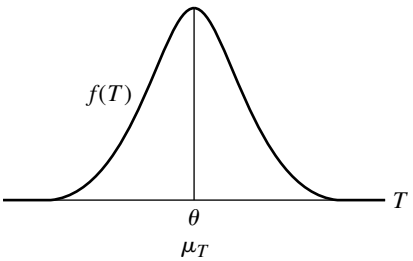
U

unbalanced data– Same as *nonorthogonal data.*

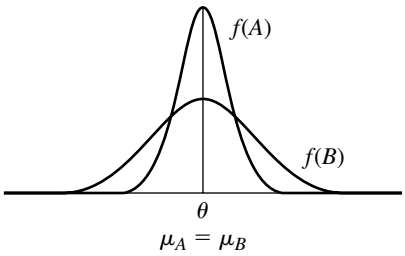
unbalanced design– Same as *nonorthogonal design.*

unbiased confidence interval– A **confidence interval** is said to be unbiased if the **probability** of containing any value not equal to the true **parameter** value is less than or equal to $1 - \alpha$. See also *uniformly most accurate interval, uniformly most accurate unbiased interval.*

unbiased estimator– An **estimator** whose **expected value** or **mean** equals the true value of the **parameter** being estimated. Thus, an unbiased estimator on the **average** assumes a value equal to the true **population parameter.** An unbiased estimator neither systematically overestimates or underestimates the value of the parameter in question. Compare *biased estimator.*



Sampling distribution of an unbiased estimator of θ



Sampling distribution of two unbiased estimators of θ with different variances

unbiased linear estimator– An **unbiased estimator** that is a **linear function** of **observations.**

unbiasedness– A term used to describe the property of an **unbiased estimator.** Compare *biasedness.* See also *biased estimator.*

uncertainty– A term denoting the lack of certainty inherent in a **random phenomenon.** When a coin is tossed it is uncertain whether it will turn up a head or tail. There is 0.5 **probability** that it will show up a head or tail. It is probability that is measurable, but not the uncertainty.

uncertainty analysis– A method of analysis carried out to determine the **variability** in the final **outcome** of a **variable** that is due to **uncertainty** inherent in the values of one or more input **parameters.** See also *sensitivity analysis.*

unconditional probability– A measure of the likelihood that a particular **event** occurs, irrespective of whether another event occurs or not. Compare *conditional probability.*

uncontrolled clinical trial– Same as *uncontrolled trial.*

uncontrolled trial– A **clinical trial** that has no **control group.** Compare *controlled trial.*

unequal probability sampling– A **sampling design** in which each **sampling unit** in a **population** has a different **probability** of being included in the **sample.** Unequal probability sampling can often reduce the **variance** of an **estimator** by sampling each unit with probability proportional to a measure of the size of the unit. **Cluster sampling** provides an ideal situation in which the unequal probability sampling, with probability proportional to the number of elements in the **cluster,** results in reducing the bound on the **error of estimation.**

unequal variance *t* test– When the **sample variances** (S_1^2 and S_2^2) suggest that there may be a problem in assuming that the two **population variances** are equal, we can modify the usual ***t* statistic** to obtain an approximate ***t* test** or *t* **confidence interval.** B. L. Welch in 1938 showed that the distribution of the **statistic**

$$t' = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$$

can be approximated by a ***t* distribution** with **degrees of freedom** (df) given by

$$\text{df} = \frac{(n_1 - 1)(n_2 - 1)}{(n_2 - 1)c^2 + (1 - c)^2(n_1 - 1)} \qquad \text{where } c = \frac{S_1^2/n_1}{S_1^2/n_1 + S_2^2/n_2}$$

unexplained variation– In an **analysis of variance,** the **variation** of the **sample data** within each of the **samples** about the respective **sample means;** it is attributed to **chance** and equals the **total sum of squares** minus **explained variation.** In a **regression analysis,** it is the **sum of the squares** of all the unexplained **deviations.** It is also referred to as **error** or **residual sum of squares.**

unfair gamble– In the **theory of games,** a game of chance in which the **expected monetary gain** of what is being lost exceeds the expected monetary gain of what is being received. Compare *fair gamble.*

ungrouped data– Data values in their original form that have not been grouped into **class intervals** in order to reduce the number of scoring categories. Compare *grouped data.* See also *ungrouped frequency distribution.*

