WIKIPEDIA

Cumulant

In probability theory and statistics, the **cumulants** κ_n of a probability distribution are a set of quantities that provide an alternative to the moments of the distribution. The moments determine the cumulants in the sense that any two probability distributions whose moments are identical will have identical cumulants as well, and similarly the cumulants determine the moments.

The first cumulant is the mean, the second cumulant is the variance, and the third cumulant is the same as the third central moment. But fourth and higher-order cumulants are not equal to central moments. In some cases theoretical treatments of problems in terms of cumulants are simpler than those using moments. In particular, when two or more random variables are statistically independent, the n^{th}-order cumulant of their sum is equal to the sum of their n^{th}-order cumulants. As well, the third and higher-order cumulants of a normal distribution are zero, and it is the only distribution with this property.

Just as for moments, where *joint moments* are used for collections of random variables, it is possible to define *joint cumulants*.

Contents

Definition
- Alternative definition of the cumulant generating function

Uses in statistics

Cumulants of some discrete probability distributions

Cumulants of some continuous probability distributions

Some properties of the cumulant generating function

Some properties of cumulants
- Invariance and equivariance
- Homogeneity
- Additivity
- A negative result
- Cumulants and moments
- Cumulants and set-partitions
- Cumulants and combinatorics

Joint cumulants
- Conditional cumulants and the law of total cumulance

Relation to statistical physics

History

Cumulants in generalized settings
- Formal cumulants
- Bell numbers
- Cumulants of a polynomial sequence of binomial type
- Free cumulants

See also

References

External links

Definition

The cumulants of a random variable X are defined using the **cumulant-generating function** $K(t)$, which is the natural logarithm of the moment-generating function:

$$K(t) = \log \operatorname{E}\left[e^{tX}\right].$$

The cumulants κ_n are obtained from a power series expansion of the cumulant generating function:

$$K(t) = \sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!} = \mu t + \sigma^2 \frac{t^2}{2} + \cdots.$$

This expansion is a Maclaurin series, so the n-th cumulant can be obtained by differentiating the above expansion n times and evaluating the result at zero:[1]

$$\kappa_n = K^{(n)}(0).$$

If the moment-generating function does not exist, the cumulants can be defined in terms of the relationship between cumulants and moments discussed later.

Alternative definition of the cumulant generating function

Some writers[2][3] prefer to define the cumulant-generating function as the natural logarithm of the characteristic function, which is sometimes also called the ***second* characteristic function,**[4][5]

$$H(t) = \log \operatorname{E}\left[e^{itX}\right] = \sum_{n=1}^{\infty} \kappa_n \frac{(it)^n}{n!} = \mu i t - \sigma^2 \frac{t^2}{2} + \cdots$$

An advantage of $H(t)$—in some sense the function $K(t)$ evaluated for purely imaginary arguments—is that $\operatorname{E}[e^{itX}]$ is well defined for all real values of t even when $\operatorname{E}[e^{tX}]$ is not well defined for all real values of t, such as can occur when there is "too much" probability that X has a large magnitude. Although the function $H(t)$ will be well defined, it will nonetheless mimic $K(t)$ in terms of the length of its Maclaurin series, which may not extend beyond (or, rarely, even to) linear order in the argument t, and in particular the number of cumulants that are well defined will not change. Nevertheless, even when $H(t)$ does not have a long Maclaurin series, it can be used directly in analyzing and, particularly, adding random variables. Both the Cauchy distribution (also called the Lorentzian) and more generally, stable distributions (related to the Lévy distribution) are examples of distributions for which the power-series expansions of the generating functions have only finitely many well-defined terms.

Uses in statistics

Working with cumulants can have an advantage over using moments because for statistically independent random variables X and Y,

$$\begin{aligned}K_{X+Y}(t)&=\log \mathrm{E}\left[e^{t(X+Y)}\right]\\&=\log\left(\mathrm{E}\left[e^{tX}\right]\mathrm{E}\left[e^{tY}\right]\right)\\&=\log \mathrm{E}\left[e^{tX}\right]+\log \mathrm{E}\left[e^{tY}\right]\\&=K_X(t)+K_Y(t),\end{aligned}$$

so that each cumulant of a sum of independent random variables is the sum of the corresponding cumulants of the addends. That is, when the addends are statistically independent, the mean of the sum is the sum of the means, the variance of the sum is the sum of the variances, the third cumulant (which happens to be the third central moment) of the sum is the sum of the third cumulants, and so on for each order of cumulant.

A distribution with given cumulants κ_n can be approximated through an Edgeworth series.

