Chapter 2
Variables, a Short Taxonomy

Variables

A phenomenon that changes over time, space, objects, or living things is called a variable. If a variable's changes could be measured, that is, expressed by numbers or quantities, then we have a *quantitative variable*; otherwise, we have a *qualitative variable*. Blood pressure is an example of a quantitative variable—because it takes different values from individual (living thing) to individual. It can be measured, and has its own scale of measurement. Tree Hight is another example of a quantitative variable. It varies from tree to tree and can be measured (in centimeters or meters in the metric system and in inches and feet in the English system)[1] and expressed by a numerical value.

Changes in qualitative variables can only be *observed*. Gender is an example of a qualitative variable. Color of automobiles is another example of a qualitative variable. It varies from automobile to automobile. Qualitative variables taking only two different values are known as *dichotomous* or *binary* variables. Those taking more than two values are called *multinomial*. Gender is an example of a dichotomous variable while color is an example of a multinomial variable. In this book we do not deal with any qualitative variables.

The U.S. Gross Domestic Product (GDP) varies from year to year. The Bureau of Economic Analysis (BEA), an agency of the Department of Commerce, produces some of the most closely watched economic statistics, including annual GDP. Besides annual estimates, BEA also produces semi-annual and quarterly estimate of GDP and its growth rate. The Bureau of Labor Statistics (BLS) of the US Department of Labor measures the size of the labor force and number of people unemployed every month and reports the monthly Unemployment rate (U-rate). The GDP and U-rate are examples of quantitative variables measured and reported for a certain time. In each case the measurement of the variable is conducted over an equally spaced time interval, a year for GDP and a month for the unemployment rate.

---

[1] There are many other measurement systems in the world. Metric and English systems are by far the most widely used.

S. Vali, *Principles of Mathematical Economics*,
Mathematics Textbooks for Science and Engineering 3,
DOI: 10.2991/978-94-6239-036-2_2, © Atlantis Press and the authors 2014

Data generated by measuring values of a quantitative variable that changes over time are called *time series* data. Frequency of measurement of a variable or the time interval over which measurement is made is called the *periodicity* of the variable. There are variables with different periodicities, as the following examples demonstrate. The size of a country's population is measured every 10 years, through a census of the population. It is a decennial variable and its values generate decennial data. Publicly traded corporations are required to report, among other things, their revenue for each quarter. The Labor Department reports jobless claims (the number of newly laid-off workers filing claims for unemployment benefits) each week. Department stores also report their sales volume weekly. At the end of a business day the Dow Jones Industrial Average (DJIA) for closing prices of the 30 major US corporations that comprise the index, is widely reported. During a trading day, the DJIA is measured, and posted on the New York Stock Exchange (NYSE) big board almost every 15 seconds. Meteorological instruments of various weather stations continuously measure and record changes in temperature and atmospheric pressure in various parts of the country. Maintaining heat inside a blast furnace is an important factor in the production of high quality steel. In modern steel making processes, infra red instruments continuously measure the heat inside the furnace and display it on a screen in the control room of the factory.

As the above examples should indicate, there are two types of time series variables; *discrete* and *continuous*. In the continuous case values of the variable are measured or observed at every moment of time, i.e. over infinitesimally small time intervals to the extent that periodicity of the variable approaches zero; however, most time series variables, particularly in economics, are discrete. Values of a discrete time series variable are recorded or measured at predetermined equal time interval. The most common periodicities of economic variables are annual, semi-annual, quarterly, monthly, and weekly. Changes in a time series variable may happen with an irregular time interval or periodicity. A good example is the change in price of shares of a company traded on a stock exchange. These changes may occur over different unequal time intervals, but the price of the company's stock is still considered a time series variable. There are numerous examples of time series variables, with or without regular periodicity. What they have in common is that they all evolve over time and are all indexed by time, irrespective of whether the index is equally spaced or not.[2]

Most macroeconomic variables are time series variables. A very good, updated and comprehensive source of macroeconomics time series is the Annual Report of the President, prepared by the Council of Economic Advisors (CEA). Appendix B of the latest report issued (at the time of writing this passage) February 2011 contains a collection of about 130 pages of statistical tables produced by various departments and agencies of the US government. The statistical tables of the report cover historical and the latest available data related to income, population, employment, wages and productivity, production, prices, money stock, corporate profit and finance,

---

[2] For a formal presentation of time series variables see Chap. 11 *"Economics Dynamics and Difference Equations"*.

agriculture, and international trade and finance. The entire report can be accessed at the US Government Printing Office site http://www.gpo.gov.

