

# Scientific Research

João Gilberto Corrêa da Silva

Department of Mathematics and Statistics, Federal University of Pelotas (retired). Pelotas, RS, Brazil.

Orcid ID [0000-0003-2985-0925](https://orcid.org/0000-0003-2985-0925).

E-mail: [jgcs1804@yahoo.com.br](mailto:jgcs1804@yahoo.com.br).

**Abstract:** *The scientific method is the strategy of science to generate knowledge. Its implementation is carried out by scientific research. Scientific research is employed in the three phases of the scientific method: in the initial phase of synthesis, to identify scientific problems; in the analysis phase, to try to solve these problems; and in the final synthesis phase, to incorporate the new knowledge generated into the existing body of knowledge. Thus, scientific research comprises systemic research, which focuses on global systems, in the initial and final phases, and analytical research, which focuses on subsystems, in the analysis phase. New scientific knowledge is achieved by analytical research, which begins with the formulation of a scientific problem. Once the initial question is posed, the systematic structure of the scientific approach becomes immediately evident. This is because research in science is very precisely structured and logically organized. This form has been developed over centuries of scientific problem-solving experience. This text summarizes the essential conceptual and methodological basis of scientific research.*

## 1. Introduction

Science aims to increase knowledge and improve understanding of natural phenomena, with the purpose of their control and, in the failure of this, prediction, in order to allow the increasing dominion of man over nature. The strategy of science to produce knowledge is the scientific method. Its implementation is carried out by scientific research.

The contribution of scientific research to the advancement of knowledge and for humanity to reach the present stage of civilization is recognized. In fact, one can clearly identify the close relationship between man's growing dominion over nature and the scientific effort and the evolution of the methodology and instruments of science in recent centuries.

Scientific research is not a highly controlled production process. There is no set of rules that guarantee the success of a scientific research. However, one can clearly detect the indispensability of solid basic training of the researcher, which includes the understanding of the conceptual and methodological foundations of scientific research, and the knowledge of techniques consolidated by the experience acquired by the scientific research work carried out in the past.

This article provides a brief review of scientific research for researchers in the various fields of science. The topics that seem most important are considered: the basic concepts, the two essential elements of scientific research - observation and reasoning, the process and organization of the scientific research, the objectives of a scientific research, the stages and orderly structure of an analytical scientific research, research work and the role of researcher in science, and Statistics in scientific research. This review is a sequel to Silva [1] and is based primarily on contributions from Bertalanffy [2], Bunge [3], Castro & Cobbe [4], Christensen [5], Cox [6], Fisher [7], Green & Tull [8], Heath [9], Kempthorne [10], Kish [11, 12], Lastrucci [13], Mitchell & Jolley [14], Silva [15, 16, 17], Smith & Sugden [18], Urquhart [19] and Wilson [20].

## 2. Basic Concepts

### 2.1. Scientific research

The implementation of the scientific method to produce knowledge is carried out by scientific research:

- **Scientific research** is the process of investigating natural phenomena using the scientific method for the purpose of discovering new facts and developing scientific theory.

Scientific research also aims at the review of facts, laws and theories in view of newly discovered facts, and the practical applications of such facts, laws and theories. Therefore, scientific research is the continuous search for knowledge and understanding of reality carried out through the scientific method. Its result is scientific knowledge.

Scientific research can have a purely cognitive objective, that is, the generation of scientific knowledge without an immediate application purpose, or a practical objective, that is, the generation of knowledge for immediate application:

- *Scientific research with purely cognitive objective is called **pure research** or **basic research**, and with practical objective, **applied research** or **technological research**.*

The progress of scientific knowledge basically consists of the permanent and progressive deepening of knowledge of the complex interconnection of natural phenomena, which, in turn, generally comprise an extremely complex set of other more elementary phenomena, also closely related.

## 2.2. Characteristic

One basic property of phenomena in nature is their heterogeneity, which characterizes what is commonly called "natural variability".

Scientific research focuses on a class of phenomena with common attributes and properties:

- *The attributes and properties of phenomena are the **characteristics** of these phenomena. Each characteristic can manifest under different alternatives or **levels**.*

For example, sex is a characteristic of a sheep herd and of the animals of this herd. This characteristic can manifest itself in each of these animals in one of two levels - male and female. Body weight at slaughter is another characteristic of this herd and its units, which, for each animal, can assume any numerical value in a certain interval of real numbers.

Characteristics are qualitative or quantitative, depending on whether their levels are related qualitatively or quantitatively. Qualitative characteristics classify units into classes or categories, according to their levels; quantitative characteristics distinguish units quantitatively. Thus, in the example in the previous paragraph, sex is a qualitative characteristic that classifies animals into two classes; body weight at slaughter is a quantitative characteristic that distinguishes animals according to the magnitude of their weight.

## 2.3. Unity or system

Each scientific research focuses on a class of phenomena with common characteristics. Generally, a technical-operational characterization of the elementary constituent of such a class of phenomena, which usually receives the designation of unit or system, and of its elements and attributes, is convenient:

- *A **unit** or **system** is a set of related entities that constitute a globally organized whole dynamically related to the external environment, and which together perform some function.*

A unit is specified by describing the following (Figure 1):

- a) function or objectives of the unit,
- b) inputs - elements that enter the unit,
- c) products - elements that leave the unit,
- d) components - internal elements that transform inputs into products,
- e) flow - movement of elements between the components of the unit, and
- f) border or limit - imaginary line that demarcates the scope of the unit, which includes all its components and elements.

The following examples are illustrative: a) in research of the incidence of brucellosis in animals in a region, the unit can be an animal, a herd, a farm or a county; b) in peach fruit fly control research, the unit may be a plant or an orchard.

The definition of the unit depends on the research objective. It must be established in the initial phase of the research, that is, in the formulation of the problem. Sometimes the definition may be obvious, as in research for the recommendation of wheat cultivars, where the unit is a crop, and in research of the effectiveness of a dewormer for the control of helminths in dairy cows, where the unit is an animal. However, very often, the definition or choice of unit is not so obvious.

The difficulty in defining the unit or system for a research stems from the fact that in nature there is a hierarchy of systems, that is, systems within systems, in a decreasing order of amplitude, such that a given system is a subsystem with respect to a higher hierarchical level, and, in turn, contains subsystems at a lower level. In these circumstances, the definition of the unit must establish the demarcation of its limit. The units of a class of phenomena are usually complex and heterogeneous entities. For this reason, the designation of system is more suggestive than that of unit, which, however, is more usual.

Units (systems) are identified and distinguished by a set of characteristics, whose description constituted a useful working definition of unit:

- *The **units** of a class of phenomena comprise a set of characteristics that identify and distinguish them.*

Often, the unit is identified by one of its characteristics or aggregate of characteristics. However, it must be clearly understood that the complete definition of unit must encompass the whole set of characteristics that comprise it. For example, in research with sheep, the unit comprises the characteristics of animals (race, sex, weight, wealth, ...), environment (installations, climate, pasture, diseases, predators, ...), management (feed, water, wormers, ...) and measurement.

