Moment (mathematics)

From Wikipedia, the free encyclopedia

In mathematics, a moment is a specific quantitative measure, used in both mechanics and statistics, of the shape of a set of points. If the points represent mass, then the zeroth moment is the total mass, the first moment divided by the total mass is the center of mass, and the second moment is the rotational inertia. If the points represent probability density, then the zeroth moment is the total probability (i.e. one), the first moment is the mean, the second central moment is the variance, the third moment is the skewness, and the fourth moment (with normalization and shift) is the kurtosis. The mathematical concept is closely related to the concept of moment in physics.

For a bounded distribution of mass or probability, the collection of all the moments (of all orders, from 0 to ∞) uniquely determines the distribution.

Contents

Significance of the moments

The n-th moment of a real-valued continuous function $f(x)$ of a real variable about a value c is

$$\mu_n = \int_{-\infty}^{\infty} (x - c)^n \, f(x) \, dx.$$

It is possible to define moments for random variables in a more general fashion than moments for real values—see moments in metric spaces. The moment of a function, without further explanation, usually refers to the above expression with $c = 0$.

For the second and higher moments, the central moments (moments about the mean, with c being the mean) are usually used rather than the moments about zero, because they provide clearer information about the distribution's shape.

Other moments may also be defined. For example, the n-th inverse moment about zero is $\mathrm{E}[X^{-n}]$ and the n-th logarithmic moment about zero is $\mathrm{E}[\ln^n(X)]$.

The n-th moment about zero of a probability density function $f(x)$ is the expected value of X^n and is called a *raw moment* or *crude moment*.[1] The moments about its mean μ are called *central* moments; these describe the shape of the function, independently of translation.

If f is a probability density function, then the value of the integral above is called the n-th moment of the probability distribution. More generally, if F is a cumulative probability distribution function of any probability distribution, which may not have a density function, then the n-th moment of the probability distribution is given by the Riemann–Stieltjes integral

$$\mu'_n = \mathrm{E}[X^n] = \int_{-\infty}^{\infty} x^n \, dF(x)$$

where X is a random variable that has this cumulative distribution F, and E is the expectation operator or mean.

When

$$\mathrm{E}[|X^n|] = \int_{-\infty}^{\infty} |x^n| \, dF(x) = \infty,$$

then the moment is said not to exist. If the n-th moment about any point exists, so does the $(n-1)$-th moment (and thus, all lower-order moments) about every point.

The zeroth moment of any probability density function is 1, since the area under any probability density function must be equal to one.

Significance of moments (raw, central, standardised) and cumulants (raw, standardised), in connection with named properties of distributions

Moment number	Raw moment	Central moment	Standardised moment	Raw cumulant	Standardised cumulant
1	mean	0	0	mean	N/A
2	–	variance	1	variance	1
3	–	–	skewness	–	skewness
4	–	–	(non-excess or historical) kurtosis (or flatness)	–	excess kurtosis
5	–	–	hyperskewness	–	–
6	–	–	hyperflatness	–	–
7+	–	–	-	–	–

Mean

The first raw moment is the mean.

Variance

The second central moment is the variance. Its positive square root is the standard deviation σ.

Normalised moments

The *normalised* n-th central moment or standardized moment is the n-th central moment divided by σ^n; the normalised n-th central moment of

$$x = \frac{\mathrm{E}[(x - \mu)^n]}{\sigma^n}.$$

These normalised central moments are dimensionless quantities, which represent the distribution independently of any linear change of scale.

For an electric signal, the first moment is its DC level, and the 2nd moment is proportional to its average power.[2][3]

Skewness

The third central moment is the measure of the lopsidedness of the distribution; any symmetric distribution will have a third central moment, if defined, of zero. The normalised third central moment is called the skewness, often γ. A distribution that is

skewed to the left (the tail of the distribution is longer on the left) will have a negative skewness. A distribution that is skewed to the right (the tail of the distribution is longer on the right), will have a positive skewness.

For distributions that are not too different from the normal distribution, the median will be somewhere near $\mu - \gamma\sigma/6$; the mode about $\mu - \gamma\sigma/2$.