BLS not only measures and reports the national unemployment rate each month, but also provides the same information for each state and a large number of major metropolitan areas like New York City, Philadelphia, and Los Angeles. In this case the variable unemployment rate takes different values over space or geographic areas. This is an example of a *cross section* variable. Another example would be the second quarter earnings reported by companies in *Standard and Poor's* (S&P) 500 index. Here variable 'earning' takes 500 different values, one for each of the 500 companies in the index at the same time (quarter). Data generated by changes in a cross section variable are called "cross section" data.

The third class of variables, called *mixed variables*, are *mixed* of time series and cross section or mixed of cross sections. Typical examples are monthly unemployment or inflation rates of 50 states of the United States, annual growth rates of the G7 (group of 7 most advanced industrialized countries) in the last 10 years, and average income of households of different size over different geographic areas. Mixed variables are sometimes referred to as *panel variables* and their values called *panel data*. This designation is due to the fact that presentation of the values of a mixed variable is best achieved in a rectangular array (panel) format, i.e. a table. For example, to present the growth rate of the G7 countries for the last 10 years we must construct a table with 7 rows, one row for each country, and 10 columns, one column for each year. This table contains $7 * 10 = 70$ cells. Each cell that is at the intersection of a row (country) and a column (year) will contain the growth rate of that country for that year.

Before we go any further in our discussion of variables we must distinguish between *flow* and *stock* variables. This distinction is very important in economics. Values of a flow variable can only be measured between two points in time, that is, over a specific time interval, while values of stock variables can be measured at every moment in time. Examples of flow variables are GDP, income, revenue, output, and rainfall. Monthly household income is the amount a household earns over a month, and the GDP is the market value of all finished goods and services produced by labor and property located in a country in a given period of time, generally one year.

Examples of stock variables are capital stock of a firm, balance of your saving account, wealth of an individual, and market value of assets of a corporation. All these variables must be measured at one point in time. Measuring these variables at different time may lead to a different value for the variable. Unemployment rate is a stock and inflation rate is a flow variable. All of the items appearing on the *balance sheet* of a company are stock variables, while items appearing on a company's *income statement* are flow variables.

Since variables can take a range of values, they must be represented by symbols. In fact Bertrand Russell—a giant of philosophy, mathematics and literature of the twentieth century—defined mathematics as "a game played with symbols". Use of the first letter, or an abbreviation, of a variable's name for denoting that variable is very common. In economics, for example, it is a well established convention that price, quantity of output, cost, profit, and revenue (just name a few) are represented by P, Q, C, Π, and R. In general and abstracting from any specific context, the uppercase letters of the English (or Greek) alphabet like X, Y, and Z are used to denote or name a variable. The corresponding lowercase letters x, y, and z are used to denote specific values that they may take.

Values that a quantitative variable assumes are *numbers*, so we need to say a few words about the number system.

The Decimal (base 10) number system starts with 10 digits 0, 1, 2, ..., 9.[3] The *Natural* or "counting" numbers start from 1 and extend beyond the last digit 9 to 10, 11, and so on with the help of an algorithm for making these new numbers. The symbol commonly used to denote the structure of the natural numbers is N. Addition and multiplication are fundamental Arithmetic operations in N. Subtraction, as long as the first number is greater than the second, and division with remainder are considered implied arithmetic operations in N, because they are actually inverse of addition and multiplication ($10 - 2 = 8$ because $2 + 8 = 10$ and $20/5 = 4$ because $5 * 4 = 20$).

By adding zero and negative integers (or nonzero natural numbers with a "negative sign") to N, we arrive at the collection or set[4] of *integer numbers*. This set is generally denoted by Z. The rules and operations in Z are a little more complicated. First, we have a new order. In N, 50 is greater than 40. In Z, -50 is smaller than -40, and -20 is less than 2. Second, we have *algebraic* operations, which mean that we must consider the signs of the numbers.

While 2 and 3 are in Z, 2.5 is not. How can we divide 3 apples between two children? If we conduct this task (division in Z) we'll be left with an undivided (remainder) apple. It is most prudent to cut the third apple in half so each kid can have 1.5 apples. We need a new set containing numbers like 1.5. This set is the set of *rational numbers*. Rational here is *ratio-nal*, which means all the numbers in this set can be expressed as a ratio of two integer numbers. 1.5 is 3/2 or 6/4. Integer numbers are in this set too. 5 is 10/2. Commonly Q is used as a symbol to denote the set of rational numbers. Z is included in or a subset of Q, as N is a subset of Z.