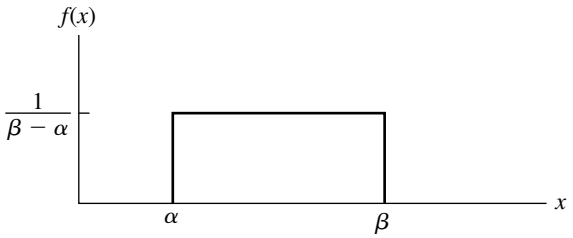
ungrouped frequency distribution– A **frequency distribution** that lists **frequencies** for individual **scores** that have not been grouped into **class intervals.** Compare *grouped frequency distribution.*

uniform distribution– A **distribution** that, in the continuous case, has a constant **density** over a given interval and, in the discrete case, assigns the same **probability** to each value within its domain. The **probability density function** of a **continuous random variable** is

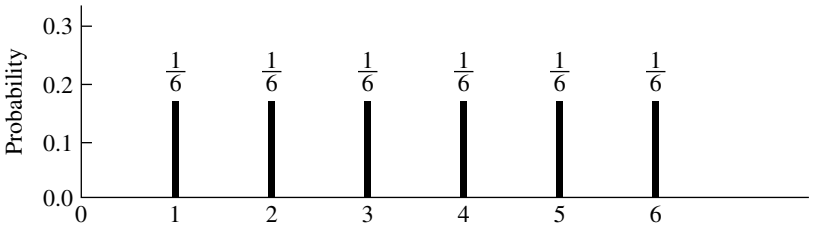
$$f(x) = \frac{1}{\beta - \alpha} \qquad \alpha < x < \beta$$

The **probability function** of a **discrete random variable** is

$$p(x) = \frac{1}{k} \qquad x = x_1, x_2, \ldots, x_k$$



Probability distribution of a continuous random variable having a uniform distribution



Probability distribution of a discrete random variable having a uniform distribution

uniformly most accurate interval– A **confidence interval** is said to be uniformly most accurate if the interval has a smaller **probability** of containing a value not equal to the true **parameter** value than any other interval with the same **confidence coefficient.**

uniformly most accurate unbiased interval– A **confidence interval** is said to be uniformly most accurate unbiased if it is uniformly most accurate within the class of all **unbiased confidence intervals.**

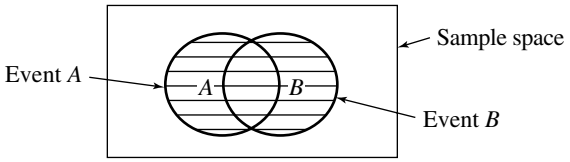
uniformly most powerful one-sided test– Same as *uniformly most powerful test.*

uniformly most powerful test– A test of a **hypothesis** against a composite alternative that is at least as powerful as any other test for all the values of the alternative and more powerful against one of the alternatives. In most situations, uniformly most powerful tests exist when the **alternative hypothesis** is constrained in some way; for example in testing $H_0{:}\theta = \theta_0$ against $H_1{:}\theta < \theta_0$ or $\theta > \theta_0$, but not both. If the test is uniformly most powerful for either one of the two sets of alternatives, it is called a uniformly most powerful one-sided test.

uniformly shortest length interval– A confidence interval is said to be uniformly shortest length if it has a shorter expected length than, or the shortest expected length of, any other interval with the same **confidence coefficient.**

unimodal distribution– A **distribution** having only one **mode.** See also *bimodal distribution, multimodal distribution, trimodal distribution.*

union of two events– The union of two **events** A and B, denoted by $A \cup B$, is the event that consists of all **outcomes** that belong to A, to B, or to both.



