

2.2: Events and Random Variables

The purpose of this section is to study two basic types of objects that form part of the model of a random experiment. If you are a new student of probability, just ignore the measure-theoretic terminology and skip the technical details.

Sample Spaces

The Set of Outcomes

Recall that in a random experiment, the outcome cannot be predicted with certainty, before the experiment is run. On the other hand:

We assume that we can identify a fixed set S that includes all possible outcomes of a random experiment. This set plays the role of the universal set when modeling the experiment.

For simple experiments, S may be precisely the set of possible outcomes. More often, for complex experiments, S is a mathematically convenient set that includes the possible outcomes and perhaps other elements as well. For example, if the experiment is to throw a standard die and record the score that occurs, we would let $S = \{1, 2, 3, 4, 5, 6\}$, the set of possible outcomes. On the other hand, if the experiment is to capture a cicada and measure its body weight (in milligrams), we might conveniently take $S = [0, \infty)$, even though most elements of this set are impossible (we hope!). The problem is that we may not know exactly the outcomes that are possible. Can a light bulb burn without failure for one thousand hours? For one thousand days? for one thousand years?

Often the outcome of a random experiment consists of one or more real measurements, and thus, the S consists of all possible measurement sequences, a subset of \mathbb{R}^n for some $n \in \mathbb{N}_+$. More generally, suppose that we have n experiments and that S_i is the set of outcomes for experiment $i \in \{1, 2, \dots, n\}$. Then the Cartesian product $S_1 \times S_2 \times \dots \times S_n$ is the natural set of outcomes for the compound experiment that consists of performing the n experiments in sequence. In particular, if we have a basic experiment with S as the set of outcomes, then S^n is the natural set of outcomes for the compound experiment that consists of n replications of the basic experiment. Similarly, if we have an infinite sequence of experiments and S_i is the set of outcomes for experiment $i \in \mathbb{N}_+$, then $S_1 \times S_2 \times \dots$ is the natural set of outcomes for the compound experiment that consists of performing the given experiments in sequence. In particular, the set of outcomes for the compound experiment that consists of indefinite replications of a basic experiment is $S^\infty = S \times S \times \dots$. This is an essential special case, because (classical) probability theory is based on the idea of replicating a given experiment.

Events

Consider again a random experiment with S as the set of outcomes. Certain subsets of S are referred to as *events*. Suppose that $A \subseteq S$ is a given event, and that the experiment is run, resulting in outcome $s \in S$.

1. If $s \in A$ then we say that A occurs.
2. If $s \notin A$ then we say that A does not occur.

Intuitively, you should think of an event as a meaningful *statement* about the experiment: every such statement translates into an event, namely the set of outcomes for which the statement is true. In particular, S itself is an event; by definition it *always* occurs. At the other extreme, the empty set \emptyset is also an event; by definition it *never* occurs.

For a note on terminology, recall that a *mathematical space* consists of a set together with other mathematical structures defined on the set. An example you may be familiar with is a *vector space*, which consists of a set (the vectors) together with the operations of addition and scalar multiplication. In probability theory, many authors use the term *sample space* for the set of outcomes of a random experiment, but here is the more careful definition:

The *sample space* of an experiment is (S, \mathcal{S}) where S is the set of outcomes and \mathcal{S} is the collection of events.

Details

Sometimes not every subset of S can be allowed as an event, but the collection of events \mathcal{S} is required to be a σ -algebra, so that the sample space (S, \mathcal{S}) is a measurable space. The axioms of a σ -algebra ensure that new sets that are constructed in a

reasonable way from given events, using the set operations, are themselves valid events. Most of the sample spaces that occur in elementary probability fall into two general categories.

1. *Discrete*: S is countable and $\mathcal{S} = \mathcal{P}(S)$ is the collection of all subsets of S . In this case, the sample space (S, \mathcal{S}) is *discrete*.
2. *Euclidean*: S is a measurable subset of \mathbb{R}^n for some $n \in \mathbb{N}_+$ and \mathcal{S} is the collection of measurable subsets of S .

In (b), the measurable subsets of \mathbb{R}^n include all of the sets encountered in calculus and in standard applications of probability theory, and many more besides. Nonetheless, for technical reasons, certain very weird subsets must be excluded. Typically S is a set defined by a finite number of inequalities involving elementary functions.

The Algebra of Events

The standard algebra of sets leads to a grammar for discussing random experiments and allows us to construct new events from given events. In the following results, suppose that S is the set of outcomes of a random experiment, and that A and B are events.

$A \subseteq B$ if and only if the occurrence of A *implies* the occurrence of B .

Proof

Recall that \subseteq is the subset relation. So by definition, $A \subseteq B$ means that $s \in A$ implies $s \in B$.

$A \cup B$ is the event that occurs if and only if A occurs *or* B occurs.

Proof

Recall that $A \cup B$ is the *union* of A and B . So by definition, $s \in A \cup B$ if and only if $s \in A$ or $s \in B$.

$A \cap B$ is the event that occurs if and only if A occurs *and* B occurs.

Proof

Recall that $A \cap B$ is the *intersection* of A and B . So by definition, $s \in A \cap B$ if and only if $s \in A$ and $s \in B$.

A and B are disjoint if and only if they are *mutually exclusive*; they cannot both occur on the same run of the experiment.

Proof

By definition, A and B *disjoint* means that $A \cap B = \emptyset$.

A^c is the event that occurs if and only if A does *not* occur.

Proof

Recall that A^c is the *complement* of A , so $s \in A^c$ if and only if $s \notin A$.

$A \setminus B$ is the event that occurs if and only if A occurs *and* B does *not* occur.

Proof

Recall that $A \setminus B = A \cap B^c$. Hence $s \in A \setminus B$ if and only if $s \in A$ and $s \notin B$.

$(A \cap B^c) \cup (B \cap A^c)$ is the event that occurs if and only if *one but not both* of the given events occurs.

Proof

The events in the union are disjoint. So for s is in the given event if and only if either $s \in A$ and $s \notin B$, or $s \in B$ and $s \notin A$.

Recall that the event in (10) is the *symmetric difference* of A and B , and is sometimes denoted $A \Delta B$. This event corresponds to *exclusive or*, as opposed to the ordinary union $A \cup B$ which corresponds to *inclusive or*.

$(A \cap B) \cup (A^c \cap B^c)$ is the event that occurs if and only if *both or neither* of the given events occurs.

