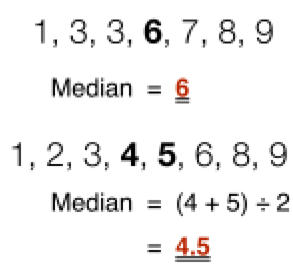
Median



Finding the median in sets of data with an odd and even number of values

In statistics and probability theory, a **median** is a value separating the higher half from the lower half of a data sample, a population or a probability distribution.

For a data set, it may be thought of as "the middle" value. For example, the basic advantage of the median in describing data compared to the mean (often simply described as the "average") is that it is not skewed so much by a small proportion of extremely large or small values, and so it may give a better idea of a "typical" value. For example, in understanding statistics like household income or assets, which vary greatly, the mean may be skewed by a small number of extremely high or low values. Median income, for example, may be a better way to suggest what a "typical" income is. Because of this, the median is of central importance in robust statistics, as it is the most resistant statistic, having a breakdown point of 50%: so long as no more than half the data are contaminated, the median will not give an arbitrarily large or small result.

Finite data set of numbers

The median of a finite list of numbers is the "middle" number, when those numbers are listed in order from smallest to greatest.

If there is an odd number of numbers, the middle one is picked. For example, consider the list of numbers

1, 3, 3, **6**, 7, 8, 9

This list contains seven numbers. The median is the fourth of them, which is 6.

If there is an even number of observations, then there is no single middle value; the median is then usually defined to be the mean of the two middle values.[1][2] For example, in the data set

1, 2, 3, **4, 5**, 6, 8, 9

the median is the mean of the middle two numbers: this is $(4+5)/2$, which is 4.5. (In more technical terms, this interprets the median as the fully trimmed mid-range). With this convention, the median can be described in a caseless formula, as follows:

$$\operatorname{median}(x) = \frac{1}{2}\left(x_{\lfloor (n+1)/2 \rfloor} + x_{\lceil (n+1)/2 \rceil}\right)$$

where x is an ordered list of n numbers, and $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the floor and ceiling functions, respectively.

Comparison of common averages of values [1, 2, 2, 3, 4, 7, 9]

Type	Description	Example	Result
Arithmetic mean	Sum of values of a data set divided by number of values: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	(1 + 2 + 2 + 3 + 4 + 7 + 9) / 7	**4**
Median	Middle value separating the greater and lesser halves of a data set	1, 2, 2, **3**, 4, 7, 9	**3**
Mode	Most frequent value in a data set	1, **2**, **2**, 3, 4, 7, 9	**2**

Formal definition

Formally, a median of a population is any value such that at most half of the population is less than the proposed median and at most half is greater than the proposed median. As seen above, medians may not be unique. If each set contains less than half the population, then some of the population is exactly equal to the unique median.

The median is well-defined for any ordered (one-dimensional) data, and is independent of any distance metric. The median can thus be applied to ranked but not numerical classes (e.g. working out a median grade when students are graded from A to F), although the result might be halfway between classes if there is an even number of cases.

A geometric median, on the other hand, is defined in any number of dimensions. A related concept, in which the outcome is forced to correspond to a member of the sample, is the medoid.

There is no widely accepted standard notation for the median, but some authors represent the median of a variable *x* either as $\tilde{x}$ or as $\mu_{1/2}$[1] sometimes also *M*.[3][4] In any of these cases, the use of these or other symbols for the median needs to be explicitly defined when they are introduced.

The median is a special case of other ways of summarising the typical values associated with a statistical distribution: it is the 2nd quartile, 5th decile, and 50th percentile.

Uses

The median can be used as a measure of location when one attaches reduced importance to extreme values, typically because a distribution is skewed, extreme values are not known, or outliers are untrustworthy, i.e., may be measurement/transcription errors.

For example, consider the multiset

1, 2, 2, 2, 3, 14.

The median is 2 in this case, (as is the mode), and it might be seen as a better indication of the center than the arithmetic mean of 4, which is larger than all-but-one of the values! However, the widely cited empirical relationship that the mean is shifted "further into the tail" of a distribution than the median is not generally true. At most, one can say that the two statistics cannot be "too far" apart;

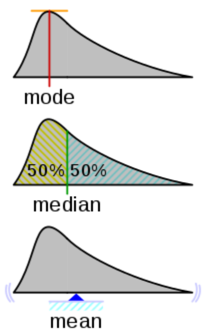
see § Inequality relating means and medians below.[5]

As a median is based on the middle data in a set, it is not necessary to know the value of extreme results in order to calculate it. For example, in a psychology test investigating the time needed to solve a problem, if a small number of people failed to solve the problem at all in the given time a median can still be calculated.[6]

Because the median is simple to understand and easy to calculate, while also a robust approximation to the mean, the median is a popular summary statistic in descriptive statistics. In this context, there are several choices for a measure of variability: the range, the interquartile range, the mean absolute deviation, and the median absolute deviation.