Cumulants of some discrete probability distributions

- The constant random variables $X = \mu$. The cumulant generating function is $K(t) = \mu t$. The first cumulant is $\kappa_1 = K'(0) = \mu$ and the other cumulants are zero, $\kappa_2 = \kappa_3 = \kappa_4 = ... = 0$.
- The Bernoulli distributions, (number of successes in one trial with probability p of success). The cumulant generating function is $K(t) = \log(1 - p + pe^t)$. The first cumulants are $\kappa_1 = K'(0) = p$ and $\kappa_2 = K''(0) = p\cdot(1 - p)$. The cumulants satisfy a recursion formula

$$\kappa_{n+1}=p(1-p)\frac{d\kappa_n}{dp}.$$

- The geometric distributions, (number of failures before one success with probability p of success on each trial). The cumulant generating function is $K(t) = \log(p / (1 + (p - 1)e^t))$. The first cumulants are $\kappa_1 = K'(0) = p^{-1} - 1$, and $\kappa_2 = K''(0) = \kappa_1 p^{-1}$. Substituting $p = (\mu + 1)^{-1}$ gives $K(t) = -\log(1 + \mu(1-e^t))$ and $\kappa_1 = \mu$.
- The Poisson distributions. The cumulant generating function is $K(t) = \mu(e^t - 1)$. All cumulants are equal to the parameter: $\kappa_1 = \kappa_2 = \kappa_3 = ... = \mu$.
- The binomial distributions, (number of successes in n independent trials with probability p of success on each trial). The special case $n = 1$ is a Bernoulli distribution. Every cumulant is just n times the corresponding cumulant of the corresponding Bernoulli distribution. The cumulant generating function is $K(t) = n \log(1 - p + pe^t)$. The first cumulants are $\kappa_1 = K'(0) = np$ and $\kappa_2 = K''(0) = \kappa_1(1 - p)$. Substituting $p = \mu\cdot n^{-1}$ gives $K'(t) = ((\mu^{-1} - n^{-1})\cdot e^{-t} + n^{-1})^{-1}$ and $\kappa_1 = \mu$. The limiting case $n^{-1} = 0$ is a Poisson distribution.
- The negative binomial distributions, (number of failures before n successes with probability p of success on each trial). The special case $n = 1$ is a geometric distribution. Every cumulant is just n times the corresponding cumulant of the corresponding geometric distribution. The derivative of the cumulant generating function is $K'(t) = n\cdot((1 - p)^{-1}\cdot e^{-t}-1)^{-1}$. The first cumulants are $\kappa_1 = K'(0) = n\cdot(p^{-1}-1)$, and $\kappa_2 = K''(0) = \kappa_1\cdot p^{-1}$. Substituting $p = (\mu\cdot n^{-1}+1)^{-1}$ gives $K'(t) = ((\mu^{-1} + n^{-1})e^{-t} - n^{-1})^{-1}$ and $\kappa_1 = \mu$. Comparing these formulas to those of the binomial distributions explains the name 'negative binomial distribution'. The limiting case $n^{-1} = 0$ is a Poisson distribution.

Introducing the variance-to-mean ratio

$$\varepsilon=\mu^{-1}\sigma^2=\kappa_1^{-1}\kappa_2,$$

the above probability distributions get a unified formula for the derivative of the cumulant generating function:

$$K'(t) = \mu \cdot (1 + \varepsilon \cdot (e^{-t} - 1))^{-1}.$$

The second derivative is

$$K''(t) = g'(t) \cdot (1 + e^{t} \cdot (\varepsilon^{-1} - 1))^{-1}$$

confirming that the first cumulant is $\kappa_1 = K'(0) = \mu$ and the second cumulant is $\kappa_2 = K''(0) = \mu\varepsilon$. The constant random variables $X = \mu$ have $\varepsilon = 0$. The binomial distributions have $\varepsilon = 1 - p$ so that $0 < \varepsilon < 1$. The Poisson distributions have $\varepsilon = 1$. The negative binomial distributions have $\varepsilon = p^{-1}$ so that $\varepsilon > 1$. Note the analogy to the classification of conic sections by eccentricity: circles $\varepsilon = 0$, ellipses $0 < \varepsilon < 1$, parabolas $\varepsilon = 1$, hyperbolas $\varepsilon > 1$.

Cumulants of some continuous probability distributions

- For the normal distribution with expected value μ and variance σ^2, the cumulant generating function is $K(t) = \mu t + \sigma^2 t^2/2$. The first and second derivatives of the cumulant generating function are $K'(t) = \mu + \sigma^2 \cdot t$ and $K''(t) = \sigma^2$. The cumulants are $\kappa_1 = \mu$, $\kappa_2 = \sigma^2$, and $\kappa_3 = \kappa_4 = \ldots = 0$. The special case $\sigma^2 = 0$ is a constant random variable $X = \mu$.
- The cumulants of the uniform distribution on the interval [−1, 0] are $\kappa_n = B_n/n$, where B_n is the n-th Bernoulli number.
- The cumulants of the exponential distribution with parameter λ are $\kappa_n = \lambda^{-n}\,(n - 1)!$.

