PCCP

COMMENT

View Article Online
View Journal | View Issue

Open Access Article. Published on 22 November 2023. Downloaded on 12/4/2024 7:44:33 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Check for updates

Cite this: *Phys. Chem. Chem. Phys.*, 2023, **25**, 32723

Comment on "Cumulant mapping as the basis of multi-dimensional spectrometry" by Leszek J. Frasinski, *Phys. Chem. Chem. Phys.*, 2022, 24, 20776–20787†

Åke Andersson

Received 31st May 2023,
Accepted 2nd November 2023

DOI: 10.1039/d3cp02525j

rsc.li/pccp

I state a general formula for the n-variate joint cumulant of the first order and prove that it satisfies the desired properties listed in Section 3.3 of *Phys. Chem. Chem. Phys.*, 2022, **24**, 20776–20787.

Motivation

A recent article by Frasinski[1] develops a theory of cumulant mapping, which extends covariance mapping[2] to any number of fragments. The central object of this theory is the n-variate joint cumulant of the first order, abbreviated nth cumulant. In Section 3.3, Frasinski lists what properties the nth cumulant should satisfy, and then gives explicit expressions for up to the 6th cumulant. How these can be found in practice is not elaborated upon.

The purpose of this comment is to show how we can find the nth cumulant—in theory and practice. I will do this by providing a general formula and describe how to evaluate it. Additionally, I will use the general formula to prove that cumulants fulfill some useful properties.

The general formula

Let $X_1,\ldots,X_n$ be random variables. Their multivariate cumulant-generating function is[3]

$$K_{X_1,\ldots,X_n}(t_1,\ldots,t_n) = \ln\left\langle \exp\left(\sum_i t_i X_i\right)\right\rangle, \tag{1}$$

From this function the nth cumulant is defined as[3]

$$\chi_n(X_1,\ldots,X_n) = \frac{\partial^n K_{X_1,\ldots,X_n}}{\partial t_1\cdots\partial t_n}(0,\ldots,0). \tag{2}$$

This definition is simple and useful for proving properties, but difficult to evaluate. Later in this comment I will derive the more explicit expression

$$\chi_n(\ldots) = \sum_{k=1}^{\left\lfloor\frac{n}{2}\right\rfloor} (-1)^{k+1}(k-1)! \sum_{S_1\sqcup\cdots\sqcup S_k = I} \prod_{j=1}^{k} \left\langle \prod_{i\in S_j} X_i - \langle X_i\rangle \right\rangle, \tag{3}$$

where $S_1\sqcup\cdots\sqcup S_k = I$ means all partitions of $\{1,\ldots,n\}$ into k sets.

How to evaluate it

Let us say we want to find the 7th cumulant. The main thing we should do is to find the partitions of 7 into k numbers that each are 2 or greater. For k = 1 there is the trivial 7; for k = 2 there is 5 + 2 and 4 + 3; and for k = 3 there is 3 + 2 + 2. From each nontrivial partition we then create a sum over all congruent products of covariances. Hopefully the rule becomes apparent by looking at the result

$$\chi_7(\ldots) = \langle 1234567\rangle - 1!\sum^{21}\langle 12\,345\rangle\langle 67\rangle - 1!\sum^{35}\langle 1234\rangle\langle 567\rangle + 2!\sum^{105}\langle 123\rangle\langle 45\rangle\langle 67\rangle, \tag{4}$$

where I have used i as a shorthand for $X_i - \langle X_i\rangle$ inside $\langle\,\rangle$. The prefactor of each sum is simply $(-1)^{k+1}(k-1)!$. The number of products in a sum can be calculated as

$$\frac{n!}{\prod_{m=1}^{n}(\#m)!(m!)^{\#m}}, \tag{5}$$

where $\#m$ is the number of parts of size m. Matlab and Python code implementing the nth cumulant is available as ESI.†

Department of Physics, University of Gothenburg, 412 96 Gothenburg, Sweden. E-mail: ake.andersson@physics.gu.se

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3cp02525j

Open Access Article. Published on 22 November 2023. Downloaded on 12/4/2024 7:44:33 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Useful properties

The four desired properties listed in Section 3.3 of the original article[1] are

- $\chi_n(\ldots) \neq 0$ only if all arguments are collectively correlated;
- $\chi_n(\ldots)$ has units of the product of all arguments;
- $\chi_n(\ldots)$ is linear in the arguments;
- $\chi_n(\ldots)$ is invariant under interchange of any two arguments.