For practical purposes, different measures of location and dispersion are often compared on the basis of how well the corresponding population values can be estimated from a sample of data. The median, estimated using the sample median, has good properties in this regard. While it is not usually optimal if a given population distribution is assumed, its properties are always reasonably good. For example, a comparison of the efficiency of candidate estimators shows that the sample mean is more statistically efficient when — and only when — data is uncontaminated by data from heavy-tailed distributions or from mixtures of distributions. Even then, the median has a 64% efficiency compared to the minimum-variance mean (for large normal samples), which is to say the variance of the median will be ~50% greater than the variance of the mean.[7][8]

Probability distributions



Geometric visualisation of the mode, median and mean of an arbitrary probability density function[9]

For any real-valued probability distribution with cumulative distribution function *F*, a median is defined as any real number *m* that satisfies the inequalities

$$\int_{(-\infty,m]} dF(x) \geq \frac{1}{2} \text{ and } \int_{[m,\infty)} dF(x) \geq \frac{1}{2}$$

.

An equivalent phrasing uses a random variable *X* distributed according to *F*:

$$\mathrm{P}(X \leq m) \geq \frac{1}{2} \text{ and } \mathrm{P}(X \geq m) \geq \frac{1}{2}$$

Note that this definition does not require *X* to have an absolutely continuous distribution (which has a probability density function *f*), nor does it require a discrete one. In the former case, the inequalities can be upgraded to equality: a median satisfies

$$\mathrm{P}(X \leq m) = \int_{-\infty}^{m} f(x)\,dx = \frac{1}{2} = \int_{m}^{\infty} f(x)\,dx = \mathrm{P}(X \geq m)$$

.

Any probability distribution on **R** has at least one median, but in pathological cases there may be more than one median: if F is constant 1/2 on an interval (so that f=0 there), then any value of that interval is a median.

Medians of particular distributions

The medians of certain types of distributions can be easily calculated from their parameters; furthermore, they exist even for some distributions lacking a well-defined mean, such as the Cauchy distribution:

- The median of a symmetric unimodal distribution coincides with the mode.

- The median of a symmetric distribution which possesses a mean μ also takes the value μ.
    - The median of a normal distribution with mean μ and variance σ^2 is μ. In fact, for a normal distribution, mean = median = mode.
    - The median of a uniform distribution in the interval [a, b] is (a + b) / 2, which is also the mean.
- The median of a Cauchy distribution with location parameter x_0 and scale parameter y is x_0, the location parameter.

- The median of a power law distribution x^{-a}, with exponent $a > 1$ is $2^{1/(a-1)}x_{\min}$, where $x_{\min}$ is the minimum value for which the power law holds[10]
- The median of an exponential distribution with rate parameter λ is the natural logarithm of 2 divided by the rate parameter: $\lambda^{-1}\ln 2$.
- The median of a Weibull distribution with shape parameter k and scale parameter λ is $\lambda(\ln 2)^{1/k}$.

Populations

Optimality property

...

The *mean absolute error* of a real variable *c* with respect to the random variable *X* is

$$E(|X-c|)$$

Provided that the probability distribution of *X* is such that the above expectation exists, then *m* is a median of *X* if and only if *m* is a minimizer of the mean absolute error with respect to *X*.[11] In particular, *m* is a sample median if and only if *m* minimizes the arithmetic mean of the absolute deviations.[12]

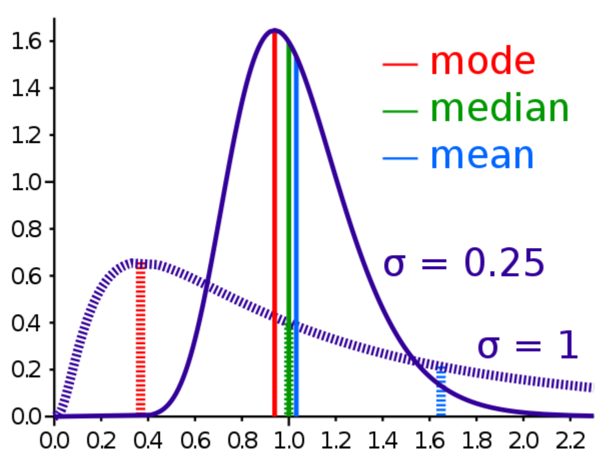
More generally, a median is defined as a minimum of

$$E(|X-c|-|X|),$$

as discussed below in the section on multivariate medians (specifically, the spatial median).

This optimization-based definition of the median is useful in statistical data-analysis, for example, in *k*-medians clustering.

Inequality relating means and medians



Comparison of mean, median and mode of two log-normal distributions with different skewness

If the distribution has finite variance, then the distance between the median $\tilde{X}$ and the mean $\bar{X}$ is bounded by one standard deviation.

This bound was proved by Mallows,[13] who used Jensen's inequality twice, as follows. Using |·| for the absolute value, we have

$$\begin{aligned} |\mu - m| = |\mathrm{E}(X-m)| &\leq \mathrm{E}(|X-m|) \\ &\leq \mathrm{E}(|X-\mu|) \\ &\leq \sqrt{\mathrm{E}\left((X-\mu)^2\right)} = \sigma. \end{aligned}$$

The first and third inequalities come from Jensen's inequality applied to the absolute-value function and the square function, which are each convex. The second inequality comes from the fact that a median minimizes the absolute deviation function $a \mapsto \mathrm{E}(|X-a|)$.

