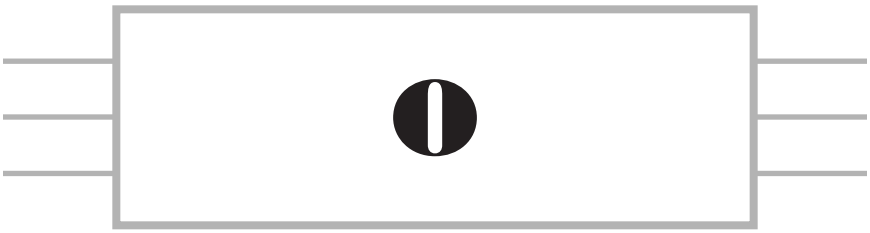


**objective probability–** In contrast to **subjective probability,** an **estimate** of **probability** based on **empirical** evidence from observable **events** or phenomena interpreted in the **frequency** sense. See also *classical probability, empirical probability.*

**oblique rotation–** See *factor rotation.*

**observation–** Process of a study or investigation; a **measurement, score,** or **datum** obtained from an **experiment.**

**observational study–** An epidemiologic study that does not involve an intervention or manipulation on the part of the investigator. In an observational study, the nature is allowed to take its course and differences in one characteristic are investigated in relation to differences in other characteristics without any human intervention. **Sample surveys** and most of the epidemiologic studies belong to this class. It is called a **case-control, cross-sectional,** or **cohort study,** according to the choice of the **study design.** See also *experimental study, prospective study, retrospective study.*

**observational unit–** The unit in an **experiment** on which an **observation** is made or recorded. An observational unit, however, may differ from an **experimental unit.** For example, in a **household survey,** a household may be an experimental unit but an individual within the household could be an observational unit.

**observed frequency–** In a **contingency table** the number of actual **observations** counted in each **cell** or category. In general, the number of times a particular **event** or phenomenon occurs. Compare *expected frequency.*

**observed significance level–** Same as *p value.*

**observed variable–** A synonym for **manifest variable.**

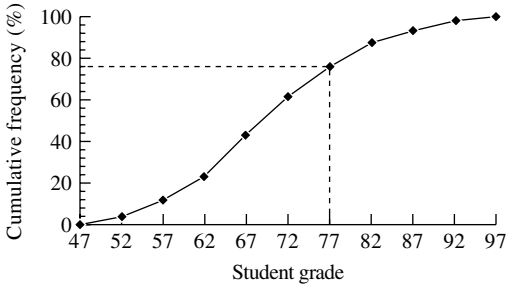
**occupational death rate–** The **death rate** calculated for a specific occupational or professional group or category. See also *age-specific death rate, cause-specific death rate.*

**occupational mortality rate–** Same as *occupational death rate.*

**odds–** The **ratio** of the **probability** that an **event** will occur to the probability that the event will not occur. It is calculated by the formula odds $= p/(1 - p)$, where $p$ is the probability of the event. The odds are used to convey the idea of probability, although for rare events the two are nearly the same. For a common event, such as getting a head or tail when a coin is tossed, the probability is 0.5 or 50%, but the odds are 1 (50:50). See also *odds ratio.*

**odds ratio–** A measure of the **relative risk** estimated in a **case-control study.** It is the **ratio** of the **odds** that a **case** was exposed to a given **risk factor** to the odds that a **control** was exposed to the risk factor. In general, the odds ratio represents the ratio of the odds favoring the occurrence of an **event** to that of another event. It is a measure of **association** between two **variables.** An odds ratio of 1 indicates that there is no relationship between the variables. An odds ratio less than 1 indicates an inverse or **negative relation** and an odds ratio greater than 1 indicates a direct or **positive relation.** In a **2 × 2 contingency table** it is calculated by the formula $(ad)/(bc)$, where $a$, $b$, $c$, and $d$ are the appropriate **cell counts.** The odds ratio is related to the **risk ratio** or relative risk in that, when the **probability** of the occurrence of the event is small (i.e., $a \ll b$ and $c \ll d$), the odds ratio equals the risk ratio. However, the odds ratio is a useful measure of difference in **risks** between two groups, irrespective of whether it approximates the relative risk or not, and arises in many important **statistical models** such as **logistic regression.** It is also called cross-product ratio.