I will now prove that the cumulant has these desired properties, starting with the interchange of arguments.

Property 1 (symmetric). The nth cumulant is invariant under permutation of its arguments:

$$\chi_n(X_{\pi(1)},\ldots,X_{\pi(n)}) = \chi_n(X_1,\ldots,X_n). \tag{6}$$

Proof. Commutativity of addition, and of differentiation. □

The desired properties about linearity and units are combined into one, because the former implies the latter.

Property 2 (multilinear). The nth cumulant is linear in each of its arguments:

$$\chi_n(aX + bY, Z_2,\ldots) = a\chi_n(X,Z_2,\ldots) + b\chi_n(Y,Z_2,\ldots). \tag{7}$$

Proof. Because of symmetry we only have to prove linearity in the first argument. By expanding the expression

$$K_{aX+bY,\ldots}(t_1,\ldots) - aK_{X,\ldots}(t_1,\ldots) - bK_{Y,\ldots}(t_1,\ldots) \tag{8}$$

in its first argument t_1, we find that the first-order terms cancel out. Differentiating with respect to t_1 (among others) and evaluating at the origin will therefore give

$$\chi_n(aX + bY, \ldots) - a\chi_n(X, \ldots) - b\chi_n(Y, \ldots,) = 0. \quad \square \tag{9}$$

Next, I phrase the desired property about correlation conversely.

Property 3 (discerning). Let $(A_i)_{i=1}^m$ and $(B_j)_{j=m+1}^n$ be nonempty tuples of random variables such that A_i and B_j are independent. Then the nth cumulant of $A \cup B$ vanishes:

$$\chi_n(A_1,\ldots,A_m, B_{m+1},\ldots,B_n) = 0. \tag{10}$$

Proof. Because $\exp\left(\sum_i t_i A_i\right)$ and $\exp\left(\sum_j t_j B_j\right)$ are independent, we can separate the generating function like

$$K_{A_1,\ldots,A_m,B_{m+1},\ldots,B_n}(\ldots) = K_{A_1,\ldots,A_m}(\ldots) + K_{B_{m+1},\ldots,B_n}(\ldots). \tag{11}$$

Differenting with respect to t_n and t_1 (among others) will annihilate both terms. □

Finally, I note an important property that follows from the last two. It tells us that independent signals simply add their contributions to a cumulant.

Property 4 (additive). Let $(A_i)_{i=1}^n$ and $(B_j)_{j=1}^n$ be equal-length tuples of random variables such that A_i and B_j are independent. Then the nth cumulant distributes over the addition of these tuples:

$$\chi_n(A_1 + B_1,\ldots,A_n + B_n) = \chi_n(A_1,\ldots,A_n) + \chi_n(B_1,\ldots,B_n). \tag{12}$$

Proof. By repeatedly using linearity, we can expand the left hand side into 2^n terms. The mixed terms containing both some A_i and some B_j vanish because of the discerning property. □

Explicit expression

Our strategy will be to approximate the generating function with Taylor series, starting from the inside. The exponential can be truncated by removing terms containing any t_i^2

$$\begin{aligned}\exp\left(\sum_i t_i X_i\right) &= \prod_i \exp(t_i X_i) \\ &= \prod_i \left(1 + t_i X_i + \mathcal{O}\left(t_i^2\right)\right) \\ &= \prod_i (1 + t_i X_i) + \sum_i \mathcal{O}\left(t_i^2\right).\end{aligned} \tag{13}$$

