Probability

Animal Health Risk Analysis
Canadian Food Inspection Agency
3851 Fallowfield Road
Ottawa, Ontario, Canada K2H 8P9

Analyse des risques de la santé des animaux
Agence canadienne d'inspection des aliments
3851, chemin Fallowfield
Ottawa (Ontario) Canada K2H 8P9

Canadian Food Inspection Agency (CFIA)

Agence canadienne d'inspection des aliments (ACIA)

Basic Set Theory

Definition: A set is a well-defined collection of objects.

The capital letters A, B, C ... represent sets whereas a, b, c, ... denote their elements. We can have finite or infinite sets.

Example - If A is a set consisting of the squares of the integers 1, 4, 5, 7, it will contain 4 elements and can be written as A = {1, 16, 25, 49} (elements in any order) = $\{x^2 \mid x = 1, 3, 5, 7,\}$

A is finite whereas B = {1, 3, 5, 7, 9, ...} is infinite. We denote elements as $x \in A$. For example, $16 \in A$ but $4 \notin A$.

Definition: A set containing no elements is called the null set denoted by $\varnothing$.

Example - $\{x \mid (x-1)^2<0\} = \varnothing$

Definition: A set A is called a subset of B if every element of a set A is also an element of a set B and is denoted as $A \subset B$.

Example - For A = {1, 16, 25, 49}, $\{1\} \subset A$, $\{1, 49\} \subset A$, $\varnothing \subset A$.

Basic Set Theory

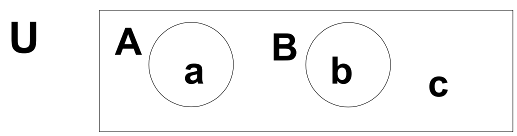
Definition: Two sets A and B are said equal if they have exactly the same elements, i.e., $A \subset B$ and $B \subset A \Rightarrow A = B$.

- U will denote the universal set.
- Any finite set of n elements has 2^n subsets.

Example - Let U = {a, b, c}
Therefore, U has $2^3 = 8$ subsets which are {a}, {b}, {c}, {a,b}, {a, c}, {b, c}, {a, b, c}, $\varnothing$.

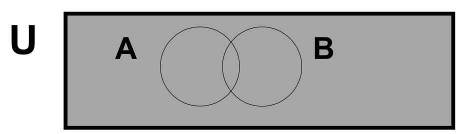
Definition: A Venn diagram consists of a rectangle representing the universal set U and circles drawn inside the rectangle representing subsets of U.

Example - If U = {a, b, c}, A = {a}, B = {b} then we have as our Vem diagram:

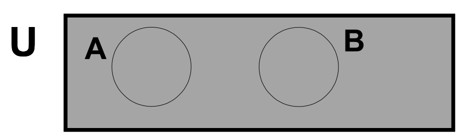


Operations on Sets

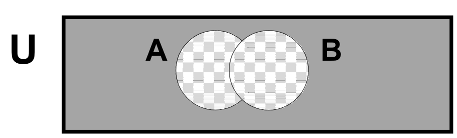
Definition: The intersection of two sets A and B, denoted by , $A \cap B$, is the set of all elements of U that are members of both A and B, i.e., $A \cap B$ - $\{x | x \in A \text{ and } x \in B\}$.



Definition: Two sets A and B are disjoint if they have no elements in common, i.e. , $A \cap B = \varnothing$.



Definition: The union of two sets A and B, denoted by $A \cup B$, is the set of all elements of U belonging either to A or to B or to both, i.e., $A \cup B = \{x | x \in A \text{ or } x \in B\}$.



Operations on Sets

<u>Definition</u>: The complement of a set A, denoted by $\bar{A}$ or A^c , is the set of all elements of U not contained in A, i.e.,

$$A^c = \{x | x \in U \text{ and } x \notin A\}$$

<u>Example</u> - If: $U = \{1, 2, 5, 8, 9, 10, 13, 15\}$ and $A = \{1, 5, 9, 10, 13\}$,

$B = \{1, 2, 5, 9, 10, 15\}$

and $C = \{2, 8, 15\}$, then:

$A \cap B = \{1, 5, 9, 10\}$

$A \cap C = \varnothing$

$A \cup B = \{1, 2, 5, 9, 10, 13, 15\}$

$A^c = \{2, 8, 15\} = C$

$(B \cup C) \cap A = \{1, 2, 5, 8, 9, 10, 15\} \cap A = \{1, 5, 9, 10\}$

The Sample Space

An event is the random outcome of an experience (covering anything from a designed experiment, like tossing a coin, to observational phenomena, such as the number of cows at slaughter detected with tuberculous lesions) . The term experiment is used to describe either uncontrolled events in nature or controlled situations in a laboratory. An experiment is the process by which an observation (or measurement) is obtained. Experiments may be quite diverse such as the following:

- **serologically testing a Holstein heifer for enzootic bovine leukosis**
- **measuring the thickness of the skin at an intradermal tuberculin injection site**
- **interviewing a swine producer by phone to determine if plate waste is being fed to his/her swine**

Each of these experiments may result in one or more outcomes or events, which are represented by capital letters, A, B, etc. Two events are mutually exclusive if, when one event occurs, the other cannot, and vice versa, e.g., a head or a tail in a toss of a coin; a calf is either male or a female. An event that cannot be decomposed is called a simple event, e.g., a head or a tail in the toss of a coin. A set of all simple events is called the sample space, e.g., a head and a tail in the toss of a coin; the results of a Complement Fixation test: positive, suspect, anti-complementary and negative.

An event is a collection of one or more simple events, e.g., the toss of two heads in a row.

Sample Space and Events

An experiment as used in statistics refers to any activity resulting in the amassing of data referring to outcomes which cannot be predicted with certainty.

Examples:

(I) Give a toxic substance to 10 rats and record the no. of deaths after four hours.
(ii) Count the number of yellow cars driven past your language lab on a specific afternoon.
(iii) Toss three coins - a nickle, a dime, a quarter - simultaneously and record whether a head or a tail appears for each.
(iv)Throw a red and a black die together and notice the values appearing on both die.