Some properties of the cumulant generating function

The cumulant generating function $K(t)$, if it exists, is infinitely differentiable and convex, and passes through the origin. Its first derivative ranges monotonically in the open interval from the infimum to the supremum of the support of the probability distribution, and its second derivative is strictly positive everywhere it is defined, except for the degenerate distribution of a single point mass. The cumulant-generating function exists if and only if the tails of the distribution are majorized by an exponential decay, that is, (*see Big O notation*,)

$$\exists c > 0,\ F(x) = O(e^{cx}), x \to -\infty; \text{ and}$$
$$\exists d > 0,\ 1 - F(x) = O(e^{-dx}), x \to +\infty;$$

where F is the cumulative distribution function. The cumulant-generating function will have vertical asymptote(s) at the infimum of such c, if such an infimum exists, and at the supremum of such d, if such a supremum exists, otherwise it will be defined for all real numbers.

If the support of a random variable X has finite upper or lower bounds, then its cumulant-generating function $y = K(t)$, if it exists, approaches asymptote(s) whose slope is equal to the supremum and/or infimum of the support,

$$y = (t + 1) \inf \operatorname{supp} X - \mu(X), \text{ and}$$
$$y = (t - 1) \sup \operatorname{supp} X + \mu(X),$$

respectively, lying above both these lines everywhere. (The integrals

$$\int_{-\infty}^{0} [t \inf \operatorname{supp} X - K'(t)] \; dt, \qquad \int_{\infty}^{0} [t \inf \operatorname{supp} X - K'(t)] \; dt$$

yield the y-intercepts of these asymptotes, since $K(0) = 0$.)

For a shift of the distribution by c, $K_{X+c}(t) = K_X(t) + ct$. For a degenerate point mass at c, the cgf is the straight line $K_c(t) = ct$, and more generally, $K_{X+Y} = K_X + K_Y$ if and only if X and Y are independent and their cgfs exist; (subindependence and the existence of second moments sufficing to imply independence.[6])

The natural exponential family of a distribution may be realized by shifting or translating $K(t)$, and adjusting it vertically so that it always passes through the origin: if f is the pdf with cgf $K(t) = \log M(t)$, and $f|\theta$ is its natural exponential family, then $f(x \mid \theta) = \frac{1}{M(\theta)} e^{\theta x} f(x)$, and $K(t \mid \theta) = K(t + \theta) - K(\theta)$.

If $K(t)$ is finite for a range $t_1 < \operatorname{Re}(t) < t_2$ then if $t_1 < 0 < t_2$ then $K(t)$ is analytic and infinitely differentiable for $t_1 < \operatorname{Re}(t) < t_2$. Moreover for t real and $t_1 < t < t_2$ $K(t)$ is strictly convex, and $K'(t)$ is strictly increasing.

Some properties of cumulants

Invariance and equivariance

The first cumulant is shift-equivariant; all of the others are shift-invariant. This means that, if we denote by $\kappa_n(X)$ the n-th cumulant of the probability distribution of the random variable X, then for any constant c:

- $\kappa_1(X + c) = \kappa_1(X) + c$ and
- $\kappa_n(X + c) = \kappa_n(X)$ for $n \geq 2$.

In other words, shifting a random variable (adding c) shifts the first cumulant (the mean) and doesn't affect any of the others.

Homogeneity

The n-th cumulant is homogeneous of degree n, i.e. if c is any constant, then

$$\kappa_n(cX) = c^n \kappa_n(X).$$

Additivity

If X and Y are independent random variables then $\kappa_n(X + Y) = \kappa_n(X) + \kappa_n(Y)$.

A negative result

Given the results for the cumulants of the normal distribution, it might be hoped to find families of distributions for which $\kappa_m = \kappa_{m+1} = \cdots = 0$ for some $m > 3$, with the lower-order cumulants (orders 3 to $m - 1$) being non-zero. There are no such distributions.[7] The underlying result here is that the cumulant generating function cannot be a finite-order polynomial of degree greater than 2.

Cumulants and moments

The moment generating function is given by:

$$M(t)=1+\sum_{n=1}^{\infty}\frac{\mu'_n t^n}{n!}=\exp\left(\sum_{n=1}^{\infty}\frac{\kappa_n t^n}{n!}\right)=\exp(K(t)).$$

So the cumulant generating function is the logarithm of the moment generating function

$$K(t)=\log M(t).$$

The first cumulant is the expected value; the second and third cumulants are respectively the second and third central moments (the second central moment is the variance); but the higher cumulants are neither moments nor central moments, but rather more complicated polynomial functions of the moments.

