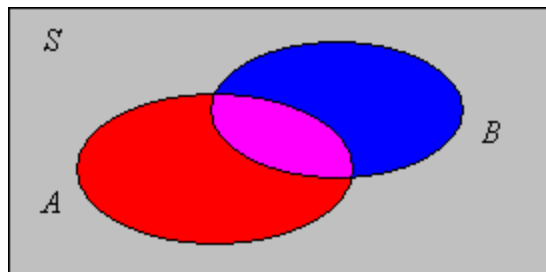


4. Conditional Probability

Definitions and Interpretations

The Basic Definition

As usual, we start with a [random experiment](#) that has a [sample space](#) S and [probability measure](#) \mathbb{P} . Suppose that we know that an event B has occurred. In general, this information should clearly alter the probabilities that we assign to other events. In particular, if A is another event then A occurs if and only if A and B occur; effectively, the sample space has been reduced to B . Thus, the probability of A , given that we know B has occurred, should be proportional to $\mathbb{P}(A \cap B)$.



However, conditional probability, given that B has occurred, should still be a probability measure, that is, it must satisfy the [axioms of probability](#). This forces the proportionality constant to be $\frac{1}{\mathbb{P}(B)}$. Thus, we are led inexorably to the following definition:

Let A and B be events in a random experiment with $\mathbb{P}(B) > 0$. The **conditional probability** of A given B is defined to be

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

The Law of Large Numbers

The previous argument was based on the axiomatic definition of probability. Let's explore the idea of conditional probability from the less formal and more intuitive notion of relative frequency (or the [law of large numbers](#)). Thus, suppose that we run the experiment repeatedly. For an arbitrary event C , let $N_n(C)$ denote the number of times C occurs in the first n runs. Note that $N_n(C)$ is a random variable in the compound experiment that consists of replicating the original experiment.

If $N_n(B)$ is large, the conditional probability that A has occurred, given that B has occurred, should be close to the **conditional relative frequency** of A given B , namely the relative frequency of A for the runs on which B occurred:

$$\frac{N_n(A \cap B)}{N_n(B)}$$

But

$$\frac{N_n(A \cap B)}{N_n(B)} = \frac{N_n(A \cap B)/n}{N_n(B)/n}$$

so by another application of the law of large numbers,

$$\frac{N_n(A \cap B)}{N_n(B)} \rightarrow \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \text{ as } n \rightarrow \infty$$

and again we are led to the same definition.

In some cases, conditional probabilities can be computed *directly*, by effectively reducing the sample space to the given event. In other cases, the formula above is better.

Conditional Distributions

Suppose that X is a random variable for our experiment that takes values in a set T . Recall that the probability distribution of X is the probability measure on T given by

$$A \mapsto \mathbb{P}(X \in A)$$

If B is an event (that is, a subset of S) with positive probability, the **conditional distribution** of X given B is the probability measure on T given by

$$A \mapsto \mathbb{P}(X \in A|B)$$

Basic Results

1. Show that for fixed B , $A \mapsto \mathbb{P}(A|B)$ is a **probability measure**.

Exercise 1 is the most important property of conditional probability because it means that any result that holds for probability measures in general holds for conditional probability, as long as the conditioning event remains fixed. In particular the results in Exercises 6-20 in the section on **Probability Measure** have analogs for conditional probability.

2. Suppose that A and B are events in a random experiment with $\mathbb{P}(B) > 0$. Prove each of the following:

- a. If $B \subseteq A$ then $\mathbb{P}(A|B) = 1$.
- b. If $A \subseteq B$ then $\mathbb{P}(A|B) = \frac{\mathbb{P}(A)}{\mathbb{P}(B)}$.
- c. If A and B are disjoint then $\mathbb{P}(A|B) = 0$