Mallows' proof can be generalized to obtain a multivariate version of the inequality[14] simply by replacing the absolute value with a norm:

$$\|\mu - m\| \leq \sqrt{\mathrm{E}\left(\|X-\mu\|^2\right)} = \sqrt{\operatorname{trace}(\operatorname{var}(X))}$$

where *m* is a spatial median, that is, a minimizer of the function

$a \mapsto \mathrm{E}(\|X - a\|).$ The spatial median is unique when the data-set's dimension is two or more.[15][16]

An alternative proof uses the one-sided Chebyshev inequality; it appears in an inequality on location and scale parameters. This formula also follows directly from Cantelli's inequality.[17]

Unimodal distributions

For the case of unimodal distributions, one can achieve a sharper bound on the distance between the median and the mean:

$$\left|\tilde{X} - \bar{X}\right| \leq \left(\frac{3}{5}\right)^{1/2} \sigma \approx 0.7746\sigma$$

.[18]

A similar relation holds between the median and the mode:

$$\left|\tilde{X} - \mathrm{mode}\right| \leq 3^{1/2}\sigma \approx 1.732\sigma.$$

Jensen's inequality for medians

Jensen's inequality states that for any random variable *X* with a finite expectation *E*[*X*] and for any convex function *f*

$$f[E(x)] \leq E[f(x)]$$

This inequality generalizes to the median as well. We say a function $f : \mathbb{R} \to \mathbb{R}$ is a **C function** if, for any *t*,

$$f^{-1}\left(\,(-\infty,t]\,\right) = \{x \in \mathbb{R} \mid f(x) \leq t\}$$

is a closed interval (allowing the degenerate cases of a single point or an empty set). Every C function is convex, but the reverse does not hold. If *f* is a C function, then

$$f(\mathrm{Median}[X]) \leq \mathrm{Median}[f(X)]$$

If the medians are not unique, the statement holds for the corresponding suprema.[19]

Medians for samples

The sample median

...

Efficient computation of the sample median

...

Even though comparison-sorting *n* items requires $\Omega(n \log n)$ operations, selection algorithms can compute the kth-smallest of n items with only $\Theta(n)$ operations. This includes the median, which is the $\frac{n}{2}$th order statistic (or for an even number of samples, the arithmetic mean of the two middle order statistics).[20]

Selection algorithms still have the downside of requiring $\Omega(n)$ memory, that is, they need to have the full sample (or a linear-sized portion of it) in memory. Because this, as well as the linear time requirement, can be prohibitive, several estimation procedures for the median have been developed. A simple one is the median of three rule, which estimates the median as the median of a three-element subsample; this is commonly used as a subroutine in the quicksort sorting algorithm, which uses an estimate of its input's median. A more robust estimator is Tukey's *ninther*, which is the median of three rule applied with limited recursion:[21] if A is the sample laid out as an array, and

$$\mathrm{med3}(A) = \mathrm{median}(A[1], A[\tfrac{n}{2}], A[n]),$$

then

$$\mathrm{ninther}(A) = \mathrm{med3}(\mathrm{med3}(A[1 \ldots \tfrac{1}{3}n]), \mathrm{m}$$

The *remedian* is an estimator for the median that requires linear time but sub-linear memory, operating in a single pass over the sample.[22]

Sampling distribution

...

The distributions of both the sample mean and the sample median were determined by Laplace.[23] The distribution of the sample median from a population with a density function $f(x)$ is asymptotically normal with mean m and variance[24]

$$\frac{1}{4nf(m)^2}$$

where m is the median of $f(x)$ and n is the sample size. A modern proof follows below. Laplace's result is now understood as a special case of the asymptotic distribution of arbitrary quantiles.

For normal samples, the density is $f(m) = 1/\sqrt{2\pi\sigma^2}$, thus for large samples the variance of the median equals $(\pi/2) \cdot (\sigma^2/n)$.[7] (See also section #Efficiency below.)

Derivation of the asymptotic distribution

E...

We take the sample size to be an odd number $N = 2n + 1$ and assume our variable continuous; the formula for the case of discrete variables is given below in § Empirical local density. The sample can be summarized as "below median", "at median", and "above median", which corresponds to a trinomial distribution with probabilities $F(v-1)$, $f(v)$ and $1 - F(v)$. For a continuous variable, the probability of multiple sample values being exactly equal to the median is 0, so one can calculate the density of at the point v directly from the trinomial distribution:

$$\Pr[\mathrm{Median} = v]\, dv = \frac{(2n+1)!}{n!n!} F(v)^n (1 - F(v))^n f(v)\, dv$$

.

Now we introduce the beta function. For integer arguments α and β, this can be expressed as $\mathrm{B}(\alpha, \beta) = \frac{(\alpha - 1)!(\beta - 1)!}{(\alpha + \beta - 1)!}$. Also, recall that $f(v)\, dv = dF(v)$. Using these relationships and setting both α and β equal to $n+1$ allows the last expression to be written as

$$\frac{F(v)^n(1-F(v))^n}{\mathrm{B}(n+1,n+1)}\,dF(v)$$

Hence the density function of the median is a symmetric beta distribution pushed forward by F. Its mean, as we would expect, is 0.5 and its variance is $1/(4(N+2))$. By the chain rule, the corresponding variance of the sample median is

$$\frac{1}{4(N+2)f(m)^2}.$$

The additional 2 is negligible in the limit.

Empirical local density

E...

