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.

Sample Spaces and Events

Sample Spaces

The **sample space** of a random experiment is a set S that includes all possible outcomes of the experiment; the sample space plays the role of the universal set when modeling the experiment. For simple experiments, the sample space may be precisely the set of possible outcomes. More often, for complex experiments, the sample space 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 outcome, the sample space is $S = \{1, 2, 3, 4, 5, 6\}$, 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 the sample space to be $S = [0, \infty)$, even though most elements of this set are practically impossible (we hope!).

Often the outcome of a random experiment consists of one or more real measurements, and thus, the sample space consists of all possible measurement *sequences*. Thus, in many cases, the sample space of a random experiment is a subset of $\mathbb{R}^n$ for some n.

If we have n experiments with sample spaces $(S_1, S_2, ..., S_n)$, then the Cartesian product $S_1 \times S_2 \times \cdots \times S_n$ is the natural sample space for the compound experiment that consists of performing the n experiments in sequence. In particular, if we have a basic experiment with sample space S, then S^n is the natural sample space for the compound experiment that consists of n replications of the basic experiment. Similarly, if we have an infinite sequence of experiments with sample spaces $(S_1, S_2, ...)$ then $S_1 \times S_2 \times \cdots$ is the natural sample space for the compound experiment that consists of performing the given experiments in sequence. In particular, the sample space for the compound experiment that consists of indefinite replications of a basic experiment is $S^\infty = S \times S \times \cdots$ This is an essential special case, because probability theory is based on the idea of replicating a given experiment.

Events

Certain subsets of the sample space of an experiment are referred to as **events**. Thus, an event is a set of outcomes of the experiment. Each time the experiment is run, a given event A either **occurs**, if the outcome of the experiment is an element of A, or **does not occur**, if the outcome of the experiment is not an element of A. Intuitively, you should think of an event as a meaningful *statement* about the experiment.

In the section on Measure Theory, we discuss some technical conditions that must be imposed on the collection of events. These need not concern you if you are new to the study of probability; we will simply assume that all subsets of the sample space that we mention are valid events.

In particular, the sample space S itself is an event; by definition it *always* occurs. At the other extreme, the empty set $\varnothing$ is also an event; by definition it *never* occurs.

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 exercises, suppose that A and B are events.

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

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

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

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

5. Show that $A \setminus B$ is the event that occurs if and only if A occurs **but not** B.

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

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

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

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

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

10. Show that the union of the collection is the event that occurs if and only if **at least one** event in the collection occurs:

$$\bigcup_{i \in I} A_i$$

11. Show that the intersection of the collection is the event that occurs if and only if **every** event in the collection occurs:

$$\bigcap_{i \in I} A_i$$

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

Suppose now that $\mathscr{A} = \{A_1, A_2, ...\}$ is a countably infinite sequence of events.

13. Show that the event below (sometimes called the **limit superior**) occurs if and only if **infinitely many** of the given events occur:

$$\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i$$

14. Show that the event below (sometimes called the **limit inferior**) occurs if and only if **all but finitely many** of the given events occur:

$$\bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i$$

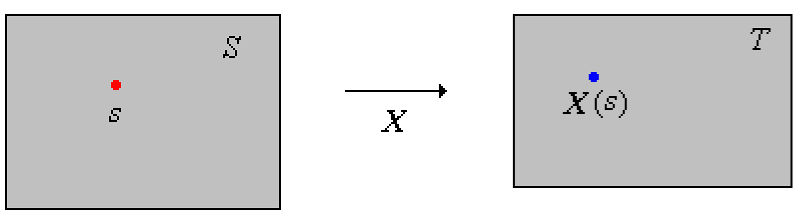
Limit superiors and inferiors are discussed in more detail in the section on convergence.

Random Variables

Suppose again that we have a random experiment with sample space S. A function X from S into another set T is called a (T-valued) **random variable**. Probability has its own notation, very different from other branches of mathematics. As a case in point, random variables are usually denoted by capital letters near the end of the alphabet.