Correlation

3. Suppose that A and B are events in a random experiment, each having positive probability. Show that

- a. $\mathbb{P}(A|B) > \mathbb{P}(A)$ if and only if $\mathbb{P}(B|A) > \mathbb{P}(B)$ if and only if $\mathbb{P}(A \cap B) > \mathbb{P}(A)\mathbb{P}(B)$
- b. $\mathbb{P}(A|B) < \mathbb{P}(A)$ if and only if $\mathbb{P}(B|A) < \mathbb{P}(B)$ if and only if $\mathbb{P}(A \cap B) < \mathbb{P}(A)\mathbb{P}(B)$

$$c. \mathbb{P}(A|B) = \mathbb{P}(A) \text{ if and only if } \mathbb{P}(B|A) = \mathbb{P}(B) \text{ if and only if } \mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B)$$

In case (a), A and B are said to be **positively correlated**. Intuitively, the occurrence of either event means that the other event is more likely. In case (b), A and B are said to be **negatively correlated**. Intuitively, the occurrence of either event means that the other event is less likely. In case (c), A and B are said to be **uncorrelated** or **independent**. Intuitively, the occurrence of either event does not change the probability of the other event. Independence is a fundamental concept that can be extended to more than two events and to random variables; these generalizations are studied in the next section on [Independence](#). A much more general version of correlation, for random variables, is explored in the section on [Covariance and Correlation](#) in the chapter on [Expected Value](#).

The Multiplication Rule

Sometimes conditional probabilities are known and can be used to find the probabilities of other events.

4. Suppose that (A_1, A_2, \dots, A_n) is a sequence of events in a random experiment whose intersection has positive probability. Prove the **multiplication rule** of probability.

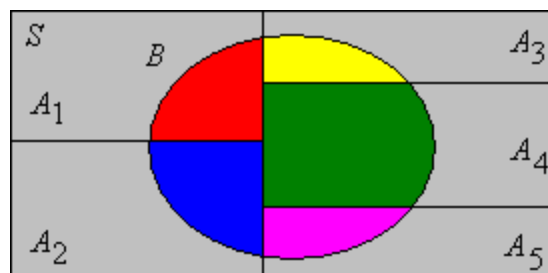
$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1) \mathbb{P}(A_2|A_1) \mathbb{P}(A_3|A_1 \cap A_2) \dots \mathbb{P}(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

Hint: Use the [definition](#) for each conditional probability on the right. You will have a collapsing product in which only the probability of the intersection of all n events survives.

The multiplication rule is particularly useful for experiments that consist of dependent stages, where A_i is an event in stage i . Compare the multiplication rule of probability with the [multiplication rule of combinatorics](#).

Conditioning and Bayes' Theorem

Suppose that $\mathcal{A} = \{A_i : i \in I\}$ is a countable collection of events that [partition](#) the sample space S , and let B be another event.



5. Prove the **law of total probability**:

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(A_i) \mathbb{P}(B|A_i)$$

Hint: Note that the i^{th} term in the sum is $\mathbb{P}(A_i \cap B)$ and recall that $\{A_i \cap B : i \in I\}$ is a partition of B .

6. Prove **Bayes' Theorem**, named after **Thomas Bayes**:

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(A_j) \mathbb{P}(B|A_j)}{\sum_{i \in I} \mathbb{P}(A_i) \mathbb{P}(B|A_i)}, \quad j \in I$$

Hint: Again the numerator is $\mathbb{P}(A_j \cap B)$ while the denominator is $\mathbb{P}(B)$ by [Exercise 5](#).

These two theorems are most useful, of course, when we know $\mathbb{P}(A_i)$ and $\mathbb{P}(B|A_i)$ for each $i \in I$. When we compute the probability of B by the law of total probability in [Exercise 5](#), we say that we are **conditioning** on the partition \mathcal{A} . Note that we can think of the sum as a **weighted average** of the conditional probabilities $\mathbb{P}(B|A_i)$ over $i \in I$, where $\mathbb{P}(A_i)$, $i \in I$ are the weight factors. In the context of Bayes theorem in [Exercise 6](#), $\mathbb{P}(A_j)$ is the **prior probability** of A_j and $\mathbb{P}(A_j|B)$ is the **posterior distribution** of A_j . We will study more general versions of conditioning and Bayes theorem in the section on [Discrete Distributions](#) in the chapter on [Distributions](#), and again in the section on [Conditional Expected Value](#) in the chapter on [Expected Value](#).

