

## APPENDIX C

## CONVOLUTIONS AND CUMULANTS

First we note some general mathematical facts which have nothing to do with probability theory. Given a set of real functions  $f_1(x), f_2(x), \dots, f_n(x)$  defined on the real line and not necessarily non-negative, suppose that their integrals (zero'th moments) and their first, second, and third moments exist:

$$\begin{aligned} Z_i &\equiv \int_{-\infty}^{\infty} f_i(x) dx < \infty, & S_i &\equiv \int_{-\infty}^{\infty} x^2 f_i(x) dx < \infty, \\ F_i &\equiv \int_{-\infty}^{\infty} x f_i(x) dx < \infty & T_i &\equiv \int_{-\infty}^{\infty} x^3 f_i(x) dx < \infty \end{aligned} \quad (\text{C-1})$$

The convolution of  $f_1$  and  $f_2$  is defined by

$$h(x) \equiv \int_{-\infty}^{\infty} f_1(y) f_2(x-y) dy \quad (\text{C-2})$$

or in condensed notation,  $h = f_1 * f_2$ . Convolution is associative:  $(f_1 * f_2) * f_3 = f_1 * (f_2 * f_3)$ , so we can write a multiple convolution as  $(h = f_1 * f_2 * f_3 * \dots * f_n)$  without ambiguity. What happens to the moments under this operation? The zero'th moment of  $h(x)$  is

$$Z_h = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f_1(y) f_2(x-y) = \int dy f_1(y) Z_2 = Z_1 Z_2 \quad (\text{C-3})$$

Therefore, if  $Z_i \neq 0$  we can multiply  $f_i(x)$  by some constant factor which makes  $Z_i = 1$ , and this property will be preserved under convolution. In the following we assume that this has been done for all  $i$ . Then the first moment of the convolution is

$$\begin{aligned} F_h &= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy f_1(y) x f_2(x-y) = \int dy f_1(y) \int_{-\infty}^{\infty} dq (y+q) f_2(q) \\ &= \int_{-\infty}^{\infty} dy f_1(y) [y Z_2 + F_2] = F_1 Z_2 + Z_1 F_2 \end{aligned} \quad (\text{C-4})$$

so the first moments are additive under convolution:

$$F_h = F_1 + F_2 \quad (\text{C-5})$$

For the second moment, we have by a similar argument

$$S_h = \int dy f_1(y) \int dq (y^2 + 2yq + q^2) f_2(q) = S_1 Z_2 + 2F_1 F_2 + Z_1 S_2 \quad (\text{C-6})$$

or,

$$S_h = S_1 + 2F_1 F_2 + S_2 \quad (\text{C-7})$$

Subtracting the square of (C-5), the cross product term cancels out and we see that there is another quantity additive under convolution:

$$[S_h - (F_h)^2] = [S_1 - (F_1)^2] + [S_2 - (F_2)^2] \quad (\text{C-8})$$

Proceeding to the third moment, we find

$$T_h = T_1 Z_2 + 3S_1 F_2 + 3F_1 S_2 + Z_1 T_2 \quad (\text{C-9})$$

and after some algebra [subtracting off functions of (C-5) and (C-7)] we can confirm that there is a third quantity, namely

$$T_h - 3 S_h F_h + 2 (F_h)^3 \quad (\text{C-10})$$

that is additive under convolution.

This generalizes at once to any number of such functions: let  $h(x) \equiv f_1 * f_2 * f_3 * \cdots * f_n$ . Then we have found the additive quantities

$$\begin{aligned} F_h &= \sum_{i=1}^n F_i \\ S_h - F_h^2 &= \sum_{i=1}^n (S_i - F_i^2) \\ T_h - 3S_h F_h + 2F_h^3 &= \sum_{i=1}^n (T_i - 3S_i F_i + 2F_i^3) \end{aligned} \quad (\text{C-11})$$

These quantities, which “accumulate” additively under convolution, are called the *cumulants*; we have developed them in this way to emphasize that the notion has nothing, fundamentally, to do with probability.

At this point we define the  $n$ 'th cumulant as the  $n$ 'th moment, with ‘correction terms’ from lower moments, so chosen as to make the result additive under convolution. Then two questions call out for solution: (1) Do such correction terms always exist?; and (2) If so, how do we find a general algorithm to construct them?

To answer them we need a more powerful mathematical method. Introduce the fourier transform of  $f_i(x)$ :

$$F_i(\alpha) \equiv \int_{-\infty}^{\infty} f_i(x) e^{i\alpha x} dx \quad f_i(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F_i(\alpha) e^{-i\alpha x} d\alpha \quad (\text{C-12})$$

Under convolution it behaves very simply:

$$\begin{aligned} H(\alpha) &= \int_{-\infty}^{\infty} h(x) e^{i\alpha x} dx = \int dy f_1(y) \int dx e^{i\alpha x} f_2(x-y) \\ &= \int dy f_1(y) \int dq e^{i\alpha(y+q)} f_2(q) \\ &= F_1(\alpha) F_2(\alpha) \end{aligned} \quad (\text{C-13})$$

In other words,  $\log F(\alpha)$  is additive under convolutions. This function has some remarkable properties in connection with the notion of the “Cepstrum” discussed later. For now, examine the power series expansions of  $F(\alpha)$  and  $\log F(\alpha)$ . The first is

$$F(\alpha) = M_0 + M_1(i\alpha) + M_2 \frac{(i\alpha)^2}{2!} + M_3 \frac{(i\alpha)^3}{3!} + \cdots \quad (\text{C-14})$$

with the coefficients

$$M_n = \left. \frac{1}{i^n} \frac{d^n F(\alpha)}{d\alpha^n} \right]_{\alpha=0} = \int_{-\infty}^{\infty} x^n f(x) dx \quad (\text{C-15})$$

which are just the  $n$ 'th moments of  $f(x)$ ; if  $f(x)$  has moments up to order  $N$ , then  $F(\alpha)$  is differentiable  $N$  times at the origin. There is a similar expansion for  $\log F(\alpha)$ :

$$\log F(\alpha) = C_0 + C_1(i\alpha) + C_2 \frac{(i\alpha)^2}{2!} + C_3 \frac{(i\alpha)^3}{3!} + \dots \quad (\text{C-16})$$