In practice, the functions f and F are often not known or assumed. However, they can be estimated from an observed frequency distribution. In this section, we give an example. Consider the following table, representing a sample of 3,800 (discrete-valued) observations:

v	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
f(v)	0.000	0.008	0.010	0.013	0.083	0.108	0.328	0.220	0.202	0.023	0.005
F(v)	0.000	0.008	0.018	0.031	0.114	0.222	0.550	0.770	0.972	0.995	1.000

Because the observations are discrete-valued, constructing the exact distribution of the median is not an immediate translation of the above expression for $\Pr(\text{Median} = v)$; one may (and typically does) have multiple instances of the median in one's sample. So we must sum over all these possibilities:

$$\Pr(\text{Median} = v) = \sum_{i=0}^{n} \sum_{k=0}^{n} \frac{N!}{i!(N-i-k)!k!} F(v-1)^i (1-F(v))^k f(v)^{N-i-k}$$

Here, *i* is the number of points strictly less than the median and *k* the number strictly greater.

Using these preliminaries, it is possible to investigate the effect of sample size on the standard errors of the mean and median. The observed mean is 3.16, the observed raw median is 3 and the observed interpolated median is 3.174.

The following table gives some comparison statistics.

Statistic / Sample size	3	9	15	21
Expected value of median	3.198	3.191	3.174	3.161
Standard error of median (above formula)	0.482	0.305	0.257	0.239
Standard error of median (asymptotic approximation)	0.879	0.508	0.393	0.332
Standard error of mean	0.421	0.243	0.188	0.159

The expected value of the median falls slightly as sample size increases while, as would be expected, the standard errors of both the median and the mean are proportionate to the inverse square root of the sample size. The asymptotic approximation errs on the side of caution by overestimating the standard error.

Estimation of variance from sample ...

data

The value of $(2f(x))^{-2}$—the asymptotic value of $n^{-\frac{1}{2}}(\nu - m)$ where ν is the population median—has been studied by several authors. The standard "delete one" jackknife method produces inconsistent results.[25] An alternative—the "delete k" method—where k grows with the sample size has been shown to be asymptotically consistent.[26] This method may be computationally expensive for large data sets. A bootstrap estimate is known to be consistent,[27] but converges very slowly (order of $n^{-\frac{1}{4}}$).[28] Other methods have been proposed but their behavior may differ between large and small samples.[29]

Efficiency

The efficiency of the sample median, measured as the ratio of the variance of the mean to the variance of the median, depends on the sample size and on the underlying population distribution. For a sample of size $N = 2n + 1$ from the normal distribution, the efficiency for large N is

$$\frac{2}{\pi}\frac{N+2}{N}$$

The efficiency tends to $\frac{2}{\pi}$ as N tends to infinity.

In other words, the relative variance of the median will be $\pi/2 \approx 1.57$, or 57% greater than the variance of the mean – the relative standard error of the median will be $(\pi/2)^{-1/2} \approx 1.25$, or 25% greater than the standard error of the mean, $\sigma/\sqrt{n}$ (see also section #Sampling distribution above.).[30]

Other estimators

...

For univariate distributions that are *symmetric* about one median, the Hodges–Lehmann estimator is a robust and highly efficient estimator of the population median.[31]

If data are represented by a statistical model specifying a particular family of probability distributions, then estimates of the median can be obtained by fitting that family of probability distributions to the data and calculating the theoretical median of the fitted distribution. Pareto interpolation is an application of this when the population is assumed to have a Pareto distribution.

Multivariate median

Previously, this article discussed the univariate median, when the sample or population had one-dimension. When the dimension is two or higher, there are multiple concepts that extend the definition of the univariate median; each such multivariate median agrees with the univariate median when the dimension is exactly one.[31][32][33][34]

Marginal median

The marginal median is defined for vectors defined with respect to a fixed set of coordinates. A marginal median is defined to be the vector whose components are univariate medians. The marginal median is easy to compute, and its properties were studied by Puri and Sen.[31][35]

Geometric median

The geometric median of a discrete set of sample points $x_1, \ldots x_N$ in a Euclidean space is the[a] point minimizing the sum of distances to the sample points.

$$\hat{\mu} = \underset{\mu \in \mathbb{R}^m}{\arg\min} \sum_{n=1}^{N} \left\| \mu - x_n \right\|_2$$

In contrast to the marginal median, the geometric median is equivariant with respect to Euclidean similarity transformations such as translations and rotations.

Centerpoint

An alternative generalization of the median in higher dimensions is the centerpoint.

Other median-related concepts

Interpolated median

When dealing with a discrete variable, it is sometimes useful to regard the observed values as being midpoints of underlying continuous intervals. An example of this is a Likert scale, on which opinions or preferences are expressed on a scale with a set number of possible responses. If the scale consists of the positive integers, an observation of 3 might be regarded as representing the interval from 2.50 to 3.50. It is possible to estimate the median of the underlying variable. If, say, 22% of the observations are of value 2 or below and 55.0% are of 3 or below (so 33% have the value 3), then the median m is 3 since the median is the smallest value of x for which $F(x)$ is greater than a half. But the interpolated median is somewhere between 2.50 and 3.50. First we add half of the interval width w to the median to get the upper bound of the median interval. Then we subtract that proportion of the interval width which equals the proportion of the 33% which lies above the 50% mark. In other words, we split up the interval width pro rata to the numbers of observations. In this case, the 33% is split into 28% below the median and 5% above it so we subtract 5/33 of the interval width from the upper bound of 3.50 to give an interpolated median of 3.35. More formally, if the values $f(x)$ are known, the interpolated median can be calculated from

$$m_{\text{int}} = m + w\left[\frac{1}{2} - \frac{F(m) - \frac{1}{2}}{f(m)}\right].$$

Alternatively, if in an observed sample there are k scores above the median category, j scores in it and i scores below it then the interpolated median is given by

$$m_{\text{int}} = m - \frac{w}{2}\left[\frac{k-i}{j}\right].$$

Pseudo-median

...