Examples and Applications

7. Suppose that A and B are events in an experiment with $\mathbb{P}(A) = \frac{1}{3}$, $\mathbb{P}(B) = \frac{1}{4}$, $\mathbb{P}(A \cap B) = \frac{1}{10}$. Find each of the following:

- $\mathbb{P}(A|B)$
- $\mathbb{P}(B|A)$
- $\mathbb{P}(A^c|B)$
- $\mathbb{P}(B^c|A)$
- $\mathbb{P}(A^c|B^c)$



8. Suppose that A , B , and C are events in a random experiment with $\mathbb{P}(A|C) = \frac{1}{2}$, $\mathbb{P}(B|C) = \frac{1}{3}$, and $\mathbb{P}(A \cap B|C) = \frac{1}{4}$. Find each of the following:

- $\mathbb{P}(A \setminus B|C)$
- $\mathbb{P}(A \cup B|C)$
- $\mathbb{P}(A^c \cap B^c|C)$



9. Suppose that A and B are events in a random experiment with $\mathbb{P}(A) = \frac{1}{2}$, $\mathbb{P}(B) = \frac{1}{3}$, and $\mathbb{P}(A|B) = \frac{3}{4}$.

- Find $\mathbb{P}(A \cap B)$
- Find $\mathbb{P}(A \cup B)$
- Find $\mathbb{P}(B \cup A^c)$
- Find $\mathbb{P}(B|A)$
- Are A and B positively correlated, negatively correlated, or independent?



10. In a certain population, 30% of the persons smoke and 8% have a certain type of heart disease. Moreover, 12% of the persons who smoke have the disease.

- What percentage of the population smoke and have the disease?
- What percentage of the population with the disease also smoke?
- Are smoking and the disease positively correlated, negatively correlated, or independent?



11. Suppose that the time X required to perform a certain job (in minutes) is uniformly distributed on the interval $(15, 60)$.

- Find the probability that the job requires more than 30 minutes.
- Given that the job is not finished after 30 minutes, find the probability that the job will require more than 15 additional minutes.
- Find the conditional distribution of X given $X > 30$.



Dice and Coins

12. Consider the experiment that consists of rolling 2 standard, fair dice and recording the sequence of scores $X = (X_1, X_2)$. Let Y denote the sum of the scores. For each of the following pairs of events, find the probability of each event and the conditional probability of each event given the other. Determine whether the events are positively correlated, negatively correlated, or independent.

- $\{X_1 = 3\}, \{Y = 5\}$
- $\{X_1 = 3\}, \{Y = 7\}$
- $\{X_1 = 2\}, \{Y = 5\}$
- $\{X_1 = 3\}, \{X_1 = 2\}$



Note that positive correlation is not a [transitive relation](#). From the previous exercise, for example, note that $\{X_1 = 3\}$ and $\{Y = 5\}$ are positively correlated, $\{Y = 5\}$ and $\{X_1 = 2\}$ are positively correlated, but $\{X_1 = 3\}$ and $\{X_1 = 2\}$ are negatively correlated (in fact, disjoint!).

13. In **dice experiment**, set $n = 2$. Run the experiment 500 times. Compute the empirical conditional probabilities corresponding to the conditional probabilities in the last exercise.

14. Consider again the experiment that consists of rolling 2 standard, fair dice and recording the sequence of scores $X = (X_1, X_2)$. Let Y denote the sum of the scores. Find the conditional distribution of X given that $Y = 7$.



15. In the die-coin experiment, a standard, fair die is rolled and then a fair coin is tossed the number of times showing on the die.

- Find the probability that all the coins show heads.

- b. Given that all coins are heads, find the probability that the die score was i for each $i \in \{1, 2, 3, 4, 5, 6\}$.