The moments can be recovered in terms of cumulants by evaluating the n-th derivative of $\exp(K(t))$ at $t=0$,

$$\mu'_n=M^{(n)}(0)=\left.\frac{\mathrm{d}^n\exp(K(t))}{\mathrm{d}t^n}\right|_{t=0}.$$

Likewise, the cumulants can be recovered in terms of moments by evaluating the n-th derivative of $\log M(t)$ at $t=0$,

$$\kappa_n=K^{(n)}(0)=\left.\frac{\mathrm{d}^n\log M(t)}{\mathrm{d}t^n}\right|_{t=0}.$$

The explicit expression for the n-th moment in terms of the first n cumulants, and vice versa, can be obtained by using Faà di Bruno's formula for higher derivatives of composite functions. In general, we have

$$\begin{aligned}\mu'_n&=\sum_{k=1}^{n}B_{n,k}(\kappa_1,\ldots,\kappa_{n-k+1})\\ \kappa_n&=\sum_{k=1}^{n}(-1)^{k-1}(k-1)!B_{n,k}(\mu'_1,\ldots,\mu'_{n-k+1}),\end{aligned}$$

where $B_{n,k}$ are incomplete (or partial) Bell polynomials.

In the like manner, if the mean is given by μ, the central moment generating function is given by

$$C(t)=\mathrm{E}[e^{t(x-\mu)}]=e^{-\mu t}M(t)=\exp(K(t)-\mu t),$$

and the n-th central moment is obtained in terms of cumulants as

$$\mu_n=C^{(n)}(0)=\left.\frac{\mathrm{d}^n}{\mathrm{d}t^n}\exp(K(t)-\mu t)\right|_{t=0}=\sum_{k=1}^{n}B_{n,k}(0,\kappa_2,\ldots,\kappa_{n-k+1}).$$

Also, for $n > 1$, the n-th cumulant in terms of the central moments is

$$\begin{aligned}\kappa_n &= K^{(n)}(0) = \frac{\mathrm{d}^n}{\mathrm{d}t^n}(\log C(t) + \mu t)\bigg|_{t=0} \\ &= \sum_{k=1}^{n}(-1)^{k-1}(k-1)!B_{n,k}(0,\mu_2,\ldots,\mu_{n-k+1}).\end{aligned}$$

The n-th moment μ'_n is an nth-degree polynomial in the first n cumulants. The first few expressions are:

$$\begin{aligned}\mu'_1 &= \kappa_1 \\ \mu'_2 &= \kappa_2 + \kappa_1^2 \\ \mu'_3 &= \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3 \\ \mu'_4 &= \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4 \\ \mu'_5 &= \kappa_5 + 5\kappa_4\kappa_1 + 10\kappa_3\kappa_2 + 10\kappa_3\kappa_1^2 + 15\kappa_2^2\kappa_1 + 10\kappa_2\kappa_1^3 + \kappa_1^5 \\ \mu'_6 &= \kappa_6 + 6\kappa_5\kappa_1 + 15\kappa_4\kappa_2 + 15\kappa_4\kappa_1^2 + 10\kappa_3^2 + 60\kappa_3\kappa_2\kappa_1 + 20\kappa_3\kappa_1^3 \\ &\quad + 15\kappa_2^3 + 45\kappa_2^2\kappa_1^2 + 15\kappa_2\kappa_1^4 + \kappa_1^6.\end{aligned}$$

The "prime" distinguishes the moments μ'_n from the central moments μ_n. To express the *central* moments as functions of the cumulants, just drop from these polynomials all terms in which κ_1 appears as a factor:

$$\begin{aligned}\mu_1 &= 0 \\ \mu_2 &= \kappa_2 \\ \mu_3 &= \kappa_3 \\ \mu_4 &= \kappa_4 + 3\kappa_2^2 \\ \mu_5 &= \kappa_5 + 10\kappa_3\kappa_2 \\ \mu_6 &= \kappa_6 + 15\kappa_4\kappa_2 + 10\kappa_3^2 + 15\kappa_2^3.\end{aligned}$$