Kurtosis

The fourth central moment is a measure of the heaviness of the tail of the distribution, compared to the normal distribution of the same variance. Since it is the expectation of a fourth power, the fourth central moment, where defined, is always positive; and except for a point distribution, it is always strictly positive. The fourth central moment of a normal distribution is $3\sigma^4$.

The kurtosis κ is defined to be the normalised fourth central moment minus 3 (Equivalently, as in the next section, it is the fourth cumulant divided by the square of the variance). Some authorities do not subtract three, but it is usually more convenient to have the normal distribution at the origin of coordinates.[4][5] If a distribution has heavy tails, the kurtosis will be high (sometimes called leptokurtic); conversely, light-tailed distributions (for example, bounded distributions such as the uniform) have low kurtosis (sometimes called platykurtic).

The kurtosis can be positive without limit, but κ must be greater than or equal to $\gamma^2 - 2$; equality only holds for binary distributions. For unbounded skew distributions not too far from normal, κ tends to be somewhere in the area of γ^2 and $2\gamma^2$.

The inequality can be proven by considering

$$\mathbf{E}\big[(T^2 - aT - 1)^2\big]$$

where $T = (X - \mu)/\sigma$. This is the expectation of a square, so it is non-negative for all a; however it is also a quadratic polynomial in a. Its discriminant must be non-positive, which gives the required relationship.

Mixed moments

Mixed moments are moments involving multiple variables.

Some examples are covariance, coskewness and cokurtosis. While there is a unique covariance, there are multiple co-skewnesses and co-kurtoses.

Higher moments

High-order moments are moments beyond 4th-order moments. As with variance, skewness, and kurtosis, these are higher-order statistics, involving non-linear combinations of the data, and can be used for description or estimation of further shape parameters. The higher the moment, the harder it is to estimate, in the sense that larger

samples are required in order to obtain estimates of similar quality. This is due to the excess degrees of freedom consumed by the higher orders. Further, they can be subtle to interpret, often being most easily understood in terms of lower order moments – compare the higher derivatives of jerk and jounce in physics. For example, just as the 4th-order moment (kurtosis) can be interpreted as "relative importance of tails versus shoulders in causing dispersion" (for a given dispersion, high kurtosis corresponds to heavy tails, while low kurtosis corresponds to heavy shoulders), the 5th-order moment can be interpreted as measuring "relative importance of tails versus center (mode, shoulders) in causing skew" (for a given skew, high 5th moment corresponds to heavy tail and little movement of mode, while low 5th moment corresponds to more change in shoulders).

Transformation of center

Since:

$$(x - b)^n = (x - a + a - b)^n = \sum_{i=0}^{n} \binom{n}{i} (x - a)^i (a - b)^{n-i}$$

where $\binom{n}{i}$ is the binomial coefficient, it follows that the moments about b can be calculated from the moments about a by:

$$E[(x - b)^n] = \sum_{i=0}^{n} \binom{n}{i} E[(x - a)^i](a - b)^{n-i}$$

Cumulants

The first moment and the second and third *unnormalized central* moments are additive in the sense that if X and Y are independent random variables then

$$\mu_1(X + Y) = \mu_1(X) + \mu_1(Y)$$
$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$
$$\mu_3(X + Y) = \mu_3(X) + \mu_3(Y)$$

(These can also hold for variables that satisfy weaker conditions than independence. The first always holds; if the second holds, the variables are called uncorrelated).

In fact, these are the first three cumulants and all cumulants share this additivity property.

Sample moments

For all k, the k-th raw moment of a population can be estimated using the k-th raw sample

moment

$$\frac{1}{n}\sum_{i=1}^{n}X_i^k$$

applied to a sample $X_1, ..., X_n$ drawn from the population.

It can be shown that the expected value of the raw sample moment is equal to the k-th raw moment of the population, if that moment exists, for any sample size n. It is thus an unbiased estimator. This contrasts with the situation for central moments, whose computation uses up a degree of freedom by using the sample mean. So for example an unbiased estimate of the population variance (the second central moment) is given by

$$\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

in which the previous denominator n has been replaced by the degrees of freedom $n-1$, and in which $\bar{X}$ refers to the sample mean. This estimate of the population moment is greater than the unadjusted observed sample moment by a factor of $\frac{n}{n-1}$, and it is referred to as the "adjusted sample variance" or sometimes simply the "sample variance".