In the section on Measure Theory, we will discuss a technical condition that must be imposed on random variables (and functions generally). These need not concern you if you are new to the study of probability; we will simply assume that all functions we mention are admissible.

Intuitively, you should think of a random variable X as a *measurement* of interest in the context of the random experiment. A random variable X is *random* in the sense that its value depends on the outcome of the experiment, which cannot be predicted with certainty before the experiment is run. Each time the experiment is run, an outcome $s \in S$ occurs, and a given random variable X takes on the value $X(s)$. In general, as you will see, the notation of probability suppresses references to the sample space. Indeed, sometimes the sample space is *hidden* in the sense that we don't know what it is.



Often, a random variable takes values in a subset T of $\mathbb{R}^k$ for some k. If $k > 1$ then we write

$$\boldsymbol{X} = (X_1, X_2, ..., X_k)$$

where X_i is a real-valued random variable for each i. In this case, we usually refer to $\boldsymbol{X}$ 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:

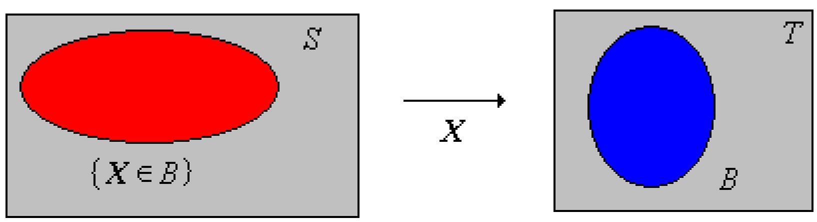
$$\boldsymbol{X} = (X_1, X_2, ..., X_k)$$

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 from the sample space S to another set T.

If X is a random variable taking values in T, and B is a subset of T, then we use the more suggestive notation for the inverse image:

$$\{X \in B\} = \{s \in S : X(s) \in B\}$$

Note that this is an event (a subset of S). In words, a *statement* about the random variable defines an *event*.



A special case of this notation is

$$\{X = x\} = \{X \in \{x\}\} = \{s \in S : X(s) = x\}, \quad x \in T$$

Suppose in particular that X is a real-valued random variable for an experiment. Another special cases of our notation is

$$\{a \le X \le b\} = \{X \in [a, b]\} = \{s \in S : a \le X(s) \le b\}$$

Of course, there are analogous results for other types of inequalities. The following exercise is a restatement of the fact that inverse images of a function preserve the set operations; only the notation changes (and is simpler).

15. Suppose that X is a random variable taking values in T, and that A and B are subsets of T. Show that

a. $\{X \in A \cup B\} = \{X \in A\} \cup \{X \in B\}$
b. $\{X \in A \cap B\} = \{X \in A\} \cap \{X \in B\}$
c. $\{X \in B \setminus A\} = \{X \in B\} \setminus \{X \in A\}$
d. $(A \subseteq B) \Rightarrow (\{X \in A\} \subseteq \{X \in B\})$
e. If A and B are disjoint, then so are $\{X \in A\}$ and $\{X \in B\}$.

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

Basic and Derived Variables

Suppose again that we have a random experiment with sample space S. The outcome of the experiment itself can be thought of as a random variable. Specifically, let X denote the identify function on S:

$$X(s) = s, \quad s \in S$$

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:

$$\{X \in A\} = A, \quad A \subseteq S$$

If Y is another random variable for the experiment, taking values in a set T, then Y is a function of X. That is, there is a function g from S into T such that Y is the composition of g with X. We use the more descriptive notation $g(X)$ instead of $g \circ X$. Thus,

$$Y(s) = g(X(s))$$

We could refer to X as the **outcome variable** and Y as a **derived variable**. In many problems of elementary probability theory, the basic object of interest is a random variable X. Whether X is the basic outcome variable or a derived variable is often irrelevant.