Similarly, the n-th cumulant κ_n is an n-th-degree polynomial in the first n non-central moments. The first few expressions are:

$$\begin{aligned}\kappa_1 &= \mu'_1 \\ \kappa_2 &= \mu'_2 - {\mu'_1}^2 \\ \kappa_3 &= \mu'_3 - 3\mu'_2\mu'_1 + 2{\mu'_1}^3 \\ \kappa_4 &= \mu'_4 - 4\mu'_3\mu'_1 - 3{\mu'_2}^2 + 12\mu'_2{\mu'_1}^2 - 6{\mu'_1}^4 \\ \kappa_5 &= \mu'_5 - 5\mu'_4\mu'_1 - 10\mu'_3\mu'_2 + 20\mu'_3{\mu'_1}^2 + 30{\mu'_2}^2\mu'_1 - 60\mu'_2{\mu'_1}^3 + 24{\mu'_1}^5 \\ \kappa_6 &= \mu'_6 - 6\mu'_5\mu'_1 - 15\mu'_4\mu'_2 + 30\mu'_4{\mu'_1}^2 - 10{\mu'_3}^2 + 120\mu'_3\mu'_2\mu'_1 \\ &\quad - 120\mu'_3{\mu'_1}^3 + 30{\mu'_2}^3 - 270{\mu'_2}^2{\mu'_1}^2 + 360\mu'_2{\mu'_1}^4 - 120{\mu'_1}^6\end{aligned}$$

To express the cumulants κ_n for $n > 1$ as functions of the central moments, drop from these polynomials all terms in which μ'_1 appears as a factor:

$$\kappa_2 = \mu_2$$

$$\begin{aligned}\kappa_3 &= \mu_3 \\ \kappa_4 &= \mu_4 - 3{\mu_2}^2 \\ \kappa_5 &= \mu_5 - 10\mu_3\mu_2 \\ \kappa_6 &= \mu_6 - 15\mu_4\mu_2 - 10{\mu_3}^2 + 30{\mu_2}^3\,.\end{aligned}$$

To express the cumulants κ_n for $n > 2$ as functions of the standardized central moments, also set μ'_2=1 in the polynomials:

$$\begin{aligned}\kappa_3 &= \mu_3 \\ \kappa_4 &= \mu_4 - 3 \\ \kappa_5 &= \mu_5 - 10\mu_3 \\ \kappa_6 &= \mu_6 - 15\mu_4 - 10{\mu_3}^2 + 30\,.\end{aligned}$$

The cumulants are also related to the moments by the following recursion formula:

$$\kappa_n = \mu'_n - \sum_{m=1}^{n-1} \binom{n-1}{m-1} \kappa_m \mu'_{n-m}.$$

Cumulants and set-partitions

These polynomials have a remarkable combinatorial interpretation: the coefficients count certain partitions of sets. A general form of these polynomials is

$$\mu'_n = \sum_{\pi \,\in\, \Pi} \prod_{B \,\in\, \pi} \kappa_{|B|}$$

where

- π runs through the list of all partitions of a set of size n;
- "$B \in \pi$" means B is one of the "blocks" into which the set is partitioned; and
- $|B|$ is the size of the set B.

Thus each monomial is a constant times a product of cumulants in which the sum of the indices is n (e.g., in the term $\kappa_3\, {\kappa_2}^2\, \kappa_1$, the sum of the indices is 3 + 2 + 2 + 1 = 8; this appears in the polynomial that expresses the 8th moment as a function of the first eight cumulants). A partition of the integer n corresponds to each term. The *coefficient* in each term is the number of partitions of a set of n members that collapse to that partition of the integer n when the members of the set become indistinguishable.

Cumulants and combinatorics

Further connection between cumulants and combinatorics can be found in the work of Gian-Carlo Rota, where links to invariant theory, symmetric functions, and binomial sequences are studied via umbral calculus.[8]

Joint cumulants

The **joint cumulant** of several random variables X_1, ..., X_n is defined by a similar cumulant generating function

$$K(t_1, t_2, \dots, t_n) = \log E(\mathrm{e}^{\sum_{j=1}^n t_j X_j}).$$

A consequence is that

$$\kappa(X_1,\dots,X_n)=\sum_\pi(|\pi|-1)!(-1)^{|\pi|-1}\prod_{B\in\pi}E\left(\prod_{i\in B}X_i\right)$$

where π runs through the list of all partitions of { 1, ..., n }, B runs through the list of all blocks of the partition π, and $|\pi|$ is the number of parts in the partition. For example,

$$\kappa(X,Y,Z)=\operatorname{E}(XYZ)-\operatorname{E}(XY)\operatorname{E}(Z)-\operatorname{E}(XZ)\operatorname{E}(Y)-\operatorname{E}(YZ)\operatorname{E}(X)+2\operatorname{E}(X)\operatorname{E}(Y)\operatorname{E}(Z).$$

If any of these random variables are identical, e.g. if $X = Y$, then the same formulae apply, e.g.

$$\kappa(X,X,Z)=\operatorname{E}(X^2Z)-2\operatorname{E}(XZ)\operatorname{E}(X)-\operatorname{E}(X^2)\operatorname{E}(Z)+2\operatorname{E}(X)^2\operatorname{E}(Z),$$

although for such repeated variables there are more concise formulae. For zero-mean random vectors,

$$\kappa(X,Y,Z)=\operatorname{E}(XYZ).$$
$$\kappa(X,Y,Z,W)=\operatorname{E}(XYZW)-\operatorname{E}(XY)\operatorname{E}(ZW)-\operatorname{E}(XZ)\operatorname{E}(YW)-\operatorname{E}(XW)\operatorname{E}(YZ).$$

The joint cumulant of just one random variable is its expected value, and that of two random variables is their covariance. If some of the random variables are independent of all of the others, then any cumulant involving two (or more) independent random variables is zero. If all n random variables are the same, then the joint cumulant is the n-th ordinary cumulant.