**ogive–** A graph of the **cumulative frequency distribution** or **cumulative relative frequency distribution.** The cumulative frequency distribution may be plotted so that each **ordinate** of the ogive expresses either the number or **proportion** of **observations** "less than" or "greater than" the corresponding **abscissa.**




Ogive based on the cumulative frequency (percentage) column in the table under frequency distribution

**ogive curve–** Same as *ogive.*

**OLS–** Acronym for *ordinary least squares.*

**one-factor analysis of variance–** Same as *one-way analysis of variance.*

**one-sample *t* test–** A **test procedure** used to compare the **mean** of a single **sample** with a hypothetical **population mean,** when the **population variance** is unknown and estimated by the **sample variance.** See also *two-sample t test.*

**one-sided alternative–** An **alternative hypothesis** that allows or holds for the **deviation** from the **null hypothesis** to be in only one particular direction. For example, if the null hypothesis asserts that the **parameter** of interest $\mu$ is equal to some specified value $\mu_0$, the

alternative $\mu > \mu_0$ is a one-sided alternative. A **test of significance** based on a one-sided alternative is called **one-sided test.**

**one-sided hypothesis–** See *directional hypothesis.*

**one-sided test–** Same as *one-tailed test.*

**one-tailed hypothesis–** Same as *one-sided hypothesis.*

**one-tailed hypothesis test–** A **test of hypothesis** in which rejection of the **null hypothesis** occurs in only one tail of the **sampling distribution** of the **test statistic.** The **critical region** of a one-tailed test is located completely at one end of the distribution of the test statistic. A one-tailed test takes into account deviations in only one direction from the value stated under the null hypothesis, either those that are greater than it or those that are less than it. The one-tail is the area of the sampling distribution that serves as basis for the rejection or nonrejection of the null hypothesis. See also *lower-tailed test, two-tailed test, upper-tailed test.*

**one-tailed test–** Same as *one-tailed hypothesis test.*

**one-way analysis of variance–** An **analysis of variance** procedure involving only one **factor** or **independent variable.** The analysis of a **completely randomized design** is an example of a one-way analysis of variance. See also *multiway analysis of variance, two-way analysis of variance, three-way analysis of variance.*

**One-way ANOVA table for equal group sizes**

| Source of variation | Degrees of freedom | Sum of squares | Mean square | Variance ratio |
|---|---|---|---|---|
| Between groups | $k-1$ | $\text{SS}_B = n\sum_{i=1}^{k}(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$ | $\text{MS}_B = \dfrac{\text{SS}_B}{k-1}$ | $\dfrac{\text{MS}_B}{\text{MS}_W}$ |
| Within groups | $k(n-1)$ | $\text{SS}_W = \sum_{i=1}^{k}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_{i\cdot})^2$ | $\text{MS}_W = \dfrac{\text{SS}_W}{k(n-1)}$ | |
| Total | $nk-1$ | $\sum_{i=1}^{k}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_{\cdot\cdot})^2$ | | |

**One-way ANOVA table for unequal group sizes**

| Source of variation | Degrees of freedom | Sum of squares | Mean square | Variance ratio |
|---|---|---|---|---|
| Between groups | $k-1$ | $\text{SS}_B = \sum_{i=1}^{k} n_i(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2$ | $\text{MS}_B = \dfrac{\text{SS}_B}{k-1}$ | $\dfrac{\text{MS}_B}{\text{MS}_W}$ |
| Within groups | $\sum_{i=1}^{k}(n_i - 1)$ | $\text{SS}_W = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{i\cdot})^2$ | $\text{MS}_W = \dfrac{\text{SS}_W}{\sum_{i=1}^{k}(n_i - 1)}$ | |
| Total | $\sum_{i=1}^{k} n_i - 1$ | $\sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij} - \bar{Y}_{\cdot\cdot})^2$ | | |

**one-way classification–** A **classification** of a set of **observations** according to a single characteristic. See also *one-way analysis of variance, single factor experiment.*

**one-way design–** Same as *one-way classification.*

**one-way layout–** Same as *one-way classification.*

**open-ended class intervals– Class intervals** that have only one stated end point, the upper or lower limit.

**open-ended interval–** In a **grouped frequency distribution,** the highest (or lowest) interval that includes all values above (or below) a particular value.

**operational research–** Same as *operations research.*

**operations research–** A collection of quantitative methods and techniques involving optimization and **stochastic models** applicable to problems and activities of a complex system, such as those arising in a large business, industrial, or governmental organization, with a view to making optimal decisions and increase efficiency.

**opinion poll–** Same as *opinion survey.*

**opinion survey–** A **survey** designed to measure opinions possessed by members of a community concerning certain social, political, or other topics of interest. Field workers are employed for this purpose where each interviews a quota of people in the streets or other public places. See also *sample survey.*

**opportunity loss–** In **decision theory,** when a decision maker maximizes benefit, the opportunity loss is the difference between (1) the optimal **payoff** for a given **event** and (2) the actual payoff achieved as a result of taking a specified course of action and the subsequent occurrence of that event. When a decision maker minimizes cost, it is the difference between (1) the actual cost incurred as a consequence of taking a specified course of action and the subsequent occurrence of an event and (2) the minimum cost achievable for that event.

**optimal strategy–** In **decision theory,** a complete plan specifying the course of actions to be taken at each possible **action point,** if the expected monetary or utility **payoff** is to be the best one available.