16. Run the **die-coin experiment** 200 times.

- Compute the empirical probability of the event that all coins are heads and compare with the probability in the previous exercise.
- For $i \in \{1, 2, 3, 4, 5, 6\}$ compute the empirical probability of the event that the die score was i given that there all coins were heads. Compare with the probability in the previous exercise.

17. Suppose that a bag contains 12 coins: 5 are fair, 4 are biased with probability of heads $\frac{1}{3}$; and 3 are two-headed. A coin is chosen at random from the bag and tossed.

- Find the probability that the coin is heads.
- Given that the coin is heads, find the conditional probability of each coin type.



Compare [Exercises 15](#) and [Exercise 17](#). In Exercise 15, we toss a coin with a *fixed* probability of heads a *random* number of times. In Exercise 17, we effectively toss a coin with a *random* probability of heads a *fixed* number of times. The random experiment of tossing a coin with a fixed probability of heads p a fixed number of times n is known as the **binomial experiment** with parameters n and p . This is a very basic and important experiment that is studied in more detail in the section on the [binomial distribution](#) in the chapter on [Bernoulli Trials](#). Thus, the experiments in Exercises 15 and 17 can be thought of as modifications of the binomial experiment in which a parameter has been randomized. In general, interesting new random experiments can often be constructed by randomizing one or more parameters in another random experiment.

18. In the coin-die experiment, a fair coin is tossed. If the coin lands tails, a fair die is rolled. If the coin lands heads, an ace-six flat die is tossed (faces 1 and 6 have probability $\frac{1}{4}$ each, while faces 2, 3, 4, and 5 have probability $\frac{1}{8}$ each).

- Find the probability that the die score is i , for $i \in \{1, 2, 3, 4, 5, 6\}$.
- Given that the die score is 4, find the conditional probability that the coin landed heads and the conditional probability that the coin lands tails.



19. Run the **coin-die experiment** 500 times.

- Compute the empirical probability of the event that the die score is i , for each i , and compare with the probability in the previous exercise
- Compute the empirical probability of the event that the coin landed heads, given that the die score is 4 and compare with the probability in the previous exercise.

Cards

20. Consider the **card experiment** that consists of dealing 2 cards from a standard deck and recording the sequence of cards dealt. For $i \in \{1, 2\}$, let Q_i be the event that card i is a queen and H_i the event that card i is a heart. For each of the following pairs of events, compute the probability of each event, and the conditional probability of each event given the other. Determine whether the events are positively correlated, negatively correlated, or independent.

- Q_1, H_1
- Q_1, Q_2
- Q_2, H_2
- Q_1, H_2



21. In the **card experiment**, set $n = 2$. Run the experiment 500 times. Compute the conditional relative frequencies corresponding to the conditional probabilities in the last exercise.

22. Consider the card experiment with $n = 3$ cards. Find the probability of the following events:

- All three cards are all hearts.
- The first two cards are hearts and the third is a spade.
- The first and third cards are hearts and the second is a spade.



23. In the **card experiment**, set $n = 3$ run the simulation 1000 times. Compute the empirical probability of each event in the previous exercise and compare with the true probability.

Buffon's Coin

24. Recall that **Buffon's coin experiment** consists of tossing a coin with radius $r \leq \frac{1}{2}$ randomly on a floor covered with square tiles of side length 1. The coordinates (X, Y) of the center of the coin are recorded relative to axes through the center of the square, parallel to the sides.

- Find $\mathbb{P}(Y > 0 | X < Y)$
- Find the conditional distribution of (X, Y) given that the coin does not touch the sides of the square.



25. Run **Buffon's coin experiment** 500 times. Compute the empirical probability that $Y > 0$ given that $X < Y$ and compare with the probability in the last exercise.

Data Analysis Exercises

26. For the **M&M data set**, find the empirical probability that a bag has at least 10 reds, given that the weight of the bag is at least 48 grams.



27. Consider the **Cicada data**.

- a. Find the empirical probability that a cicada weighs at least 0.25 grams given that the cicada is male.
- b. Find the empirical probability that a cicada weighs at least 0.25 grams given that the cicada is the tredecula species.



