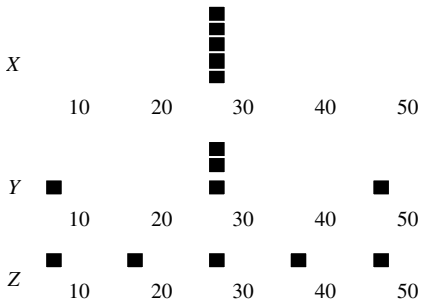
V

vague prior– A term used in **Bayesian statistics** to refer to a prior when the analyst lacks any information about the value of the unknown **parameter.** See also *informative prior, noninformative prior.*

validity– The property of a measuring instrument or test that measures the characteristic it is supposed to measure, i.e., the extent to which a **measurement** is free of any **systematic error.** It provides a measure of the **accuracy** of the concept it is intended to measure. The term is also used for a measurement or assessment that is not biased. See also *external validity, internal validity.*

validity checks– Routine checks in **data editing** to ensure that all the **data values** are correct within the allowable range. For example, an age of 193 years or a height of 123 inches clearly is not permissible a value.

variability– Variability refers to the characteristic of a set of **data points** to spread out and vary among themselves. It is the complementary quality to the **central tendency** of a **distribution.** The variability of a distribution is also referred to as its **dispersion, spread,** or **scatter.** Various **data sets** may have the same center but different amount of spreads. The **standard deviation** and **variance** are two of the most commonly used **measures of variability.**



Three data sets with the same mean and sample size but different amounts of spread

variable– Any quantity that varies; that is, an aspect or characteristic of a person, object, or situation that can assume different values. Examples of a variable are the height of men and the price on the New York Stock Exchange. The opposite of a variable is a **constant.** The term is often used as a clipped form of **random variable.**

variable sampling– A **sampling procedure** in which the characteristic of interest is measured on a numerical scale rather then merely classified by its quality or **attribute.** Compare *attribute sampling.*

variable selection– The problem of selecting the "best" possible set of **predictors** in using a **regression model.** See also *all subsets regression, backward elimination procedure, forward selection procedure, stepwise regression.*

variance– A **measure of variability** or **dispersion** of the values of a **data set** found by averaging the squared **deviations** about the **mean.** It is calculated by summing the squared deviations of the **data values** about the mean and then dividing the total by N if the data set is a **population** or by $n - 1$ if the data set is from a **sample.** See also *population variance, sample variance.*

variance analysis– Same as *analysis of variance.*

variance components– A term used to denote **variances** of **random effects** terms in an **analysis of variance** or **regression model.** Variance components are widely used in a variety of fields requiring the measurement of variance.

variance components model– In an **analysis of variance** or **regression model,** a term used to designate a **random** or a **mixed effects model.**

variance–covariance martix– Same as *covariance matrix.*

variance efficiency– See *efficiency.*

variance of the population– Same as *population variance.*

variance ratio– The **ratio** of two independent **estimates** of the **population variance.** The term is also used as an alternative name for the ***F* statistic.**

variance ratio distribution– Same as *F distribution.*

variance ratio test– Same as *F test.*

variance stabilizing transformation– The use of algebraic **transformation** on **data values** to make the **variances** constant for different groups of **data sets.** See also *data transformation.*

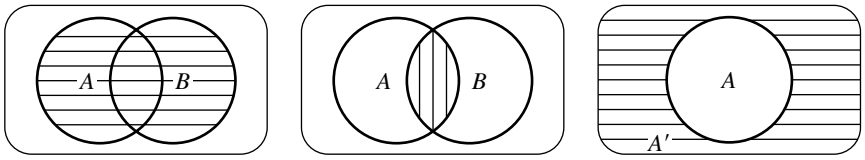
variate– A synonym for **variable.**

variate difference method– A technique for the analysis of **time-series data** whose **variations** stem from a systematic and a random component. It is based on the assumption that, if the systematic part of the series can be represented by a polynomial, then the successive differences will eliminate this component and thus help to estimate the **random variation.**

variation– The **scatter** or **variability** in **measurements** or **observations** of the same object or subject; that is, the extent to which observations are spread out. It may occur naturally or may represent an **error** in measurement. The term is often used as a synonym for **variability.**

vector– A one-dimensional **array** of numbers or mathematical objects. A row or column of a **matrix** constitutes a vector.

Venn diagram– A **graphical device** for symbolically representing the **sample space, events,** and operations of union, intersection, and complements involving events. Usually a rectangle is drawn to represent the sample space and various events are represented by circles contained within the rectangle. It is useful to demonstrate relationship or covariation among a set of events or **variables.**



$A \cup B$ $A \cap B$ Complement of event A

Some examples of Venn diagrams

vertical axis– The **ordinate** or vertical dimension of a two-dimensional **cartesian graph.** It is also called the ***y*** **axis.**

vital index– Same as *birth–death ratio.*

vital statistics– The **statistics** on **mortality** and **morbidity** such as birth, death, marriage, and divorce; so called because they have to do with life (la vita). More loosely any information about health and sickness is referred to as vital statistics.

volunteer bias– A possible source of **bias** that occurs when the subjects in a study are volunteers rather than a **representative sample** of the **study population.** Studies have shown that volunteers in a study tend to be different from nonvolunteers in terms of demographic characteristics and other psychosocial profiles.