

2.1: Random Experiments

Experiments

Probability theory is based on the paradigm of a *random experiment*; that is, an experiment whose outcome cannot be predicted with certainty, before the experiment is run. In *classical* or *frequency-based* probability theory, we also assume that the experiment can be repeated indefinitely under essentially the same conditions. The repetitions can be *in time* (as when we toss a single coin over and over again) or *in space* (as when we toss a bunch of similar coins all at once). The repeatability assumption is important because the classical theory is concerned with the long-term behavior as the experiment is replicated. By contrast, *subjective* or *belief-based* probability theory is concerned with measures of belief about what will happen when we run the experiment. In this view, repeatability is a less crucial assumption. In any event, a complete description of a random experiment requires a careful definition of precisely what information about the experiment is being recorded, that is, a careful definition of what constitutes an *outcome*.

The term *parameter* refers to a non-random quantity in a model that, once chosen, remains constant. Many probability models of random experiments have one or more parameters that can be adjusted to fit the physical experiment being modeled.

The subjects of probability and statistics have an inverse relationship of sorts. In probability, we start with a completely specified mathematical model of a random experiment. Our goal is perform various computations that help us understand the random experiment, help us predict what will happen when we run the experiment. In statistics, by contrast, we start with an incompletely specified mathematical model (one or more parameters may be unknown, for example). We run the experiment to collect data, and then use the data to draw inferences about the unknown factors in the mathematical model.

Compound Experiments

Suppose that we have n experiments (E_1, E_2, \dots, E_n) . We can form a new, compound experiment by performing the n experiments in sequence, E_1 first, and then E_2 and so on, independently of one another. The term *independent* means, intuitively, that the outcome of one experiment has no influence over any of the other experiments. We will make the term mathematically precise later.

In particular, suppose that we have a basic experiment. A fixed number (or even an infinite number) of independent replications of the basic experiment is a new, compound experiment. Many experiments turn out to be compound experiments and moreover, as noted above, (classical) probability theory itself is based on the idea of replicating an experiment.

In particular, suppose that we have a simple experiment with two outcomes. Independent replications of this experiment are referred to as Bernoulli trials, named for Jacob Bernoulli. This is one of the simplest, but most important models in probability. More generally, suppose that we have a simple experiment with k possible outcomes. Independent replications of this experiment are referred to as multinomial trials.

Sometimes an experiment occurs in well-defined stages, but in a *dependent* way, in the sense that the outcome of a given stage is influenced by the outcomes of the previous stages.

Sampling Experiments

In most statistical studies, we start with a population of objects of interest. The objects may be people, memory chips, or acres of corn, for example. Usually there are one or more numerical measurements of interest to us—the height and weight of a person, the lifetime of a memory chip, the amount of rain, amount of fertilizer, and yield of an acre of corn.

Although our interest is in the entire population of objects, this set is usually too large and too amorphous to study. Instead, we collect a random sample of objects from the population and record the measurements of interest of for each object in the sample.

There are two basic types of sampling. If we sample *with replacement*, each item is replaced in the population before the next draw; thus, a single object may occur several times in the sample. If we sample *without replacement*, objects are not replaced in the population. The chapter on Finite Sampling Models explores a number of models based on sampling from a finite population.

Sampling with replacement can be thought of as a compound experiment, based on independent replications of the simple experiment of drawing a single object from the population and recording the measurements of interest. Conversely, a compound experiment that consists of n independent replications of a simple experiment can usually be thought of as a sampling experiment.

On the other hand, sampling without replacement is an experiment that consists of dependent stages, because the population changes with each draw.

Examples and Applications

Probability theory is often illustrated using simple devices from games of chance: coins, dice, card, spinners, urns with balls, and so forth. Examples based on such devices are pedagogically valuable because of their simplicity and conceptual clarity. On the other hand, it would be a terrible shame if you were to think that probability is *only* about gambling and games of chance. Rather, try to see problems involving coins, dice, etc. as metaphors for more complex and realistic problems.

Coins and Dice

In terms of probability, the important fact about a coin is simply that when tossed it lands on one side or the other. Coins in Western societies, dating to antiquity, usually have the head of a prominent person engraved on one side and something of lesser importance on the other. In non-Western societies, coins often did not have a head on either side, but did have distinct engravings on the two sides, one typically more important than the other. Nonetheless, *heads* and *tails* are the ubiquitous terms used in probability theory to distinguish the *front* or *obverse* side of the coin from the *back* or *reverse* side of the coin.



Figure 2.1.1: Obverse and reverse sides of a Roman coin, about 241 CE, from [Wikipedia](#)

Consider the *coin experiment* of tossing a coin n times and recording the score (1 for heads or 0 for tails) for each toss.