Reliability

28. A plant has 3 assembly lines that produces memory chips. Line 1 produces 50% of the chips and has a defective rate of 4%; line 2 has produces 30% of the chips and has a defective rate of 5%; line 3 produces 20% of the chips and has a defective rate of 1%. A chip is chosen at random from the plant.

- a. Find the probability that the chip is defective.
- b. Given that the chip is defective, find the conditional probability for each line.



Genetics

Recall our earlier discussion [genetics](#) in the section on [Probability Measures](#).

In the following exercise, suppose that a certain type of pea plant has either green pods or yellow pods, and that the green-pod gene is dominant. Thus, a plant with genotype gg or gy has green pods, while a plant with genotype yy has yellow pods.

29. Suppose that a green-pod plant and a yellow-pod plant are bred together. Suppose further that the green-pod plant has a $\frac{1}{4}$ chance of carrying the recessive yellow-pod gene.

- a. Find the probability that a child plant will have green pods.
- b. Given that a child plant has green pods, find the updated probability that the green-pod parent has the recessive gene.



30. Suppose that two green-pod plants are bred together. Suppose further that with probability $\frac{1}{3}$ neither plant has the recessive gene, with probability $\frac{1}{2}$ one plant has the recessive gene, and with probability $\frac{1}{6}$ both plants have the recessive gene.

- a. Find the probability that a child plant has green pods.
- b. Given that a child plant has green pods, find the updated probability that both parents have the recessive gene.



Recall that a sex-linked hereditary disorder is associated with a gene on the X chromosome. As before, let n denote the dominant normal gene and d the recessive defective gene. Thus, a woman of gene type nn is normal; a woman of genotype nd is free of the disease, but is a carrier; and a woman of genotype dd has the disease. A man of genotype n is

normal and a man of genotype d has the disease. **Dichromatism**, a form of color-blindness, is a sex-linked hereditary disorder, and is thus more common in males than females.

31. Suppose that in a certain population, 50% are male and 50% are female. Moreover, suppose that 10% of males are color-blind but only 1% of females are color-blind.
- Find the percentage of color-blind persons in the population.
 - Find the percentage of color-blind persons that are male.



32. A man and a woman do not have a certain sex-linked hereditary disorder, but the woman has a $\frac{1}{3}$ chance of being a carrier.
- Find the probability that a son born to the couple will be normal.
 - Find the probability that a daughter born to the couple will be a carrier.
 - Given that a son born to the couple is normal, find the updated probability that the mother is a carrier.



Urn Models

33. Urn 1 contains 4 red and 6 green balls while urn 2 contains 7 red and 3 green balls. An urn is chosen at random and then a ball is chosen from the selected urn.
- Find the probability that the ball is green.
 - Given that the ball is green, find the conditional probability that urn 1 was selected.



34. Urn 1 contains 4 red and 6 green balls while urn 2 contains 6 red and 3 green balls. A ball is selected at random from urn 1 and transferred to urn 2. Then a ball is selected at random from urn 2.
- Find the probability that the ball from urn 2 is green.
 - Given that the ball from urn 2 is green, find the conditional probability that the ball from urn 1 was green.



35. An urn initially contains a red and b green balls, where a and b are positive integers. A ball is chosen at random from the urn and its color is noted. It is then replaced in the urn and k new balls of the same color are added to the urn. The process is repeated. The parameter k is an integer, and if it is negative, the balls are removed.
- Find the probability of the event that the balls 1 and 2 are red and ball 3 is green.
 - Find the probability of the event that balls 1 and 3 are red and ball 2 is green.
 - Find the probability of the event that ball 1 is green and balls 2 and 3 are red.
 - Find the probability that the second ball is red.
 - Find the probability that the first ball is red given that the second ball is red.