Proof

The events in the union are disjoint. Thus s is in the given event if and only if either $s \in A$ and $s \in B$, or $s \notin A$ and $s \notin B$.

In the Venn diagram app, observe the diagram of each of the 16 events that can be constructed from A and B .

Suppose now that $\mathcal{A} = \{A_i : i \in I\}$ is a collection of events for the random experiment, where I is a countable index set.

$\bigcup_{i \in I} A_i$ is the event that occurs if and only if *at least one* event in the collection occurs.

Proof

Note that $s \in \bigcup_{i \in I} A_i$ if and only if $s \in A_i$ for some $i \in I$.

$\bigcap_{i \in I} A_i$ is the event that occurs if and only if *every* event in the collection occurs:

Proof

Note that $s \in \bigcap_{i \in I} A_i$ if and only if $s \in A_i$ for every $i \in I$.

\mathcal{A} is a pairwise disjoint collection if and only if the events are *mutually exclusive*; at most one of the events could occur on a given run of the experiment.

Proof

By definition, $A_i \cap A_j = \emptyset$ for distinct $i, j \in I$.

Suppose now that (A_1, A_2, \dots) is an infinite sequence of events.

$\bigcup_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i$ is the event that occurs if and only if *infinitely many* of the given events occur. This event is sometimes called the *limit superior* of (A_1, A_2, \dots) .

Proof

Note that s is in the given event if and only if for every $n \in \mathbb{N}_+$ there exists $i \in \mathbb{N}_+$ with $i \geq n$ such that $s \in A_i$. In turn this means that $s \in A_i$ for infinitely many $i \in I$.

$\bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i$ is the event that occurs if and only if *all but finitely many* of the given events occur. This event is sometimes called the *limit inferior* of (A_1, A_2, \dots) .

Proof

Note that s is in the given event if and only if there exists $n \in \mathbb{N}_+$ such that $s \in A_i$ for every $i \in \mathbb{N}_+$ with $i \geq n$. In turn, this means that $s \in A_i$ for all but finitely many $i \in I$.

Limit superiors and inferiors are discussed in more detail in the section on convergence.

Random Variables

Intuitively, a *random variable* is a measurement of interest in the context of the experiment. Simple examples include the number of heads when a coin is tossed several times, the sum of the scores when a pair of dice are thrown, the lifetime of a device subject to random stress, the weight of a person chosen from a population. Many more examples are given below in the [exercises below](#). Mathematically, a random variable is a function defined on the set of outcomes.

A function X from S into a set T is a *random variable* for the experiment with values in T .

Details

The set T will also come with a σ -algebra \mathcal{T} of admissible subsets, so that (T, \mathcal{T}) is a measurable space, just like (S, \mathcal{S}) . The function X is required to be measurable, an assumption which ensures that meaningful statements involving X define events. In the discussion below, all subsets of T are assumed to be in \mathcal{T} .

Probability has its own notation, very different from other branches of mathematics. As a case in point, random variables, even though they are functions, are usually denoted by capital letters near the end of the alphabet. The use of a letter near the end of the alphabet is intended to emphasize the idea that the object is a *variable* in the context of the experiment. The use of a capital letter is intended to emphasize the fact that it is not an ordinary *algebraic* variable to which we can assign a specific value, but rather a

random variable whose value is indeterminate until we run the experiment. Specifically, when we run the experiment an outcome $s \in S$ occurs, and random variable X takes the value $X(s) \in T$.

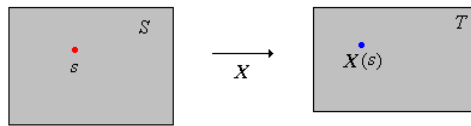


Figure 2.2.1: A random variable as a function defined on the set of outcomes.

If $B \subseteq T$, we use the notation $\{X \in B\}$ for the inverse image $\{s \in S : X(s) \in B\}$, rather than $X^{-1}(B)$. Again, the notation is more natural since we think of X as a variable in the experiment. Think of $\{X \in B\}$ as a *statement* about X , which then translates into the event $\{s \in S : X(s) \in B\}$

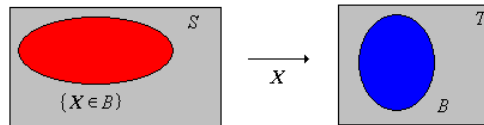


Figure 2.2.2: The event $\{X \in B\}$ corresponding to $B \subseteq T$

Again, every statement about a random variable X with values in T translates into an inverse image of the form $\{X \in B\}$ for some $B \in \mathcal{T}$. So, for example, if $x \in T$ then $\{X = x\} = \{X \in \{x\}\} = \{s \in S : X(s) = x\}$. If X is a real-valued random variable and $a, b \in \mathbb{R}$ with $a < b$ then $\{a \leq X \leq b\} = \{X \in [a, b]\} = \{s \in S : a \leq X(s) \leq b\}$.

Suppose that X is a random variable taking values in T , and that $A, B \subseteq T$. Then

1. $\{X \in A \cup B\} = \{X \in A\} \cup \{X \in B\}$
2. $\{X \in A \cap B\} = \{X \in A\} \cap \{X \in B\}$
3. $\{X \in A \setminus B\} = \{X \in A\} \setminus \{X \in B\}$
4. $A \subseteq B \implies \{X \in A\} \subseteq \{X \in B\}$
5. If A and B are disjoint, then so are $\{X \in A\}$ and $\{X \in B\}$.

Proof

This is a restatement of the fact that inverse images of a function preserve the set operations; only the notation changes (and is simpler).

1. $s \in \{X \in A \cup B\}$ if and only if $X(s) \in A \cup B$ if and only if $X(s) \in A$ or $X(s) \in B$ if and only if $s \in \{X \in A\}$ or $s \in \{X \in B\}$ if and only if $s \in \{X \in A\} \cup \{X \in B\}$.
2. The proof is exactly the same as (a), with *and* replacing *or*.
3. The proof is also exactly the same as (a), with *but not* replacing *or*.
4. If $s \in \{X \in A\}$ then $X(s) \in A$ so $X(s) \in B$ and hence $s \in \{X \in B\}$.
5. This follows from part (b).

As with a general function, the result in part (a) holds for the union of a countable collection of subsets, and the result in part (b) holds for the intersection of a countable collection of subsets. No new ideas are involved; only the notation is more complicated.

