$ and urn 2 by $U_2$, and that we pick a white ball by $W$.




We are given the probabilities $\Pr(U_1) = \Pr(U_2) = \frac{1}{2}$, since this is what it means that an urn is picked at random.

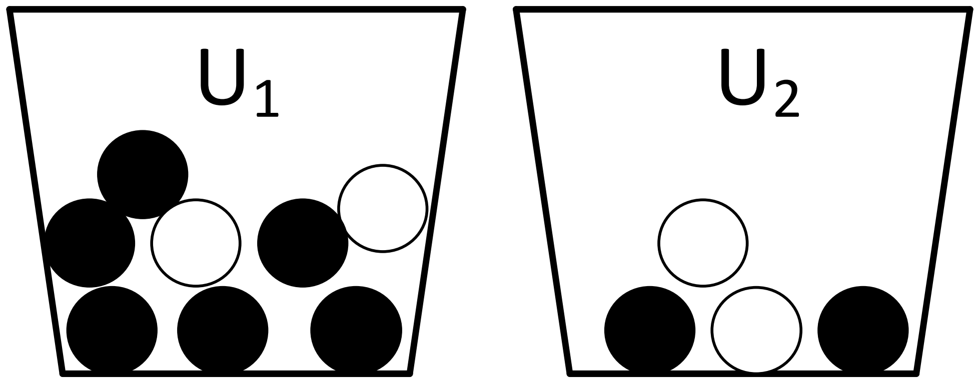
## Total Probability

Given that urn 1 is chosen, the random choice of a ball gives us the conditional probability

$$\Pr(W|U_1) = \frac{2}{8} = \frac{1}{4},$$

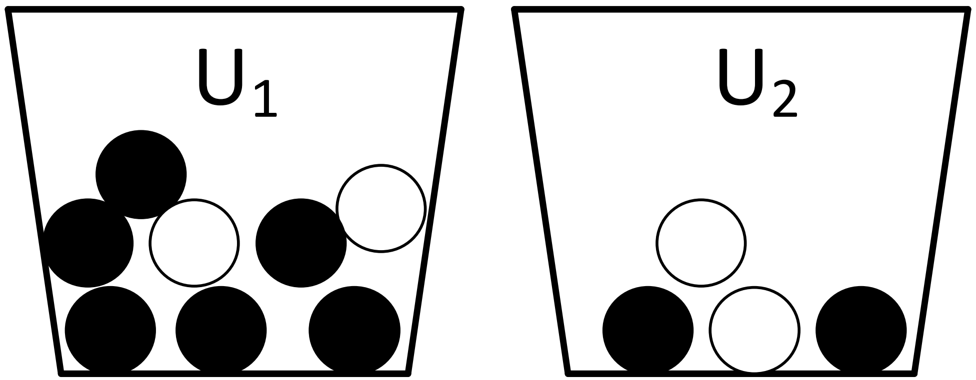
and similarly

$$\Pr(W|U_2) = \frac{2}{4} = \frac{1}{2}.$$



## Total Probability

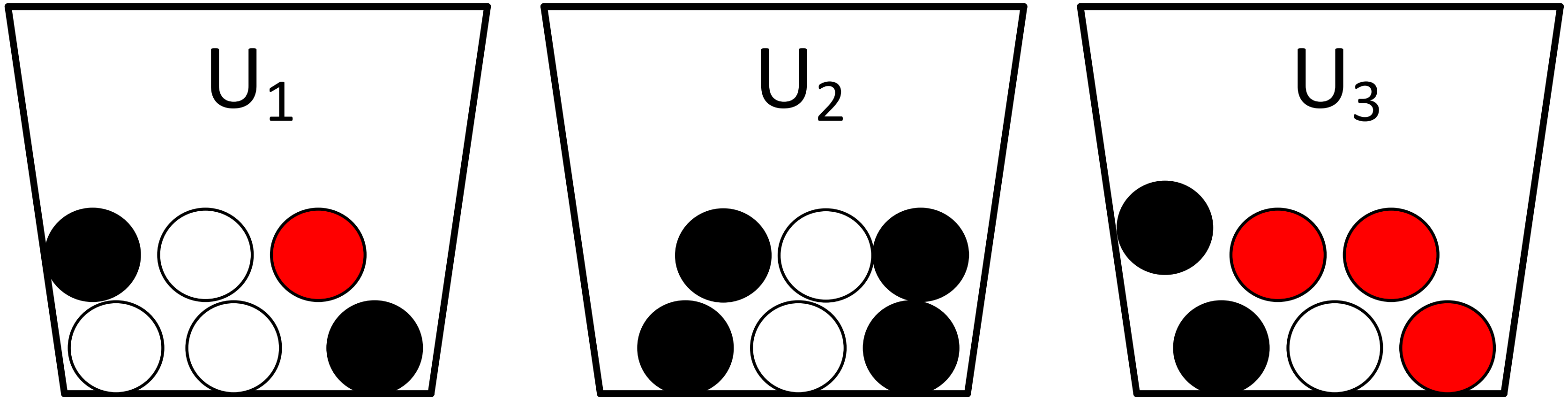
Then, by the formula of total probability