The random process in described in the previous exercise is known as **Pólya's urn scheme**, named after **George Pólya**. Note that $k = -1$ corresponds to sampling without replacement and $k = 0$ corresponds to sampling with replacement (at least with respect to the colors of the balls). We will study **Pólya's urn** in more detail in the chapter on **Finite Sampling Models**

36. An urn initially contains 6 red and 4 green balls. A ball is chosen at random from the urn and then replaced along with a ball of the other color. The process is repeated.
- Find the probability of the event that the balls 1 and 2 are red and ball 3 is green.
 - Find the probability of the event that balls 1 and 3 are red and ball 2 is green.
 - Find the probability of the event that ball 1 is green and balls 2 and 3 are red.
 - Find the probability that the second ball is red.
 - Find the probability that the first ball is red given that the second ball is red.

Diagnostic Testing

Suppose that we have a random experiment with an event A of interest. When we run the experiment, of course, event A will either occur or not occur. However, suppose that we are not able to observe the occurrence or non-occurrence of A directly. Instead we have a **diagnostic test** designed to indicate the occurrence of event A ; thus the test that can be either **positive** for A or **negative** for A . The test also has an element of randomness, and in particular can be in error. Here are some typical examples of the type of situation we have in mind:

- The event is that a person has a certain disease and the test is a blood test for the disease.
- The event is that a woman is pregnant and the test is a home pregnancy test.
- The event is that a person is lying and the test is a lie-detector test.
- The event is that a device is defective and the test consists of sensor readings.
- The event is that a missile is in a certain region of airspace and the test consists of radar signals.
- The event is that a person has committed a crime, and the test is a jury trial with evidence presented for and against the event.

Let T be the event that the test is positive for the occurrence of A . The conditional probability $\mathbb{P}(T|A)$ is called the **sensitivity** of the test. The complementary probability

$$\mathbb{P}(T^c|A) = 1 - \mathbb{P}(T|A)$$

is the **false negative** probability. The conditional probability $\mathbb{P}(T^c|A^c)$ is called the **specificity** of the test. The complementary probability

$$\mathbb{P}(T|A^c) = 1 - \mathbb{P}(T^c|A^c)$$

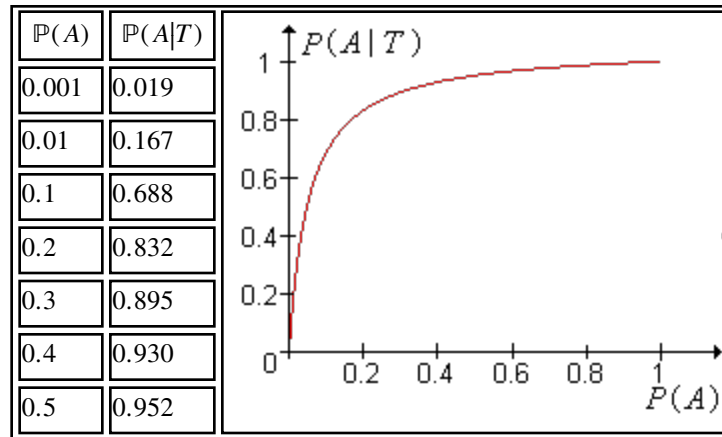
is the **false positive** probability. In many cases, the sensitivity and specificity of the test are known, as a result of the development of the test. However, the *user* of the test is interested in the opposite conditional probabilities, namely $\mathbb{P}(A|T)$, the probability of the event of interest, given a positive test, and $\mathbb{P}(A^c|T^c)$, the probability of the

complementary event, given a negative test.

37. Use Bayes' Theorem to show that

$$\mathbb{P}(A|T) = \frac{\mathbb{P}(T|A)\mathbb{P}(A)}{\mathbb{P}(T|A)\mathbb{P}(A) + \mathbb{P}(T|A^c)\mathbb{P}(A^c)}$$

38. For a concrete example, suppose that the sensitivity of the test is 0.99 and the specificity of the test is 0.95. Superficially, the test looks good. Find $\mathbb{P}(A|T)$ as a function of $\mathbb{P}(A)$ and verify the table and graph given below:



The small value of $\mathbb{P}(A|T)$ for small values of $\mathbb{P}(A)$ is striking. The moral, of course, is that $\mathbb{P}(A|T)$ depends critically on $\mathbb{P}(A)$ not just on the sensitivity and specificity of the test. Moreover, the correct comparison is $\mathbb{P}(A|T)$ with $\mathbb{P}(A)$, as in the table, not $\mathbb{P}(A|T)$ with $\mathbb{P}(T|A)$. In terms of this correct comparison, the test does indeed work well; $\mathbb{P}(A|T)$ is significantly larger than $\mathbb{P}(A)$ in all cases.