For univariate distributions that are *symmetric* about one median, the

Hodges–Lehmann estimator is a robust and highly efficient estimator of the population median; for non-symmetric distributions, the Hodges–Lehmann estimator is a robust and highly efficient estimator of the population *pseudo-median*, which is the median of a symmetrized distribution and which is close to the population median.[37] The Hodges–Lehmann estimator has been generalized to multivariate distributions.[38]

Variants of regression

The Theil–Sen estimator is a method for robust linear regression based on finding medians of slopes.[39]

Median filter

In the context of image processing of monochrome raster images there is a type of noise, known as the salt and pepper noise, when each pixel independently becomes black (with some small probability) or white (with some small probability), and is unchanged otherwise (with the probability close to 1). An image constructed of median values of neighborhoods (like 3×3 square) can effectively reduce noise in this case.

Cluster analysis

...

In cluster analysis, the k-medians clustering algorithm provides a way of defining clusters, in which the criterion of maximising the distance between cluster-means that is used in k-means clustering, is replaced by maximising the distance between cluster-medians.

Median–median line

...

This is a method of robust regression. The idea dates back to Wald in 1940 who suggested dividing a set of bivariate data into two halves depending on the value of the independent parameter x: a left half with values less than the median and a right half with values greater than the median.[40] He suggested taking the means of the dependent y and independent x variables of the left and the right halves and estimating the slope of the line joining these two points. The line could then be adjusted to fit the majority of the points in the data set.

Nair and Shrivastava in 1942 suggested a similar idea but instead advocated dividing the sample into three equal parts before calculating the means of the subsamples.[41] Brown and Mood in 1951 proposed the idea of using the medians of two subsamples rather the means.[42] Tukey combined these ideas and recommended dividing the sample into three equal size subsamples and estimating the line based on the medians of the subsamples.[43]

Median-unbiased estimators

Any *mean*-unbiased estimator minimizes the risk (expected loss) with respect to the squared-error loss function, as observed by Gauss. A *median*-unbiased estimator minimizes the risk with respect to the absolute-deviation loss function, as observed by Laplace. Other loss functions are used in statistical theory, particularly in robust statistics.

The theory of median-unbiased estimators was revived by George W. Brown in 1947:[44]

> *An estimate of a one-dimensional parameter θ will be*

> *said to be median-unbiased if, for fixed θ, the median of the distribution of the estimate is at the value θ; i.e., the estimate underestimates just as often as it overestimates. This requirement seems for most purposes to accomplish as much as the mean-unbiased requirement and has the additional property that it is invariant under one-to-one transformation.*
>
> *—page 584*

Further properties of median-unbiased estimators have been reported.[45][46][47][48] Median-unbiased estimators are invariant under one-to-one transformations.

There are methods of constructing median-unbiased estimators that are optimal (in a sense analogous to the minimum-variance property for mean-unbiased estimators). Such constructions exist for probability distributions having monotone likelihood-functions.[49][50] One such procedure is an analogue of the Rao–Blackwell procedure for mean-unbiased estimators: The procedure holds for a smaller class of probability distributions than does the Rao—Blackwell procedure but for a larger class of loss functions.[51]

History

Scientific researchers in the ancient near east appear not to have used summary statistics altogether, instead choosing values that offered maximal consistency with a broader theory that integrated a wide variety of phenomena.[52] Within the Mediterranean (and, later, European) scholarly community, statistics like the mean are fundamentally a medieval and early modern development. (The history of the median outside Europe and its predecessors remains relatively unstudied.)

The idea of the median appeared in the 13th century in the Talmud, in order to fairly analyze divergent appraisals.[53][54] However, the concept did not spread to the broader scientific community.

Instead, the closest ancestor of the modern median is the mid-range, invented by Al-Biruni.[55]:31[56] Transmission of Al-Biruni's work to later scholars is unclear. Al-Biruni applied his technique to assaying metals, but, after he published his work, most assayers still adopted the most unfavorable value from their results, lest they appear to cheat.[55]:35–8 However, increased navigation at sea during the Age of Discovery meant that ship's navigators increasingly had to attempt to determine latitude in unfavorable weather against hostile shores, leading to renewed interest in summary statistics. Whether rediscovered or independently invented, the mid-range is recommended to nautical navigators in Harriot's "Instructions for Raleigh's Voyage to Guiana, 1595".[55]:45–8

The idea of the median may have first appeared in Edward Wright's 1599 book

Certaine Errors in Navigation on a section about compass navigation. Wright was reluctant to discard measured values, and may have felt that the median — incorporating a greater proportion of the dataset than the mid-range — was more likely to be correct. However, Wright did not give examples of his technique's use, making it hard to verify that he described the modern notion of median.[52][56][b] The median (in the context of probability) certainly appeared in the correspondence of Christiaan Huygens, but as an example of a statistic that was inappropriate for actuarial practice.[52]