Definition: A sample space S of an experiment E is the set of all possible distinct outcomes. Each element of S or outcome of E is called a sample point and is denoted by e_i where i = 1, 2, ... Therefore, S ={e_1, e_2, e_3, ...}.

Examples: For the experiments E in the examples above, the respective sample spaces can be expressed as follows:

(i) S = {0, 1, 2, 3, 4, ..., 10}
(ii) S = {0, 1, 2, 3,...}
(iii)If H and T respectively represent a head and a tail and in the sequence the first letter recording the sign of the nickel, the second that of the dime and the third that of the quarter, then S = {HHH, HTH, HHT, THH, HTT, THT, TTH, TTT}.
(iv)If we use ordered pairs where the first value is the outcome of the red die and the second that of the black die, we have S= {r, b)|$1 \le r \le 6$, $1 \le b \le 6$}. (36 outcomes).

Sample Space and Events

Definition: An event is a subset of a sample space S. Therefore, capital letters are used for events and ∅ and S are always events.

Examples - Referring to the examples , we could consider the following events:

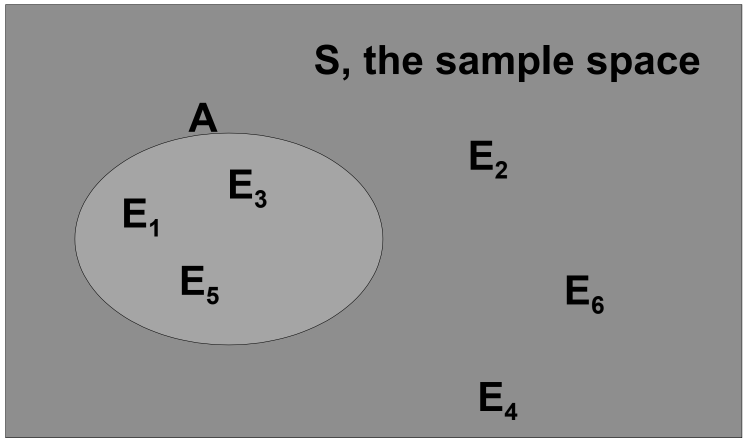
(I) A is the event that at least 2 rats died but no more than 6 died. B = {2, 3, 4, 5, 6}
(ii) B is the event that at most 6 yellow cars drove past the lab. A = {0, 1, 2, 3, 4, 5, 6}
(iii)C is the event that at least two tails occurred. C = {HTT, THT, TTH, TTT}
(iv) D is the event that all heads appeared. D = {HHH}
(v) G is the event that the black die recorded a 2. G = {(1, 2), (2, 2), (3, 2), (4, 2), (5, 2), (6, 2)}.

Definition: Two events A and B are mutually exclusive if $A \cap B = \varnothing$. In the examples above, $C \cap D = \varnothing$, thus C and D are mutually exclusive events.

Die Toss Sample Space Example

Experiment: Toss a die and observe the number that appears on the upper face.

Event A : Observe an odd number
Event B : Observe a number less than 4
Event E1: Observe a 1
Event E2: Observe a 2
Event E3: Observe a 3
Event E4: Observe a 4
Event E5: Observe a 5
Event E6: Observe a 6



Events A and B are not mutually exclusive because events occur when the number on the upper face of the die is a 1 or a 3. Event A occurs when the face is 1, 3 or 5 and can be decomposed into a collection of simpler events, E_1, E_3 and E_5 which are mutually exclusive. Event B occurs when E_1, E_2 or E_3 occurs and can be viewed as a collection of these mutually exclusive events. All six events (E_1 to E_6)form a set of mutually exclusive outcomes and are called simple events.

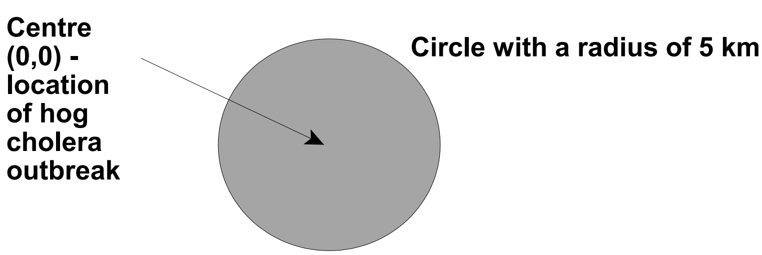
Sample Space Examples

In the sexing of chicks, the sample space consists of only two outcomes, S = {M, F}.

The sample space for serum sample serological test results may comprise negative (N), BPAT positive and ELISA positive (BE) and BPAT positive and ELISA negative (Be). Therefore, S = {N, BE, Be).

If the possible outcomes of an experiment or survey is the set of swine herds with less than 200 pigs, the sample space can be written S = {*x*|*x* is a swine herd with less than 200 pigs}.

S could be the set of all swine farms within a circle of radius of 5 km with centre at the origin, a hog cholera outbreak. S = {all swine farms within circle|$x^2 + y^2 < 25$}



The Concept of Probability

Probabilities are proportions that denote how likely the events are. A probability is a numerical value between zero and one. The number is associated with an outcome or event and specifies the likelihood of occurrence. The probability may represent the frequency with which some event occurs in a long sequence of similar trials or the degree of belief that a person has that it will occur given all the relevant information.

Odds of an event is the ratio of the number of chances in favor over the number of chances not in favor of an event. Odds is the ratio of the probability of the occurrence of an event to that of nonoccurrence, or the ratio of the probability that something is so, to the probability that it is not so. Between 1973-84 the odds of human fatalities in aircraft flights were 1 to 54 (11 in 600 flights) in Turkey and 1 in a million flights in Australia. In the U.S.A. the odds that an aircraft will fall on and kill you sometime this year is 1 in 25 million. In the tossing of a balanced coin, there are two events: "Head" and "Tail". Intuitively, the probability of a head is 0.5. The odds then are 1:1.