$$\Pr(W) = \Pr(W|U_1)\Pr(U_1) + \Pr(W|U_2)\Pr(U_2) =$$

$$= \frac{1}{4}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{3}{8}.$$

Total Probability

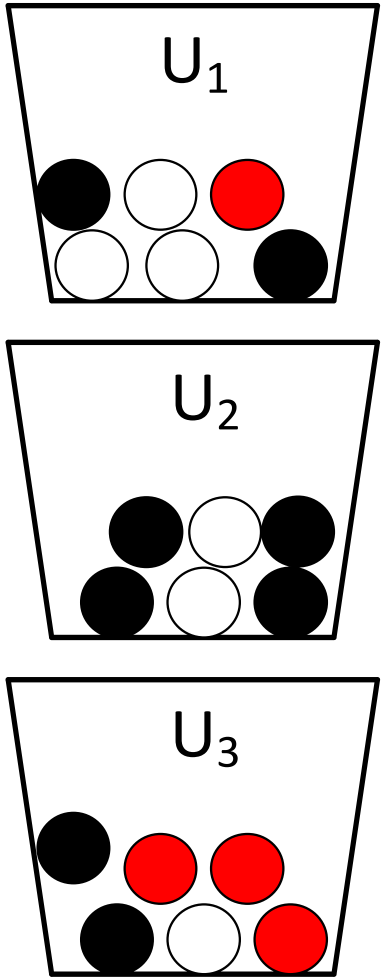



Suppose we have three urns, with
the first one containing three white, two black and one red ball,
the second one containing two white and four black balls,
and the third one – one white, two black and three red balls.

## Total Probability

We pick an urn at random and then pick two balls from the chosen urn at random (without replacement).

What is the probability of picking
- ❑ two white balls?
- ❑ two black balls?
- ❑ two red balls?
- ❑ two balls of the same color?
- ❑ two balls of different colors?



## Total Probability

Since we choose the urn randomly,

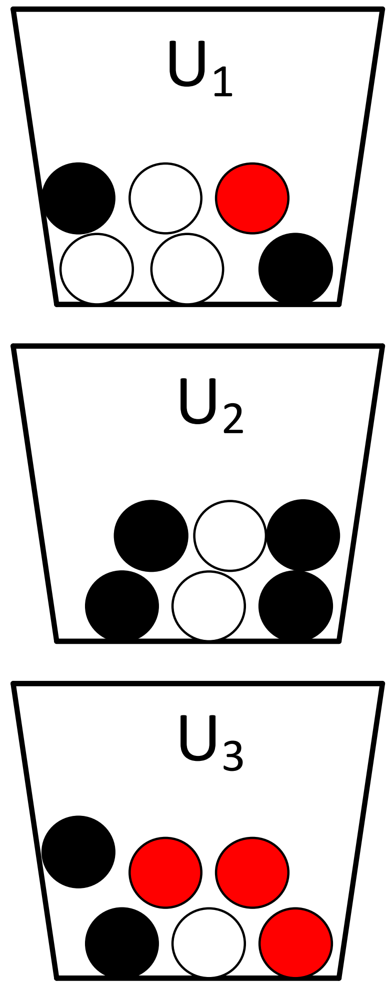
$$\Pr(U_1) = \Pr(U_2) = \Pr(U_3) = \frac{1}{3}.$$

Let's denote with

$W_i$ - white ball is obtained at *i*th withdrawal;

$B_i$ - black ball is obtained at *i*th withdrawal;

$R_i$ - red ball is obtained at *i*th withdrawal.