The shaded region depicits $A \cup B$

Figure showing union of two events A and B

unit normal (random) variable– A term used to denote a **normal random variable** with unit **standard deviation.**

unit of analysis– The level of aggregates that is being studied or investigated. The unit of analysis may be individuals, schools, hospitals, countries, and so forth.

univariable analysis– A term sometimes used in contradistinction to **univarite analysis** to refer to an analysis that contains one **independent variable** at a time.

univariate analysis– Statistical analysis involving **measurements** on only one **variable.** The term is used in contrast to **bivariate** and **multivariate analysis** involving measurements on two or more variables simultaneously.

univariate data set– A **data set** containing **measurement** values on one **variable** only.

univariate distribution– The **distribution** of a set of **scores** that measures only one **variable** at a time. This is the usual type of distribution that displays the score values for a single **random variable.**

univariate *k* statistic– See *k statistics.*

univariate normal distribution– Same as *normal distribution.*

universal set– In **set theory,** the set consisting of all elements.

universe– The aggregate values, of which the values observed in the **sample** constitute **a representative sample,** and to which the findings of the sample can be generalized. The universe may be a hypothetical or a real population of values, and it may be finite or infinite, depending on the type of sample and the nature of the information under study. In any statistical practice, the universe under study or investigation needs to be carefully circumscribed. The term is more or less synonmous with **population.**

unreplicated factorial design– A term sometimes used for a 2^k **factorial design** containing a single **replicate.**

unstandardized score– A **score** in the original unit that has not been transformed into a *z* **score** or any other **standard score.**

unweighted mean– An **arithmetic mean** of a set of **observations** in which no weights are assigned to them. Compare *weighted mean.*

unweighted means analysis– A method of analysis in two-way and higher-order **factorial designs** containing unequal numbers of **observations** in each **cell.** The procedure consists of calculating the **cell means** and then carrying out a **balanced data** analysis by assuming that the cell means constitute a single observation in each cell.

upper confidence limit– See *confidence limits.*

upper hinge– See *five-number summary.*

upper *p*th percentile– A value such that $100p\%$ of the **observations** in the **population** have **measurements** greater than this value and $100(1 - p)\%$ of the observations are less than its value.

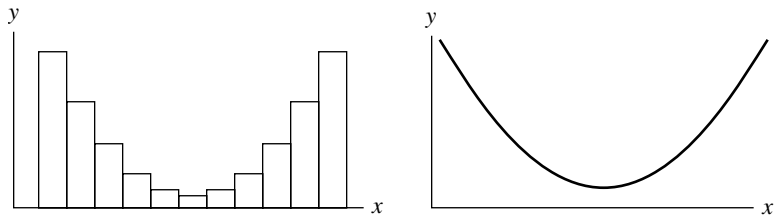
upper real limit– See *real limits.*

upper-tailed test– A **one-tailed hypothesis test** in which the entire **rejection region** is located in the upper tail of the **sampling distribution** of the **test statistic.** See also *lower-tailed test, one-tailed test, two-tailed test.*

U. S. Bureau of the Census– Same as *Bureau of the Census.*

U-shaped distribution– An asymmetrical **frequency distribution** having general resemblance to the shape of the letter U. The **distribution** has maximum **frequencies** at both ends of the distribution, which decline rapidly at first and then more slowly, reaching a minimum between them.

utility– In **decision analysis,** the term used to denote the monetary value of an **outcome.** It is the gain often expressed in terms of money derived from the decision outcome.



U-shaped distribution: **(a)** histogram and **(b)** continuous curve

utility analysis– In **decision analysis,** a method for making decisions under **uncertainty** that is based on certain axioms of rational behavior.

utility-of-money function– The relationship between alternative amounts of money a player might possess and the different utility values associated with these amounts.

utility theory– A branch of **decision theory,** in which **utilities** of different **outcomes** are assigned numerical values in order of preference, made by referring to the **expected monetary return** and the risks involved. The values range from 0 to 1, 0 being alloted to the least preferred and 1 to the most preferred outcome.