<u>Classical Probability Concept</u>
A useful method to determine the probability of an event is where the experiment has many basic Equally Likely Possibilities, such as in games of chance and tossing of a coin. If there are *n* equally likely possibilities, one of which must occur and *s* are regarded as favorable, or as a "success," then the probability of a "success" is given by the ratio *p = s/n* . In a well shuffled deck of 52 cards, the probability of one draw of a card resulting in an ace is 4/52. All cards are equally likely. Four possibilities in favor of drawing an ace are the four different aces. Problems with the classical approach: Not all events can be cast in terms of a game of chance and there is no situation with equally likely possibilities.

The Concept of Probability

<u>Definition</u>: If an experiment consists of n different equally likely outcomes and exactly m of these outcomes favour an event A, then the probability of event A denoted by P(A), is defined as

$$P(A) = \frac{m}{n}$$

<u>Example</u> - A cage contains four white, eight brown and three colour-mixed guinea pigs. If A = event that a brown guinea pig is chosen and B = event that a white or a colour-mixed guinea pig is drawn, then

$$P(A) = \frac{8}{15}, P(B) = \frac{7}{15}$$

<u>Example</u> - A card is drawn from a standard deck of 52 cards. If A = event that an ace is picked, B = event that the ace of hearts is picked, C = event that a heart is picked and D = event that a spade with value 5 or 6 is picked, then

$$P(A) = \frac{4}{52} = \frac{1}{13}, P(B) = \frac{1}{52}, P(C) = \frac{13}{52} = \frac{1}{4}, P(B) = \frac{2}{52} = \frac{1}{26}$$

The Concept of Probability

Example - A pair of dice is cast. If A = event that both die record the same value, B = event that an 8 is rolled and C = event that at least a "4" is rolled, then

$$P(A) = P\{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\} = \frac{6}{36} = \frac{1}{6},$$

$$P(B) = P\{(2,6),(3,5),(4,4),(5,3),(6,2)\} = \frac{5}{36},$$

$$P(C) = P\{(1,3),(1,4),(1,5),(1,6),(2,2),(2,3),(2,4),(2,5),(2,6)$$
$$(3,1),(3,2),(3,3),(3,4),(3,5),(3,6),(4,1),\ldots,(6,6)\} = \frac{33}{36} = \frac{11}{12}.$$

The Concept of Probability

Properties of a Probability Model: If we have an experiment E, then the probability of an event A, P(A), satisfies the following properties:

(i) $0 \le P(A) \le 1$
(ii) $P(\varnothing) = 0$
(iii) $P(S) = 1$,
where S is the event where all possible outcomes of E occur.

Note: P(A) = $\sum$ probabilities of the outcomes in A.

In many experiments, the outcomes are not equally likely. For instance, in recording the probability of a car accident on a particular road, of an electron microscope being defective in its first month of use, of a person living to age 80 or of getting a head when an unbalanced coin is tossed, the outcomes comprising the events are not equally likely. Here we must consider how frequently an event occurs when the experiment is repeatedly performed under identical conditions. In *n* repetitions or trials, if an, event A occurs *m/n* times, is called the relative frequency of the event A in *n* trials. As *n* increases, this relative frequency stabilizes and is ensured a limit *f/n* say, which we refer to as the relative frequency or *a posteriori* concept of probability.

Definition: If an experiment is performed n times and if an event A occurs m times, then the relative frequency of A is *m/n*. If we assume that as *n* increases the relative frequency *m/n* approaches a limit, then this limit is called the probability of A. In practice, we estimate P(A) by *m/n* for a large *n*.

The Concept of Probability

Example - A toxic substance is given to 300 rats and the time taken for death to occur (in hours) recorded. The following results were obtained assuming the count is done only at hour intervals:

No. of Hours Until Death Occurred	Frequency of Rats (cumulative)
1	48
2	172
3	240
4	261
5	280
≥6	300

If A = event that a rat took at most 2 hours to die,
B = event that a rat took at most 2 hours but more than 1 hour to die and
C = event that a rat took more than 5 hours to die, then

$$P(A) = \frac{172}{300}, P(B) = \frac{172-48}{300} = \frac{124}{300}, P(C) = \frac{20}{300} = \frac{1}{15}$$

by the frequency definition of probability.

The Concept of Probability

(iii) In the early 1950's, L.J. Savage introduced the personalistic concept of probability which measure the confidence rather than the likelihood that a particular individual has in the truth of a particular proposition or occurrence. The process here need not involve repeatability such as events which can only occur once.

This subjective approach, in the absence of past data, requires assigning the probability to an event based on one's best available evidence and judgment.

Example -
- **(i) probability that a cure for cancer will be discovered within the next seven years**
- **(ii) probability or chance that rain will occur tomorrow**
- **(iii)probability that Iraq's government will fall during the oncoming month**
- **(iv) probability that a woman will be elected Prime Minister of Canada**

The Concept of Probability

Despite the fact that the concept of probability has been under the scrutiny of philosophers, mathematicians, and scientists for about 300 years, it has successfully resisted reduction to a single, perfectly clear definition. To some, probability implies a degree of uncertainty about future events (Will it rain tonight?) or about whether or not some event has already taken place (Did Caesar cross hear?). When probabilities are assigned to such statements, they are called "subjective probabilities" because the probability lies in the person who holds the belief rather than in the event. For, after all, rain will or will not fall tonight, and Caesar did or did not cross here. There is no uncertainty in the event itself, only in our knowledge about it.

To others, the term "probability" may also be attached to a scientific hypothesis; one might attach a probability of 0.8 (or 0.3) to the proposition that a child's creative ability is fully determined by the time that he reaches age six. The degree of probability assigned is reached in accordance with David Hume's suggestion that "a wise man proportions his belief to the evidence." It has been argued that this kind of probability is merely another form of subjective probability, for the proposition itself can be only true or false.

To the mathematician, however, the term "probability" has different meanings. Most important is that of relative frequency. It should be emphasized that the frequency concept of probability is applicable not to events that are unique but only to those that occur repeatedly. This feature permits the mathematical development of the theory of probability underlying the entire statistical method. G. Polya provides an amusing example of "how not to interpret the frequency concept of probability."