Often, a random variable takes values in a subset T of \mathbb{R}^k for some $k \in \mathbb{N}_+$. We might express such a random variable as $\mathbf{X} = (X_1, X_2, \dots, X_k)$ where X_i is a real-valued random variable for each $i \in \{1, 2, \dots, k\}$. In this case, we usually refer to \mathbf{X} as a *random vector*, to emphasize its higher-dimensional character. A random variable can have an even more complicated structure. For example, if the experiment is to select n objects from a population and record various real measurements for each object, then the outcome of the experiment is a vector of vectors: $\mathbf{X} = (X_1, X_2, \dots, X_n)$ where X_i is the vector of measurements for the i th object. There are other possibilities; a random variable could be an infinite sequence, or could be set-valued. Specific examples are given in the computational exercises below. However, the important point is simply that a random variable is a function defined on the set of outcomes S .

The outcome of the experiment itself can be thought of as a random variable. Specifically, let $T = S$ and let X denote the identity function on S so that $X(s) = s$ for $s \in S$. Then trivially X is a random variable, and the events that can be defined in terms of X

are simply the original events of the experiment. That is, if A is an event then $\{X \in A\} = A$. Conversely, every random variable effectively defines a new random experiment.

In the general setting above, a random variable X defines a new random experiment with T as the new set of outcomes and subsets of T as the new collection of events.

Details

Technically, the σ -algebra \mathcal{T} would be the new collection of events.

In fact, often a random experiment is modeled by specifying the random variables of interest, in the language of the experiment. Then, a mathematical definition of the random variables specifies the sample space. A function (or transformation) of a random variable defines a new random variable.

Suppose that X is a random variable for the experiment with values in T and that g is a function from T into another set U . Then $Y = g(X)$ is a random variable with values in U .

Details

Technically, T and U both come with σ -algebras of admissible subsets \mathcal{T} and \mathcal{U} , respectively. The function g , just like the function X , is required to be measurable. This assumption ensures that $Y = g(X)$ is a measurable function from S into U , and hence is a valid random variable.

Note that, as functions, $g(X) = g \circ X$, the composition of g with X . But again, thinking of X and Y as *variables* in the context of the experiment, the notation $Y = g(X)$ is much more natural.

Indicator Variables

For an event A , the indicator function of A is called the *indicator variable* of A .

The value of this random variables tells us whether or not A has occurred:

$$\mathbf{1}_A = \begin{cases} 1, & A \text{ occurs} \\ 0, & A \text{ does not occur} \end{cases} \quad (2.2.1)$$

That is, as a function on S ,

$$\mathbf{1}_A(s) = \begin{cases} 1, & s \in A \\ 0, & s \notin A \end{cases} \quad (2.2.2)$$

If X is a random variable that takes values 0 and 1, then X is the indicator variable of the event $\{X = 1\}$.

Proof

Note that for $s \in S$, $X(s) = 1$ if $s \in \{X = 1\}$ and $X(s) = 0$ otherwise.

Recall also that the set algebra of events translates into the arithmetic algebra of indicator variables.

Suppose that A and B are events.

1. $\mathbf{1}_{A \cap B} = \mathbf{1}_A \mathbf{1}_B = \min \{\mathbf{1}_A, \mathbf{1}_B\}$
2. $\mathbf{1}_{A \cup B} = 1 - (1 - \mathbf{1}_A)(1 - \mathbf{1}_B) = \max \{\mathbf{1}_A, \mathbf{1}_B\}$
3. $\mathbf{1}_{B \setminus A} = \mathbf{1}_B (1 - \mathbf{1}_A)$
4. $\mathbf{1}_{A^c} = 1 - \mathbf{1}_A$
5. $A \subseteq B$ if and only if $\mathbf{1}_A \leq \mathbf{1}_B$

The results in part (a) extends to arbitrary intersections and the results in part (b) extends to arbitrary unions. If the event A has a complicated description, sometimes we use $\mathbf{1}(A)$ for the indicator variable rather than $\mathbf{1}_A$.

Examples and Applications

Recall that probability theory is often illustrated using simple devices from games of chance: coins, dice, cards, 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, remember that probability is not 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

The basic *coin experiment* consists of tossing a coin n times and recording the sequence of scores (X_1, X_2, \dots, X_n) (where 1 denotes heads and 0 denotes tails). This experiment is a generic example of n Bernoulli trials, named for Jacob Bernoulli.

Consider the coin experiment with $n = 4$, and Let Y denote the number of heads.

1. Give the set of outcomes S in list form.
2. Give the event $\{Y = k\}$ in list form for each $k \in \{0, 1, 2, 3, 4\}$.

Answer

To simplify the notation, we represent outcomes a *bit strings* rather than ordered sequences.

1. $S = \{1111, 1110, 1101, 1011, 0111, 1100, 1010, 1001, 0110, 0101, 0011, 1000, 0100, 0010, 0001, 0000\}$
2.

$\{Y = 0\} = \{0000\}$	(2.2.3)
$\{Y = 1\} = \{1000, 0100, 0010, 0001\}$	(2.2.4)
$\{Y = 2\} = \{1100, 1010, 1001, 0110, 0101, 0011\}$	(2.2.5)
$\{Y = 3\} = \{1110, 1101, 1011, 0111\}$	(2.2.6)
$\{Y = 4\} = \{1111\}$	(2.2.7)

In the simulation of the coin experiment, set $n = 4$. Run the experiment 100 times and count the number of times that the event $\{Y = 2\}$ occurs.

Now consider the general coin experiment with the coin tossed n times, and let Y denote the number of heads.

1. Give the set of outcomes S in Cartesian product form, and give the cardinality of S .
2. Express Y as a function on S .
3. Find $\#\{Y = k\}$ (as a subset of S) for $k \in \{0, 1, \dots, n\}$

Answer

1. $S = \{0, 1\}^n$ and $\#(S) = 2^n$.
2. $Y(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_n$. The set of possible values is $\{0, 1, \dots, n\}$
3. $\#\{Y = k\} = \binom{n}{k}$

The basic *dice experiment* consists of throwing n distinct k -sided dice (with faces numbered from 1 to k) and recording the sequence of scores (X_1, X_2, \dots, X_n) . This experiment is a generic example of n multinomial trials. The special case $k = 6$ corresponds to *standard dice*.