Problem of moments

The *problem of moments* seeks characterizations of sequences { $\mu'_n : n = 1, 2, 3, ...$ } that are sequences of moments of some function *f*.

Partial moments

Partial moments are sometimes referred to as "one-sided moments." The n-th order lower and upper partial moments with respect to a reference point *r* may be expressed as

$$\mu_{\bar{n}}(r) = \int_{-\infty}^{r}(r-x)^n\,f(x)\,dx,$$
$$\mu_n^{+}(r) = \int_{r}^{\infty}(x-r)^n\,f(x)\,dx.$$

Partial moments are normalized by being raised to the power 1/n. The upside potential ratio may be expressed as a ratio of a first-order upper partial moment to a normalized second-order lower partial moment. They have been used in the definition of some financial metrics, such as the Sortino ratio, as they focus purely on upside or downside.

Central moments in metric spaces

Let (M, d) be a metric space, and let B(M) be the Borel σ-algebra on M, the σ-algebra generated by the d-open subsets of M. (For technical reasons, it is also convenient to assume that M is a separable space with respect to the metric d.) Let $1 \leq p \leq \infty$.

The pth central moment of a measure μ on the measurable space (M, B(M)) about a given point $x_0 \in M$ is defined to be

$$\int_M d(x, x_0)^p \, \mathrm{d}\mu(x).$$

μ is said to have finite **p**-th central moment if the p-th central moment of μ about x_0 is finite for some $x_0 \in M$.

This terminology for measures carries over to random variables in the usual way: if $(\Omega, \Sigma, \mathbf{P})$ is a probability space and $X : \Omega \to M$ is a random variable, then the **p**-th central moment of X about $x_0 \in M$ is defined to be

$$\int_M d(x, x_0)^p \, \mathrm{d}\left(X_*(\mathbf{P})\right)(x) \equiv \int_\Omega d(X(\omega), x_0)^p \, \mathrm{d}\mathbf{P}(\omega),$$

and X has finite **p**-th central moment if the p-th central moment of X about x_0 is finite for some $x_0 \in M$.

See also

- Energy (signal processing)
- Factorial moment
- Generalised mean
- Hamburger moment problem
- Hausdorff moment problem
- Image moment
- L-moment
- Method of moments (probability theory)
- Method of moments (statistics)
- Moment-generating function
- Moment measure
- Second moment method
- Standardised moment
- Stieltjes moment problem
- Taylor expansions for the moments of functions of random variables

References

1. http://mathworld.wolfram.com/RawMoment.html Raw Moments at Math-world
2. Clive Maxfield; John Bird; Tim Williams; Walt Kester; Dan Bensky (2011). *Electrical Engineering: Know It All*. Newnes. p. 884. ISBN 978-0-08-094966-6.
3. Ha H. Nguyen; Ed Shwedyk (2009). *A First Course in Digital Communications*. Cambridge University Press. p. 87. ISBN 978-0-521-87613-1.
4. Casella, George; Berger, Roger L. (2002). *Statistical Inference* (2 ed.). Pacific Grove: Duxbury. ISBN 0-534-24312-6.
5. Ballanda, Kevin P.; MacGillivray, H. L. (1988). "Kurtosis: A Critical Review". *The American Statistician*. American Statistical Association. 42 (2): 111–119. doi:10.2307/2684482. JSTOR 2684482.

Further reading

- Spanos, Aris (1999). *Probability Theory and Statistical Inference*. New York: Cambridge University Press. pp. 109–130. ISBN 0-521-42408-9.

External links

- Hazewinkel, Michiel, ed. (2001), "Moment", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Moments at Mathworld (http://mathworld.wolfram.com/topics/Moments.html)
- Higher Moments (http://www.geo.upm.es/postgrado/CarlosLopez/geost_03 /node37.html)