The doctor shook his head as he finished examining the patient. "You have a very serious disease," said the doctor. "Of ten patients who have got this disease only one survives. But you are lucky. You will survive because you came to me. I have already had nine patients who all died of it."
From G. Polya, Patterns of Plausible Inference, (Princeton, N.J.: Princeton University Press, 1954).

The Concept of Probability

<u>The Frequency Interpretation of Probability</u>
Sometimes called the empirical approach: Based on analysis of data randomly sampled from a population of values. The probability is used in making statements about the makeup of the population, that is, in making statistical inferences. The probability of an event (happening or outcome) is the proportion of times that events of the same kind will occur in the long run.
For example: A certain serological test detects an infected animal in 95% of infected animals that are tested.

The relative frequency = frequency/*n* =*x/n* and: $P(A) = \lim_{n\to\infty} \frac{Frequency}{n}$

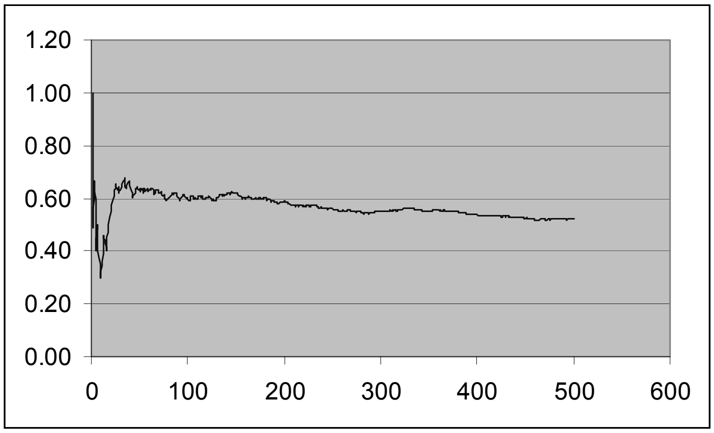
<u>Law of Large Numbers</u>

If a situation, trial, or experiment is repeated again and again, the proportion of successes will tend to approach the probability that any one outcome will be a success. The Law of Large Numbers can be illustrated with a simulation: To simulate the flipping of a balanced coin, a random number was generated between 0 and 1, X, using the Rand() function in Microsoft Excel. For X < 0.5, then a "head" occurred, otherwise a "tail" occurred. The proportion of heads as the number of flips of the coin increases was then determined. The Law of Large Numbers predicts that this proportion will get arbitrarily close to 0.5.

The Concept of Probability

The Law of Large Numbers predicts that the proportion of heads will become close to 0.5 as the number of flips increases. In the graph below, as the number of flips approaches 500 the proportion of heads approaches 50%.

Proportion of heads



Number of flips of a balanced coin

The Concept of Probability

Subjective Probability

The frequency approach has its limitations. A frequentist who has no data is paralyzed. Even in cases where data are available, the data may not be representative of or relevant to the problem at hand. Hence, statistical evaluation of data may be an insufficient basis for estimating variablility and uncertainty in a quantity. In many situations it is not possible to repeat the experience. Thus, there may be cases in which data are lacking in quantity or quality but for which an analyst has other information that can be used to construct a probabilistic representation of an input to a model. This alternative to the fequentist approach is the subjectivist approach, based theoretically on Bayesian inference.

With the subjectivist or Bayesian view, a probability of an event is the degree of belief that a person has that it will occur, given all the relevant information currently known to that person. The subjective approach is to base the probability of an event on a person's perceived probability of the event. Thus the probability is a function not only of the event, but of the state of information. Subjective probabilities obey the same axioms as objective or frequentist probabilities. When there is sufficient empirical data for the frequentist ot estimate a probability, the subjecitivist's assessment of his/her probability will converge to the frequentist's estimate of the probability of the event.

Bayesian methods heavily use subjective probabilities and likelihoods of data configurations. For example, a prior probability that a population of animals is infected with a disease agent would be updated after survey or screening test results were obtained. This prior probability would be based on subjective probability.

Gambling

Although it may be surmised that the student of probability is not likely to be a gambler (having more faith in his knowledge of relative frequencies than in luck), we cannot expect that the reader will be able to suppress completely whatever tendencies he may have to indulge in games of chance. Therefore we offer two pieces of advice to those who will succumb to temptation. The first is from John Maynard Keynes, the famous British economist, and is quite general in nature. The second is quite specific. Neither is guaranteed to increase one's fortune in the slightest.

Keynes gives a mathematical proof that "the poorer a gambler is, relatively to his opponent, the more likely he is to be ruined." Furthermore, if one's opponent has resources of an infinite amount, one's ruin is certain. Continuing the argument, he points out that "The infinitely rich gambler is the public. It is against the public that the professional gambler plays, and his ruin is therefore certain." Keynes then considers (possibly with tongue in cheek) some of the implications of this conclusion thus:

... no gambler plays, as this argument supposes, forever. At the end of any finite quantity of play, the player, even if he is not the public, may finish with winnings of any finite size. The gambler is in a worse position if his capital is smaller than his opponent's - at poker, for instance, or on the Stock Exchange. This is clear. But our desire for moral improvement outstrips our logic if we tell him that he must lose. Besides it is paradoxical to say that everybody individually must lose and that everybody collectively must win. For every individual gambler who loses there is an individual gambler or syndicate of gamblers who win. The true moral is this, that poor men should not gamble and that millionaires should do nothing else. But millionaires gain nothing by gambling with one another, and until the poor man departs from the path ofprudence the millionaire does notfind his opportunity. If it be replied that in fact most millionaires are men originally poor who departed from the path of prudence, it must be admitted that the poor man is not doomed with certainty. Thus the philosopher must draw what comfort he can from the conclusion with which his theory furnishes him, that millionaires are often fortunate fools who have thrived on unfortunate ones.