Consider the dice experiment with $n = 2$ standard dice. Let S denote the set of outcomes, A the event that the first die score is 1, and B the event that the sum of the scores is 7. Give each of the following events in the form indicated:

1. S in Cartesian product form
2. A in list form
3. B in list form
4. $A \cup B$ in list form
5. $A \cap B$ in list form
6. $A^c \cap B^c$ in predicate form

Answer

1. $S = \{1, 2, 3, 4, 5, 6\}^2$
2. $A = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)\}$

3. $B = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$
4. $A \cup B = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$
5. $A \cap B = \{(1, 6)\}$
6. $A^c \cap B^c = \{(x, y) \in S : x + y \neq 7 \text{ and } x \neq 1\}$

In the simulation of the dice experiment, set $n = 2$. Run the experiment 100 times and count the number of times each event in the previous exercise occurs.

Consider the dice experiment with $n = 2$ standard dice, and let S denote the set of outcomes, Y the sum of the scores, U the minimum score, and V the maximum score.

1. Express Y as a function on S and give the set of possible values in list form.
2. Express U as a function on S and give the set of possible values in list form.
3. Express V as a function on the S and give the set of possible values in list form.
4. Give the set of possible values of (U, V) in predicate form

Answer

Note that $S = \{1, 2, 3, 4, 5, 6\}^2$. The following functions are defined on S .

1. $Y(x_1, x_2) = x_1 + x_2$. The set of values is $\{2, 3, \dots, 12\}$
2. $U(x_1, x_2) = \min\{x_1, x_2\}$. The set of values is $\{1, 2, \dots, 6\}$
3. $V(x_1, x_2) = \max\{x_1, x_2\}$. The set of values is $\{1, 2, \dots, 6\}$
4. $\{(u, v) \in \{1, 2, 3, 4, 5, 6\}^2 : u \leq v\}$

Consider again the dice experiment with $n = 2$ standard dice, and let S denote the set of outcomes, Y the sum of the scores, U the minimum score, and V the maximum score. Give each of the following as subsets of S , in list form.

1. $\{X_1 < 3, X_2 > 4\}$
2. $\{Y = 7\}$
3. $\{U = 2\}$
4. $\{V = 4\}$
5. $\{U = V\}$

Answer

1. $\{(1, 5), (2, 5), (1, 6), (2, 6)\}$
2. $\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$
3. $\{(2, 2), (2, 3), (3, 2), (2, 4), (4, 2), (2, 5), (5, 2), (2, 6), (6, 2)\}$
4. $\{(4, 1), (1, 4), (2, 4), (4, 2), (4, 3), (3, 4), (4, 4)\}$
5. $\{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$

In the dice experiment, set $n = 2$. Run the experiment 100 times. Count the number of times each event in the previous exercise occurred.

In the general dice experiment with n distinct k -sided dice, let Y denote the sum of the scores, U the minimum score, and V the maximum score.

1. Give the set of outcomes S and find $\#(S)$.
2. Express Y as a function on S , and give the set of possible values in list form.
3. Express U as a function on S , and give the set of possible values in list form.
4. Express V as a function on S , and give the set of possible values in list form.
5. Give the set of possible values of (U, V) in predicate form.

Answer

1. $S = \{1, 2, \dots, k\}^n$ and $\#(S) = k^n$
2. $Y(x_1, x_2, \dots, x_n) = x_1 + x_2 + \dots + x_n$. The set of possible values is $\{n, n+1, \dots, nk\}$
3. $U(x_1, x_2, \dots, x_n) = \min\{x_1, x_2, \dots, x_n\}$. The set of possible values is $\{1, 2, \dots, k\}$

4. $V(x_1, x_2, \dots, x_n) = \max\{x_1, x_2, \dots, x_n\}$. The set of possible values is $\{1, 2, \dots, k\}$
5. $\{(u, v) \in \{1, 2, \dots, k\}^2 : u \leq v\}$

The set of outcomes of a random experiment depends of course on what information is recorded. The following exercise is an illustration.

An experiment consists of throwing a pair of standard dice repeatedly until the sum of the two scores is either 5 or 7. Let A denote the event that the sum is 5 rather than 7 on the final throw. Experiments of this type arise in the casino game craps.

1. Suppose that the pair of scores on each throw is recorded. Define the set of outcomes of the experiment and describe A as a subset of this set.
2. Suppose that the pair of scores on the final throw is recorded. Define the set of outcomes of the experiment and describe A as a subset of this set.

Answer

Let $D_5 = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$ $D_7 = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$ $D = D_5 \cup D_7$, and $C = D^c$

1. $S = D \cup (C \times D) \cup (C^2 \times D) \cup \dots$, $A = D_5 \cup (C \times D_5) \cup (C^2 \times D_5) \cup \dots$
2. $S = D$, $A = D_5$

Suppose that 3 standard dice are rolled and the sequence of scores (X_1, X_2, X_3) is recorded. A person pays \$1 to play. If some of the dice come up 6, then the player receives her \$1 back, plus \$1 for each 6. Otherwise she loses her \$1. Let W denote the person's net winnings. This is the game of chuck-a-luck and is treated in more detail in the chapter on Games of Chance.

1. Give the set of outcomes S in Cartesian product form.
2. Express W as a function on S and give the set of possible values in list form.

Answer

1. $S = \{1, 2, 3, 4, 5, 6\}^3$
2. $W(x_1, x_2, x_3) = \mathbf{1}(x_1 = 6) + \mathbf{1}(x_2 = 6) + \mathbf{1}(x_3 = 6) - \mathbf{1}(x_1 \neq 6, x_2 \neq 6, x_3 \neq 6)$. The set of possible values is $\{-1, 1, 2, 3\}$

Play the chuck-a-luck experiment a few times and see how you do.

In the *die-coin experiment*, a standard die is rolled and then a coin is tossed the number of times shown on the die. The sequence of coin scores \mathbf{X} is recorded (0 for tails, 1 for heads). Let N denote the die score and Y the number of heads.

1. Give the set of outcomes S in terms of Cartesian powers and find $\#(S)$.
2. Express N as a function on S and give the set of possible values in list form.
3. Express Y as a function on S and give the set of possible values in list form.
4. Give the event A that all tosses result in heads in list form.

Answer

1. $S = \bigcup_{n=1}^6 \{0, 1\}^n$, $\#(S) = 126$
2. $N(x_1, x_2, \dots, x_n) = n$ for $(x_1, x_2, \dots, x_n) \in S$. The set of values is $\{1, 2, 3, 4, 5, 6\}$
3. $Y(x_1, x_2, \dots, x_n) = \sum_{i=1}^n x_i$ for $(x_1, x_2, \dots, x_n) \in S$. The set of possible values is $\{0, 1, 2, 3, 4, 5, 6\}$
4. $A = \{1, 11, 111, 1111, 11111, 111111\}$

Run the simulation of the die-coin experiment 10 times. For each run, give the values of the random variables \mathbf{X} , N , and Y of the previous exercise. Count the number of times the event A occurs.