## 2.4. Target population

Scientific research aims to solve problems concerning the units of a collection of units of specific interest:

- *In scientific research, the **target population**, or more simply, **population**, is the well-defined collection of units (systems) of interest for which it is desired to infer. The number of units is called the **size of the population**.*

The target population is defined by specifying its units or the conditions for units to be part of it. The specification of the population, as well as of its units, is determined by the research objectives and must be established in the formulation of the research problem. Populations existing in nature are **finite populations**, whose size is expressed by a natural number, very often large and unknown. Furthermore, such populations have a dynamic constitution, as a result of the mutability of the systems that comprise them over time. In some research, the target population consists of units existing at the time the research is carried out. A population in these circumstances, whose units can be identified, is a **real population**. Very often, however, the target population comprises units that do not exist at the time the research is being carried out, but which, presumably, may exist in the future. A population whose units are not identifiable, but defined by the conditions to belong to it, is a **conceptual population**. In wheat breeding research, for example, the units of the target population are not the wheat crops existing in the region of interest at the time the research is carried out, but the crops that will exist in that region in the future.

## 2.5. Sample and sampling

In some situations, the research is carried out on all units of the target population. Research in these circumstances is called **census**. Very often, however, it is impracticable or inconvenient to conduct the research on all units of the target population. In these situations, research is carried out on a set of units chosen from the target population or constructed to represent it:

*- A subset of the units of a target population or a set of units constructed to represent it is a **sample** of this population. The number of sample units is the **sample size**. The process of choosing or building a sample is called **sampling**.*

The characteristics of the sample are the same as those of the target population. The levels of each sample characteristic can be the same as the target population or a subset of it.

The choice of the sample can be determined by an objective and random or subjective and arbitrary sampling process:

*- The sample is a **random sample** if it is determined by random sampling, that is, by a process that assigns to each unit of the target population a known (non-zero) probability of constituting the sample.*

The execution of a research through a sample aims to obtain information about the target population that is supposed to be represented by this sample. Inferences from the sample usually do not coincide with inferences that would be derived directly from the target population:

*- The discrepancies between the inferences derived from a particular sample and those that would be obtained from the target population, that is, the **inference error** resulting from the observation of a sample instead of the target population, constitute the **sampling error**.*

It is desirable that the average of the inferences of all possible samples of a target population coincide with the inferences that would be provided by the population itself:

*- Inferences derived from a sample to the target population are **valid (unbiased)** if the average of the inferences that would be provided by all samples of the same size from that population are the same as those that would be provided by the population itself.*

In general, two sample properties are desirable: validity and representativeness:

*- The sample is a **valid sample** if it provides valid inferences for the target population; it is a **representative sample** if the distributions of the relevant characteristics in the sample are equal to the respective distributions in the target population, except for random variation due to sampling.*

The sample is valid if it is a random sample; it is representative if it has a structure consistent with the structure of the target population and supplies plausible inferences. A valid and representative sample is critical for inferences. When the sample is constituted by units of the target population, these properties can be achieved using adequate design and random sampling. Very often, however, random sampling is unfeasible or undesirable, due to the impossibility or inconvenience of considering all units of the target population in the sample selection process. In these circumstances, the sample is determined by non-random sampling, in which the researcher uses judgment in order to achieve the best possible representation. This is obviously the case with conceptual populations, whose units do not exist at the time of sample selection. In these circumstances, in some situations, the sample units are not units of the target population, but units built to simulate them especially for carrying out the research.

It is important to characterize the scope for which the sample inferences are valid:

*- The collection of units for which the sample provides valid inferences is called the **sampled population**.*

Inferences from the sample to the target population are valid when the sampled population is the target population itself. This is the case for random sampling research. However, in non-random sampling situations, the validity of inferences for the target population must be subjectively evaluated.

In general, the process of inference from the sample to the target population comprises two steps: generalization from the sample to the sampled population and generalization from this population to the target population. Generalizations referring to the first step, i.e., from the sample to the sampled population, are conceptually valid and can be derived by objective statistical procedures. Extensions of these generalizations to the target population are valid if the disparity between the sampled population and the target population, that is, the sampling error, is irrelevant. This is the case in research with random sampling. However,

in research with non-random sampling the judgment of the validity of this second step of generalization to the target population is necessarily subjective.

### 3. Observation and Reasoning

The generation of knowledge through scientific research essentially takes place through observation and reasoning:

- **Observation** is the sensory perception by which knowledge of phenomena is obtained. **Reasoning** is the mental elaboration that conceives and discovers the meanings of these phenomena, their interrelationships and relations with the existing body of scientific knowledge, insofar as the present knowledge and the researcher's skill allow.

A basic postulate of science is that all information concerning phenomena is derived from sense impressions. The data of science are mental impressions of sensory experiences, that is, ideas derived from the senses, mainly from sight, hearing, smell, taste and touch. Man reacts to sense impressions by mental manipulations, and these reactions or ideas constitute the foundation of knowledge. Knowledge generation in science is essentially a process that comprises:

- a) stimulus reception of the phenomenon that is studied,
- b) mental manipulation of the impressions of this stimulus to interpret it,
- c) combining these impressions with other earlier impressions and their interpretations with the aid of memory, and
- d) deduction of a conclusive interpretation of the phenomenon.

This process involves two basic elements: the sensory organs, aided by instruments, and the mind. The sensory organs employed depend on the properties of the phenomenon, while mental manipulations are related to the researcher's knowledge, mental ability, awareness, interest and attitude.

#### 3.1. Observation

Man has limited power of observation. For this reason, it is necessary to limit the observations to the relevant characteristics. Some characteristics can be observed with the use of simple instruments to aid the senses, such as a scale, ruler and test tube. However, others require the use of more sophisticated observation instruments, such as a stethoscope, microscope and telescope.

Before proceeding with observation, the researcher must clarify and establish basic definitions relevant to scientific observation, particularly the following:

- a) the phenomena that must be observed, the relevant behavior of these phenomena and the significant facts that must be addressed,
- b) the places, moments and conditions in which the observation must be carried out, and the way of structuring and describing the observed situation,
- c) the relevant characteristics of the phenomena to be observed,
- d) the operational definitions and symbolic representations of these characteristics, and the procedures and instruments that should be used for measurement,
- e) the stabilization and standardization of observation conditions to allow verification.

The observation will be truly scientific only when all these definitions are established. However, it will only be successful if it achieves the necessary accuracy, that is, if it portrays the phenomenon as it really is. Accuracy demands some basic requisites:

- a) sensitivity,
- b) objectivity,
- c) use of sensory aids and
- d) definition of the relevant characteristics to be observed.

The sensitivity to observe is a combination of experience and insight. Experience directs the observation of phenomena to their empirical properties and interprets them rationally; insight refers to the ability to perceive these properties through discernment and understanding of their inner nature and reality. Objectivity is achieved using calibrated and impartial observation instruments and with impersonal attitudes, actions and judgments. The use of sensory aids is essential to clarify and discern what is observed. The definition of relevant characteristics implies questions of conceptualization, validity and relevance. Decisions on these issues must be made in accordance with the research objectives and are closely related to the problem and the scientific hypothesis.

The observation of a characteristic of a set of units comprises the description and representation of that characteristic through symbols that express its levels and allow classifying or quantifying these units. This process is called measurement.