## Total Probability

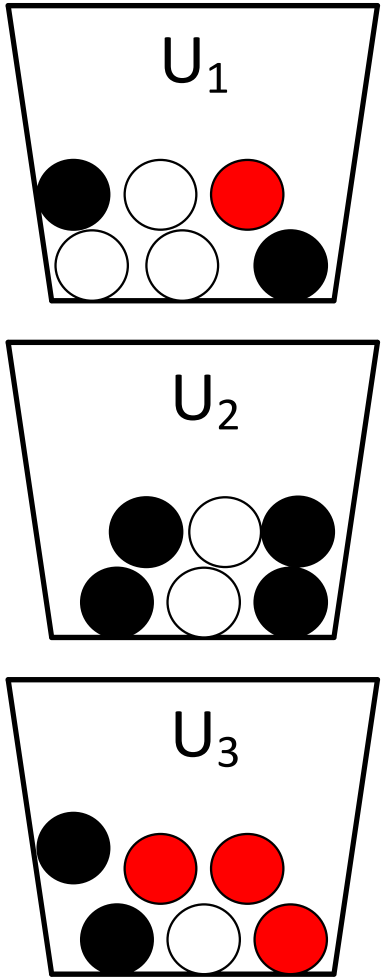
First, let's find the probability of picking 2 white balls, i.e.

$$\Pr(W_2 \cap W_1) = \Pr(W_1) \cdot \Pr\left(W_2 \middle| W_1\right)$$

$$U_1: \Pr(W_2 \cap W_1) = \frac{3}{6} \cdot \frac{2}{5} = \frac{1}{5};$$

$$U_2: \Pr(W_2 \cap W_1) = \frac{2}{6} \cdot \frac{1}{5} = \frac{1}{15};$$

$$U_3: \Pr(W_2 \cap W_1) = \frac{1}{6} \cdot \frac{0}{5} = 0.$$



## Total Probability

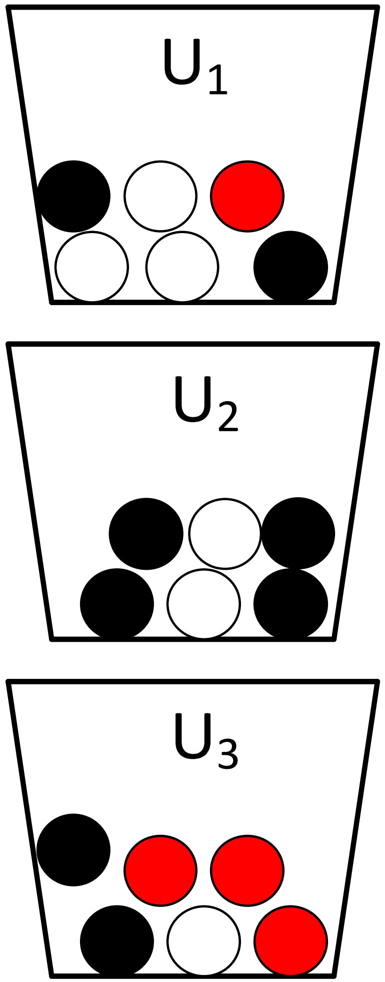
By the formula of total probability

$$\Pr(W_2 \cap W_1) = \frac{1}{5} \cdot \frac{1}{3} + \frac{1}{15} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} = \frac{4}{45}.$$

Similar reasoning allows us to obtain

$$\Pr(B_2 \cap B_1) = \frac{1}{15} \cdot \frac{1}{3} + \frac{2}{5} \cdot \frac{1}{3} + \frac{1}{15} \cdot \frac{1}{3} = \frac{8}{45}.$$

$$\Pr(R_2 \cap R_1) = 0 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{5} \cdot \frac{1}{3} = \frac{3}{45}.$$