In the *coin-die experiment*, we have a coin and two distinct dice, say one red and one green. First the coin is tossed, and then if the result is heads the red die is thrown, while if the result is tails the green die is thrown. The coin score X and the score of the chosen die Y are recorded. Suppose now that the red die is a standard 6-sided die, and the green die a 4-sided die.

1. Give the set of outcomes S in list form.
2. Express X as a function on S .

- Express Y as a function on S .
- Give the event $\{Y \geq 3\}$ as a subset of S in list form.

Answer

- $\{(0, 1), (0, 2), (0, 3), (0, 4), (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)\}$
- $X(i, j) = i$ for $(i, j) \in S$
- $Y(i, j) = j$ for $(i, j) \in S$
- $\{(0, 3), (0, 4), (1, 3), (1, 4), (1, 5), (1, 6)\}$

Run the coin-die experiment 100 times, with various types of dice.

Sampling Models

Recall that many random experiments can be thought of as *sampling experiments*. For the general finite sampling model, we start with a population D with m (distinct) objects. We select a sample of n objects from the population. If the sampling is done in a random way, then we have a random experiment with the sample as the basic outcome. Thus, the set of outcomes of the experiment is literally the *set of samples*; this is the historical origin of the term *sample space*. There are four common types of sampling from a finite population, based on the criteria of *order* and *replacement*. Recall the following facts from the section on combinatorial structures:

Samples of size n chosen from a population with m elements.

- If the sampling is with replacement and with regard to order, then the set of samples is the Cartesian power D^n . The number of samples is m^n .
- If the sampling is without replacement and with regard to order, then the set of samples is the set of all *permutations* of size n from D . The number of samples is $m^{(n)} = m(m-1) \cdots [m-(n-1)]$.
- If the sampling is without replacement and without regard to order, then the set of samples is the set of all *combinations* (or *subsets*) of size n from D . The number of samples is $\binom{m}{n}$.
- If the sampling is with replacement and without regard to order, then the set of samples is the set of all *multisets* of size n from D . The number of samples is $\binom{m+n-1}{n}$.

If we sample with replacement, the sample size n can be any positive integer. If we sample without replacement, the sample size cannot exceed the population size, so we must have $n \in \{1, 2, \dots, m\}$.

The basic [coin and dice experiments](#) are examples of sampling *with* replacement. If we toss a coin n times and record the sequence of scores (where as usual, 0 denotes tails and 1 denotes heads), then we generate an ordered sample of size n with replacement from the population $\{0, 1\}$. If we throw n (distinct) standard dice and record the sequence of scores, then we generate an ordered sample of size n with replacement from the population $\{1, 2, 3, 4, 5, 6\}$.

Suppose that the sampling is without replacement (the most common case). If we record the *ordered* sample $\mathbf{X} = (X_1, X_2, \dots, X_n)$, then the *unordered* sample $\mathbf{W} = \{X_1, X_2, \dots, X_n\}$ is a random variable (that is, a function of \mathbf{X}). On the other hand, if we just record the unordered sample \mathbf{W} in the first place, then we cannot recover the ordered sample. Note also that the number of ordered samples of size n is simply $n!$ times the number of unordered samples of size n . No such simple relationship exists when the sampling is with replacement. This will turn out to be an important point when we study probability models based on random samples, in the next section.

Consider a sample of size $n = 3$ chosen without replacement from the population $\{a, b, c, d, e\}$.

- Give T , the set of unordered samples in list form.
- Give in list form the set of all ordered samples that correspond to the unordered sample $\{b, c, e\}$.
- Note that for every unordered sample, there are 6 ordered samples.
- Give the cardinality of S , the set of ordered samples.

Answer

- $T = \{\{a, b, c\}, \{a, b, d\}, \{a, b, e\}, \{a, c, d\}, \{a, c, e\}, \{a, d, e\}, \{b, c, d\}, \{b, c, e\}, \{b, d, e\}, \{c, d, e\}\}$
- $\{(b, c, e), (b, e, c), (c, b, e), (c, e, b), (e, b, c), (e, c, b)\}$
- 60

Traditionally in probability theory, an *urn* containing *balls* is often used as a metaphor for a finite population.

Suppose that an urn contains 50 (distinct) balls. A sample of 10 balls is selected from the urn. Find the number of samples in each of the following cases:

1. Ordered samples with replacement
2. Ordered samples without replacement
3. Unordered samples without replacement
4. Unordered samples with replacement

Answer

1. 97 656 250 000 000 000
2. 37 276 043 023 296 000
3. 10 272 278 170
4. 62 828 356 305

Suppose again that we have a population D with m (distinct) objects, but suppose now that each object is one of two types—either type 1 or type 0. Such populations are said to be *dichotomous*. Here are some specific examples:

- The population consists of persons, each either *male* or *female*.
- The population consists of voters, each either *democrat* or *republican*.
- The population consists of devices, each either *good* or *defective*.
- The population consists of balls, each either *red* or *green*.

Suppose that the population D has r type 1 objects and hence $m - r$ type 0 objects. Of course, we must have $r \in \{0, 1, \dots, m\}$. Now suppose that we select a sample of size n without replacement from the population. Note that this model has three parameters: the population size m , the number of type 1 objects in the population r , and the sample size n .

Let Y denote the number of type 1 objects in the sample. Then

1. $\#\{Y = k\} = \binom{n}{k} r^{(k)} (m - r)^{(n-k)}$ for each $k \in \{0, 1, \dots, n\}$, if the event is considered as a subset of S , the set of ordered samples.
2. $\#\{Y = k\} = \binom{r}{k} \binom{m-r}{n-k}$ for each $k \in \{0, 1, \dots, n\}$, if the event is considered as a subset of T , the set of unordered samples.
3. The expression in (a) is $n!$ times the expression in (b).