Gambling

The second piece of advice concerns what is known as the "gambler's fallacy." This fallacy is well known to all students of probability and statistics; yet surprisingly few gamblers seem to be aware of it, and they greet the notion with considerable skepticism. The gambler's fallacy is the belief that the probability of an event's occurrence changes with the occurrence of the preceding events. For example, a roulette player may observe that on the last six turns of the wheel the ball has fallen in a red pocket. "Aha," he muses, "we are about ready for a black." Perhaps he is cautious. He waits for another turn of the wheel. Red again! Now he has no doubts whatever. Red has appeared seven times in a row. On the next turn the ball must fall in the black! He splurges and bets all his chips on the black. Does he win? We cannot answer that, unfortunately. But we can say that the gambler's hopes were based on a mistaken premise - for the odds against a black on the crucial turn of the wheel are exactly what they are on any turn of the wheel. Each event, each turn of the wheel (assuming that it is honest) involves exactly the same probability that red or black will appear as does any other turn, for each turn is independent of every other. The probability of the occurrence of an event is not dependent upon the outcome of previous events, provided that the events occur at random. And randomness, of course, is the essential element of gambling.

From John Maynard Keynes, A Treatise on Probability (New York: St. Martin's Press, 1921).

Probability Rules

Simple Events
The requirements for simple-event probabilities, are that each probability must lie between 0 and 1 and the sum of the probabilities for all simple events in S equals 1. For every event A, $0 \leq P(A) \leq 1$. The impossible event has probability zero, $P(\varnothing) = 0$.

Definition: The probability of an event A is equal to the sum of the probabilities of the simple events contained in A.

Calculating the probability of an event:

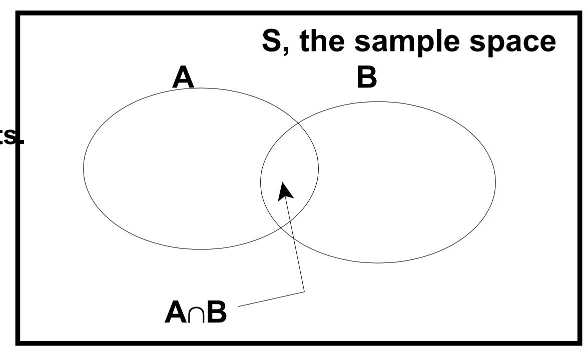
1. **List all simple events in the sample space.**
2. **Assign an appropriate probability to each simple event.**
3. **Determine which simple events results in the event of interest.**
4. **Sum the probabilities of the simple events that result in the event of interest.**

Be careful to satisfy two conditions in your calculation:

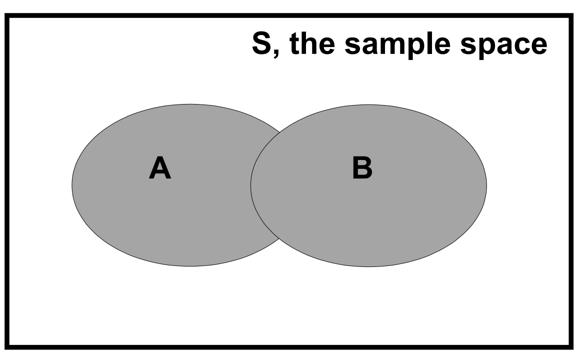
- **Include all simple events in the sample space.**
- **Assign realistic probabilities to the simple events.**

Probability Rules

Compound Events
Compound events can be formed by unions or intersections of other events.



Definition: The intersection of events A and B, denoted by $A \cap B$, is the event that A or B occur.



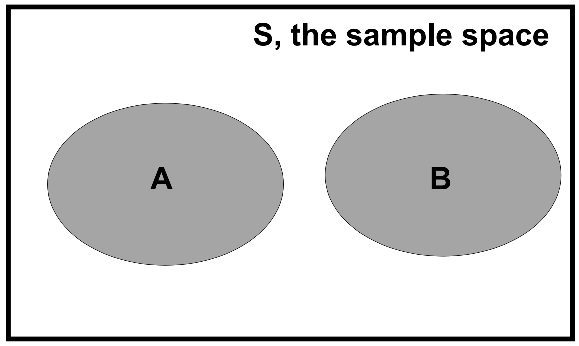
Definition: The union of events A and B, denoted by $A \cup B$, is the event that A or B or both occur.

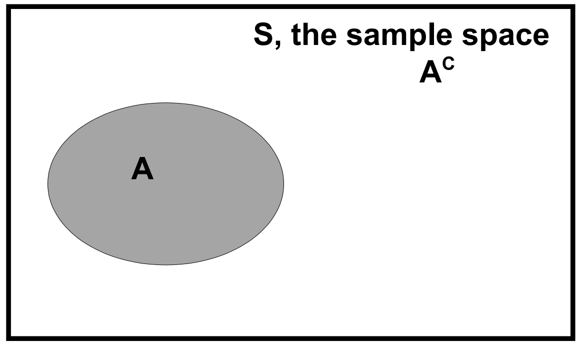
Probability Rules

Definition: When two events A and B are mutually exclusive, it means that when A occurs, B cannot, and vice versa. Mutually exclusive events are also referred to as disjoint events. When events A and B are mutually exclusive: $P(A \cap B) = 0$ and $P(A \cup B) = P(A) + P(B)$.

If P(A) and P(B) are known, $(A \cup B)$ do not need to be broken down into simple events, simply sum them. The generalized rule for mutually exclusive events: if *k* events are mutually exclusive, the probability that one of them will occur equals the sum of their individual probabilities, i.e., $P(A_1 \cup A_2 \cup \ldots \cup A_k) = P(A_1)+P(A_2)+\ldots+P(A_k)$ for mutually exclusive events $A_1, A_2, \ldots A_k$.

Definition: The complement of an event A, denoted A^C, consists of all the simple events in the sample space S that are not in A. $P(A) + (A^C) = 1$ and $P(A^C) = 1 - P(A)$.





Probability Rules

The conditional probability of A, given that B has occurred, is denoted as P(A | B), where the vertical bar is read "given" and the events appearing to the right of the bar are those that are known to have occurred. A and B are two events of the sample space S.