##### 3.1.1. Measurement

- The **measurement** of a characteristic is a process of description and representation of this characteristic that expresses its levels by symbols that present among themselves the same relevant relationships existing between those levels.

Thus, measurement demands the definition of a correspondence rule between the levels of the characteristic and the symbols that correspond to them:

- A rule or function that establishes a one-to-one correspondence between the set of levels of a characteristic and a set of symbols is called **variable**. Each of these symbols is a **level** of this variable.

Usually, the variable receives the same name as the characteristic it expresses, and often these two terms are used interchangeably. However, characteristic and variable are different concepts. In fact, the same characteristic can be expressed by many alternative variables, depending on the measurement level adopted and the set of symbols chosen. The choice of one of these alternatives is arbitrary and depends on convenience and availability of measurement resources. For example: a) In a research with sheep, the sex characteristic can be expressed by a variable that makes an animal correspond to the initial letter of its sex, that is, the symbol m or f, depending on whether the animal is male or female, or a variable with levels 0 and 1, which assigns the symbol 0 to one of the sexes and 1 to the other, or by any variable with two levels; the characteristic amount of parasites in the feces can be expressed by a variable that assumes values of a subset of non-negative integers with zero lower end; and the body weight of the animal at slaughter by a variable with values of an interval of real numbers ( $a$  ;  $b$ ), where  $a$  and  $b$  are, respectively, numbers lower and higher than the expected minimum and maximum weights. b) In wheat research, the characteristic degree of giberela infection can be expressed by a two-level variable - without and with infection, or a variable that expresses the degree of infection, such as no infection, weak, medium, strong and high infection; or a variable that expresses percentage of infection with values in the interval  $[0; 100]$ .

The measurement of a characteristic involves two problems: what to measure and how to measure it. The first problem, that is, what to measure must be solved according to the research objectives and is closely related to the scientific problem and hypothesis. A characteristic is an abstract notion or concept of a property of the units of a target population, which can be simple or complex, objective and easily interpretable or subjective and of dubious interpretation. Therefore, the measurement of a characteristic requires the translation of a theoretical concept into an operational equivalent. Disagreements may arise from different operational definitions, presumably of the same characteristic. This can occur when a concept expresses a complex characteristic. Suppose, for example, that a theory holds that the productivity of a cultivated plant species is positively related to the quality of the seed. The difficulty is finding agreement regarding an operational equivalent of seed quality. Several useful and convincing operational definitions can be proposed, but there may be no logical way to determine which of these definitions is the most valid.

The second problem of measurement, that is, how to measure, is related to the first: what to measure. The problem now is how classify or quantify, according to a characteristic, units presumably already classified for that characteristic. This classification or quantification is obtained by representing the characteristic through a variable.

Observation depends on some degree of subjective judgment. For this reason, it is important to establish the conditions of observation in order to avoid observer bias. Elaborate strategies must often be established to enable the observer to avoid his own bias and obtain the correct record of the facts.

### Measurement scales

Measurement of a characteristic is an attempt to find a correspondence between the alternatives of the characteristic and the levels of a variable, that is, a set of symbols or numbers that properly considers the relationships between these alternatives and the operations that can be performed on them. Variables are classified into four measurement levels (**measurement scales**) according to the relationships between their levels and the operations allowed between these levels: nominal, ordinal, interval and ratio.

A **nominal variable** classifies units into classes or categories as to the characteristic it represents, without establishing any relationship of magnitude or order. This attribute only differentiates units that belong to different classes. Two units are assigned the same symbol in relation to a variable if they have the same level of that characteristic. Thus, all units of a class are equivalent with respect to this characteristic. This is the property of equivalence or equality of classes or levels. Examples are sex (male, female), breed of sheep (romney march, corriedale, suffolk, ...), numbers of boxes in a sheep facility (1, 2, 3, ...). Note that the levels of a nominal variable are just symbols and, even if symbols are numbers, there is no order relationship between them. Therefore, arithmetic operations on these numbers have no meaning with respect to the objects they identify. The nominal scale allows only a few elementary operations, as the count of the number of units (frequencies) of the classes (number of animals of each breed, for example) and determine the most numerous class, that is, the mode of distribution. Statistical inferences regarding the distribution of these frequencies can be carried out. However, because a nominal scale does not infer degree or quantity, it only classifies the units, classes cannot be manipulated mathematically (eg, by adding or subtracting numerical equivalents from those classes). Consequently, most of the usual statistics such as mean and standard deviation are meaningless, as operations for their determination are not permissible.

Nominal variables have the following three properties regarding the equality or equivalence of units in the classes: 1) reflexivity - a unit is equivalent to itself, that is, if  $a$  is the value of a unit of class A, then  $a=a$ ; 2) symmetry - if unit A is equivalent to unit B, then unit B is equivalent to unit A, that is, if  $a$  and  $b$  are values of two units A and B, respectively, then  $a=b$  implies  $b=a$ ; 3) transitivity - if A and B are equivalent units and B and C are equivalent units, then A and C are equivalent units, i.e. if  $a$ ,  $b$  and  $c$  are the values of three units A, B and C, respectively, then  $a=b$  and  $b=c$  implies  $a=c$ .

An **ordinal variable** classifies units into classes or categories as to the characteristic it represents, establishing an order relationship between the units. For example, animals in a herd can be classified by the degree of tick infestation, assigning the number zero to indicate no infestation and the numbers 1, 2 and 3 to indicate increasing levels of infestation, such as low, medium and high. Thus, for example, if a unit belongs to category 2, it is known that its degree of infestation by ticks is lower than that of a unit belonging to category 3, as there is the same order relationship between these categories. Examples of ordinal variables are also: age with child, youth, adult and elderly levels, and academic performance grade with levels A, B, C, ...

The ordinal scale because it has both identity and magnitude.

The ordinal scale maintains the properties of the nominal scale regarding the equivalence of units assigned to the same order: reflexivity, symmetry and transitivity. In addition, the ordinal scale has the property of magnitude of classes, which imply the property of asymmetry of classes. Therefore, units can be designated not only as equivalent to other units, but also as non-equivalent. For example, an ordinal scale can designate that class A is higher than class B, and therefore that class B is smaller than class A.

The additional properties of the ordinal scale characterize its superiority over the nominal scale. However, arithmetic operations cannot be performed with symbols that only characterize order and vaguely designate quantity. Consequently, statistical descriptions are still limited. Position measures are restricted to the median, quartile, percentile, and other measures dealing with orders. Statistical inference procedures are restricted to those specifically appropriate for order data.

An **interval variable** orders the units according to the measured characteristic with numeric values. Thus, it allows the determination of differences between measured units. However, the zero value of a variable with an interval scale does not mean the absence of the measured characteristic. Consequently, the origin or zero point of this scale is arbitrary. In addition, ratios and proportions are meaningless.