The combinatorial meaning of the expression of moments in terms of cumulants is easier to understand than that of cumulants in terms of moments:

$$\operatorname{E}(X_1\cdots X_n)=\sum_\pi\prod_{B\in\pi}\kappa(X_i:i\in B).$$

For example:

$$\operatorname{E}(XYZ)=\kappa(X,Y,Z)+\kappa(X,Y)\kappa(Z)+\kappa(X,Z)\kappa(Y)+\kappa(Y,Z)\kappa(X)+\kappa(X)\kappa(Y)\kappa(Z).$$

Another important property of joint cumulants is multilinearity:

$$\kappa(X+Y,Z_1,Z_2,\dots)=\kappa(X,Z_1,Z_2,\dots)+\kappa(Y,Z_1,Z_2,\dots).$$

Just as the second cumulant is the variance, the joint cumulant of just two random variables is the covariance. The familiar identity

$$\operatorname{var}(X+Y)=\operatorname{var}(X)+2\operatorname{cov}(X,Y)+\operatorname{var}(Y)$$

generalizes to cumulants:

$$\kappa_n(X+Y)=\sum_{j=0}^n\binom{n}{j}\kappa(\underbrace{X,\dots,X}_j,\underbrace{Y,\dots,Y}_{n-j}).$$

Conditional cumulants and the law of total cumulance

The law of total expectation and the law of total variance generalize naturally to conditional cumulants. The case $n = 3$, expressed in the language of (central) moments rather than that of cumulants, says

$$\mu_3(X)=\operatorname{E}(\mu_3(X\mid Y))+\mu_3(\operatorname{E}(X\mid Y))+3\operatorname{cov}(\operatorname{E}(X\mid Y),\operatorname{var}(X\mid Y)).$$

In general,[9]

$$\kappa(X_1,\ldots,X_n)=\sum_\pi \kappa(\kappa(X_{\pi_1}\mid Y),\ldots,\kappa(X_{\pi_b}\mid Y))$$

where

- the sum is over all partitions π of the set { 1, ..., n } of indices, and
- π_1, ..., π_b are all of the "blocks" of the partition π; the expression $\kappa(X_{\pi_m})$ indicates that the joint cumulant of the random variables whose indices are in that block of the partition.

Relation to statistical physics

In statistical physics many extensive quantities – that is quantities that are proportional to the volume or size of a given system – are related to cumulants of random variables. The deep connection is that in a large system an extensive quantity like the energy or number of particles can be thought of as the sum of (say) the energy associated with a number of nearly independent regions. The fact that the cumulants of these nearly independent random variables will (nearly) add make it reasonable that extensive quantities should be expected to be related to cumulants.

A system in equilibrium with a thermal bath at temperature T can occupy states of energy E. The energy E can be considered a random variable, having the probability density. The partition function of the system is

$$Z(\beta)=\langle\exp(-\beta E)\rangle,$$

where β = 1/(kT) and k is Boltzmann's constant and the notation $\langle A\rangle$ has been used rather than $\mathbf{E}[A]$ for the expectation value to avoid confusion with the energy, E. The Helmholtz free energy is then

$$F(\beta)=-\beta^{-1}\log Z$$

and is clearly very closely related to the cumulant generating function for the energy. The free energy gives access to all of the thermodynamics properties of the system via its first second and higher order derivatives, such as its internal energy, entropy, and specific heat. Because of the relationship between the free energy and the cumulant generating function, all these quantities are related to cumulants e.g. the energy and specific heat are given by

$$E=\langle E\rangle_c$$
$$C=dE/dT=k\beta^2\langle E^2\rangle_c=k\beta^2(\langle E^2\rangle-\langle E\rangle^2)$$

and $\langle E^2\rangle_c$ symbolizes the second cumulant of the energy. Other free energy is often also a function of other variables such as the magnetic field or chemical potential μ, e.g.

$$\Omega=-\beta^{-1}\log(\langle\exp(-\beta E-\beta\mu N)\rangle),$$

where N is the number of particles and Ω is the grand potential. Again the close relationship between the definition of the free energy and the cumulant generating function implies that various derivatives of this free energy can be written in terms of joint cumulants of E and N.