What if a number cannot be expressed as a ratio of two integers? An example is a number like $\sqrt{2} = 1.4142...$, which is non-terminating and non-repeating decimal. Another classic example is the number π, the ratio of the circumference of a circle to its diameter. To four decimal places π is 3.1415, but it has a non-terminating and non-repeating decimal part. Numbers like $\sqrt{2}$ and π, that cannot be expressed as ratio of two integers, comprise the set of *irrational numbers*. Irrational numbers fill in the gaps between rational numbers.

There is no commonly used symbol for the set of irrationals. The set of irrationals does not have a structure similar to N or Z. An important property of N and Z is that they are *closed* under addition and multiplication. *Closure* under addition means that sum of any two numbers in Z is also a number in $Z(10 + 5 = 15)$. The product of any two numbers in Z is another number in $Z(10 * 5 = 50)$, so it is also closed under multiplication. The set of irrationals does not have these properties. For example, if

[3] The Octal (base 8) system has 8 digits 0, 1, ..., 7. The Binary (base 2) system has only two digits, 0 and 1.

[4] See Chap. 3 for a formal discussion of set and related topics.

we add two irrational numbers $2 + \pi$ and $3 - \pi$ we get 5 which is not an irrational number. Similarly if we multiply $2\sqrt{2}$ and $5\sqrt{2}$, the resultant number 20 is an integer. For this reason no specific symbol is devoted to the irrationals.

A union of the set of rational numbers and the set of irrational numbers creates the set of *real numbers*, universally denoted by R. R has all the nice properties of a well-structured set. It is not only closed under operations like addition, subtraction, multiplication, and division (except by zero), it is also closed under operations like integer exponentiation or raising to the power of a whole number ($2.53^3 = 15.625$).

Is R closed under fractional exponentiation or 'taking roots'? It is not. $\sqrt{-5} = (-5)^{1/2}$ does not exist in R, nor does the even root of any negative number. Note that we can express $\sqrt{-5}$ as $\sqrt{5} * \sqrt{-1}$. Denoting $\sqrt{-1}$ by i, so $i^2 = -1$, then we can write $\sqrt{-5} = \sqrt{5}\ i$, which is an example of an *imaginary number*. In general a number of the form $a + bi$, where a and b are real numbers, is called a complex number. In $a + bi$, a is the real part and bi is the imaginary part of the number. For all values of a and b, $a + bi$ generates the set of *complex* numbers, denoted by C. R is a subset of C. In this book we only deal with variables that take on real numbers.

Following our short discussion of the number system, we can introduce another classification of quantitative variables; *discrete* and *continuous*. A discrete variable can take only whole numbers or integer values, that is, the set of possible values that this variable can assume is a subset of Z. Examples of discrete variables are the number of children in a family, the number of passengers on a bus, the number of tourists visiting Grand Canyon during summer, and the number of employees of a company.

Unlike a discrete variable, values that a continuous variable can take is not restricted to integer numbers, it can also take rational and irrational numbers. The set of all possible values of a continuous variable is a subset of the set of all real numbers R (see Fig. 2.1).

By a time-honored convention, all economic variables are treated as continuous. It may be a little counter intuitive, but there are reasonably good justifications for this treatment. If a variable is measured in dollars then the smallest unit must be a cent. How could it be continuous, you might ask? Let's consider an example: the exchange rate between Dollar and Euro. Assume that the exchange rate between Euro and dollar is 0.67120, this is how much a dollar is worth in Euro. We can express the same relationship by specifying one Euro in terms of units of US currency, that is as \$1.48987. It must be clear that if you want to exchange one Euro for dollar, you can only get \$1.48 (don't think of rounding it to \$1.49!) How about if you want to exchange 1000 Euros? Obviously you expect to receive \$1489.87, \$9.87 more than trading your Euro for \$1.48 each! Exchanging 10,000,000 Euro generates \$14,898,700. This is \$98,700 more than exchange on the basis of one $Euro \simeq$ \$1.48 (here the symbol $\simeq$ is used to indicate approximately "equivalent").