Examples of interval variables are temperature measured in degrees Celsius or Fahrenheit. A day with zero temperature measurement on the Celsius or Fahrenheit scale is not a day without temperature. Each of these two scales assigns an arbitrary zero value, which is just a reference point. Furthermore, if in one day the temperature is 15 degrees centigrade and another 30 degrees, this does not mean that on the second day the heat is twice as high as the other day, because this ratio is only verified on this measurement scale. The Fahrenheit scale is related to the Celsius according to the expression  $T_F = (9/5)T_C + 32$ . Thus, to the value of zero degrees on the Celsius scale corresponds 32 on the Fahrenheit scale, and to the values 15 and 30 degrees on the Celsius scale correspond 59 and 86 on the Fahrenheit scale, that is, to a difference of 15 degrees on the Celsius scale corresponds a difference of 27 degrees on the Fahrenheit scale.

The interval scale allows arithmetic operations. Consequently, it enables most descriptive statistical measures such as mean, standard deviation, and correlation coefficient. For example, if the average temperature of a city in a month is determined in degrees Celsius or degrees Fahrenheit, the days of the month with higher-than-average temperature are the same under each of these two scales. However, some statistical measures, such as the coefficient of variation, can be misleading when applied to interval-scale variable data.

A **ratio variable** ranks the units in terms of the measured characteristic, has a constant unit of measurement and an absolute zero point, which is a unique origin. So, unlike the interval scale, the ratios have meaning. This is the most elaborate of the measurement scales, in the sense that it allows all arithmetic operations. It is the most common measurement scale in the physical sciences, such as scales to measure length, width, height, weight and volume. As the designation suggests, equal ratios between values on the ratio scale correspond to equal ratios between the measured units. Thus, ratio scales are invariant under positive proportion transformations, that is, transformations of the form  $y = c x$ ,  $c > 0$ . For example, if one plant is 3m high and the other 1m, it can be said that the first plant is 3 times taller than the second. This is because, if the heights of the two plants are converted into centimeters, their measurements will be, respectively, 30cm and 10cm, which are in the same 3:1 ratio. One can convert measurements from one ratio scale to another ratio scale merely by multiplying by an appropriate constant.

The ratio scale contains all the information of the lower order scales, namely, class equality, order and difference equality, and more. All descriptive statistics can be determined for data of one variable expressed in ratio scale.

### Other classifications and variable designations

Variables are often classified into two classes: qualitative and quantitative. Variables with nominal and ordinal measurement scales are often called qualitative variables because their levels are qualitatively different. They are also called categorical variables, as their levels designate categories or classes in which the units are classified according to the characteristic represented. Interval and ratio scale variables are called quantitative variables because their levels are quantitatively related.

Variables can also be classified as discrete and continuous. A variable that takes on isolated values from a set of real numbers, that is, values from a discrete set of real numbers, is a discrete variable; a variable that can take on any value in an interval of real numbers is a continuous variable. Qualitative or categorical variables, that is, of nominal or ordinal measurement scale, are, by definition, discrete and finite, that is, they assume a finite set of distinct values. Variables with an interval or ratio measurement scale can be discrete, assuming a finite or infinite set of distinct values, or continuous. For example, variables that express the amount of fruit on a tree and the number of piglets in a litter are discrete and finite; variables that express the weight and height of an animal or a plant are continuous variables.

Note that measurement with an interval or ratio scale variable can be considered as a refined form of classification. If lambs in a herd are weighed with a scale to the nearest grams, the result is a distribution of the animals among discrete weight categories, where neighboring categories differ by one gram. If weighing is carried out on a scale to the nearest milligram, the number of possible categories increases considerably, but the principle remains the same. Thus, in practice, the measurement process, even of conceptually continuous characteristics, constitutes a classification of units.

### Choice of measurement scale

For some characteristics the choice of measurement scale is limited in nature. More often, however, a characteristic may alternatively be expressed by more than one variable with measurement scales of different levels of precision. For example, a) the sex and race of an animal are necessarily expressed on a nominal scale; b) The seed size of a cultivar can be expressed by a two-level nominal variable: 0 - abnormal, 1 - normal, or by an ordinal variable: 1 - very small, 2 - small, 3 - medium, 4 - large and 5 - very large, or by a ratio variable whose levels are the numbers of a certain range of real numbers. If a characteristic can be alternatively expressed by variables of nominal, ordinal, interval and ratio scales, the ordering of the amounts of information provided by individual measures by these variables is as follows: nominal < ordinal < interval < ratio.

The choice of measurement scale must take into account the importance of the characteristic and theoretical and practical considerations. From a theoretical point of view, the highest-level measurement scale should always be used, that is, the ratio scale. However, practical considerations must also be considered, particularly the importance of the variable and the availability of resources, such as measurement instruments and qualified personnel to perform the measurement. Statistical considerations are also important, as data from the research will have to be subjected to statistical analysis. In general, the statistical methodology is more developed and popularized for normally distributed variables, which are continuous variables with a ratio scale. In many studies, intrinsically quantitative characteristics are expressed by ordinal scale variables whose levels are subjectively assigned by an evaluator. This is often the case for characteristics that comprise the degree or intensity of a property as measured by visual assessment, such as degree of disease infection or pest infestation. To represent such a characteristic, it is common to define a variable with the lowest level equal to zero to express the null degree or intensity, such as the absence of the symptom; increasing intensities are expressed by the following integers, i.e. 1, 2,...; for example, 0 - no infection, 1 - mild infection, 2 - medium infection, 3 - high infection, and 4 - very high infection. These variables are generally inconvenient, as they express a continuous characteristic very imprecisely and because of the difficulties that arise for statistical inference procedures. Continuous characteristics should be expressed, whenever possible, by continuous variables, even in subjective measurements. For example, the evaluation of the intensity of rust infection in wheat leaves can be done in percentage, with a continuous measurement scale of values between 0 and 100, using as patterns figures of leaves with infection intensities corresponding to arbitrary points on this scale. Infection intensities are measured by comparing the appearance of leaf samples with these standards and interpolation on this continuous scale.

### Precision and accuracy of a measurement process

The **precision** of a measurement process of a characteristic refers to the proximity of measurements of that characteristic carried out repeatedly in the same unit. The measurement of a characteristic is more precise the closer its different measurements referring to the same unit are located. Thus, low precision means large variation of repeated measurements; high precision means small variation of repeated measurements. For example, if repeated measurements of the weight of an object carried out with scale A are quite close and those carried out with scale B vary considerably, the measurement process with balance A have higher precision than with balance B.

The precision of a measurement process does not refer to correct assessment: a measurement process of a characteristic can be highly precise, determining successive measurements that are very close to each other, but far from the true value of that characteristic. A measurement process in these circumstances and the measures it determines are said to be **biased**. The difference between the mean value of repeated measures of a characteristic in one unit and the true value of that characteristic is called **bias**. A biased measurement process systematically overestimates or underestimates the true value of the characteristic.

The **accuracy** of a measurement process of a characteristic refers to the proximity of measurements of that characteristic carried out repeatedly in the same unit in relation to the actual value of that characteristic. A measurement process is more accurate the closer to its true magnitude are repeated measurements performed on a unit by this process.

Accuracy is higher the higher the precision and the lower the bias. High accuracy corresponds to little or no bias and high precision. Low accuracy can result from a large bias and high or low precision, or a null or low bias and low precision. For example, weighing performed with a low precision balance has low accuracy, regardless of the presence or absence of bias; weighing performed with a high-bias balance also has low accuracy, regardless of the precision level of the weighing. Weighing performed with a balance has high accuracy only if it has high precision and low bias.