The earliest recommendation of the median dates to 1757, when Roger Joseph Boscovich developed a regression method based on the L^1 norm and therefore implicitly on the median.[52][57] In 1774, Laplace made this desire explicit: he suggested the median be used as the standard estimator of the value of a posterior PDF. The specific criterion was to minimize the expected magnitude of the error; $|\alpha - \alpha^*|$ where α^* is the estimate and α is the true value. To this end, Laplace determined the distributions of both the sample mean and the sample median in the early 1800s.[23][58] However, a decade later, Gauss and Legendre developed the least squares method, which minimizes $(\alpha - \alpha^*)^2$ to obtain the mean. Within the context of regression, Gauss and Legendre's innovation offers vastly easier computation. Consequently, Laplaces' proposal was generally rejected until the rise of computing devices 150 years later (and is still a relatively uncommon algorithm).[59]

Antoine Augustin Cournot in 1843 was the first[60] to use the term *median* (*valeur médiane*) for the value that divides a probability distribution into two equal halves. Gustav Theodor Fechner used the median (*Centralwerth*) in sociological and psychological phenomena.[61] It had earlier been used only in astronomy and related fields. Gustav Fechner popularized the median into the formal analysis of data, although it had been used previously by Laplace,[61] and the median appeared in a textbook by F. Y. Edgeworth.[62] Francis Galton used the English term *median* in 1881,[63][64] having earlier used the terms *middle-most value* in 1869, and the *medium* in 1880.[65][66]

Statisticians encouraged the use of medians intensely throughout the 19th century for its intuitive clarity and ease of manual computation. However, the notion of median does not lend itself to the theory of higher moments as well as the arithmetic mean does, and is much harder to compute by computer. As a result, the median was steadily supplanted as a notion of generic average by the arithmetic mean during the 20th century.[52][56]

See also

- Medoids which are a generalisation of the median in higher dimensions
- Central tendency
  - Mean
  - Mode
- Absolute deviation

- Bias of an estimator
- Concentration of measure for Lipschitz functions
- Median (geometry)
- Median graph
- Median search
- Median slope
- Median voter theory
- Weighted median

Notes

a. *The geometric median is unique unless the sample is collinear.*[36]

b. *Subsequent scholars appear to concur with Eisenhart that Boroughs' 1580 figures, while suggestive of the median, in fact describe an arithmetic mean.;*[55]:62–3 *Boroughs is mentioned in no other work.*

References

1. *Weisstein, Eric W. "Statistical Median" . MathWorld.*
2. *Simon, Laura J.; "Descriptive statistics" Archived 2010-07-30 at the Wayback Machine, Statistical Education Resource Kit, Pennsylvania State Department of Statistics*

3. *David J. Sheskin (27 August 2003). Handbook of Parametric and Nonparametric Statistical Procedures: Third Edition . CRC Press. pp. 7–. ISBN 978-1-4200-3626-8. Retrieved 25 February 2013.*
4. *Derek Bissell (1994). Statistical Methods for Spc and Tqm . CRC Press. pp. 26–. ISBN 978-0-412-39440-9. Retrieved 25 February 2013.*
5. *"Journal of Statistics Education, v13n2: Paul T. von Hippel" . amstat.org.*

6. *Robson, Colin (1994). Experiment, Design and Statistics in Psychology. Penguin. pp. 42–45. ISBN 0-14-017648-9.*
7. *Williams, D. (2001). Weighing the Odds . Cambridge University Press. p. 165 . ISBN 052100618X.*
8. *Maindonald, John; Braun, W. John (2010-05-06). Data Analysis and Graphics Using R: An Example-Based Approach . Cambridge University Press. p. 104. ISBN 978-1-139-48667-5.*

9. *"AP Statistics Review - Density Curves and the Normal Distributions" . Archived from the original on 8 April 2015. Retrieved 16 March 2015.*
10. *Newman, Mark EJ. "Power laws, Pareto distributions and Zipf's law." Contemporary physics 46.5 (2005): 323–351.*
11. *Stroock, Daniel (2011). Probability Theory . Cambridge University Press. pp. 43 . ISBN 978-0-521-13250-3.*

12. *André Nicolas (https://math.stackexchange.com/users/6312/andr%c3%a9-nicolas), The Median Minimizes the Sum of Absolute Deviations (The $ {L}_{1} $ Norm), URL (version: 2012-02-25): https://math.stackexchange.com/q/113336*
13. *Mallows, Colin (August 1991). "Another comment on O'Cinneide". The American Statistician. **45** (3): 257. doi:10.1080/00031305.1991.10475815 .*

14. *Piché, Robert (2012). Random Vectors and Random Sequences. Lambert Academic Publishing. ISBN 978-3659211966.*
15. *Kemperman, Johannes H. B. (1987). Dodge, Yadolah (ed.). "The median of a finite measure on a Banach space: Statistical data analysis based on the L1-norm and related methods". Papers from the First International Conference Held at Neuchâtel, August 31–September 4, 1987. Amsterdam: North-Holland Publishing Co.: 217–230. MR 0949228 .*

16. *Milasevic, Philip; Ducharme, Gilles R. (1987). "Uniqueness of the spatial median". Annals of Statistics.* ***15*** *(3): 1332–1333. doi:10.1214/aos/1176350511 . MR 0902264 .*
17. *K.Van Steen Notes on probability and statistics*
18. *Basu, S.; Dasgupta, A. (1997). "The Mean, Median, and Mode of Unimodal Distributions:A Characterization". Theory of Probability and Its Applications.* ***41*** *(2): 210–223. doi:10.1137/S0040585X97975447 . S2CID 54593178 .*