History

The history of cumulants is discussed by Anders Hald.[10][11]

Cumulants were first introduced by Thorvald N. Thiele, in 1889, who called them *semi-invariants*.[12] They were first called *cumulants* in a 1932 paper[13] by Ronald Fisher and John Wishart. Fisher was publicly reminded of Thiele's work by Neyman, who also notes previous published citations of Thiele brought to Fisher's attention.[14] Stephen Stigler has said that the name *cumulant* was suggested to Fisher in a letter from Harold Hotelling. In a paper published in 1929,[15] Fisher had called them *cumulative moment functions*. The partition function in statistical physics was introduced by Josiah Willard Gibbs in 1901. The free energy is often called Gibbs free energy. In statistical mechanics, cumulants are also known as Ursell functions relating to a publication in 1927.

Cumulants in generalized settings

Formal cumulants

More generally, the cumulants of a sequence $\{ m_n : n = 1, 2, 3, \ldots \}$, not necessarily the moments of any probability distribution, are, by definition,

$$1+\sum_{n=1}^\infty \frac{m_n t^n}{n!} = \exp\left(\sum_{n=1}^\infty \frac{\kappa_n t^n}{n!}\right),$$

where the values of κ_n for $n = 1, 2, 3, \ldots$ are found formally, i.e., by algebra alone, in disregard of questions of whether any series converges. All of the difficulties of the "problem of cumulants" are absent when one works formally. The simplest example is that the second cumulant of a probability distribution must always be nonnegative, and is zero only if all of the higher cumulants are zero. Formal cumulants are subject to no such constraints.

Bell numbers

In combinatorics, the n-th Bell number is the number of partitions of a set of size n. All of the cumulants of the sequence of Bell numbers are equal to 1. The Bell numbers are the moments of the Poisson distribution with expected value 1.

Cumulants of a polynomial sequence of binomial type

For any sequence $\{ \kappa_n : n = 1, 2, 3, \ldots \}$ of scalars in a field of characteristic zero, being considered formal cumulants, there is a corresponding sequence $\{ \mu' : n = 1, 2, 3, \ldots \}$ of formal moments, given by the polynomials above. For those polynomials, construct a polynomial sequence in the following way. Out of the polynomial

$$\begin{aligned}\mu'_6 = {} & \kappa_6 + 6\kappa_5\kappa_1 + 15\kappa_4\kappa_2 + 15\kappa_4\kappa_1^2 + 10\kappa_3^2 + 60\kappa_3\kappa_2\kappa_1 + 20\kappa_3\kappa_1^3 \\ & + 15\kappa_2^3 + 45\kappa_2^2\kappa_1^2 + 15\kappa_2\kappa_1^4 + \kappa_1^6\end{aligned}$$

make a new polynomial in these plus one additional variable x:

$$\begin{aligned}p_6(x) = {} & \kappa_6\,x + (6\kappa_5\kappa_1 + 15\kappa_4\kappa_2 + 10\kappa_3^2)\,x^2 + (15\kappa_4\kappa_1^2 + 60\kappa_3\kappa_2\kappa_1 + 15\kappa_2^3)\,x^3 \\ & + (45\kappa_2^2\kappa_1^2)\,x^4 + (15\kappa_2\kappa_1^4)\,x^5 + (\kappa_1^6)\,x^6,\end{aligned}$$

and then generalize the pattern. The pattern is that the numbers of blocks in the aforementioned partitions are the exponents on x. Each coefficient is a polynomial in the cumulants; these are the Bell polynomials, named after Eric Temple Bell.

This sequence of polynomials is of binomial type. In fact, no other sequences of binomial type exist; every polynomial sequence of binomial type is completely determined by its sequence of formal cumulants.

Free cumulants

In the above moment-cumulant formula

$$E(X_1 \cdots X_n) = \sum_\pi \prod_{B \in \pi} \kappa(X_i : i \in B)$$

for joint cumulants, one sums over *all* partitions of the set { 1, ..., n }. If instead, one sums only over the noncrossing partitions, then, by solving these formulae for the κ in terms of the moments, one gets **free cumulants** rather than conventional cumulants treated above. These free cumulants were introduced by Roland Speicher and play a central role in free probability theory.[16][17] In that theory, rather than considering independence of random variables, defined in terms of tensor products of algebras of random variables, one considers instead free independence of random variables, defined in terms of free products of algebras[17].

The ordinary cumulants of degree higher than 2 of the normal distribution are zero. The *free* cumulants of degree higher than 2 of the Wigner semicircle distribution are zero.[17] This is one respect in which the role of the Wigner distribution in free probability theory is analogous to that of the normal distribution in conventional probability theory.