1. Identify a parameter of the experiment.
2. Interpret the experiment as a compound experiment.
3. Interpret the experiment as a sampling experiment.
4. Interpret the experiment as n Bernoulli trials.

Answer

1. The number of coins n is the parameter.
2. The experiment consists of n independent replications of the simple experiment of tossing the coin one time.
3. The experiment can be thought of as selecting a sample of size n with replacement from the population $\{0, 1\}$.
4. There are two outcomes on each toss and the tosses are independent.

In the simulation of the coin experiment, set $n = 5$. Run the simulation 100 times and observe the outcomes.

Dice are randomizing devices that, like coins, date to antiquity and come in a variety of sizes and shapes. Typically, the faces of a die have numbers or other symbols engraved on them. Again, the important fact is that when a die is thrown, a unique face is chosen (usually the upward face, but sometimes the downward one). For more on dice, see the introductory section in the chapter on Games of Chance.

Consider the *dice experiment* of throwing a k -sided die (with faces numbered 1 to k), n times and recording the scores for each throw.

1. Identify the parameters of the experiment.
2. Interpret the experiment as a compound experiment.
3. Interpret the experiment as a sampling experiment.
4. Identify the experiment as n multinomial trials.

Answer

1. The parameters are the number of dice n and the number of faces k .
2. The experiment consists of n independent replications of the simple experiment of throwing one die.
3. The experiment can be thought of as selecting a sample of size n with replacement from the population $\{1, 2, \dots, k\}$.
4. The same k outcomes occur for each die the throws are independent.

In reality, most dice are *Platonic solids* (named for Plato of course) with 4, 6, 8, 12, or 20 sides. The six-sided die is the *standard die*.



Figure 2.1.2: Blue Platonic dice

In the simulation of the dice experiment, set $n = 5$. Run the simulation 100 times and observe the outcomes.

In the *die-coin experiment*, a standard die is thrown and then a coin is tossed the number of times shown on the die. The sequence of coin scores is recorded (1 for heads and 0 for tails). Interpret the experiment as a compound experiment.

Answer

The first stage consists rolling the die and the second stage consists of tossing the coin. The stages are dependent because the number of tosses depends on the outcome of the die throw.

Note that this experiment can be obtained by *randomizing* the parameter n in the basic coin experiment in (1).

Run the simulation of the die-coin experiment 100 times and observe the outcomes.

In the *coin-die experiment*, a coin is tossed. If the coin lands heads, a red die is thrown and if the coin lands tails, a green die is thrown. The coin score (1 for heads and 0 for tails) and the die score are recorded. Interpret the experiment as a compound experiment.

Answer

The first stage consists of tossing the coin and the second stage consists of rolling the die. The stages are dependent because different dice (that may behave differently) are thrown, depending on the outcome of the coin toss.

Run the simulation of the coin-die experiment 100 times and observe the outcomes.

Cards

Playing cards, like coins and dice, date to antiquity. From the point of view of probability, the important fact is that a playing card encodes a number of properties or attributes on the front of the card that are hidden on the back of the card. (Later in this chapter, these properties will become random variables.) In particular, a standard *card deck* can be modeled by the Cartesian product set

$$D = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, j, q, k\} \times \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\} \quad (2.1.1)$$

where the first coordinate encodes the *denomination* or *kind* (ace, 2–10, jack, queen, king) and where the second coordinate encodes the *suit* (clubs, diamonds, hearts, spades). Sometimes we represent a card as a *string* rather than an ordered pair (for example $q\heartsuit$ rather than (q, \heartsuit) for the queen of hearts). Some other properties, derived from the two main ones, are *color* (diamonds and hearts are red, clubs and spades are black), *face* (jacks, queens, and kings have faces, the other cards do not), and *suit order* (from least to highest rank: $(\clubsuit, \diamondsuit, \heartsuit, \spadesuit)$).

Consider the card experiment that consists of dealing n cards from a standard deck (without replacement).

1. Identify a parameter of the experiment.
2. Interpret the experiment as a compound experiment.
3. Interpret the experiment as a sampling experiment.

Answer

1. The parameter is n , the number of cards dealt.
2. At each stage, we draw a card from a deck, but the deck changes from one draw to the next, so the stages are dependent.
3. The experiment is to select a sample of size n from the population D , without replacement.

In the simulation of the card experiment, set $n = 5$. Run the simulation 100 times and observe the outcomes.

The special case $n = 5$ is the *poker experiment* and the special case $n = 13$ is the *bridge experiment*.