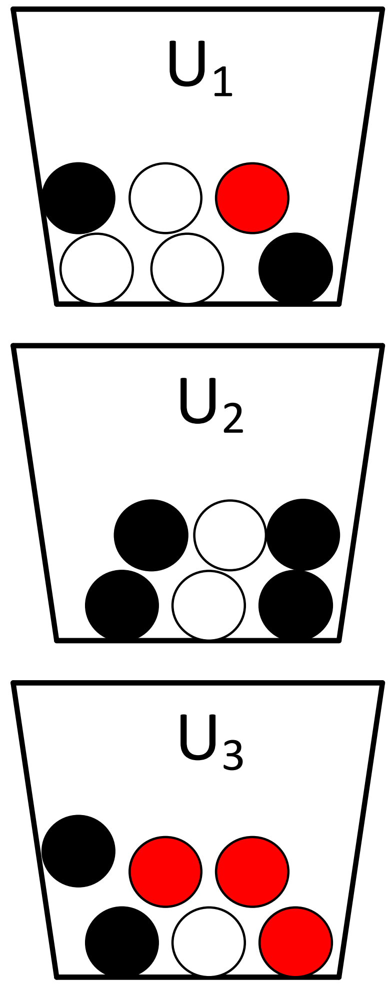
## Total Probability

The union of disjoint events $W_2 \cap W_1$, $B_2 \cap B_1$ and $R_2 \cap R_1$ results in the event $S$ consisting in picking 2 balls of the same color with probability of

$$\Pr(S) = \frac{4+8+3}{45} = \frac{15}{45} = \frac{1}{3}.$$

The event $D = S^C$ consists in picking 2 balls of different colors with probability of

$$\Pr(D) = 1 - \Pr(S) = \frac{2}{3}.$$



## Bayes' Theorem

Total probability formula allows computing probabilities for the events yet to come, provided that *a priori* (prior) probabilities of causes for these events are known.

However, if the fact of event occurrence has been established, we might recalculate the probabilities of its causes, thus obtaining *a posteriori* (posterior) probabilities.

Bayes' Theorem

Let $H_1, H_2, \dots, H_k$ be a partition of the sample space $\Omega$, and $A$ – some fixed event. Then

$$\Pr(H_i|A) = \frac{\Pr(A|H_i) \cdot \Pr(H_i)}{\sum_{i=1}^{k} \Pr(A|H_i) \cdot \Pr(H_i)}$$

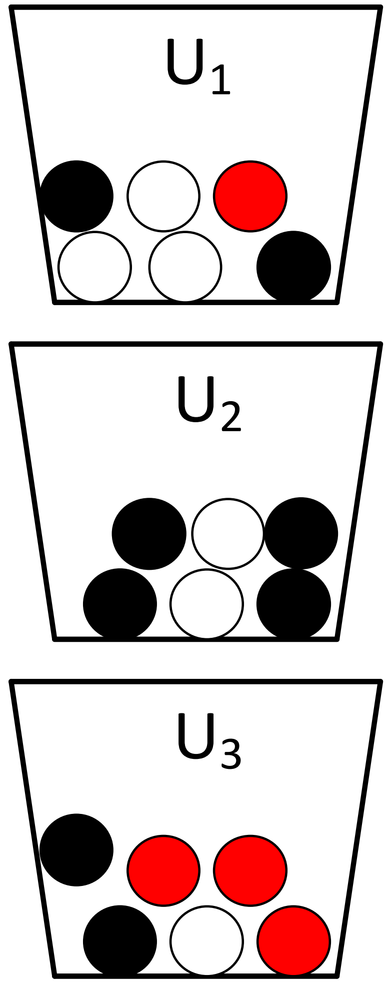
**Rev. Thomas Bayes**
**1701-1761**

The denominator represents total probability for the event $A$.

## Bayes’ Theorem

Let’s get back to the previous balls-and-urns example.