Given the astronomical volume of dollars and Euros traded every day in the international foreign exchange market, it is clear that there is a great deal of financial motivation to treat the exchange rate as a continuous variable and to express it to 6 or 7 significant digits (an alternative way of saying to 5 or 6 decimal places). But if this works for variables measured in monetary terms, does it work for variables

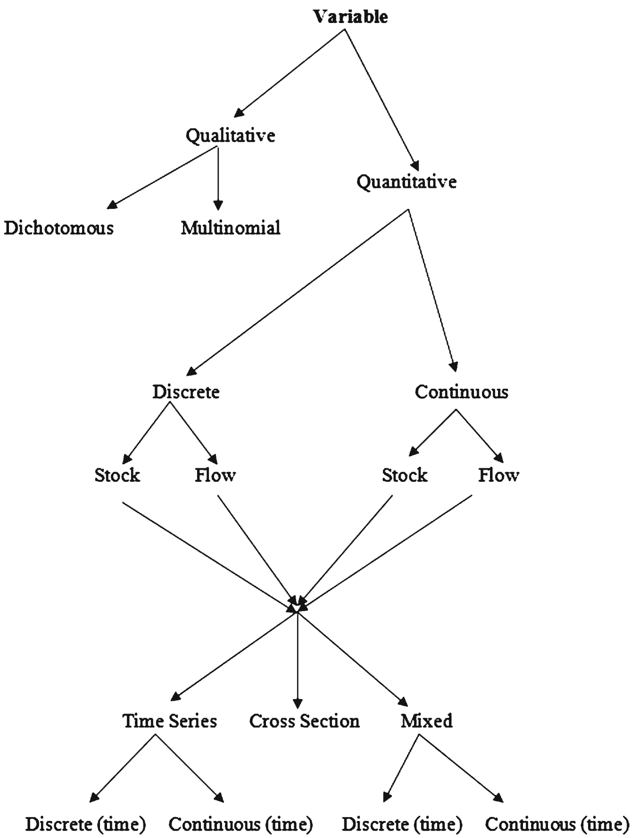

Fig. 2.1 Variable chart

expressed in physical terms? How can we treat the volume of output of an automobile manufacturer as continuous variable?

A typical automobile has 1000s of components that must be assembled on an assembly line. An assembly line is an arrangement of machinery and equipments used for the continuous flow of work pieces in mass production operations. Given that at different stages on an assembly line for automobile production, a group of components are added, then an automobile at various stages of its assembly must be considered a different fraction of a complete car, and by extension, output of an automobile assembly plant can be considered a continuous variable.

Another justification for treating economic variables as continuous is purely "mathematical". Mathematical economics employs tools and techniques of calculus of variation, a part of differential calculus, for solving a series of optimization (finding maximum or minimum) or equilibrium problems. A continuous variable lends itself much easier to these techniques.

2.1 Exercises

1. Classify the following as quantitative or qualitative variables

   (a) Height (b) Ethnicity (c) Number of passengers on a bus

   (d) Weight (e) Language (f) Religion

   (g) Party affiliation (h) US government budget deficit

   (i) US trade deficit with China (j) Time takes to run 100 meter

   (k) Record of the top 100 runners in New York Marathon

   (l) Consumer credit outstanding (m) Corporate profit

   (n) Closing price of stocks listed on NYSE on January 28, 2008

   (o) Population of major cities in the US based on 2000 census.

   (p) Unemployment rate of 27 members of European Union in the last 10 years.

   (q) Unemployment rate of 27 members of European Union in 2007

   (r) Closing price of stocks listed on NYSE from Jan. 2, 1990 to Jan. 2, 2008.

   (s) Record of the top 100 runners in New York Marathon in the last 5 years.

2. Classify the qualitative variables in Exercise 1 as dichotomous or multinomial.
3. Classify the quantitative variables in Exercise 1 as discrete or continuous.
4. Classify the economic variables in Exercise 1 as time-series, cross-section, or mixed.
5. Classify k and s in Problem 1 as time-series, cross-section, or mixed.
6. Classify the economic variables in Exercise 1 as stock or flow.
7. Determine the size of the table—number of rows, columns, and cells—you need in order to present the variables that you identified as mixed in Exercise 1 (for (r) assume 3500 stocks listed on NYSE and 252 trading days).
8. How would you present the unemployment rate of white and non-white population in the last 20 years for 50 states of the United States?

Springer
http://www.springer.com/978-94-6239-035-5

Principles of Mathematical Economics
Vali, S.
2014, XXVIII, 491 p. 121 illus., 60 illus. in color.,
Hardcover
ISBN: 978-94-6239-035-5
A product of Atlantis Press