Open each of the following to see depictions of card playing in some famous paintings.

1. Cheat with the Ace of Clubs by Georges de La Tour
2. The Cardsharps by Michelangelo Carravagio
3. The Card Players by Paul Cézanne
4. His Station and Four Aces by CM Coolidge
5. Waterloo by CM Coolidge

Urn Models

Urn models are often used in probability as simple metaphors for sampling from a finite population.

An urn contains m distinct balls, labeled from 1 to m . The experiment consists of selecting n balls from the urn, without replacement, and recording the sequence of ball numbers.

1. Identify the parameters of the experiment.
2. Interpret the experiment as a compound experiment.
3. Interpret the experiment as a sampling experiment.

Answer

1. The parameters are the number of balls m and the sample size n .
2. At each stage, we draw a ball from the urn, but the contents of the urn change from one draw to the next, so the stages are dependent
3. The experiment is to select a sample of size n from the balls in the urn (the population), without replacement.

Consider the basic urn model of the previous exercise. Suppose that r of the m balls are red and the remaining $m - r$ balls are green. Identify an additional parameter of the model. This experiment is a metaphor for sampling from a general *dichotomous* population

Answer

The parameters are the population size m , the sample size n , and the number of red balls r .

In the simulation of the urn experiment, set $m = 100$, $r = 40$, and $n = 25$. Run the experiment 100 times and observe the results.

An urn initially contains m balls; r are red and $m - r$ are green. A ball is selected from the urn and removed, and then replaced with k balls of the same color. The process is repeated. This is known as *Pólya's urn model*, named after George Pólya.

1. Identify the parameters of the experiment.
2. Interpret the case $k = 0$ as a sampling experiment.
3. Interpret the case $k = 1$ as a sampling experiment.

Answer

1. The parameters are the population size m , the initial number of red balls r , and the number new balls added k .
2. When $k = 0$, each ball drawn is removed and no new balls are added, so the experiment is to select a sample of size n from the urn, without replacement.
3. When $k = 1$, each ball drawn is replaced with another ball of the same color. So at least in terms of the colors of the balls, the experiment is equivalent to selecting a sample of size n from the urn, with replacement.

Open the image of the painting *Allegory of Fortune* by Dosso Dossi. Presumably the young man has chosen lottery tickets from an urn.

Buffon's Coin Experiment

Buffon's coin experiment consists of tossing a coin with radius $r \leq \frac{1}{2}$ on a floor covered with square tiles of side length 1. The coordinates of the center of the coin are recorded, relative to axes through the center of the square, parallel to the sides. The experiment is named for comte de Buffon.

1. Identify a parameter of the experiment
2. Interpret the experiment as a compound experiment.
3. Interpret the experiment as sampling experiment.

Answer

1. The parameter is the coin radius r .
2. The experiment can be thought of as selecting the coordinates of the coin center independently of one another.
3. The experiment is equivalent to selecting a sample of size 2 from the population $[-\frac{1}{2}, \frac{1}{2}]$, with replacement.

In the simulation of Buffon's coin experiment, set $r = 0.1$. Run the experiment 100 times and observe the outcomes.

Reliability

In the usual model of *structural reliability*, a system consists of n components, each of which is either *working* or *failed*. The states of the components are uncertain, and hence define a random experiment. The system as a whole is also either working or failed, depending on the states of the components and how the components are connected. For example, a *series system* works if and only if each component works, while a *parallel system* works if and only if at least one component works. More generally, a *k out of n system* works if at least k components work.

Consider the k out of n reliability model.

1. Identify two parameters.
2. What value of k gives a series system?
3. What value of k gives a parallel system?

Answer

1. The parameters are k and n
2. $k = n$ gives a series system.
3. $k = 1$ gives a parallel system.

The reliability model above is a *static* model. It can be extended to a *dynamic* model by assuming that each component is initially working, but has a random time until failure. The system as a whole would also have a random time until failure that would depend on the component failure times and the structure of the system.

Genetics

In ordinary sexual reproduction, the genetic material of a child is a random combination of the genetic material of the parents. Thus, the birth of a child is a random experiment with respect to outcomes such as eye color, hair type, and many other physical traits. We are often particularly interested in the random transmission of traits and the random transmission of genetic disorders.

For example, let's consider an overly simplified model of an *inherited trait* that has two possible states (*phenotypes*), say a pea plant whose pods are either green or yellow. The term *allele* refers to alternate forms of a particular gene, so we are assuming that there is a gene that determines pod color, with two alleles: g for green and y for yellow. A pea plant has two alleles for the trait (one from each parent), so the possible *genotypes* are

- gg , alleles for green pods from each parent.
- gy , an allele for green pods from one parent and an allele for yellow pods from the other (we usually cannot observe which parent contributed which allele).
- yy , alleles for yellow pods from each parent.