Proof

1. $\binom{n}{k}$ is the number of ways to pick the coordinates (in the ordered sample) where the type 1 objects will go, $r^{(k)}$ is the number of ways to select a permutation of k type 1 objects, and $(m - r)^{(n-k)}$ is the number of ways to select a permutation of $n - k$ type 0 objects. The result follows from the multiplication principle.
2. $\binom{r}{k}$ is the number of ways to select a combination of k type 1 objects and $\binom{m-r}{n-k}$ is the number of ways to select a combination of $n - k$ type 0 objects. The result again follows from the multiplication principle.
3. This result can be shown algebraically, but a combinatorial argument is better. For every combination of size n there are $n!$ permutations of those objects.

A batch of 50 components consists of 40 good components and 10 defective components. A sample of 5 components is selected, without replacement. Let Y denote the number of defectives in the sample.

1. Let S denote the set of ordered samples. Find $\#(S)$.
2. Let T denote the set of unordered samples. Find $\#(T)$.
3. As a subset of T , find $\#\{Y = k\}$ for each $k \in \{0, 1, 2, 3, 4, 5\}$.

Answer

1. 254 251 200
2. 2 118 760
3. $\#\{Y = 0\} = 658 008$ $\#\{Y = 1\} = 913 900$ $\#\{Y = 2\} = 444 600$ $\#\{Y = 3\} = 93 600$ $\#\{Y = 4\} = 8 400$,
 $\#\{Y = 5\} = 252$

Run the simulation of the ball and urn experiment 100 times for the parameter values in the last exercise: $m = 50$, $r = 10$, $n = 5$. Note the values of the random variable Y .

Cards

Recall that a standard *card deck* can be modeled by the Cartesian product set

$$D = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, j, q, k\} \times \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\} \quad (2.2.8)$$

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).

Most card games involve **sampling** without replacement from the deck D , which plays the role of the population. Thus, the basic *card experiment* consists of dealing n cards from a standard deck without replacement; in this special context, the sample of cards is often referred to as a *hand*. Just as in the general sampling model, if we record the *ordered* hand $\mathbf{X} = (X_1, X_2, \dots, X_n)$, then the *unordered* hand $\mathbf{W} = \{X_1, X_2, \dots, X_n\}$ is a random variable (that is, a function of \mathbf{X}). On the other hand, if we just record the unordered hand \mathbf{W} in the first place, then we cannot recover the ordered hand. Finally, recall that $n = 5$ is the *poker experiment* and $n = 13$ is the *bridge experiment*. The game of poker is treated in more detail in the chapter on Games of Chance.

Suppose that a single card is dealt from a standard deck. Let Q denote the event that the card is a queen and H the event that the card is a heart. Give each of the following events in list form:

1. Q
2. H
3. $Q \cup H$
4. $Q \cap H$
5. $Q \setminus H$

Answer

1. $Q = \{q\clubsuit, q\diamondsuit, q\heartsuit, q\spadesuit\}$
2. $H = \{1\heartsuit, 2\heartsuit, \dots, 10\heartsuit, j\heartsuit, q\heartsuit, k\heartsuit\}$
3. $Q \cup H = \{1\heartsuit, 2\heartsuit, \dots, 10\heartsuit, j\heartsuit, q\heartsuit, k\heartsuit, q\clubsuit, q\diamondsuit, q\spadesuit\}$
4. $Q \cap H = \{q\heartsuit\}$
5. $Q \setminus H = \{q\clubsuit, q\diamondsuit, q\spadesuit\}$

In the card experiment, set $n = 1$. Run the experiment 100 times and count the number of times each event in the previous exercise occurs.

Suppose that two cards are dealt from a standard deck and the sequence of cards recorded. Let S denote the set of outcomes, and let Q_i denote the event that the i th card is a queen and H_i the event that the i th card is a heart for $i \in \{1, 2\}$. Find the number of outcomes in each of the following events:

1. S
2. H_1
3. H_2
4. $H_1 \cap H_2$
5. $Q_1 \cap H_1$
6. $Q_1 \cap H_2$
7. $H_1 \cup H_2$

Answer

1. 2652
2. 663
3. 663
4. 156
5. 51

6. 51
7. 1170

Consider the general card experiment in which n cards are dealt from a standard deck, and the ordered hand \mathbf{X} is recorded.

1. Give cardinality of S , the set of values of the ordered hand \mathbf{X} .
2. Give the cardinality of T , the set of values of the unordered hand \mathbf{W} .
3. How many ordered hands correspond to a given unordered hand?
4. Explicitly compute the numbers in (a) and (b) when $n = 5$ (poker).
5. Explicitly compute the numbers in (a) and (b) when $n = 13$ (bridge).

Answer

3. $\#(S) = 52^{(n)}$
4. $\#(T) = \binom{52}{n}$
5. $n!$
6. 311 875 2002 598 960
7. 3 954 242 643 911 239 680 000 501 359 600

Consider the bridge experiment of dealing 13 cards from a deck and recording the unordered hand. In the most common *point counting system*, an ace is worth 4 points, a king 3 points, a queen 2 points, and a jack 1 point. The other cards are worth 0 points. Let S denote the set of outcomes of the experiment and V the point value of the hand.

1. Find the set of possible values of V .
2. Find the cardinality of the event $\{V = 0\}$ as a subset of S .

Answer

1. $\{0, 1, \dots, 37\}$
2. $\#\{V = 0\} = 2\,310\,789\,600$

In the card experiment, set $n = 13$ and run the experiment 100 times. For each run, compute the value of each of the random variable V in the previous exercise.

Consider the poker experiment of dealing 5 cards from a deck. Find the cardinality of each of the events below, as a subset of the set of unordered hands.

1. A : the event that the hand is a *full house* (3 cards of one kind and 2 of another kind).
2. B : the event that the hand has *4 of a kind* (4 cards of one kind and 1 of another kind).
3. C : the event that all cards in the hand are in the same suit (the hand is a *flush* or a *straight flush*).

Answer

1. $\#(A) = 3744$
2. $\#(B) = 624$
3. $\#(C) = 5148$

Run the poker experiment 1000 times. Note the number of times that the events A , B , and C in the previous exercise occurred.

Consider the bridge experiment of dealing 13 cards from a standard deck. Let S denote the set of unordered hands, Y the number of hearts in the hand, and Z the number of queens in the hand.

1. Find the cardinality of the event $\{Y = y\}$ as a subset of S for each $y \in \{0, 1, \dots, 13\}$.
2. Find the cardinality of the event $\{Z = z\}$ as a subset of S for each $z \in \{0, 1, 2, 3, 4\}$.