### 3.2. Reasoning

Facts are the essential building blocks of science. However, they must be arranged in useful and interrelated structures. The most essential tool of science, along with verified fact, is the system of valid logical reasoning about facts that allows reliable conclusions to be derived from them. These conclusions are propositions about interrelationships of facts, which constitute scientific principles, laws and theories.

At the heart of logical reasoning about facts is a system of rules and prescriptions whose correct use is fundamental to every scientific endeavor:

*- The reasoning process used in scientific research starts with one or more propositions and proceeds to one or more other propositions, whose veracity is believed to be implied by the veracity of the first set of propositions. This psychological process is called **inference**.*

The validity of this implication depends on the logical relationship between the propositions. The absence of an implicative relationship between the propositions can lead to false inference. Two processes of inference are fundamentally distinguished: deductive inference and inductive inference:

- **Deductive inference** is the process of reasoning in which one starts from a general proposition or set of propositions and proceeds to a specific proposition or set of propositions. Or, put another way, it is the process of deriving knowledge of a specific member of a class from general knowledge concerning all members of the same class.

- **Inductive inference** is the process of reasoning in which one starts from a specific proposition or set of propositions and proceeds to a general proposition or set of propositions. That is, it is the reasoning process by which knowledge of some members of a class is applied or extended to all unknown members of the same class.

The discovery of the nature of inference came relatively late in human culture. Aristotle (384-322 BC) was the first to stress the systematic nature of science, holding that it can be developed by reason alone. He established the logical scheme of deductive reasoning called **syllogism**. This method of deductive inference starts with two propositions called the main premise and secondary premise that are logically related in such a way that from them a third proposition, called the conclusion, can be derived. For example:

Main premise: All living plants absorb water (induction).

Secondary premise: This peach tree is a living plant (observation).

Conclusion: Therefore, this peach tree absorbs water (deduction).

The inductive method is used to generalize from known and verified phenomena of a given class to unknown and not yet verified phenomena of the same class. Induction involves formulating a generalization. This generalization step is called “inductive jump”.

Inductive reasoning is commonly employed whenever a judgment is made about a situation based on experience with a presumably similar previous situation. For example, when lightning is observed and then the sound of thunder is expected, and when a cultivar is selected for planting on the assumption that it will have similar yield to the previous years. The bases of these claims are the postulates of uniformity and permanence of nature. Although strictly every event is unique and therefore not repeatable in the future, from a practical point of view many events show similarities in some essential characteristics. Thus, the observation of the succession of events provides a basis for the evaluation, in terms of chance, of the possibility of the occurrence of an event when other events that preceded it in the past occur with known frequency.

In summary, induction in science is based on incomplete evidence, due to the impossibility of considering all units of the target population. The conclusions are only probable, to a greater or lesser extent, depending on the number of units considered and the precaution taken in their selection. The relevant fact is that, in well-planned and executed research, evidence of inductive inference can be established on a probabilistic basis by statistical methods.

#### 4. Scientific Research Process

The generation of knowledge about the units (systems) of a target population through the scientific method comprises of three phases: synthesis, analysis and synthesis. The initial synthesis phase consists of the identification of scientific problems through the successive decomposition of these systems into sufficiently simple subsystems to which important specific problems correspond that may be susceptible to attempts at solutions based on existing scientific knowledge and with the available resources. In the analysis phase, for each of these scientific problems, a conjecture of solution is chosen, which can be verified empirically by scientific research. In the final synthesis phase, the new knowledge generated is incorporated into the body of knowledge about the systems.

Scientific research encompasses all phases of the scientific method: it is used to identify scientific problems, to try to solve these problems, and to incorporate the new knowledge generated into the existing body of knowledge. Thus, it comprises systemic research, which focuses on global systems, in the initial and final phases, and analytical research, which focuses on subsystems, in the analysis phase.

Analytical research is necessarily carried out in the scientific research of the units of a target population, but systemic research is little used. The identification of scientific problems and the integration of new knowledge generated by analytical research into the existing body of knowledge is often accomplished by alternative means.

The scientific research process is dynamic and reiterated over time. When new knowledge generated from research is incorporated into current systems, new research problems are identified, and the process is reiterated.

#### 5. Objectives of a Scientific Research

The objective of a synthetic research, in the phases of synthesis of the scientific method, is exploratory. The objective of an analytical scientific research is the verification of a conjecture of solution of a particular scientific problem. Thus, depending on the purpose of the scientific problem that gives rise to it, the objective of an analytical research can be descriptive or explanatory. In practice, these three objectives or functions are not mutually exclusive, and scientific research can often combine more than one of these objectives. The essential characteristics of these three objectives are explained below.

### 5.1. Exploratory research

- **Exploratory research** focuses on systems globally for the examination and understanding of their structure and constitution, and the identification of their relevant characteristics and their interactions. For this reason, it is called **systemic research**

Exploratory research is suitable for the initial and final synthesis phases of a scientific method cycle. In the initial synthesis phase, the systems are globally focused to understand them as a complex whole in order to identify relevant characteristics and problems that may be amenable to analytical research, to obtain suggestions of conjectures of solutions (hypothesis) of these problems, and to investigate the feasibility of research procedures and techniques. At this phase, exploratory research is also useful for identifying and establishing research priorities and alternative courses of action for fruitful research. In the final synthesis phase, exploratory research is employed for practical verification of responses to technologies generated by analytical research when integrated into real systems.

Exploratory research is particularly important in emerging or developing areas of science, particularly in areas where fruitful theories have not yet been formulated, important characteristics have not been identified, and research demand is unknown. It can also be important in the initial phase of a large-scale research program to indicate its feasibility and chance of success.

For focusing on systems globally, exploratory research has a precarious theoretical basis and is poorly structured. Particularly, it does not require the formulation of scientific problem and hypothesis.

### 5.2. Descriptive research

- **Descriptive research** aims to characterize the units of a target population, through their description and representation.

Descriptive research can also aim to identify and describe relationships between characteristics. Its function is fundamentally to portray and report. It focuses on characteristics of subsystems of the systems that constitute the object of research.

Descriptive research can only focus on measuring characteristics and determining their important properties, such as means and variances, or frequency distributions. It can also aim at identifying association relationships, that is, covariation or correlation of characteristics. In some research, this information is used for prediction purposes. However, descriptive research may provide a basis for prediction, but is not appropriate for explaining the nature of characteristic relationships.

Unlike exploratory research, descriptive research is characterized by the a priori formulation of the research problem. Usually, the researcher already has substantial knowledge to formulate the research problem, possibly as a result of exploratory research. This knowledge allows him the clear definition of the research objectives, particularly the relevant characteristics that must be considered.

### 5.3. Explanatory research

Explanatory research focuses on the relationship between two subsets of characteristics of units in a target population, one of which expresses the performance or behavior of the units and the other supposedly affects that performance or behavior. In this causal relationship these two subsets of characteristics are supposed to constitute the effect and the cause.