**optimization methods–** A loosely defined term often used to designate procedures and techniques useful in finding optimal solutions of a given problem, which generally involve finding the maxima or minima of functions of several variables.

**optimum allocation–** In a **stratified random sampling,** the method of allocation of total **sample size** to various **strata** so as to maximize **precision** for a fixed cost. See also *proportional allocation.*

**ordered alternative–** An **alternative hypothesis** that specifies an order for a set of **parameters** being tested. For example, in a **one-way analysis of variance** problem with means $\mu_1, \mu_2, \ldots, \mu_k$, the null and ordered alternative hypotheses are: $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$ versus $H_1 : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$.

**ordered alternative hypothesis–** Same as *ordered alternative.*

**ordered array–** Same as *array.*

**ordered logistic regression–** A **logistic regression** method involving an **ordinal variable** as the **dependent variable.** See also *polytomous logistic regression.*

**order statistics–** A **sample** of **variate** values arranged in ascending order of magnitude are known as order statistics. For a sample of $n$ **measurements** with values $x_1, x_2, \ldots, x_n$ the order statistics are denoted by $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$. The $i$th largest value is called the $i$th order statistic and is denoted by $x_{(i)}$.

**ordinal contingency table–** A **contingency table** in which either row or column, or both, follow an ordinal **ranking.** If both row and column follow an ordinal ranking, the table is known as a doubly ordinal contingency table.

**ordinal data– Data** obtained by using an **ordinal level of measurement.**

**ordinal kappa statistic–** See *kappa statistic.*

**ordinal level of measurement–** Same as *ordinal scale.*

**ordinal scale–** Ordered scales in which the categories are defined in relationship to one another by the algebra of inequalities (less than and greater than). It is the process of rank ordering objects or persons with respect to some attribute from the "smallest" to the "largest." The measure assigned allows the items to be rank ordered with respect to a criterion. An ordinal scale has two properties: It classifies **observations** into classes in terms of the relationship of greater than or less than. There is no provision made for specifying the degree to which the observations differ from one another. Ordinal scaling produces ordinal or rank-ordered data. Some examples of an ordinal scale are social class and the Apgar score used to appraise the status of newborn infants.

**ordinal variable–** A **variable** measured on an **ordinal scale.** See also *categorical variable, continuous variable, discrete variable.*

**ordinary least squares–** Same as *least squares estimation.*

**ordinate–** The **vertical axis** or ***y* axis** on a graph using the **cartesian coordinate** system. Compare *abscissa.*

**orthogonal contrasts–** See *contrast.*

**orthogonal data– Experimental data** obtained by using an **orthogonal design.** Compare *nonorthogonal data.*

**orthogonal design–** A term used to denote an **analysis of variance** design with two or more **factors** having an equal number of **observations** in each **cell** or **level** of a factor, and where each **treatment** occurs the same number of times at all the levels. Compare *nonorthogonal design.*

**orthogonal matrix–** A **square matrix A** is said to be an orthogonal matrix if $\mathbf{A}\mathbf{A}' = \mathbf{I}$ where $\mathbf{A}'$ is the transpose of **A** and **I** is the **identity matrix.** Thus, for an orthogonal matrix, its transpose is equal to its inverse.

**orthogonal rotation–** See *factor rotation.*

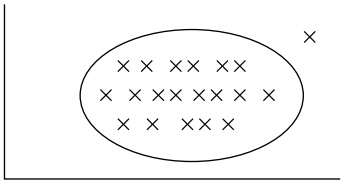
**orthogonal variables– Variables** that do not have any relationship to each other. More specifically, two variables are said to be orthogonal if they are statistically independent of each other.

**outcome–** A general term for the result of any **experiment** or **trial** measured as a **response value** of a **variable.** The term is also used as a synonym for **elementary event.**

**outcome space–** Same as *sample space.*

**outcome variable–** Same as *criterion variable.*

**outlier–** An **observation** that is so extreme that it stands apart from the rest of the observations; that is, it differs so greatly from the remaining observations that it give rises to the question whether it is from the same **population** or involves **measurement error. Statistical tests** are normally used to determine whether such an observation is indeed an outlier. The presence of outliers violates the assumption of **normality** and it may be necessary to transform the **data** or use **nonparametric methods.**



Example of an outlier

**overfitted model–** A term used for a **model** in which the number of **parameters** being fitted is larger than can be accommodated by the **data.** See also *overparametrized model.*

**overmatching–** A term used in the context of **matching** in **case-control studies** when the **cases** and **controls** are matched for **variables** that are not **confounding factors.** The use of overmatching results in a loss of **efficiency** of the design.

**overparametrized model–** A term used for a **model** in which the number of **parameters** being fitted is larger than the number of **observations** available for **estimation.** See also *overfitted model.*

**overviews–** An alternative term for **meta-analysis.**