
DR. RANDOM
PROFESSOR OF PROBABILITY
IN
OUT
PLEASE USE THIS COIN
Sumanta
sumanta.baruah@gmail.com

MAT 2379 - Introduction to Biostatistics
Chapter 1. Interpreting Probabilities

Statistics can refer to the mathematical techniques used to analyze data. Usually the data comes from a random experiment.

Definition. A *random experiment* is an experiment whose outcome is determined by chance and cannot be predicted with certainty.

Definition. The *sample space* of a random experiment is the set of all possible outcomes and is denoted by S.

Example 1. Here are examples of random experiments. Give the corresponding sample space.

(i) Flip a coin.

$$S = \{T, H\}$$

(ii) The roll of a die.

$$S = \{1, 2, 3, 4, 5, 6\}$$

(iii) diagnose H_1N_1 in a person.

$$S = \{sick, healthy\}$$

(iv) Survival time (in months) after the diagnosis of a particular disease.

$$S = [0, \infty)$$

Definition. An *event* is a sub-collection of possible outcomes of a random experiment.

Example 2 (Diffusion of Molecules)**.** Suppose that there are 10 molecules in a cell and that we observe the number of molecules remaining in the cell after 15 minutes. The sample space is

$$S = \{0, 1, \cdots, 10\}.$$

Here are examples of events.

- $A =$"less than 3 molecules", *i.e.* $A = \{0, 1, 2\}$.
- $B =$"more than 7 molecules", *i.e.* $B = \{8, 9, 10\}$.
- $C =$"at most 4 molecules", *i.e.* $C = \{0, 1, 2, 3, 4\}$.
- $D =$"at least 6 molecules", *i.e.* $D = \{6, 7, 8, 9, 10\}$.

Remark.

- If the outcome of the random experiment is x and x belongs to the event A, the we will say A has occurred.
- The sample space S is called a certain event.
- The empty event $\emptyset$, *i.e.* no outcomes are found in $\emptyset$, is called an impossible event.

Definition. The *probability* of an event is a numeric value which is associated with the chance that the event will take place. It is expressed as a percentage, or a number between 0 and 1.

There are 3 different methods of assigning probabilities:

1. The personal method. The probability of an event is a subjective quantity expressing a person's degree of belief that the event will happen. It is usually associated with *one-shot* situations. This method is not accurate and will not be used in this course.

Example 3. A doctor has diagnosed a disease. After considering a few factors the doctor must assess the likelihood that the disease can be cured. The doctor forms a personal opinion and says that there is a 50% chance that the patient will be cured.

Remark.

- The advantage of the personal approach is that it can always be used.
- The disadvantage of the personal approach is that it is very subjective. We can ask different people for their personal opinions and obtain different answers.

2. The classical method. This is appropriate when we have *equally likely* events. This is the method that we use in this course. When using this method, the probability of an event A is defined as:

$$P(A) = \frac{n(A)}{n(S)},$$

where $n(A)$ is the number of possible outcomes of the event A.

Example 4. The familiar chance operation of flipping a coin. If the coin is fair, the probability that it falls heads is 1/2.

Remark.

- In an exercise, if it is stated that all the outcomes of the experiment are equally likely, we will use the above classical approach to calculate probabilities.
- The classical method requires that the experiment be broken down into equally likely outcomes. It is not always possible to do this, and It is not always clear when outcomes should be considered as equally likely.

3. The relative frequency method. It is associated with events that occur as results of *repeatable* experiments. How it is calculated? We run the experiment n times and we observe that the event that we are interested in appeared $f(A)$ times. The probability of the event A is calculated as:

$$P(A) = \frac{f(A)}{n}.$$

Example 5. A medical team conducted a study of a therapy for the treatment of a disease. A 100 patients were considered, out of which 77 gave a "satisfactory/good" response. The probability that the therapy is successful is 77/100=77%.

Example 6. Some trees in a forest were showing signs of disease. A random sample of 200 trees of various sizes was examined yielding the following results: the probability

Type	Disease free	Doubtful	Diseased	Total
Large	42	19	9	70
Medium	40	28	17	85
Small	24	12	9	45
Total	106	59	35	200

that a tree selected at random is large is $70/200 = 35\%$, and the probability that a tree selected at random is diseased is $35/200 = 17.5\%$.

Example 7. A population of 150 fruit flies is maintained in a lab. In the population, 30 are black because of a mutation, and the rest have normal gray color. One fly is chosen at random. The chance that it is a black fly is 30/150=20%.

Remark.
- This approach allows us to consider experiments with outcomes which are not necessarily equally likely.

- To determine the probability of an event, we must repeat the experiment an infinite number of times or at least a very large number of times to obtain a good approximation.

Remark. The modern theory of probability is based on three basic principles, called the axioms of probability. Both the classical and relative frequency approaches satisfy these axioms. Thus this approach is more general and any consequences of the axioms will hold true for the other two models.