Answer

1. $\#(Y = y) = \binom{13}{y} \binom{39}{13-y}$ for $y \in \{0, 1, \dots, 13\}$
2. $\#(Z = z) = \binom{4}{z} \binom{48}{4-z}$ for $z \in \{0, 1, 2, 3, 4\}$

Geometric Models

In the experiments that we have considered so far, the sample spaces have all been discrete (so that the set of outcomes is finite or countably infinite). In this subsection, we consider Euclidean sample spaces where the set of outcomes S is *continuous* in a sense that we will make clear later. The experiments we consider are sometimes referred to as *geometric models* because they involve selecting a point *at random* from a Euclidean set.

We first consider *Buffon's coin experiment*, which 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 in which the coin lands. Buffon's experiments are studied in more detail in the chapter on Geometric Models and are named for Comte de Buffon

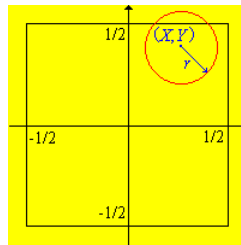


Figure 2.2.3: Buffon's coin experiment

In Buffon's coin experiment, let S denote the set of outcomes, A the event that the coin does not touch the sides of the square, and let Z denote the distance from the center of the coin to the center of the square.

1. Describe S as a Cartesian product.
2. Describe A as a subset of S .
3. Describe A^c as a subset of S .
4. Express Z as a function on S .
5. Express the event $\{X < Y\}$ as a subset of S .
6. Express the event $\{Z \leq \frac{1}{2}\}$ as a subset of S .

Answer

1. $S = [-\frac{1}{2}, \frac{1}{2}]^2$
2. $A = [r - \frac{1}{2}, \frac{1}{2} - r]^2$
3. $A^c = \{(x, y) \in S : x < r - \frac{1}{2} \text{ or } x > \frac{1}{2} - r \text{ or } y < r - \frac{1}{2} \text{ or } y > \frac{1}{2} - r\}$
4. $Z(x, y) = \sqrt{x^2 + y^2}$ for $(x, y) \in S$
5. $\{X < Y\} = \{(x, y) \in S : x < y\}$
6. $\{Z < \frac{1}{2}\} = \{(x, y) \in S : x^2 + y^2 < \frac{1}{4}\}$

Run Buffon's coin experiment 100 times with $r = 0.2$. For each run, note whether event A occurs and compute the value of random variable Z .

A point (X, Y) is chosen at random in the circular region of radius 1 in \mathbb{R}^2 centered at the origin. Let S denote the set of outcomes. Let A denote the event that the point is in the inscribed square region centered at the origin, with sides parallel to the coordinate axes. Let B denote the event that the point is in the inscribed square with vertices $(\pm 1, 0), (0, \pm 1)$.

1. Describe S mathematically and sketch the set.
2. Describe A mathematically and sketch the set.
3. Describe B mathematically and sketch the set.
4. Sketch $A \cup B$
5. Sketch $A \cap B$
6. Sketch $A \cap B^c$

Answer

1. $S = \{(x, y) : x^2 + y^2 \leq 1\}$

2. $A = \left\{ (x, y) : -\frac{1}{\sqrt{2}} \leq x \leq \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \leq y \leq \frac{1}{\sqrt{2}} \right\}$
3. $B = \{(x, y) \in S : -1 \leq |x + y| \leq 1, -1 \leq |y - x| \leq 1\}$

Reliability

In the simple model of *structural reliability*, a system is composed of n components, each of which is either *working* or *failed*. The state of component i is an indicator random variable X_i , where 1 means working and 0 means failure. Thus, $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is a vector of indicator random variables that specifies the states of all of the components, and therefore the set of outcomes of the experiment is $S = \{0, 1\}^n$. The system as a whole is also either working or failed, depending only on the states of the components and how the components are connected together. Thus, the state of the system is also an indicator random variable and is a function of \mathbf{X} . The state of the system (working or failed) as a function of the states of the components is the *structure function*.

A *series system* is working if and only if each component is working. The state of the system is

$$U = X_1 X_2 \cdots X_n = \min \{X_1, X_2, \dots, X_n\} \quad (2.2.9)$$

A *parallel system* is working if and only if at least one component is working. The state of the system is

$$V = 1 - (1 - X_1)(1 - X_2) \cdots (1 - X_n) = \max \{X_1, X_2, \dots, X_n\} \quad (2.2.10)$$

More generally, a k out of n system is working if and only if at least k of the n components are working. Note that a parallel system is a 1 out of n system and a series system is an n out of n system. A k out of $2k$ system is a *majority rules system*.

The state of the k out of n system is $U_{n,k} = \mathbf{1}(\sum_{i=1}^n X_i \geq k)$. The structure function can also be expressed as a polynomial in the variables.

Explicitly give the state of the k out of 3 system, as a polynomial function of the component states (X_1, X_2, X_3) , for each $k \in \{1, 2, 3\}$.

Answer

1. $U_{3,1} = X_1 + X_2 + X_3 - X_1 X_2 - X_1 X_3 - X_2 X_3 + X_1 X_2 X_3$
2. $U_{3,2} = X_1 X_2 + X_1 X_3 + X_2 X_3 - 2 X_1 X_2 X_3$
3. $U_{3,3} = X_1 X_2 X_3$

In some cases, the system can be represented as a *graph* or *network*. The *edges* represent the components and the *vertices* the connections between the components. The system functions if and only if there is a working path between two designated vertices, which we will denote by a and b .

Find the state of the *Wheatstone bridge network* shown below, as a function of the component states. The network is named for Charles Wheatstone.

Answer

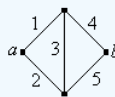


Figure 2.2.4: The Wheatstone bridge network

Not every function $u : \{0, 1\}^n \rightarrow \{0, 1\}$ makes sense as a structure function. Explain why the following properties might be desirable:

1. $u(0, 0, \dots, 0) = 0$ and $u(1, 1, \dots, 1) = 1$
2. u is an increasing function, where $\{0, 1\}$ is given the ordinary order and $\{0, 1\}^n$ the corresponding product order.
3. For each $i \in \{1, 2, \dots, n\}$, there exist \mathbf{x} and \mathbf{y} in $\{0, 1\}^n$ all of whose coordinates agree, except $x_i = 0$ and $y_i = 1$, and $u(\mathbf{x}) = 0$ while $u(\mathbf{y}) = 1$.