Assuming that as a result of the experiment two white balls were extracted, what is the probability that these balls were withdrawn from the first urn? the second urn?



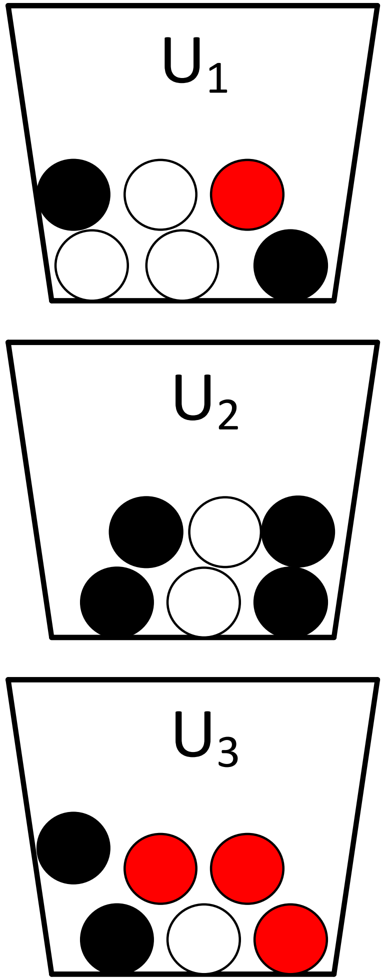
## Bayes' Theorem

Prior probabilities are

$$\Pr(U_1) = \Pr(U_2) = \Pr(U_3) = \frac{1}{3}.$$

Total probability of event $W_2 \cap W_1$ was determined as

$$\Pr(W_2 \cap W_1) = \frac{4}{45}.$$

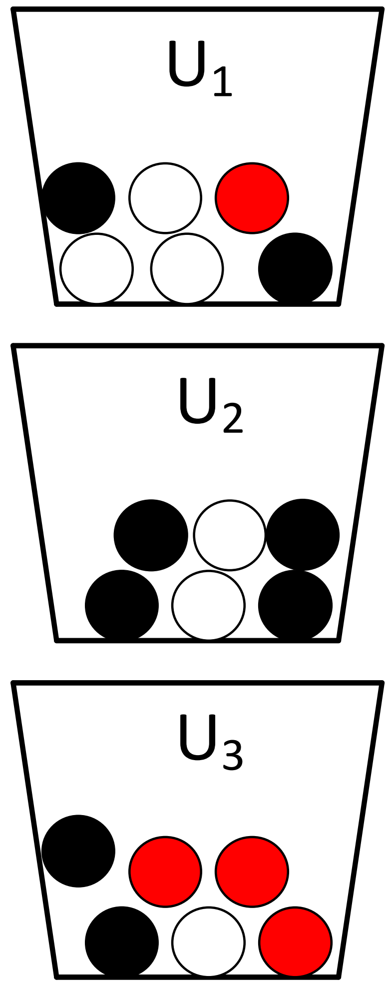


## Bayes’ Theorem

Conditional probabilities were also calculated:

$$\Pr\left(W_2 \cap W_1 \middle| U_1\right) = \frac{1}{5};$$

$$\Pr\left(W_2 \cap W_1 \middle| U_2\right) = \frac{1}{15}.$$

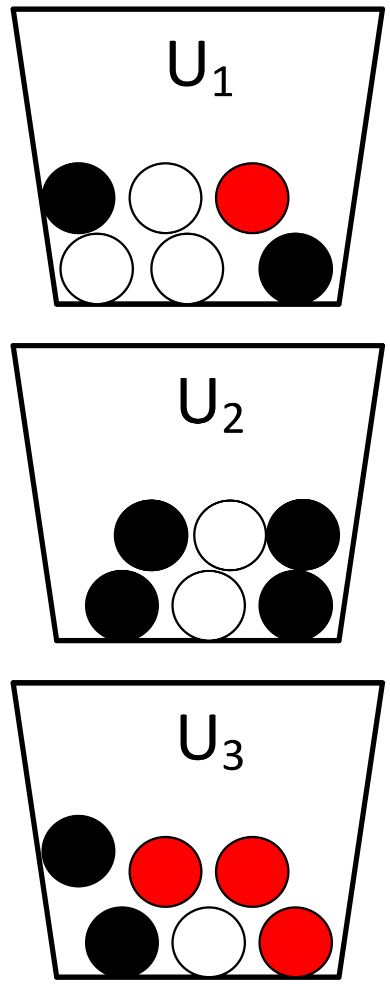


## Bayes' Theorem

Then, by Bayes' theorem we get:

$$\Pr\left(U_1 \middle| W_2 \cap W_1\right) = \frac{\Pr\left(W_2 \cap W_1 \middle| U_1\right) \cdot \Pr(U_1)}{\Pr(W_2 \cap W_1)} = \frac{\frac{1}{5} \cdot \frac{1}{3}}{\frac{4}{45}} = \frac{3}{4}.$$

$$\Pr\left(U_2 \middle| W_2 \cap W_1\right) = \frac{\Pr\left(W_2 \cap W_1 \middle| U_2\right) \cdot \Pr(U_2)}{\Pr(W_2 \cap W_1)} = \frac{\frac{1}{15} \cdot \frac{1}{3}}{\frac{4}{45}} = \frac{1}{4}.$$



## Bayes' Theorem

sensitivity of the test

A certain blood test for a disease gives a positive result 90% of the time among patients having the disease. But it also gives a positive result 25% of the time among people who do not have the disease.

100 – 25 = 75%

specificity of the test

It is believed that 30% of the population has this disease.

What is the probability that a person with a positive test result indeed has the disease?

prior probability

false positive

$A$ = "the person has the disease";
$B$ = "the blood test gives a positive result for the person."

Then, by Bayes' theorem,

$$\Pr(A|B) = \frac{\Pr(B|A) \cdot \Pr(A)}{\Pr(B|A) \cdot \Pr(A) + \Pr\left(B\middle|A^C\right) \cdot \Pr(A^C)} =$$

$$= \frac{0{,}9 \cdot 0{,}3}{0{,}9 \cdot 0{,}3 + 0{,}25 \cdot (1 - 0{,}3)} \approx 0{,}607.$$

Let's reevaluate the result by increasing the sensitivity to 99% and the specificity to 90% (10% - false positive).

At the same time, let's assume only 0,1% of the population has the disease.

Then, by Bayes' theorem,

$$\Pr(A|B) = \frac{0{,}99 \cdot 0{,}001}{0{,}99 \cdot 0{,}001 + 0{,}1 \cdot (1 - 0{,}001)} \approx 0{,}01.$$

## Independence

In our everyday language, we say that two events $A$ and $B$ are independent if the occurrence of $A$ has no effect on the occurrence of $B$ and vice versa.

Independence of events corresponds to lack of probabilistic information in one event about some other event; i.e., even if knowledge that some event $A$ has occurred was available, it would not cause us to modify the chances of the event $B$.

## Independence

An event $A$ is called independent of $B$ if

$$\Pr(A|B) = \Pr(A).$$

By application of the multiplication rule and the definition of conditional probability one may prove that independence is a *mutual property.*

## Independence

To show that $A$ and $B$ are independent it suffices to prove just one of the following:

$$\Pr(A|B) = \Pr(A),$$
$$\Pr(B|A) = \Pr(B),$$
$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B),$$

where $A$ may be replaced by $A^C$ and $B$ replaced by $B^C$, or both. If one of these statements holds, all of them are true.

Remark 1

Since the word *independence* has several meanings, one sometimes uses the terms stochastic or statistical independence to avoid ambiguity.

Remark 2

Independence of two events should not be confused with their *mutual exclusivity*.

In fact, if $A$ and $B$ are disjoint events with nonzero probabilities, then they cannot be independent: as soon as $A$ occurs, $B$ becomes impossible!

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{\Pr(\emptyset)}{\Pr(A)} = 0 \neq \Pr(B)$$

Events $A_1, A_2, \dots, A_k$ are called independent if

$$\Pr\left(\bigcap_{i=1}^{k} A_i\right) = \prod_{i=1}^{k} \Pr(A_i)\,.$$

This statement also holds when any number of the events $A_i$ are replaced by their complements throughout the formula.