- **Explanatory scientific research** aims at inferences about the relationship between a subset of the class of characteristics of the units of a target population that express the performance of these units (**response characteristics**) and the class of characteristics whose control and alteration can, supposedly, imply the improvement of that performance (**explanatory characteristics**), in the presence of the third class of characteristics of the units (**extraneous characteristics**).

In other words, the function of explanatory research is to assess and express significant and predictable causal relationships of a subset of characteristics of the units of a target population with another subset of characteristics. It must be remembered that in science these relationships can only be expressed in terms of probable relationships, not in terms of exact relationships.

## 6. Stages of an Analytical Scientific Research

Analytical scientific research starts from a particular problem or question concerning the units of a target population. Once the initial question is established, the systematic nature of the scientific method dictates the structuring of the research in a precise and logically arranged form in several sequential stages. From the initial question to its final answer, an analytical scientific research proceeds through eight main stages of operation, listed below:

- First stage: Identification and formulation of the scientific problem or research problem.
- Second stage: Formulation of the scientific hypothesis or research hypothesis.
- Third stage: Literature review.
- Fourth stage: Construction of the research plan.
- Fifth stage: Data collection.
- Sixth stage: Analysis and interpretation of data.
- Seventh stage: Derivation of conclusions, which may lead to the confirmation or rejection of the original hypothesis, and confirmation or questioning of the results of another research.
- Eighth stage: Presentation of results through a report, and dissemination of these results.

Some relevant observations regarding this structure of analytical research process deserve to be highlighted: a) It is arbitrary; research methodology texts list different numbers of stages. However, an examination of such lists indicates that all eight stages

are involved. b) Stages do not necessarily occur in all research in the order indicated. An experienced researcher may be studying the literature at the same time as he is formulating his hypothesis and designing his research plan. c) Stages are not always rigidly established at the beginning of the research. Well-planned research can allow for adjustments or changes to be made as it progresses. d) All stages are equally important in contributing to the results of the research, but do not involve the same amount of time, cost and effort.

The eight stages are inherent to the analytical function of science and constitute a common framework that ensures the basic attributes of the scientific method. However, scientific research with a descriptive objective may comprise only part of the eight stages.

In the following sections, the first two steps are considered, which define the research objectives.

### 6.1. Identification and formulation of the scientific problem

Analytical research begins with the identification and formulation of a scientific problem, that is, a specific problem concerning a subsystem of the systems of the target population that can be solved by scientific research with available resources:

*- A **scientific problem** or **research problem** is a problem or question concerning characteristics of a target population and their relationships that is answerable through the scientific method with the techniques, procedures and resources available.*

The identification and formulation of research problems can originate from research in the initial phase of synthesis of a cycle of the scientific method, from the literature and from the observation of the systems object of the research. It can also be formulated as a theory-based prediction.

The first consideration for choosing a scientific problem must be its relative importance; the second, the possibilities of its resolution considering the available resources.

A research problem is formulated for a descriptive or explanatory purposes. A **descriptive problem** refers to the lack of, or demand for, knowledge about the units of a target population. For example: “what technologies are used in the vineyards of region A?”, “which pathogenic microorganisms contaminate the bovine milk produced in the dairy basin of the region B?”, “what are the most relevant problems of wheat crops in the region C?”. An **explanatory problem** is a question concerning causal relationships between characteristics of the units a target population. For example, “what is the relationship between the incidence of mastitis in dairy cows and the age of the animal?”, “what is the relationship between the incidence of wheat blight and relative humidity?”, “what is the relationship between productivity and season?”

The correct formulation of the scientific problem is crucial, not only because it constitutes the definition of the research objective, but also because of its intrinsic difficulty. The importance of this first stage of scientific research is highlighted by the frequency of problems that are imperfectly or inadequately formulated as a result of faulty information and reasoning. These failures can be avoided by a careful, detailed and complete formulation of the problem, preceded by adequate knowledge of its background, particularly its origin and importance, and the destination of the research results that it will give rise to.

Although there are no general rules for formulating a scientific problem, certain requirements must be satisfied for a problem to be considered a well-formulated scientific problem. Some of these requirements are set out below:

- a) The problem must be inserted in an accessible body of scientific knowledge,
- b) complex problems must be decomposed into a set of specific problems that give rise to a research program,
- c) all relevant characteristics related to the problem must be identified and considered,
- d) the existence of the problem must be ensured by a clear definition of the attributes of the problem and an explicit delimitation of its area of existence,
- e) the problem must be well conceived and delimited,
- f) the problem must be understood before it is formulated,
- g) it must be treatable by existing scientific research techniques and available resources,
- h) must be answerable in objective terms,
- i) must be formulated clearly so that it has the same meaning for all intelligent and well-informed people,
- j) the formulation of the problem must delimit its scope, which demands a clear characterization of the target population and its units,
- k) alternative or substitute formulations should be considered if the original formulation is not feasible,
- l) among the alternative formulations of the problem, the one corresponding to the most feasible and most efficient solution must be chosen,
- m) the problem must be formulated in a systematic way,
- n) the formulation of the problem should not influence the researched phenomena.

### 6.2. Formulation of the scientific hypothesis

After the formulation of the problem, the next step is the inquiry regarding the nature and connections of characteristics that lead to the idealization of one or more paths for the solution or answer to the problem. Each of these idealized paths constitutes a research hypothesis.

The research hypothesis can be suggested by the literature. Less often, it comes from reasoning based on occasional observation of events. More commonly, however, it is formulated as a prediction based on theory. Using a theory to answer a question or solve a problem requires deducing inferences from it. These inferences are scientific hypotheses:

- *A **scientific hypothesis** or **research hypothesis** is a proposition of solution or answer to a research problem that is derived from a theory by deductive inference and that allows an empirical confirmation test through scientific research, with available techniques, procedures and resources.*

The function of the scientific hypothesis is to extend scientific knowledge beyond the present frontiers of theoretical knowledge. So, the scientific hypothesis is more than a connection between speculation and verification; it is the essential factor in the growth of scientific knowledge.

It should be noted that exploratory research, in the initial and final phases of a cycle of the scientific method, when the relevant characteristics of the target population units and their relationships are not yet known, do not include problem and hypothesis formulation. Its objective is precisely the identification of relevant problems (scientific problems) for research and fertile conjectures for the solution of these problems (scientific hypotheses).

In line with the corresponding scientific problem, the function of a scientific hypothesis can be either descriptive or explanatory. A **descriptive hypothesis** is stated for a descriptive problem. It declares the existence of an empirical uniformity, that is, of a fact or relationship that is commonly known but has not yet been empirically verified, or which is not yet known but whose existence is suspected. As science attempts to build a systematic body of verified knowledge, it is natural that a great deal of its effort is spent on verifying hypotheses that appear obviously true but have not yet been empirically verified. An **explanatory hypothesis** refers to a causal relationship of characteristics. Seeks to find out if changing one or more characteristics implies changing one or more other characteristics.

Some simple research concerning descriptive problems can be conducted without reference to a hypothesis framework. However, most descriptive research implies an underlying hypothesis.

For the same problem, one or more hypotheses can be formulated that differ in the level of complexity of the problem and in the proposed solution. A simple hypothesis may be a mere generalization of a particular empirical observation. More complex hypotheses may postulate connections between events, or elaborated chains of chance relationships. The formulation of an adequate hypothesis depends on the knowledge and experience of the researcher. However, creativity is of the utmost importance for the construction of useful and fertile hypotheses.