19. *Merkle, M. (2005). "Jensen's inequality for medians". Statistics & Probability Letters.* ***71*** *(3): 277–281. doi:10.1016/j.spl.2004.11.010 .*
20. *Alfred V. Aho and John E. Hopcroft and Jeffrey D. Ullman (1974). The Design and Analysis of Computer Algorithms . Reading/MA: Addison-Wesley. ISBN 0-201-00029-6. Here: Section 3.6 "Order Statistics", p.97-99, in particular Algorithm 3.6 and Theorem 3.9.*

21. *Bentley, Jon L.; McIlroy, M. Douglas (1993). "Engineering a sort function" . Software—Practice and Experience.* ***23*** *(11): 1249–1265. doi:10.1002/spe.4380231105 .*

22. *Rousseeuw, Peter J.; Bassett, Gilbert W. Jr. (1990). "The remedian: a robust averaging method for large data sets" (PDF). J. Amer. Statist. Assoc.* ***85*** *(409): 97–104. doi:10.1080/01621459.1990.10475311 .*

23. *Stigler, Stephen (December 1973). "Studies in the History of Probability and Statistics. XXXII: Laplace, Fisher and the Discovery of the Concept of Sufficiency". Biometrika.* ***60*** *(3): 439–445. doi:10.1093/biomet/60.3.439 . JSTOR 2334992 . MR 0326872 .*
24. *Rider, Paul R. (1960). "Variance of the median of small samples from several special populations". J. Amer. Statist. Assoc.* ***55*** *(289): 148–150. doi:10.1080/01621459.1960.10482056 .*

25. *Efron, B. (1982). The Jackknife, the Bootstrap and other Resampling Plans. Philadelphia: SIAM. ISBN 0898711797.*
26. *Shao, J.; Wu, C. F. (1989). "A General Theory for Jackknife Variance Estimation". Ann. Stat.* ***17*** *(3): 1176–1197. doi:10.1214/aos/1176347263 . JSTOR 2241717 .*
27. *Efron, B. (1979). "Bootstrap Methods: Another Look at the Jackknife". Ann. Stat.* ***7*** *(1): 1–26. doi:10.1214/aos/1176344552 . JSTOR 2958830 .*

28. *Hall, P.; Martin, M. A. (1988). "Exact Convergence Rate of Bootstrap Quantile Variance Estimator". Probab Theory Related Fields. **80** (2): 261–268. doi:10.1007/BF00356105 . S2CID 119701556 .*

29. *Jiménez-Gamero, M. D.; Munoz-García, J.; Pino-Mejías, R. (2004). "Reduced bootstrap for the median" . Statistica Sinica. **14** (4): 1179–1198.*

30. *Maindonald, John; John Braun, W. (2010-05-06). Data Analysis and Graphics Using R: An Example-Based Approach . ISBN 9781139486675.*

31. *Hettmansperger, Thomas P.; McKean, Joseph W. (1998). Robust nonparametric statistical methods. Kendall's Library of Statistics.* ***5****. London: Edward Arnold. ISBN 0-340-54937-8. MR 1604954 .*
32. *Small, Christopher G. "A survey of multidimensional medians." International Statistical Review/Revue Internationale de Statistique (1990): 263–277. doi:10.2307/1403809 JSTOR 1403809*
33. *Niinimaa, A., and H. Oja. "Multivariate median." Encyclopedia of statistical sciences (1999).*

34. *Mosler, Karl. Multivariate Dispersion, Central Regions, and Depth: The Lift Zonoid Approach. Vol. 165. Springer Science & Business Media, 2012.*
35. *Puri, Madan L.; Sen, Pranab K.; Nonparametric Methods in Multivariate Analysis, John Wiley & Sons, New York, NY, 197l. (Reprinted by Krieger Publishing)*

36. *Vardi, Yehuda; Zhang, Cun-Hui (2000). "The multivariate L_1-median and associated data depth" . Proceedings of the National Academy of Sciences of the United States of America.* ***97*** *(4): 1423–1426 (electronic). Bibcode:2000PNAS...97.1423V . doi:10.1073/pnas.97.4.1423 . MR 1740461 . PMC 26449 . PMID 10677477 .*

37. *Pratt, William K.; Cooper, Ted J.; Kabir, Ihtisham (1985-07-11). Corbett, Francis J (ed.). "Pseudomedian Filter". Architectures and Algorithms for Digital Image Processing II.* ***0534****: 34. Bibcode:1985SPIE..534...34P . doi:10.1117/12.946562 . S2CID 173183609 .*
38. *Oja, Hannu (2010). Multivariate nonparametric methods with R: An approach based on spatial signs and ranks. Lecture Notes in Statistics.* ***199****. New York, NY: Springer. pp. xiv+232. doi:10.1007/978-1-4419-0468-3 . ISBN 978-1-4419-0467-6. MR 2598854 .*

39. *Wilcox, Rand R. (2001), "Theil–Sen estimator", Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy , Springer-Verlag, pp. 207–210, ISBN 978-0-387-95157-7.*
40. *Wald, A. (1940). "The Fitting of Straight Lines if Both Variables are Subject to Error"* (PDF). *Annals of Mathematical Statistics.* ***11*** *(3): 282–300. doi:10.1214/aoms/1177731868 . JSTOR 2235677 .*

41. *Nair, K. R.; Shrivastava, M. P. (1942). "On a Simple Method of Curve Fitting". Sankhyā: The Indian Journal of Statistics.* ***6*** *(2): 121–132. JSTOR 25047749 .*