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:

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

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

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

17. Suppose that A and B are events. Show that

a. $\mathbf{1}(A \cap B) = \mathbf{1}(A)\,\mathbf{1}(B) = \min\{\mathbf{1}(A), \mathbf{1}(B)\}$
b. $\mathbf{1}(A \cup B) = 1 - (1 - \mathbf{1}(A))\,(1 - \mathbf{1}(B)) = \max\{\mathbf{1}(A), \mathbf{1}(B)\}$
c. $\mathbf{1}(B \setminus A) = \mathbf{1}(B)\,(1 - \mathbf{1}(A))$
d. $\mathbf{1}(A^c) = 1 - \mathbf{1}(A)$

e. $A \subseteq B$ if and only if $\mathbf{1}(A) \le \mathbf{1}(B)$

The results in part (a) extends to arbitrary intersections and the results in part (b) extends to arbitrary unions.

Examples and Applications

Recall that 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, 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

18. The **coin experiment** consists of tossing n (distinct) coins and recording the sequence of scores $(X_1, X_2, ..., X_n)$ (where 1 denotes heads and 0 denotes tails). This experiment is a generic example of n Bernoulli trials, named for Jacob Bernoulli. Let Y denote the number of heads.

a. Show that the sample space of the experiment is $S = \{0, 1\}^n$ and that $\#(S) = 2^n$.
b. Express Y as a function on the sample space S.
c. Show that $\#(\{Y = k\}) = \binom{n}{k}$ for $k \in \{0, 1, ..., n\}$
d. With $n = 5$, explicitly list the elements in the event $\{Y = 3\}$.

19. In the simulation of the coin experiment, set $n = 5$. Run the experiment 100 times and count the number of times that the event $\{Y = 3\}$ occurs.

20. Recall that the basic **dice experiment** consists of throwing n distinct, k sided dice (with sides numbered from 1 to k) and recording the sequence of scores $(X_1, X_2, ..., X_n)$. This experiment is a generic example of n multinomial trials. The special case $k = 6$ corresponds to **standard dice**. Let Y denote the sum of the scores, U the minimum score, and V the maximum score.

a. Show that the sample space of the experiment is $S = \{1, 2, ..., k\}^n$ and that $\#(S) = k^n$.
b. Express Y as a function on the sample space, and give the set of possible values.
c. Express U as a function on the sample space, and give the set of possible values.
d. Express V as a function on the sample space, and give the set of possible values.
e. Give the set of possible values of (U, V).

21. Consider the dice experiment with 2 standard dice. Let A denote the event that the first die score is 1 and B the event that the sum of the scores is 7. Explicitly list the elements of the following events:

a. A
b. B
c. $A \cup B$
d. $A \cap B$
e. $A^c \cap B^c$

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

23. Consider the dice experiment of Exercise 20 with 2 standard dice. Explicitly list the elements of the following events:

a. $\{X_1 < 3, X_2 > 4\}$
b. $\{Y = 7\}$
c. $\{U = V\}$

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

25. 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.

a. Suppose that the pair of scores on each throw is recorded. Define the sample space of the experiment and describe A as a subset of this sample space.
b. Suppose that the pair of scores on the final throw are recorded. Define the sample space of the experiment and describe A as a subset of this sample space.

26. 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 and then receives \$1 for each die that lands on the number 6. 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.

a. Give the sample space S of the experiment.
b. Express W as a function on S.

27. 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 $\boldsymbol{X}$ is recorded (0 for tails, 1 for heads). Let N denote the die score and Y the number of heads.

a. Give the sample space S of the experiment and find $\#(S)$.
b. Express N as a function on the sample space S.
c. Express Y as a function on the sample space S.
d. Explicitly list the elements of the event $\{Y = 2\}$

28. Run the simulation of the die-coin experiment 10 times. For each run, give the values of the random variables X, N, and Y of the previous exercise. Count the number of times the event $\{Y = 2\}$ occurs.