Definition: The conditional probability of B, given that A has occurred, is

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \; if \; P(A) \neq 0$$

The conditional probability of A, given that B has occurred, is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \; if \; P(B) \neq 0$$

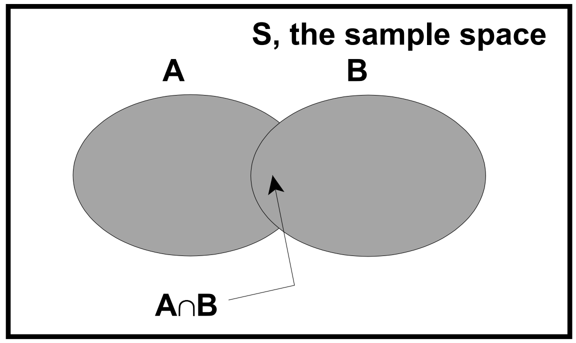
Definition: Two events A and B are said to be independent if and only if either P(A| B) = P(A) or P(B | A) = P(B), otherwise, the events are said to be dependent.

Two events are independent if the occurrence or nonoccurrence of one of the events does not change the probability of the occurrence of the other event.

Probability Rules

Additive Rule of Probability

Given two events of a finite sample space S, A and B, the probability of their union, $A \cup B$, is equal to $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. If A and B are mutually exclusive, then $P(A \cap B) = 0$ and $P(A \cup B) = P(A) + P(B)$.



Multiplicative Rule of Probability

The probability that both of the two events of a finite sample space S, A and B, occur is $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$.

If A and B of a finite sample space are independent, $P(A \cap B) = P(A)P(B)$ and $P(A|B) = P(A)$ or $P(B|A) = P(B)$.

Similarly, if A, B, and C are mutually independent events, then the probability that A, B, and C occur is $P(A \cap B \cap C) = P(A)P(B)P(C)$.

Bayes' Rule

Thomas Bayes (1702-1761) was a Presbyterian minister who introduced subjective/reverse probability thinking which caused him great persecution and no recognition by mathematicians of his day. But today a whole field of statistics called Bayesian inference has been developed from his original concept and is widely used, gaining a greater foothold with the years even though some statisticians still fiercely oppose it. One of his basic results is known today as Bayes' Theorem.

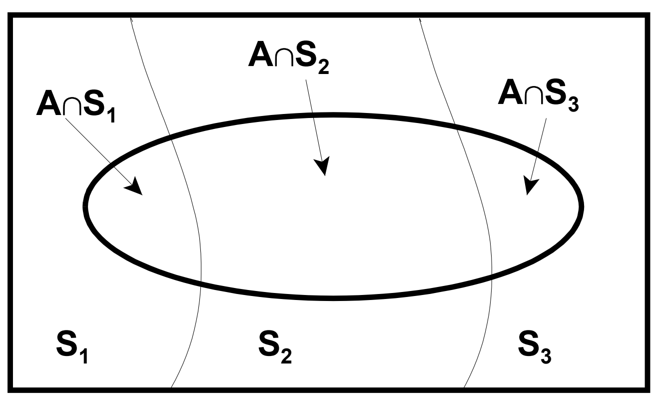
Definition: Mutually exclusive and exhaustive events are non-overlapping events that, taken together, make up the entire sample space.

In the adjacent figure, the sample space is divided into 3 subpopulations.

$A = (A \cap S_1) \cup ((A \cap S_2) \cup (A \cap S_3)$

$A = \cup S_i$ for $i = 1$ to k and $S_i \cap S_j = \varnothing$, $i \neq j$

S, the sample space



Law of Total Probability

Given a set of events $S_1, S_2, S_3, \ldots, S_k$ that are mutually exclusive and exhaustive and an event A, the probability of the event A can be expressed as:

$P(A) = P(S_1)(PA|S_1) = P(S_2)P(A|S_2) + P(S_3)P(A|S_3) + \ldots + P(S_k)P(A|S_k)$.

Bayes' Rule

Let $S_1, S_2, ..., S_k$ represent k mutually exclusive and exhaustive subpopulations with prior probabilities $P(S_1), P(S_2), ..., P(S_k)$. If an event A occurs, the posterior probability of S_i given A is the conditional probability

$$Pr(S_i \mid A) = \frac{P(S_i \cap A)}{P(A)} = \frac{P(S_i)P(A|S_i)}{P(A)} = \frac{P(S_i)P(A|S_i)}{\sum_{j=1}^{k} P(S_j)P(A|S_j)}$$

for i = 1, 2, ..., k, and $A = \cup S_i$ for *i* = 1 to *k* and $S_i \cap S_j = \varnothing$, $i \neq j$.

The probabilities $P(S_1), P(S_2), ..., P(S_k)$ are called prior probabilities. When the prior probabilities are unknown, it is possible to assume that all subpopulations are equally likely, so that $P(S_1) = ... = P(S_k) = 1/k$. The conditional probabilities P(Si|A) are called the conditional probabilities because these are the probabilities that result after taking account of the sample information contained in the event A.

Bayes' Rule

Other wording for Bayes' Rule is as follows:

Definition: If $A_1, A_2, ..., A_n$ are events of a sample space S such that $A_1 \cup A_2 \cup ... A_n = S$, then $A_1, A_2, ..., A_n$ are called *exhaustive* events of S.

Bayes' Theorem: Let $A_1, A_2, ..., A_n$ be pairwise mutually exclusive and exhaustive events of a sample space S where $P(A_i) \neq 0$ for *i* = 1, 2, ..., *n* . Let E be any event of S such that $P(E) \neq 0$. Then

$$P(A_i|E) = \frac{P(E|A_i(PA_i)}{\sum_{j=1}^{n} P(E|A_j)P(A_j)}$$ for i = 1, 2, ..., n.