42. *Brown, G. W.; Mood, A. M. (1951). "On Median Tests for Linear Hypotheses". Proc Second Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, CA: University of California Press. pp. 159–166. Zbl 0045.08606 .*

43. *Tukey, J. W. (1977). Exploratory Data Analysis . Reading, MA: Addison-Wesley. ISBN 0201076160.*

44. *Brown, George W. (1947). "On Small-Sample Estimation". Annals of Mathematical Statistics.* ***18*** *(4): 582–585. doi:10.1214/aoms/1177730349 . JSTOR 2236236 .*

45. *Lehmann, Erich L. (1951). "A General Concept of Unbiasedness". Annals of Mathematical Statistics.* ***22*** *(4): 587–592. doi:10.1214/aoms/1177729549 . JSTOR 2236928 .*

46. *Birnbaum, Allan (1961). "A Unified Theory of Estimation, I". Annals of Mathematical Statistics.* ***32*** *(1): 112–135. doi:10.1214/aoms/1177705145 . JSTOR 2237612 .*

47. *van der Vaart, H. Robert (1961). "Some Extensions of the Idea of Bias". Annals of Mathematical Statistics.* ***32*** *(2): 436–447. doi:10.1214/aoms/1177705051 . JSTOR 2237754 . MR 0125674 .*
48. *Pfanzagl, Johann; with the assistance of R. Hamböker (1994). Parametric Statistical Theory. Walter de Gruyter. ISBN 3-11-013863-8. MR 1291393 .*
49. *Pfanzagl, Johann. "On optimal median unbiased estimators in the presence of nuisance parameters." The Annals of Statistics (1979): 187–193.*

50. *Brown, L. D.; Cohen, Arthur; Strawderman, W. E. (1976). "A Complete Class Theorem for Strict Monotone Likelihood Ratio With Applications" . Ann. Statist.* ***4*** *(4): 712–722. doi:10.1214/aos/1176343543* .

51. *Page; Brown, L. D.; Cohen, Arthur; Strawderman, W. E. (1976). "A Complete Class Theorem for Strict Monotone Likelihood Ratio With Applications" . Ann. Statist.* ***4*** *(4): 712–722. doi:10.1214/aos/1176343543* .

52. *Bakker, Arthur; Gravemeijer, Koeno P. E. (2006-06-01). "An Historical Phenomenology of Mean and Median". Educational Studies in Mathematics.* ***62*** *(2): 149–168. doi:10.1007/s10649-006-7099-8 . ISSN 1573-0816 . S2CID 143708116 .*

53. *Adler, Dan (31 December 2014). "Talmud and Modern Economics" . Jewish American and Israeli Issues. Archived from the original on 6 December 2015. Retrieved 22 February 2020.*

54. *Modern Economic Theory in the Talmud by Yisrael Aumann*

55. *Eisenhart, Churchill (24 August 1971). The Development of the Concept of the Best Mean of a Set of Measurements from Antiquity to the Present Day (PDF) (Speech). 131st Annual Meeting of the American Statistical Association. Colorado State University.*
56. *"How the Average Triumphed Over the Median" . Priceonomics. Retrieved 2020-02-23.*
57. *Stigler, S. M. (1986). The History of Statistics: The Measurement of Uncertainty Before 1900 . Harvard University Press. ISBN 0674403401.*

58. *Laplace PS de (1818) Deuxième supplément à la Théorie Analytique des Probabilités, Paris, Courcier*
59. *Jaynes, E.T. (2007). Probability theory : the logic of science (5. print. ed.). Cambridge [u.a.]: Cambridge Univ. Press. p. 172. ISBN 978-0-521-59271-0.*
60. *Howarth, Richard (2017). Dictionary of Mathematical Geosciences: With Historical Notes. Springer. p. 374.*

61. *Keynes, J.M. (1921) A Treatise on Probability. Pt II Ch XVII §5 (p 201) (2006 reprint, Cosimo Classics, ISBN 9781596055308 : multiple other reprints)*
62. *Stigler, Stephen M. (2002). Statistics on the Table: The History of Statistical Concepts and Methods . Harvard University Press. pp. 105–7. ISBN 978-0-674-00979-0.*
63. *Galton F (1881) "Report of the Anthropometric Committee" pp 245–260. Report of the 51st Meeting of the British Association for the Advancement of Science*

64. *David, H. A. (1995). "First (?) Occurrence of Common Terms in Mathematical Statistics". The American Statistician.* ***49*** *(2): 121–133. doi:10.2307/2684625 . ISSN 0003-1305 . JSTOR 2684625 .*

65. *encyclopediaofmath.org*

66. *personal.psu.edu*

External links

- Hazewinkel, Michiel, ed. (2001) [1994], "Median (in statistics)" , *Encyclopedia of Mathematics*, Springer Science+Business Media B.V. / Kluwer

Academic Publishers, ISBN 978-1-55608-010-4

- Median as a weighted arithmetic mean of all Sample Observations
- On-line calculator
- Calculating the median
- A problem involving the mean, the median, and the mode.
- Weisstein, Eric W. "Statistical Median" . *MathWorld*.
- Python script for Median computations and income inequality metrics
- Fast Computation of the Median by Successive Binning

- 'Mean, median, mode and skewness' , A tutorial devised for first-year psychology students at Oxford University, based on a worked example.
- The Complex SAT Math Problem Even the College Board Got Wrong : Andrew Daniels in *Popular Mechanics*

title=Median&oldid=965007172"