The genotypes gg and yy are called *homozygous* because the two alleles are the same, while the genotype gy is called *heterozygous* because the two alleles are different. Typically, one of the alleles of the inherited trait is *dominant* and the other *recessive*. Thus, for example, if g is the dominant allele for pod color, then a plant with genotype gg or gy has green pods, while a plant with genotype

yy has yellow pods. Genes are passed from parent to child in a random manner, so each new plant is a random experiment with respect to pod color.

Pod color in peas was actually one of the first examples of an inherited trait studied by Gregor Mendel, who is considered the father of modern genetics. Mendel also studied the color of the flowers (yellow or purple), the length of the stems (short or long), and the texture of the seeds (round or wrinkled).

For another example, the *ABO blood type* in humans is controlled by three alleles: *a*, *b*, and *o*. Thus, the possible genotypes are *aa*, *ab*, *ao*, *bb*, *bo* and *oo*. The alleles *a* and *b* are *co-dominant* and *o* is *recessive*. Thus there are four possible blood types (phenotypes):

- Type *A*: genotype *aa* or *ao*
- Type *B*: genotype *bb* or *bo*
- Type *AB*: genotype *ab*
- type *O*: genotype *oo*

Of course, blood may be typed in much more extensive ways than the simple *ABO* typing. The *RH* factor (positive or negative) is the most well-known example.

For our third example, consider a *sex-linked* hereditary disorder in humans. This is a disorder due to a defect on the *X chromosome* (one of the two chromosomes that determine gender). Suppose that *h* denotes the healthy allele and *d* the defective allele for the gene linked to the disorder. Women have two X chromosomes, and typically *d* is *recessive*. Thus, a woman with genotype *hh* is completely normal with respect to the condition; a woman with genotype *hd* does not have the disorder, but is a *carrier*, since she can pass the defective allele to her children; and a woman with genotype *dd* has the disorder. A man has only one X chromosome (his other sex chromosome, the *Y chromosome*, typically plays no role in the disorder). A man with genotype *h* is normal and a man with genotype *d* has the disorder. Examples of sex-linked hereditary disorders are *dichromatism*, the most common form of color-blindness, and *hemophilia*, a bleeding disorder. Again, genes are passed from parent to child in a random manner, so the birth of a child is a random experiment in terms of the disorder.

Point Processes

There are a number of important processes that generate “random points in time”. Often the random points are referred to as *arrivals*. Here are some specific examples:

- times that a piece of radioactive material emits elementary particles
- times that customers arrive at a store
- times that requests arrive at a web server
- failure times of a device

To formalize an experiment, we might record the number of arrivals during a specified interval of time or we might record the times of successive arrivals.

There are other processes that produce “random points in space”. For example,

- flaws in a piece of sheet metal
- errors in a string of symbols (in a computer program, for example)
- raisins in a cake
- misprints on a page
- stars in a region of space

Again, to formalize an experiment, we might record the number of points in a given region of space.

Statistical Experiments

In 1879, Albert Michelson constructed an experiment for measuring the speed of light with an interferometer. The velocity of light data set contains the results of 100 repetitions of Michelson's experiment. Explore the data set and explain, in a general way, the variability of the data.

Answer

The variability is due to measurement and other experimental errors beyond the control of Michelson.

In 1998, two students at the University of Alabama in Huntsville designed the following experiment: purchase a bag of M&Ms (of a specified advertised size) and record the counts for red, green, blue, orange, and yellow candies, and the net weight (in grams). Explore the M&M data set and explain, in a general way, the variability of the data.

Answer

The variability in weight is due to measurement error on the part of the students and to manufacturing errors on the part of the company. The variability in color counts is less clear and may be due to purposeful randomness on the part of the company.

In 1999, two researchers at Belmont University designed the following experiment: capture a cicada in the Middle Tennessee area, and record the body weight (in grams), the wing length, wing width, and body length (in millimeters), the gender, and the species type. The cicada data set contains the results of 104 repetitions of this experiment. Explore the cicada data and explain, in a general way, the variability of the data.

Answer

The variability in body measurements is due to differences in the three species, to all sorts of environmental factors, and to measurement errors by the researchers.

On June 6, 1761, James Short made 53 measurements of the parallax of the sun, based on the transit of Venus. Explore the Short data set and explain, in a general way, the variability of the data.

Answer

The variability is due to measurement and other experimental errors beyond the control of Short.