29. In the **coin-die experiment**, we have a coin and two standard dice, 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.

a. Define the sample space S and find $\#(S)$.
b. Express X as a function on the sample space.
c. Express Y as a function on the sample space.
d. Explicitly list the elements of the event $\{Y \ge 4\}$

30. Run the coin-die experiment, with the default settings, 100 times. Count the number of times that event $\{Y \ge 4\}$ occurred.

Cards

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

$$D = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \text{j}, \text{q}, \text{k}\} \times \{♣, ♦, ♥, ♠\}$$

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♥).

31. The basic **card experiment** consists of dealing n cards from a standard deck without replacement and recording the (ordered) sequence of cards: $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ where $X_i \in D$ is the i^{th} card. Let $\boldsymbol{W} = \{X_1, X_2, ..., X_n\}$ denote the (unordered) set of cards in the hand. 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.

a. Show that the sample space S consists of all permutations of size n from D and $\#(S) = 52^{(n)}$
b. Show that the set T of possible values of $\boldsymbol{W}$ consists of all combinations of size n from D and

$$\#(T) = \binom{52}{n}$$

c. Explicitly compute the numbers in (a) and (b) when $n = 5$ (poker).
d. Explicitly compute the numbers in (a) and (b) when $n = 13$ (bridge).

32. Consider the card experiment with $n = 1$. Let Q denote the event that the card is a queen and H the event that the card is a heart. Explicitly list the elements in the events below:

a. Q
b. H
c. $Q \cup H$
d. $Q \cap H$
e. $Q \setminus H$

33. 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.

34. Consider the bridge experiment of dealing 13 cards from a deck. 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 V denote the point value of the hand.

a. Describe V as a function of $\boldsymbol{W}$ (the unordered hand), and find the set of possible values of V.
b. Find $\#(\{V = 0\})$, where the event is considered as a subset of T (the set of unordered bridge hands).

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

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

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

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

Buffon's Coin Experiment

38. Recall that Buffon's coin experiment, named for Compte de Buffon, consists of tossing a coin with radius $r \le \frac{1}{2}$ 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. Let A denote the event that the coin does not touch the sides of the square and let Z denote the distance form the center of the coin to the center of the square.

a. Describe the sample space S mathematically.
b. Describe A as a subset of S.
c. Describe A^c as a subset of S.
d. Express Z as a function on S.
e. Express the event $\{X < Y\}$ as a subset of S.
f. Express the event $\left\{Z < \frac{1}{2}\right\}$ as a subset of S.

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

Urn Models

40. An urn contains m balls with distinct labels. A sample of n balls is chosen from the urn without replacement, and the sequence of ball labels $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ is recorded. Let $\boldsymbol{W} = \{X_1, X_2, ..., X_n\}$ denote the unordered set of balls. This model is a metaphor for the sampling without replacement from a finite population.

a. Show that the sample space S of the experiment consists of all permutations of size n from the population of balls and $\#(S) = m^{(n)}$
b. Show that the set T of possible values of $\boldsymbol{W}$ consists of all combinations of size n from the population of balls and that $\#(T) = \binom{m}{n}$

41. In the basic urn model above, suppose that r of the m balls are red and the remaining $m - r$ are green. Again, a sample of n balls is chosen without replacement from the urn. This model is metaphor for the experiment of sampling without replacement from a **dichotomous population**. Let Y denote the number of red balls in the sample.

a. Describe Y as a function of $\boldsymbol{W}$, the unordered sample of balls, and give the set of values of Y
b. Show that $\#(\{Y = k\}) = \binom{n}{k} r^{(k)} (m - r)^{(n-k)}$ for each k, if the event is considered as a subset of S, the set of ordered samples.
c. Show that $\#(\{Y = k\}) = \binom{r}{k} \binom{m-r}{n-k}$ for each k, if the event is considered as a subset of T, the set of unordered samples.
d. Write the expressions in (a) and (b) in terms of factorials, and note that the expression in (a) is $n!$ times

the expression in (b), as should be the case.