See also

- Entropic value at risk
- Cumulant generating function from a multiset
- Cornish–Fisher expansion
- Edgeworth expansion
- Polykay
- k-statistic, a minimum-variance unbiased estimator of a cumulant
- Ursell function
- Total position spread tensor as an application of cumulants to analyse the electronic wave function in quantum chemistry.

References

1. Weisstein, Eric W. "Cumulant". From MathWorld – A Wolfram Web Resource. http://mathworld.wolfram.com/Cumulant.html
2. Kendall, M. G., Stuart, A. (1969) *The Advanced Theory of Statistics*, Volume 1 (3rd Edition). Griffin, London. (Section 3.12)
3. Lukacs, E. (1970) *Characteristic Functions* (2nd Edition). Griffin, London. (Page 27)
4. Lukacs, E. (1970) *Characteristic Functions* (2nd Edition). Griffin, London. (Section 2.4)
5. Aapo Hyvarinen, Juha Karhunen, and Erkki Oja (2001) *Independent Component Analysis*, John Wiley & Sons. (Section 2.7.2)
6. Hamedani, G. G.; Volkmer, Hans; Behboodian, J. (2012-03-01). "A note on sub-independent random variables and a class of bivariate mixtures". *Studia Scientiarum Mathematicarum Hungarica*. **49** (1): 19–25. doi:10.1556/SScMath.2011.1183 (https://doi.org/10.1556%2FSScMath.2011.1183).
7. Lukacs, E. (1970) Characteristic Functions (2nd Edition), Griffin, London. (Theorem 7.3.5)

8. Rota, G.-C.; Shen, J. (2000). "On the Combinatorics of Cumulants". *Journal of Combinatorial Theory*. Series A. **91** (1–2): 283–304. doi:10.1006/jcta.1999.3017 (https://doi.org/10.1006%2Fjcta.1999.3017).
9. Brillinger, D.R. (1969). "The Calculation of Cumulants via Conditioning". *Annals of the Institute of Statistical Mathematics*. **21**: 215–218. doi:10.1007/bf02532246 (https://doi.org/10.1007%2Fbf02532246).
10. Hald, A. (2000) "The early history of the cumulants and the Gram–Charlier series" *International Statistical Review*, 68 (2): 137–153. (Reprinted in Steffen L. Lauritzen (http://www.stats.ox.ac.uk/~steffen/), ed. (2002). *Thiele: Pioneer in Statistics*. Oxford U. P. (http://www.oup.com/uk/catalogue/?ci=9780198509721) ISBN 978-0-19-850972-1. External link in `|publisher=` (help))
11. Hald, Anders (1998). *A History of Mathematical Statistics from 1750 to 1930*. New York: Wiley. ISBN 978-0-471-17912-2.
12. H. Cramér (1946) Mathematical Methods of Statistics, Princeton University Press, Section 15.10, p. 186.
13. Fisher, R.A. , John Wishart, J.. (1932) *The derivation of the pattern formulae of two-way partitions from those of simpler patterns* (http://plms.oxfordjournals.org/content/s2-33/1/195.full.pdf+html), Proceedings of the London Mathematical Society, Series 2, v. 33, pp. 195–208 doi: 10.1112/plms/s2-33.1.195 (https://doi.org/10.1112%2Fplms%2Fs2-33.1.195)
14. Neyman, J. (1956): 'Note on an Article by Sir Ronald Fisher,' *Journal of the Royal Statistical Society*, Series B (Methodological), 18, pp. 288–94.
15. Fisher, R. A. (1929). "Moments and Product Moments of Sampling Distributions" (https://digital.library.adelaide.edu.au/dspace/bitstream/2440/15200/1/74pt2.pdf) (PDF). *Proceedings of the London Mathematical Society*. **30**: 199–238. doi:10.1112/plms/s2-30.1.199 (https://doi.org/10.1112%2Fplms%2Fs2-30.1.199). hdl:2440/15200 (https://hdl.handle.net/2440%2F15200).
16. Speicher, Roland (1994). "Multiplicative functions on the lattice of non-crossing partitions and free convolution". *Mathematische Annalen*. **298** (4): 611–628. doi:10.1007/BF01459754 (https://doi.org/10.1007%2FBF01459754).
17. Novak, Jonathan; Śniady, Piotr (2011). "What Is a Free Cumulant?". *Notices of the American Mathematical Society*. **58** (2): 300–301. ISSN 0002-9920 (https://www.worldcat.org/issn/0002-9920).

External links

- Weisstein, Eric W. "Cumulant" (http://mathworld.wolfram.com/Cumulant.html). *MathWorld*.
- cumulant (http://jeff560.tripod.com/c.html) on the Earliest known uses of some of the words of mathematics (http://jeff560.tripod.com/mathword.html)