The last product turns into 2^n terms, one for each subset of $I = \{1,\ldots,n\}$. The term corresponding to S contains t_i if and only if S contains i. Explicitly,

$$\prod_i (1 + t_i X_i) = \sum_{S\subseteq I} \prod_{i\in S} t_i X_i. \tag{14}$$

Applying the expectation value is straightforward linearity

$$\begin{aligned}\left\langle \exp\left(\sum_i t_i X_i\right)\right\rangle &= \left\langle \sum_{S\subseteq I} \prod_{i\in S} t_i X_i + \sum_i \mathcal{O}\left(t_i^2\right)\right\rangle \\ &= \sum_{S\subseteq I} \left\langle \prod_{i\in S} t_i X_i \right\rangle + \sum_i \mathcal{O}\left(t_i^2\right).\end{aligned} \tag{15}$$

In anticipation of taking the logarithm, we extract a 1 from our expectation value taking out the term where S is the empty set

$$\begin{aligned}x &= \left\langle \exp\left(\sum_i t_i X_i\right)\right\rangle - 1 \\ &= \sum_{\varnothing\subset S\subseteq I} \left\langle \prod_{i\in S} t_i X_i \right\rangle + \sum_i \mathcal{O}\left(t_i^2\right).\end{aligned} \tag{16}$$

Now, we plug this x into the Taylor series of the logarithm

$$\ln(1 + x) = \sum_{k=1}^{\infty} (-1)^{k+1} x^k/k = x - x^2/2 + \cdots. \tag{17}$$

Recall that the cumulant is obtained by differentiating the above expression with respect to each variable and evaluating at the origin. This can be thought of as extracting the coefficient of the $\prod_i t_i$ term. Hence we will focus on contributions to it.

The first term, x, contains one subterm for each nonempty subset S of I. Only the subterm corresponding to $S = I$ will be proportional to $\prod_i t_i$. Its coefficient will then be $\left\langle \prod_i X_i \right\rangle$.

The second term, $-x^2/2$, contains when expanded one subterm for every ordered pair of nonempty subsets (S_1,S_2) of I. In order to get a term proportional to $\prod_i t_i$, each index i must be an element in exactly one of S_1 and S_2. In other words, $\{S_1,S_2\}$ must be a partition of I. For every partition of I there are 2! matching ordered pairs, each contributing $-1/2 \cdot \left\langle \prod_{i \in S_1} X_i \right\rangle \left\langle \prod_{i \in S_2} X_i \right\rangle$ to the coefficient.

By now the rule for the kth term is clear. It will contain subterms corresponding to each partition of I into k nonempty sets. Each subterm will be a product of the prefactor $(-1)^{k+1}(k-1)!$ and k expectation values.

$$\chi_n(X_1, \ldots, X_n) = \sum_{k=1}^{n} (-1)^{k+1}(k-1)! \sum_{S_1 \sqcup \cdots \sqcup S_k = I} \prod_{j=1}^{k} \left\langle \prod_{i \in S_j} X_i \right\rangle \tag{18}$$

We can make single-variable expectations vanish by replacing X_i with $X_i - \langle X_i \rangle$ everywhere. Using the additive property with $A_i = X_i$ and $B_i = -\langle X_i \rangle$, we see that this change preserves the cumulant.

Conflicts of interest

There are no conflicts to declare.

References

1 L. J. Frasinski, *Phys. Chem. Chem. Phys.*, 2022, **24**, 20776–20787.
2 V. Zhaunerchyk, L. Frasinski, J. H. Eland and R. Feifel, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2014, **89**, 053418.
3 A. Stuart and K. Ord, *Kendall's advanced theory of statistics, distribution theory*, John Wiley & Sons, 2010, vol. 1.

Open Access Article. Published on 22 November 2023. Downloaded on 12/4/2024 7:44:33 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.