39. A woman initially believes that there is an even chance that she is or is not pregnant. She takes a home pregnancy test with sensitivity 0.95 and specificity 0.90. The test is positive. Find the updated probability that she is pregnant.



40. Suppose that 70% of defendants brought to trial for a certain type of crime are guilty. Moreover, historical data show that juries convict guilty persons 80% of the time and convict innocent persons 10% of the time. Find the probability that a person convicted of a crime of this type is guilty.



41. The "Check Engine" light on your car has turned on. Without the information from the light, you believe that there is a 10% chance that your car has a serious engine problem. You learn that if the car has such a problem, the light will come on with probability 0.99, but if the car does not have a serious problem, the light will still come on, under circumstances similar to yours, with probability 0.3. Find the updated probability that you have an engine problem.



42. The **ELISA test** for HIV has a sensitivity and specificity of 0.999. Suppose that a person is selected at random from a population in which 1% are infected with HIV. If the person has a positive test, find the probability the person has HIV.



Diagnostic testing is closely related to a general statistical procedure known as **hypothesis testing**. A separate chapter on [hypothesis testing](#) explores this procedure in detail.

Exchangeability

In this subsection, we will discuss briefly a somewhat specialized topic, but one that is still very important.

Exchangeable Events

Suppose that $\{A_i : i \in I\}$ is a collection of events in random experiment, where I is a countable index set. The collection is said to be **exchangeable** if the probability of the intersection of a finite number of the events depends only on the number of events. That is, if J and K are finite subsets of I and $\#(J) = \#(K)$ then

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \mathbb{P}\left(\bigcap_{k \in K} A_k\right)$$

43. Clearly, exchangeability has the basic inheritance property. Suppose that \mathcal{A} is a collection of events.
- Show that if \mathcal{A} is exchangeable then \mathcal{B} is exchangeable for every $\mathcal{B} \subseteq \mathcal{A}$
 - Conversely, show that if \mathcal{B} is exchangeable for every finite $\mathcal{B} \subseteq \mathcal{A}$ then \mathcal{A} is exchangeable.

For a collection of exchangeable events, the [inclusion exclusion law](#) for the probability of a union is much simpler than the general version.

44. Suppose that $\{A_1, A_2, \dots, A_n\}$ is an exchangeable collection of events. For $J \subseteq \{1, 2, \dots, n\}$ with $\#(J) = k$, let $p_k = \mathbb{P}\left(\bigcap_{j \in J} A_j\right)$. Show that

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \binom{n}{k} p_k$$

45. In [Pólya's urn scheme](#), let A_i denote the event that the i^{th} ball chosen is red. Show that $\{A_1, A_2, \dots\}$ is an exchangeable collection of events.

Exchangeable Random Variables

The concept of exchangeability can be extended to random variables in the natural way. Suppose that \mathcal{C} is a collection of random variables for the experiment, each taking values in a set T . The collection \mathcal{C} is said to be **exchangeable** if for any $\{X_1, X_2, \dots, X_n\} \subseteq \mathcal{C}$, the distribution of the random vector (X_1, X_2, \dots, X_n) depends only on n . Thus, the distribution of the random vector is unchanged if the coordinates are permuted.

46. Suppose that \mathcal{A} is a collection of events for a random experiment, and let $\mathcal{C} = \{\mathbf{1}(A) : A \in \mathcal{A}\}$ denote the corresponding collection of indicator random variables. Show that \mathcal{A} is an exchangeable collection of events if and only if \mathcal{C} is exchangeable collection of random variables.