Examples for Bayes' Theorem

Cage I contains four brown guinea pigs and six white ones while cage II has three brown and two white guinea pigs. We select a cage at random and then choose one guinea pig randomly from that cage.
i)What is the probability that cage I was chosen given that the brown guinea pig was drawn?
Let A1 = event that cage I is chosen, A2 = event that cage II is chosen, E = event that a brown guinea pig is drawn. Then we have as given that $P(A_1) + P(A_2) + \frac{1}{2}, P(E|A_1) = \frac{4}{10}$
$P(E|A_2 = \frac{3}{5}, P(\overline{E}|A_1 = \frac{6}{10} = \frac{3}{5}, P(\overline{E}|A_2) = \frac{2}{5}$

and we are asked to find $P(A_1|E)$.

Bayes' Rule

$$P(A_1|E) = \frac{P(E|A_1)P(A_1)}{P(E|A_1)P(A_1) + P(E|A_2)P(A_2)} = \frac{\left(\frac{2}{5}\right)\left(\frac{1}{2}\right)}{\left(\frac{2}{5}\right)\left(\frac{1}{2}\right) + \left(\frac{3}{5}\right)\left(\frac{1}{2}\right)} = \frac{2}{5}.$$

ii) What is the probability that cage II was chosen if a brown guinea pig was drawn?

$$P(A_2|E) = \frac{P(E|A_2)P(A_2)}{P(E|A_1)P(A_1) + P(E|A_2)P(A_2)} = \frac{\left(\frac{3}{5}\right)\left(\frac{1}{2}\right)}{\left(\frac{2}{5}\right)\left(\frac{1}{2}\right) + \left(\frac{3}{5}\right)\left(\frac{1}{2}\right)} = \frac{3}{5} = 1 - P(A_1|E)$$

iii) What is the probability that a brown guinea pig is drawn?

$$P(E) = P\left[(E \cap A_1) \cup (E \cap A_2)\right] = P\left(E \cap A_1\right) + P\left(E \cap A_2\right)$$

$$= P(E|A_1)P(A_1) + P(E|A_2)P(A_2) = \left(\frac{2}{5}\right)\left(\frac{1}{2}\right) + \left(\frac{3}{5}\right)\left(\frac{1}{2}\right) = \frac{1}{2}$$

iv) What is the probability that cage I was chosen knowing that a white guinea pig was randomly drawn?

$$P(A_1|\overline{E}) = \frac{P(\overline{E}|A_1)P(A_1)}{P(\overline{E}|A_1)P(A_1) + P(\overline{E}|A_2)P(A_2)} = \frac{\left(\frac{3}{5}\right)\left(\frac{1}{2}\right)}{\left(\frac{3}{5}\right)\left(\frac{1}{2}\right) + \left(\frac{2}{5}\right)\left(\frac{1}{2}\right)} = \frac{3}{5} \qquad \therefore\ P(A_2|\overline{E}) = 1 - P(A_1|\overline{E}) = \frac{2}{5}$$

Bayes' Rule

<u>Example</u> A healthy rat may be infected with three types of organisms A, B and C in a laboratory experiment. If we assign an equal number of rats to each possible organism and notice that a rat has a probability and of becoming ill when introduced to organism A, B and C respectively.

(i) A rat is chosen at random after the experiment is concluded and found to be ill. What is the probability that he received organism A? What is the probability that he was infected with either organism A or organism B?

(ii) What is the probability that he received organism C if the rat is found healthy?

Let A_1 = event that the rat is infected with organism A, A_2 = event that the rat is infected with organism B, A_3 = event that the rat is infected with organism C and E = the event that the rat is found ill.

$$\therefore\ P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$$

$$P(E|A_1) = \frac{1}{2}, P(E|A_2) = \frac{1}{3}, P(E|A_3) = \frac{1}{4}$$

$$\Rightarrow\ P(\overline{E}|A_1) = \frac{1}{2}, P(\overline{E}|A_2) = \frac{2}{3}, P(\overline{E}|A_3) = \frac{3}{4}$$

Bayes' Rule

(i) $$P(A_1|E) = \frac{P(E|A_1)P(A_1)}{P(E|A_1)P(A_1) + P(E|A_2)P(A_2) + P(E|A_3)P(A_3)}$$

$$= \frac{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{4}\right)\left(\frac{1}{3}\right)} = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{9} + \frac{1}{12}} = \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{3} + \frac{1}{4}} = \frac{\frac{1}{2}}{\frac{13}{12}} = \frac{6}{13}$$

$$P(A_1 \cup A_2|E) = P(A_1|E) + P(A_2|E) \because A_1 \cap A_2 = \Phi$$

$$P(A_2|E) = \frac{P(E|A_2)P(A_2)}{\sum_{i=1}^{3} P(E|A_i)P(A_i)} = \frac{\left(\frac{1}{3}\right)\left(\frac{1}{3}\right)}{\frac{1}{6} + \frac{1}{9} + \frac{1}{12}} = \frac{\frac{1}{3}}{\frac{13}{12}} = \frac{4}{13}$$

$$\therefore P(A_1 \cup A_2|E) = \frac{6}{13} + \frac{4}{13} = \frac{10}{13}$$

(ii) $$P(A_3|\overline{E}) = \frac{P(\overline{E}|A_3)P(A_3)}{P(\overline{E}|A_1)P(A_1) + P(\overline{E}|A_2)P(A_2) + P(\overline{E}|A_3)P(A_3)}$$

$$= \frac{\left(\frac{3}{4}\right)\left(\frac{1}{3}\right)}{\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) + \left(\frac{2}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{3}{4}\right)\left(\frac{1}{3}\right)} = \frac{\frac{3}{4}}{\frac{1}{2} + \frac{2}{3} + \frac{3}{4}} = \frac{\frac{3}{4}}{\frac{23}{12}} = \frac{9}{23}$$

Bayes' Rule

The use of Bayes' Rule and diagnostic testing is commonplace as follows:
P(D+) is the prior probability of a disease in a population.

P(D-) is the prior probability of non-disease in a population. D+ and D- are mutually exclusive of each other and P(D+) + P(D-) = 1.
P(T+|D+) is the probability of a positive test result (T+) given that the animal is diseased (D+). It represents the test sensitivity.

P(T+|D-) is the probability of a positive test result (T+) given that the animal is not diseased (D-). It is equal to 1 -specificity.

