Gaussian calculus

Let X be a Gaussian random variable with mean m and variance σ^2. Show that, with k a real number

$$\boxed{\langle e^{kX}\rangle \,=\, e^{km+\sigma^2k^2/2}}. \tag{1}$$

Note in passing that $\langle e^{kX}\rangle \geq e^{k\langle X\rangle}$. Why is that so?

In the remainder, we deal with the multivariate case. We denote by $\vec{x}$ a random vector with components $(x_1, x_2, \ldots x_n)$ and probability density function (pdf)

$$p(x_1, \ldots x_n) \,=\, \mathcal{N} \exp\left[-\frac{1}{2}\sum_{i,j=1}^{n} A_{ij}(x_i - m_i)(x_j - m_j)\right] \tag{2}$$

where $\vec{m} = (m_1, \ldots m_n)$ is some constant vector, A is a $n \times n$ matrix and $\mathcal{N}$ a normalization factor.

1) What is the reason why we can restrict to symmetric matrices A ? Which property should it exhibit?

2) What is the meaning of $\vec{m}$?

3) Normalize the distribution and show that $\mathcal{N} \,=\, \dfrac{\sqrt{\det A}}{(2\pi)^{n/2}}$.

4) Whenever $\vec{m} \neq \vec{0}$, it is convenient to work with the shifted variable $\vec{y} = \vec{x} - \vec{m}$. What is the pdf of $\vec{y}$? We consider hereafter that $\vec{m} = \vec{0}$.

5) Show that the two-point correlation function reads

$$\boxed{\langle x_i x_j\rangle \,=\, (A^{-1})_{ij}}. \tag{3}$$

To this end, one can start from $\sum_j \langle x_i x_j\rangle A_{jk}$ and integrate by parts (a little trick due to Onsager).

6) **Novikov relation**. For $f(\vec{x})$ an (essentially) arbitrary function, show that

$$\langle x_i f(\vec{x})\rangle \,=\, \sum_j \langle x_i x_j\rangle \left\langle \frac{\partial f}{\partial x_j}\right\rangle. \tag{4}$$

7) **Wick's theorem**. Repeated use of Eq. (4) enables us to show Wick's theorem: while odd moments vanish, all even moments of a Gaussian distribution factorize into second moments as

$$\langle x_{i_1} x_{i_2} \ldots x_{i_{2\ell}}\rangle \,=\, \sum_{\text{all complete pairings of } 2\ell \text{ elements}} \;\; \prod_{\text{each pair } (k,p)} \langle x_k x_p\rangle\,. \tag{5}$$

In the simplest case $\ell = 2$, this means

$$\langle x_i x_j x_k x_p\rangle \,=\, \langle x_i x_j\rangle \langle x_k x_p\rangle + \langle x_i x_k\rangle \langle x_j x_p\rangle + \langle x_i x_p\rangle \langle x_j x_k\rangle\,. \tag{6}$$

For $\ell = 3$, how many products of the type $\langle x_i x_j\rangle \langle x_k x_l\rangle \langle x_m x_n\rangle$ would be summed in the development of $\langle x_i x_j x_k x_l x_m x_n\rangle$? How many terms for ℓ arbitrary? How do these results transpose to the scalar case? Show in particular that then,

$$\langle x^4\rangle \,=\, 3\left\langle x^2\right\rangle^2, \tag{7}$$

which is an important characteristics of a centered scalar Gaussian variable (implying the vanishing of the fourth cumulant). And, after all these prolegomena, show relation (6) from Eq. (4).

8) Application. Let the pdf of a Gaussian couple (x, y) be $p(x, y) = \mathcal{N} \exp(-x^2/2 - xy - 2y^2)$. Compute $\langle x^2\rangle$, $\langle y^2\rangle$, $\langle xy\rangle$, $\langle x^2y^2\rangle$, $\langle x^4y^2\rangle$.

9) Characteristic function. A convenient way of calculating moments/cumulants of a pdf is to work out its characteristic function. Show that

$$\boxed{\left\langle e^{\,i\,\vec{k}\cdot\vec{x}}\right\rangle \;=\; \exp\left(-\frac{1}{2}\sum_{i,j} k_i\, A^{-1}_{ij}\, k_j\right)}, \tag{8}$$

where $\vec{k} = (k_1, \dots k_n)$. What if $\vec{m} \neq \vec{0}$? The cumulant generating function is thus quadratic[1] in the k_i.

At this point, we see that taking the inverse Fourier transform of any function of the form (8) yields a pdf of Gaussian type (very similar calculation as that which leads to (8)). We thus have the equivalence (assuming that the relevant coupling matrices A/A^{-1} are invertible): *Gaussian pdf ⇔ quadratic cumulant generating function.*

10) Study of the marginals: the characteristic function is useful. Define the marginal distribution

$$p_{\backslash m}(x_1 \dots x_{m-1}, x_{m+1}, \dots x_n) \;=\; \int dx_m\, p(x_1, \dots x_n). \tag{9}$$

Show that its characteristic function simply stems from the original one, with $k_m = 0$. How is the new coupling matrix $A_{\backslash m}$ obtained? *Remember this result*: unlike A, A^{-1} directly encodes information on the marginal distributions. Application to question 8: what are the (scalar) x and y marginals?

11) Linear combinations of components. Consider a modified vector $\vec{z} = B\vec{x}$ where B is an arbitrary matrix. Show that $\vec{z}$ remains a Gaussian vector, by considering its characteristic function.

12) In general, "independent" and "2-point uncorrelated" are not synonymous. Why? Yet, prove that for a Gaussian distribution, the two notions coincide.

13) A simple random walk. Consider a random walk on the line, starting from position $x_0 = 0$. At each time step, a Gaussian random variable η (independent from step to step, with mean 0 and unit variance) is chosen: the position at step n is thus

$$x_n \;=\; \eta_1 + \eta_2 \dots + \eta_n. \tag{10}$$

a) Argue that $(x_1, x_2, \dots, x_n)$ forms a Gaussian multivariate distribution.

b) Compute $p(x_1)$, $p(x_2|x_1)$, $p(x_1, x_2)$ etc. to get finally the joint distribution $p(x_1, \dots, x_n)$. What is the coupling matrix A here?

c) Show directly that the correlation function is given by

$$C_{ij} \;=\; \langle x_i x_j \rangle \;=\; \min(i, j). \tag{11}$$

Check that A and C are inverse matrices, as they should.

d) From the matrix C, discuss the marginal distributions.

Gaussian random variables, from which Gaussian processes are defined, are entirely specified by their first and second moment (a vector for the mean, a matrix for the correlations). These variables constitute an important class, met in many areas of physics and beyond (Brownian motion, modelling of phenomena involving noise, learning etc.). The central limit theorem is at the root of their ubiquity (mostly, but arguably not exclusively...).

---

[1] a theorem due to Marcinkiewicz (1939) establishes that a cumulant generating function cannot be a polynomial of the k_i of degree larger than 2. As a consequence, either all cumulants but the first two vanish, or there is an infinite number of non vanishing cumulants. This can be kept in mind: a theory that truncates a cumulant expansion at a non vanishing order strictly greater than 2 cannot be self-consistent, and yield an underlying positive pdf. Yet, such an expansion can be useful in practice, see P. Hänggi and P. Talkner, *J. Stat. Phys.* **22**, 65 (1980).