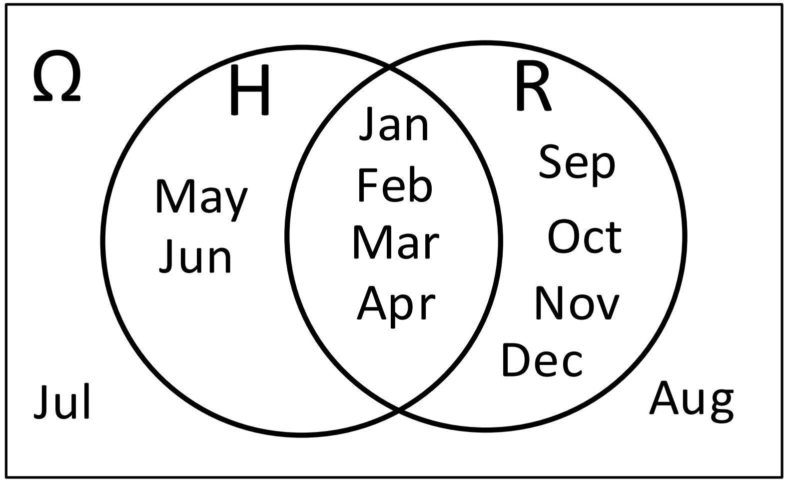
Recall the birthday events $L =$ "*born in a long month*" and $R =$ "*born in a month with the letter r*."

Let $H$ be the event "*born in the first half of the year*."

Figure out whether these pairs of events independent or not:

- $R$ and $H$;
- $L$ and $H$.

Obviously, for the event $H$, $\Pr(H) = \Pr(H^C) = \frac{1}{2}$.



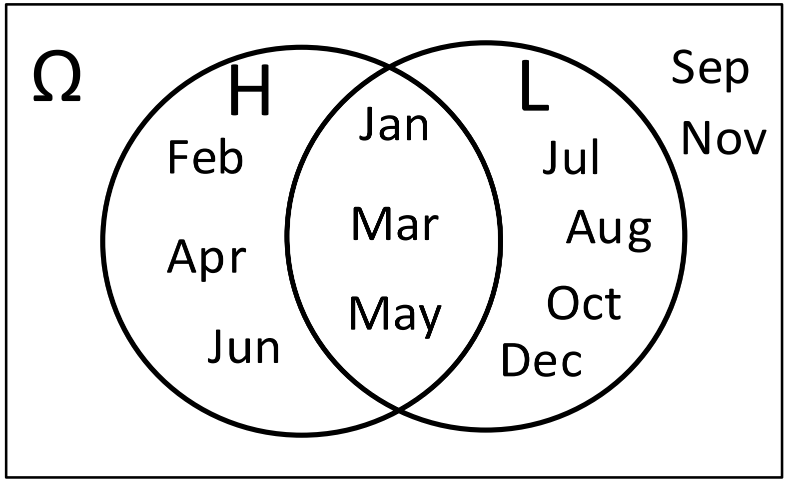

Also, $\Pr(H|R) = \frac{1}{2}$.

Since $\Pr(H|R) = \Pr(H)$, the events $R$ and $H$ are independent.

## Independence

We know that $\Pr(L \cap H) = \frac{3}{12} = \frac{1}{4}$ and $\Pr(L) = \frac{7}{12}$.

If events $L$ and $H$ are independent, $\Pr(L \cap H) = \Pr(L) \cdot \Pr(H)$.




We have $\frac{7}{12} \cdot \frac{1}{2} = \frac{7}{24} \neq \frac{1}{4}$, hence the events $L$ and $H$ are dependent.

Lecture 2

## Textbook Assignment

Géza Schay. *Introduction to Probability…*

❖Chapter 4. 75-103 pp.

❖Ex. 4.5.6 and 4.5.12

F.M. Dekking et al. *A Modern Introduction to…*

❖Chapter 3. 25-40 pp.

❖Ex. 3.5, 3.15 and 3.16