P(D+|T+) is the posterior probability of disease (D+) given a positive test result (T+). It is the predictive value of a positive test.

$$P(D+|T+) = \frac{P(D+\cap T+)}{P(D+\cap T+) + P(D-\cap T+)}$$

$$= \frac{P)(D+)(P(T+|D+)}{P(D+)(P(T+|D+)) + P(D-)(P(T+|D-))}$$

A false positive is the event that the test is positive for a given condition, given that the person does not have the condition.
A false negative is the event that the test is negative for a given condition, given that the person has the condition.
<u>Prior probability</u>: Probabilities of subpopulations (also called states of nature) occurring prior to observing sample information.
<u>Posterior probability</u>: Probabilities of the subpopulations that have been updated after observing the sample information contained in an event.

Bayes' Rule

	Test positive	Test negative	
Disease positive	P(D+)P(T+|D+) True positives (TP)	P(D+)((1-P(T+|D+)) False Negatives (FN)	P(D+) (TP + FN)
Disease negative	P(D-)P(T+|D-) False Positives (FP)	P(D-)((1-P(T+|D-)) True Negatives (TN)	P(D-) (FP + TN)
	P(D+)P(T+|D+) + P(D-)P(T+|D-) (TP + FP)	P(D+)((1-P(T+|D+)) + P(D-)((1-P(T+|D-)) (FN + TN)	P(D+) + P(D-) = 1

$$Se = \frac{TP}{TP + FN} \times 100$$

$$Sp = \frac{TN}{TN + FP} \times 100$$

$$PVP = \frac{TP}{TP + FP} \times 100$$

Bayes' Rule

	Test positive	Test negative	
Disease positive	*p* Se	*p* (1-Se)	*p*
Disease negative	(1-*p*)(1-Sp)	(1-*p*)Sp	1-*p*
	p Se + (1-*p*)(1-Sp)	*p* (1-Se) + (1-*p*)Sp	1

In relation to Bayes' Rule the prior probability is represented by the disease prevalence (*p*). The posterior probability is represented by the predictive value of a positive test (PVP) or P(D+|T+). In the table above in which words are substituted for notation, the PVP can be expressed as follows:

$$PVP = \frac{pSe}{pSe + (1-p)(1-Sp)}$$

Bayes' Rule

The following expressions illustrate the application of Bayes' Rule and other probabilities associated with the diagnostic test:

Sensitivity of the test

$$P(T+|D+) = \frac{P(T+)(P(D+|T+))}{P(T+)(P(D+|T+)) + P(T-)(P(D+|T-))}$$

$$1 - Se = P(T-|D+)$$

$$Specificity\,(Sp) = P(T-|D-)\ and\ 1 - Sp = P(T+|D-)$$

Predictive value of a positive test (PVP)

$$P(D+|T+) = \frac{P(D+\cap T+)}{P(D+\cap T+) + P(D-\cap T+)}$$

$$= \frac{P(D+)P(T+|D+)}{P(D+)(P(T+|D+)) + P(D-)(P(T+|D-))}$$

Predictive value of a negative test (PVN)

$$P(D-|T-) = \frac{(1 - P(D+))P(T-|D-)}{(1 - P(D+))(P(T-|D-)) + P(D+)(P(T-|D+))}$$

$$= \frac{P(D-\cap T-)}{P(D-\cap T-) + P(D+\cap T-)}$$

$$= \frac{(1 - TP)Sp}{(1 - TP)Sp + TP(1 - Se)}$$

Bayes' Rule

Apparent prevalence (AP)

$$P(T+) = P(D+\cap T+) + P(D-\cap T+)$$

$$= P(D+)(P(T+|D+)) + (1 - P(D+))P(T+|D-)$$

$$= TP(Se) + (1 - TP)(1 - Sp)$$

True prevalence (TP)

$$P(D+) = \frac{P(T+) - P(T+|D-)}{1 - (P(T+|D-) + P(T-|D+))}$$

$$= \frac{AP - (1 - Sp)}{1 - ((1 - Sp) + (1 - Se))}$$

Probability of a negative test

$$P(T-) = P(D+)(P(T-|D+)) + (1 - P(D+))(P(T-|D-))$$

$$= TP(1 - Se) + (1 - TP)(Sp)$$

Bayes' Rule

Proof for True Prevalence

$$\because P(T+) - P(T+|D-) = \left[1 - P(T+|D-)\right] - \left[1 - P(T+)\right]$$

$$= P(T-|D-) - P(T-)$$

$$= P(T-|D-) - \left[P(T- \cap D-) + P(T- \cap D+)\right]$$

$$= P(T-|D-) - \left[P(D-)P(T-|D-) + P(T- \cap D+)\right]$$

$$= P(T-|D-)\left[1 - P(D-)\right] - P(T- \cap D+)$$

$$= P(T-|D-)P(D+) - P(D+)P(T-|D+)$$

$$= P(D+)\left[P(T-|D-) - P(T-|D+)\right]$$

$$= P(D+)\left[1 - P(T+|D-) - P(T-|D+)\right]$$

$$= P(D+)\{1 - \left[P(T+|D-) + P(T-|D+)\right]\}$$

$$\therefore P(D+) = \frac{P(T+) - P(T+|D-)}{1 - \left[P(T+|D-) + P(T-|D+\right]} = \frac{AP - (1 - S_P)}{1 - \left[(1 - S_P) + (1 - S_e)\right]}$$

Bayes' Rule

Probability that no infected animals exist in a group given that all animals test negative

$$P(I = 0|R = 0) = (PVN)^n$$

$$= \left(\frac{((1 - p)Sp}{(1 - p)Sp + p(1 - Se)}\right)^n$$

Where,
I = number of infected animals in the group
R = number of animals that test positive in the group
n = number of animals in the group

Probability that at least one infected animal exists in a group given that all animals test negative

$$P(I \geq 1|R = 0) = 1 - (PVN)^n$$

$$= 1 - \left(\frac{((1 - p)Sp}{(1 - p)Sp + p(1 - Se)}\right)^n$$