There is no set of rules that can guarantee the formulation of the most appropriate and fruitful hypothesis. However, the following attributes are essential for a conjecture to be a scientific hypothesis:

- a) the hypothesis must have meaning and be formally correct,
- b) it must be based on prior knowledge and be compatible with the body of scientific knowledge,
- c) must be logically related to the theory from which it is derived,
- d) must respond to the problem that gave rise to it,
- e) it must be conceptually clear and unambiguous,
- f) it must be plausible, that is, it must be logically possible,
- g) must be expressed in objective and operational terms,
- h) should be as specific and simple as possible,
- i) it must be empirically verifiable, that is, it must suggest research by which it can be tested,
- j) must allow a decision regarding the problem,
- k) it must allow a reliable means of predicting unknown events,
- l) it must include the criterion of its confirmation, that is, it must be stated in terms that do not leave doubts regarding the answers considered as confirmation.

Very often, more than one hypothesis can be formulated as a likely solution to a research problem. In these circumstances, these requirements should be employed as criteria to suggest which hypothesis should be chosen as the most appropriate and fruitful.

A scientific hypothesis is never definitively and absolutely proved. An empirically confirmed scientific hypothesis is a scientific fact, which is not necessarily the exact expression of reality. A hypothesis is only confirmed or not confirmed, depending on the degree of confirmation established. A scientific fact, that is, a confirmed hypothesis, is a new hypothesis, presumably with a higher level of reliability than the original.

## **7. Research Work and the Role of the Researcher in Science**

The scientist or researcher is the individual who rigorously employs the scientific method in the pursuit of knowledge and who ultimately builds science. The researcher must actively seek out and investigate nature to discover orderly relationships. To do so, he must have some important qualities, among them: curiosity, patience, objectivity, collaboration, communication, creativity, discrimination, enthusiasm, accuracy, firmness, honesty, imagination, liberality, morality, perseverance and concern for professional development.

Originally, research work was an individual initiative. The advancement of scientific knowledge has made the work of scientific research highly complex. As a result, scientific research began to require a multidisciplinary team that must be made up of

specialists with cooperative action in the different phases of its execution. Naturally, the more intense participation of each member is more required in specific phases.

It is highly relevant to the success of research, particularly in applied science, that researchers are in contact with real problems and that they establish appropriate strategies so that research is relevant and has its results applied in practice. In fact, the fundamental objective of scientific research is to serve the community in which it operates. For this purpose, knowledge and experience of the research team is needed, particularly:

- In-depth knowledge of the research area, in particular the target population units. In agriculture, for example, knowledge of production systems, the economic and social conditions of agricultural companies, market requirements, changes in product costs and value, appreciation of product quality, etc.,
- knowledge of the results of research conducted previously in the region or in other similar regions; in particular, knowledge of the existing literature, and
- knowledge of the experience of the most progressive producers, acquired over years of observation and experience.

This knowledge is essential for a broader view of the problems that demand research and their possible solutions, and for the basis of research planning. These and other requirements may prevent the execution of fruitless research.

Much research is ineffective and fail to achieve their objectives as a result of limited resources and structural conditions for the research activity and deficiencies in the very orientation and discipline of this activity. On the other hand, research institutions have accumulated many unanalyzed, uninterpreted and undisclosed research results. Furthermore, it is not uncommon for experiments to be carried out repetitively for a long time, without the necessary evaluation of their development.

The logical deduction, all too often recognized but rarely taken seriously, is the little attention that has been devoted to planning and analyzing research findings. In fact, it is generally observed that researchers devote very little time and effort to research planning, which is the starting point, and the usefulness of the research results depends on it. Most technological research is intended to generate well-defined recommendations. It is therefore essential that such recommendations are derived with a high level of security. These goals can only be achieved by rational planning, which considers the research objectives, uses existing resources more efficiently and adopts the most appropriate available methodology to achieve these objectives.

## **8. Statistics in Scientific Research**

### **8.1. Scientific method and Statistics**

The success of the physics of Isaac Newton (1643-1727), supported by applied mathematics, developed and strengthened the deterministic view of science in the 17th, 18th and 19th centuries. However, with the advent of quantum mechanics by Max Karl Ernst Ludwig Planck (1858-1947) and the uncertainty principle of Werner Karl Heisenberg (1901-1976), even physics, the leader of the exact sciences, ceased to be deterministic.

Statistics is part of the modern scientific approach to uncertainty. This approach assumes that "laws" can only predict "expectations" and that actual observations can differ from these by "random errors". It is the study of these errors that enables predictions under uncertainty. Prediction under uncertainty is made possible by Applied Statistics, based on probability calculus. Applied Statistics is vitally dependent on concepts from the areas to which it is applied. For this reason, it is commonly identified with these application areas, assuming specific designations, such as experimental statistics, biometrics, biostatistics, econometrics and sociometry.

Statistics is an integral part of the scientific method. The development of Statistics has resulted from the demand from scientific and technological progress. It is closely related to the development of the scientific method and the advancement of different areas of science. The progress of science in the last two centuries has been the source for the extraordinary development of Statistics.

It is notable that the development of modern Statistics was greatly influenced by agricultural research. In the second decade of the 20th century, research at the Experimental Station in Rothamsted, England, which began in 1843, had generated a considerable amount of data. The mathematician Ronald Aylmer Fisher (1890-1962) was assigned to analyze this numerical information. During just 14 years at Rothamsted, Fisher developed the theory and methods he needed and achieved the practical application of his theory of statistical inference, which became the foundation of modern Statistics. His work led to conclusions relevant to scientific research, including that the amount of information generated by the inferences of a research cannot be greater than that contained in the data. Therefore, the research planning and data generation process took on a fundamental importance.

From Fisher's contributions to agricultural scientific research, the new statistical methods began to apply to other areas of science and technology. Scientific developments in different areas, in turn, demanded new statistical methodologies that also became, in general, applicable to other areas.

The development of the conceptual and methodological basis of scientific research, of plans appropriate to the circumstances of each research, and of more objective and efficient methods of data analysis, have increased the possibilities, speed and reliability of scientific research. However, the effectiveness of scientific research depends not only on specialized knowledge in the research area, but also, fundamentally, on knowledge of the scientific method, which includes the statistical method.

This methodological development has resulted, in large part, from the vast increase in computing capacity that became available, mainly from the mid-20th century onwards. The growing computing facilities have allowed researchers greater possibilities for decisions regarding the best use of available research resources; in particular, to better plan the research and analyze its data more adequately, achieving a more efficient exploitation of the information provided by the research.

These developments have had a considerable impact on researchers' activities, the effective exercise of which demands changes in interest and attitude. A fundamental requirement is permanent training and updating, particularly for understanding the methods and procedures implemented in computing "packages". In fact, the powerful facilities available for data analysis cannot be used properly without understanding the methods implemented and, if used incorrectly, can lead to misleading results.