42. 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.

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

43. Run the simulation of the ball and urn experiment 100 times for the parameter values in the last exercise. Note the values of the random variable Y.

Reliability

In a 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, $(X_1, X_2, ..., X_n)$ is a vector of indicator random variables that specifies the states of all of the components. Thus, the sample space of the experiment is $\{0, 1\}^n$. The system as a whole is also either working or failed, depending only on the states of the components. Thus, the state of the system is an indicator random variable defined on this state space: $Y(X_1, X_2, ..., X_n)$. The state of the system (working or failed) as a function of the states of the components is the **structure function**.

44. A **series system** is working if and only if each component is working. Show that the state of the system is

$$U = X_1 X_2 \cdots X_n = \min\{X_1, X_2, ..., X_n\}$$

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

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

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 $2k - 1$ system is a **majority rules system**.

46. Show that the state of the k out of n system is given by the formula below. The structure function can also be expressed as a polynomial in the variables.

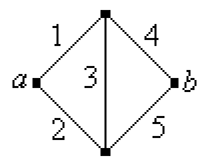
$$U_{n,k} = 1 \text{ if and only if } \sum_{i=1}^{n} X_i \ge k$$

47. 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 $k \in \{1, 2, 3\}$.

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.

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

a. Note that if component 3 works, then the system works if and only if component 1 or 2 works, and component 4 or 5 works.
b. Note that if component 3 does not work, then the system works if and only if components 1 and 4 work, or components 2 and 5 work.



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

a. $s(0, 0, ..., 0) = 0$ and $s(1, 1, ..., 1) = 1$
b. s is an increasing function, where $\{0, 1\}$ is given the ordinary order and $\{0, 1\}^n$ the corresponding product order.
c. For each $i \in \{1, 2, ..., n\}$, there exist x and y in $\{0, 1\}^n$ all of whose coordinates agree, except coordinate i, but $s(x) = 0$ while $s(y) = 1$

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 $i \in \{1, 2, ..., n\}$. Thus, the basic outcome of the experiment is the random vector of failure times: $(T_1, T_2, ..., T_n)$ and the sample space of the experiment is $S = [0, \infty)^n$.

50. Show that the state of component i at time $t \geq 0$ is $X_i(t) = \mathbf{1}(T_i > t)$. If the structure function is s, show that the state of the system at time t is

$$X(t) = s(X_1(t), X_2(t), ..., X_n(t))$$

51. Suppose that s is a valid structure function. Show that the time to failure of the system is

$$T = \min\{t \geq 0 : X(t) = 0\}$$

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 $(T_1, T_2, ...)$ and the sample space of the experiment is $S = \{(t_1, t_2, ...) : 0 < t_1 < t_2 < \cdots\}$.

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

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

a. $N_t = \max\{n \in \{1, 2, ...\} : T_n \le t\}$
b. $N_t \ge n$ if and only if $T_n \le t$

54. 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

55. 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.

a. Define a sample space for the basic experiment.
b. Let F be the event that a cicada is female. Describe F as a subset of the sample space.
c. Determine whether F occurs for each cicada in the data set.
d. Let V denote the ratio of wing length to wing with. Compute V for each cicada.
e. Give the sample space for the compound experiment that consists of 104 repetitions of the basic experiment.

56. 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.

a. Define a sample space for the basic experiment.
b. Let A be the event that a bag contains at least 57 candies. Describe A as a subset of the sample space.
c. Determine whether A occurs for each bag in the data set.
d. Let N denote the total number of candies. Compute N for each bag in the data set.
e. Give the sample space for the compound experiment that consists of 30 repetitions of the basic experiment.