In 1954, two massive field trials were conducted in an attempt to determine the effectiveness of the new vaccine developed by Jonas Salk for the prevention of polio. In both trials, a *treatment group* of children were given the vaccine while a *control group* of children were not. The incidence of polio in each group was measured. Explore the polio field trial data set and explain, in a general way, the underlying random experiment.

Answer

The basic random experiment is to observe whether a given child, in the treatment group or control group, comes down with polio in a specified period of time. Presumably, a lower incidence of polio in the treatment group compared with the control group would be evidence that the vaccine was effective.

Each year from 1969 to 1972 a lottery was held in the US to determine who would be drafted for military service. Essentially, the lottery was a *ball and urn model* and became famous because many believed that the process was not sufficiently random. Explore the Vietnam draft lottery data set and speculate on how one might judge the degree of randomness.

Answer

This is a difficult problem, but presumably in a sufficiently random lottery, one would not expect to see dates in the same month clustered too closely together. Observing such clustering, then, would be evidence that the lottery was not random.

Deterministic Versus Probabilistic Models

One could argue that some of the examples discussed above are inherently *deterministic*. In tossing a coin, for example, if we know the initial conditions (involving position, velocity, rotation, etc.), the forces acting on the coin (gravity, air resistance, etc.), and the makeup of the coin (shape, mass density, center of mass, etc.), then the laws of physics should allow us to predict precisely how the coin will land. This is true in a technical, theoretical sense, but false in a very real sense. Coins, dice, and many more complicated and important systems are *chaotic* in the sense that the outcomes of interest depend in a very sensitive way on the initial conditions and other parameters. In such situations, it might well be impossible to ever know the initial conditions and forces accurately enough to use deterministic methods.

In the coin experiment, for example, even if we strip away most of the real world complexity, we are still left with an essentially random experiment. Joseph Keller in his article "[The Probability of Heads](#)" deterministically analyzed the toss of a coin under a number of ideal assumptions:

1. The coin is a perfect circle and has negligible thickness

2. The center of gravity of the coin is the geometric center.
3. The coin is initially heads up and is given an initial upward velocity u and angular velocity ω .
4. In flight, the coin rotates about a horizontal axis along a diameter of the coin.
5. In flight, the coin is governed only by the force of gravity. All other possible forces (air resistance or wind, for example) are neglected.
6. The coin does not bounce or roll after landing (as might be the case if it lands in sand or mud).

Of course, few of these ideal assumptions are valid for real coins tossed by humans. Let $t = u/g$ where g is the acceleration of gravity (in appropriate units). Note that the t just has units of time (in seconds) and hence is independent of how distance is measured. The scaled parameter t actually represents the time required for the coin to reach its maximum height.

Keller showed that the regions of the parameter space (t, ω) where the coin lands either heads up or tails up are separated by the curves

$$\omega = \left(2n \pm \frac{1}{2}\right) \frac{\pi}{2t}, \quad n \in \mathbb{N} \quad (2.1.2)$$

The parameter n is the total number of revolutions in the toss. A plot of some of these curves is given below. The largest region, in the lower left corner, corresponds to the event that the coin does not complete even one rotation, and so of course lands heads up, just as it started. The next region corresponds to one rotation, with the coin landing tails up. In general, the regions alternate between heads and tails.

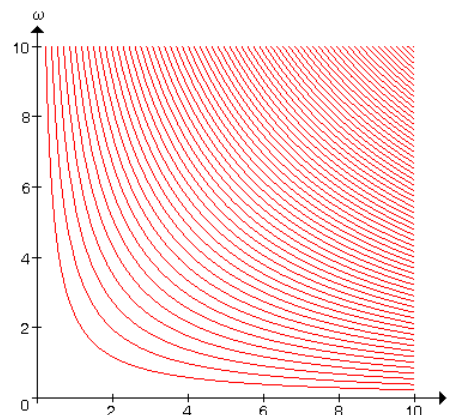


Figure 2.1.3: Regions of heads and tails

The important point, of course, is that for even moderate values of t and ω , the curves are very close together, so that a small change in the initial conditions can easily shift the outcome from heads up to tails up or conversely. As noted in Keller's article, the probabilist and statistician Persi Diaconis determined experimentally that typical values of the initial conditions for a real coin toss are $t = \frac{1}{4}$ seconds and $\omega = 76\pi \approx 238.6$ radians per second. These values correspond to $n = 19$ revolutions in the toss. Of course, this parameter point is far beyond the region shown in our graph, in a region where the curves are exquisitely close together.

This page titled [2.1: Random Experiments](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist](#) ([Random Services](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.