A notorious fact is that, despite the consolidation of the scientific method, particularly the statistical method, basic concepts are still not the domain of many researchers. Consequently, the impact of this methodological development is still not felt as widely as might be expected. This delay is not explained only by the natural delay in incorporating theoretical results into practical activity. The divorce that persists between theory and practice has several origins, among them: failures in teaching and in the transmission of knowledge through literature, lack of infrastructure and material resources in research institutions and lack of vocation of researchers. From the first and the last origin, another no less important cause arises: the ritualistic attitude of many researchers who tend to adhere to, copy and use ideas, concepts, methods and procedures mechanically, without a critical attitude.

## **8.2. Use of Statistics in scientific research**

Statistical methods cannot prove a proposition or the result of scientific research. What statistical methods provide is the assessment of the probability of error of a proposition, or of the confidence that can be placed on the result of a scientific research.

Suppose, for example, that scientific research is carried out to estimate the effect of dietary supplementation on the growth of lambs. The research may indicate that the supplement increases an animal's body weight at weaning by an average of 7.5 kg. This result, however, does not provide any proof that similar growth, or even any growth, would be observed if the research were repeated under similar conditions. Considerable doubt will remain about this result. The use of a statistical-based planned research followed by the appropriate statistical analysis of the results will not eliminate the doubt of the research result, but it will allow to assign a value to the reliability of this result and to estimate how often such a result will arise if the effect does not exist, and to use it as a basis for conclusions. For example, if it can be shown that differences as extreme or more extreme than observed would occur less than once in a thousand if the supplement had no effect, then there would be reasonable safety in disregarding the possibility of ineffectiveness of supplementation and concluding that it is effective. However, the possibility will remain that it is ineffective, but now some measure of the possibility will be available and can be used to demonstrate the plausibility of the existence of a supplement effect.

It is usually recognized that the use of statistical methods can greatly contribute to the efficiency of scientific research and, particularly, to the validity of the inferences it derives. However, the lack of knowledge of the fundamentals of statistical methods, especially about the requirements for the validity of their applications, often leads to their misuse. Statistics is often used without attention to the assumptions required for the validity of the procedures employed. For these and other reasons, it is necessary for the researcher to clearly understand the techniques he employs.

## **8.3. Knowledge of Statistics by the researcher**

Researchers from different areas of research must have mastery of knowledge of scientific research methodology that is essential for understanding and making decisions and actions related to the usual aspects of planning, conducting and statistical analysis of research under their responsibility. This domain of knowledge should cover statistical methods and computational resources, including "packages" for statistical analysis.

It is up to statistical specialists to complement the domain of the necessary knowledge about methods that require specialized training and that require a more advanced theoretical base in Statistics. Particularly, it is the responsibility of the statistician to carry out studies and research aimed at adapting and developing new methods and techniques.

## **References**

1. Silva, J.G.C. (2022). Science and Scientific Method. *International Journal of Science and Research*, 11(4): 621 - 633. Access 26 April 2022. Available: [https://www.ijsr.net/get\\_abstract.php?paper\\_id=SR22412084104](https://www.ijsr.net/get_abstract.php?paper_id=SR22412084104).
2. Bertalanffy, L.von. *Teoria geral dos sistemas*. Petrópolis: Vozes, 1973. 351p.
3. Bunge, M. *La investigación científica, su estrategia y su filosofía*. 4. ed. Barcelona: Editorial Ariel, 1975. 955p.
4. Castro, A.M.G.; Cobbe, R.V. Aplicação do enfoque sistêmico na gestão de C&T. In: Silva, E. C. (editor) *Métodos quantitativos e planejamento na EMBRAPA com enfoque na informação e na tecnologia da informação*. Brasília: Empresa Brasileira de Pesquisa Agropecuária, Departamento de Pesquisa e Difusão de Tecnologia, 1995. p.19-41.
5. Christensen, L.B. *Experimental methodology*. 7.ed. Boston: Allyn and Bacon, 1997. 590p.
6. Cox, D.R. *Planning of experiments*. New York: John Wiley, 1958. 308p.
7. Fisher, R.A. The place of the design of experiments in the logic of scientific inference. *Colloques Internationaux du Centre National de la Recherche Scientifique*, Paris, n.110, p.13-19, 1962.

8. Green, P.E.; Tull, D.S. Research for marketing decisions. 2. ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1970. 644p.
9. Heath, O.V.S. A Estatística na pesquisa científica. São Paulo: Universidade de São Paulo, 1981. 95p.
10. Kempthorne, O. Sampling inference, experimental inference and observation inference. *Sankhya*, 40, Serie B, Parts 3 and 4, p.115-145, 1979.
11. Kish, L. Sample surveys versus experiments, controlled observations, census, registers, and local studies. *Australian Journal of Statistics*, v.27, n.2, p.111-122, 1985.
12. Kish, L. Statistical design for research. New York: John Wiley, 1987. 267p.
13. Lastrucci, C.L. The scientific approach, basic principles of the scientific method. Cambridge, Massachusetts: Schenkman Publishing Company, 1963. 257p.
14. Mitchell, M.; Jolley, J. Research design explained. 3. ed. Fort Worth, Texas: Harcourt Brace, 1996. 624p.
15. Silva, J.G.C. O ensino da estatística no Sistema Nacional de Pesquisa Agropecuária. In: Encontro Nacional de Métodos Quantitativos, III, Anais. Brasília, 20-22 junho, 1995. Empresa Brasileira de Pesquisa Agropecuária, Brasília, 1995. p.89-107. Portuguese. Accessed 02 May 2022. Available: [https://www.researchgate.net/publication/281974746\\_O\\_ensino\\_da\\_Estatistica\\_no\\_Sistema\\_Nacional\\_de\\_Pesquisa\\_Agropecuaria#fullTextFileContent](https://www.researchgate.net/publication/281974746_O_ensino_da_Estatistica_no_Sistema_Nacional_de_Pesquisa_Agropecuaria#fullTextFileContent).
16. Silva, J.G.C. Métodos quantitativos no Sistema Nacional de Pesquisa Agropecuária (SNPA): Um programa de capacitação e reciclagem de pesquisadores agrícolas. In: Silva, E. C. (editor) Métodos quantitativos e planejamento na EMBRAPA com enfoque na informação e na tecnologia da informação. Brasília: Empresa Brasileira de Pesquisa Agropecuária, Departamento de Pesquisa e Difusão de Tecnologia, 1995. p.69-87. Portuguese. Accessed 02 May 2022. Available: <https://pt.scribd.com/document/393431262/Metodos-quantitativos>
17. Silva, J.G.C. Pesquisa científica. Boletim Técnico 7. Instituto de Matemática e Estatística. Pelotas: Universidade Federal de Pelotas. 2005. Portuguese. Accessed 24 May 2022. Available: [https://www.researchgate.net/publication/266387738\\_Scientific\\_Research](https://www.researchgate.net/publication/266387738_Scientific_Research).
18. Smith, T.M.F.; Sugden, R.A. Sampling and assignment mechanisms in experiments, surveys and observational studies. *International Statistical Review*, v.56, n.2, p.165-180, 1988.
19. Urquhart, N.S. The anatomy of a study. *Hortscience*, v.16, n.5, p.621-627, 1981.
20. Wilson, E.B. An introduction to scientific research. New York: McGraw-Hill, 1952. 373p.