Answer

1. This means that if all components have failed then the system has failed, and if all components are working then the system is working.
2. This means that if a particular component is changed from failed to working, then the system may also go from failed to working, but not from working to failed. That is, the system can only improve.
3. This means that every component is relevant to the system, that is, there exists a configuration in which changing component i from failed to working changes the system from failed to working.

The model just discussed is a *static model*. We can extend it to a *dynamic model* by assuming that component i is initially working, but has a random *time to failure* T_i , taking values in $[0, \infty)$, for each $i \in \{1, 2, \dots, n\}$. Thus, the basic outcome of the experiment is the random vector of failure times (T_1, T_2, \dots, T_n) , and so the set of outcomes is $[0, \infty)^n$.

Consider the dynamic reliability model for a system with structure function u (valid in the sense of the previous exercise).

1. The state of component i at time $t \geq 0$ is $X_i(t) = \mathbf{1}(T_i > t)$.
2. The state of the system at time t is $X(t) = s[X_1(t), X_2(t), \dots, X_n(t)]$.
3. The time to failure of the system is $T = \min\{t \geq 0 : X(t) = 0\}$.

Suppose that we have two devices and that we record (X, Y) , where X is the failure time of device 1 and Y is the failure time of device 2. Both variables take values in the interval $[0, \infty)$, where the units are in hundreds of hours. Sketch each of the following events:

1. The set of outcomes S
2. $\{X < Y\}$
3. $\{X + Y > 2\}$

Answer

1. $S = [0, \infty)^2$, the first quadrant of the coordinate plane.
2. $\{X < Y\} = \{(x, y) \in S : x < y\}$. This is the region below the diagonal line $x = y$.
3. $\{X + Y > 2\} = \{(x, y) \in S : x + y > 2\}$. This is the region above (or to the right) of the line $x + y = 2$.

Genetics

Please refer to the discussion of genetics in the section on random experiments if you need to review some of the definitions in this section.

Recall first that the *ABO blood type* in humans is determined by three alleles: a , b , and o . Furthermore, o is recessive and a and b are co-dominant.

Suppose that a person is chosen at random and his genotype is recorded. Give each of the following in list form.

1. The set of outcomes S
2. The event that the person is type A
3. The event that the person is type B
4. The event that the person is type AB
5. The event that the person is type O

Answer

1. $S = \{aa, ab, ao, bb, bo, oo\}$
2. $A = \{aa, ao\}$
3. $B = \{bb, bo\}$
4. $AB = \{ab\}$
5. $O = \{oo\}$

Suppose next that pod color in certain type of pea plant is determined by a gene with two alleles: g for green and y for yellow, and that g is dominant.

Suppose that n (distinct) pea plants are collected and the sequence of pod color genotypes is recorded.

1. Give the set of outcomes S in Cartesian product form and find $\#(S)$.
2. Let N denote the number of plants with green pods. Find $\#(N = k)$ (as a subset of S) for each $k \in \{0, 1, \dots, n\}$.

Answer

1. $S = \{gg, gy, yy\}^n$, $\#(S) = 3^n$
2. $\binom{n}{k} 2^k$

Next consider a sex-linked hereditary disorder in humans (such as colorblindness or hemophilia). Let h denote the healthy allele and d the defective allele for the gene linked to the disorder. Recall that d is recessive for women.

Suppose that n women are sampled and the sequence of genotypes is recorded.

1. Give the set of outcomes S in Cartesian product form and find $\#(S)$.
2. Let N denote the number of women who are completely healthy (genotype hh). Find $\#(N = k)$ (as a subset of S) for each $k \in \{0, 1, \dots, n\}$.

Answer

1. $S = \{hh, hd, dd\}^n$, $\#(S) = 3^n$
2. $\binom{n}{k} 2^{n-k}$

Radioactive Emissions

The emission of elementary particles from a sample of radioactive material occurs in a random way. Suppose that the time of emission of the i th particle is a random variable T_i taking values in $(0, \infty)$. If we measure these *arrival times*, then basic outcome vector is (T_1, T_2, \dots) and so the set of outcomes is $S = \{(t_1, t_2, \dots) : 0 < t_1 < t_2 < \dots\}$.

Run the simulation of the gamma experiment in single-step mode for different values of the parameters. Observe the arrival times.

Now let N_t denote the number of emissions in the interval $(0, t]$. Then

1. $N_t = \max \{n \in \mathbb{N}_+ : T_n \leq t\}$.
2. $N_t \geq n$ if and only if $T_n \leq t$.

Run the simulation of the Poisson experiment in single-step mode for different parameter values. Observe the arrivals in the specified time interval.

Statistical Experiments

In the basic cicada experiment, a cicada in the Middle Tennessee area is captured and the following measurements recorded: body weight (in grams), wing length, wing width, and body length (in millimeters), species type, and gender. The cicada data set gives the results of 104 repetitions of this experiment.

1. Define the set of outcomes S for the basic experiment.
2. Let F be the event that a cicada is female. Describe F as a subset of S . Determine whether F occurs for each cicada in the data set.
3. Let V denote the ratio of wing length to wing width. Compute V for each cicada.
4. Give the set of outcomes for the compound experiment that consists of 104 repetitions of the basic experiment.

Answer

For gender, let 0 denote female and 1 male, for species, let 1 denote tredecula, 2 tredecim, and 3 tredecassini.

1. $S = (0, \infty)^4 \times \{0, 1\} \times \{1, 2, 3\}$
2. $F = \{(x_1, x_2, x_3, x_4, y, z) \in S : y = 0\}$
5. S^{104}

In the basic M&M experiment, a bag of M&Ms (of a specified size) is purchased and the following measurements recorded: the number of red, green, blue, yellow, orange, and brown candies, and the net weight (in grams). The M&M data set gives the

results of 30 repetitions of this experiment.

1. Define the set of outcomes S for the basic experiment.
2. Let A be the event that a bag contains at least 57 candies. Describe A as a subset of S .
3. Determine whether A occurs for each bag in the data set.
4. Let N denote the total number of candies. Compute N for each bag in the data set.
5. Give the set of outcomes for the compound experiment that consists of 30 repetitions of the basic experiment.

Answer

1. $S = \mathbb{N}^6 \times (0, \infty)$
2. $A = \{(n_1, n_2, n_3, n_4, n_5, n_6, w) \in S : n_1 + n_2 + \cdots + n_6 > 57\}$
5. S^{30}

This page titled [2.2: Events and Random Variables](#) 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.