Evidently, all its coefficients

$$C_n = \left. \frac{1}{i^n} \frac{d^n}{d\alpha^n} \log F(\alpha) \right]_{\alpha=0} \quad (\text{C-17})$$

are additive under convolution, and are therefore cumulants. The first few are

$$C_0 = \log F(0) = \log \int f(x) dx = \log Z \quad (\text{C-18})$$

$$C_1 = \frac{1}{i} \frac{\int ix f(x) dx}{\int f(x) dx} = \frac{F}{Z} \quad (\text{C-19})$$

$$C_2 = \frac{d^2}{d(i\alpha)^2} \log F(\alpha) = \frac{d}{d(i\alpha)} \frac{\int x f(x) e^{i\alpha x}}{\int f(x) e^{i\alpha x} dx} = \frac{\int f \int x^2 f - (\int x f)^2}{(\int f)^2}$$

or,

$$C_2 = \frac{S}{Z} - \left( \frac{F}{Z} \right)^2 \quad (\text{C-20})$$

which we recognize as just the cumulants found directly above; likewise, after some tedious calculation  $C_3$  and  $C_4$  prove to be equal to the third and fourth cumulants (C-10). Have we then found in (C-17) all the cumulants of a function, or are there still more that cannot be found in this way? We would argue that if all the  $C_i$  exist (*i.e.*  $f(x)$  has moments of all orders, so  $F(\alpha)$  is an entire function) then the  $C_i$  uniquely determine  $F(\alpha)$  and therefore  $f(x)$ , so they must include all the algebraically independent cumulants; any others must be linear functions of the  $C_i$ . But if  $f(x)$  does not have moments of all orders, the answer is not obvious, and further investigation is needed.

### Relation of Cumulants and Moments

While adhering to our convention  $Z = 1$ , let us go to a more general notation for the  $n$ 'th moment of a function:

$$M_n \equiv \int_{-\infty}^{\infty} x^n f(x) dx = \left. \frac{d^n}{d(i\alpha)^n} \int f(x) e^{i\alpha x} dx \right]_{\alpha=0} = i^{-n} F^{(n)}(0), \quad n = 0, 1, 2, \dots \quad (\text{C-21})$$

It is often convenient to use also the notation

$$M_n = \overline{x^n} \quad (\text{C-22})$$

indicating an average of  $x^n$  with respect to the function  $f(x)$ . We stress that these are not in general probability averages; we are indicating some general algebraic relations in which  $f(x)$  need not be nonnegative. For probability averages we always reserve the notation  $\langle x \rangle$  or  $E(x)$ .

If a function  $f(x)$  has moments of all orders, then its fourier transform has a power series expansion

$$F(\alpha) = \sum_{n=0}^{\infty} M_n (i\alpha)^n \quad (\text{C-23})$$

Evidently, the first cumulant is the same as the first moment:

$$C_1 = M_1 = \bar{x} \quad (\text{C-24})$$

while for the second cumulant we have  $C_2 = M_2 - M_1^2$ ; but this is

$$C_2 = \int [x - M_1]^2 f(x) dx = \overline{(x - \bar{x})^2} = \overline{x^2} - \bar{x}^2, \quad (\text{C-25})$$

the second moment of  $x$  about its mean value, called the *second central moment* of  $f(x)$ . Likewise, the third central moment is

$$\int (x - \bar{x})^3 f(x) dx = \int [x^3 - 3\bar{x}x^2 + 3\bar{x}^2x - \bar{x}^3] f(x) dx \quad (\text{C-26})$$

but this is just the third cumulant (C-11):

$$C_3 = M_3 - 3M_1M_2 + 2M_1^3 \quad (\text{C-27})$$

and at this point we might conjecture that all the cumulants are just the corresponding central moments. However, this turns out not to be the case: we find that the fourth central moment is

$$\overline{(x - \bar{x})^4} = M_4 - 4M_3M_1 + 6M_2M_1^2 - 3M_1^4 \quad (\text{C-28})$$

but the fourth cumulant is

$$C_4 = M_4 - 4M_3M_1 - 3M_2^2 + 12M_2M_1^2 - 6M_1^4. \quad (\text{C-29})$$

So they are related by

$$\overline{(x - \bar{x})^4} = C_4 + 3C_2^2. \quad (\text{C-30})$$

Thus the fourth central moment is not a cumulant; it is not a linear function of cumulants. However, we have found it true that, for  $n = 1, 2, 3, 4$  the moments up to order  $n$  and the cumulants up to order  $n$  uniquely determine each other; we leave it for the reader to see, from examination of the above relations, whether this is or is not true for all  $n$ .

If our functions  $f(x)$  are probability densities, many useful approximations are written most efficiently in terms of the first few terms of a cumulant expansion.

### Examples

What are the cumulants of a gaussian distribution? Let

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (\text{C-31})$$

Then we find the fourier transform

$$F(\alpha) = \exp(i\alpha\mu - \alpha^2\sigma^2/2) \quad (\text{C-32})$$

so that

$$\log F(\alpha) = i\alpha\mu - \alpha^2\sigma^2/2 \quad (\text{C-33})$$

and so

$$C_0 = 0, \quad C_1 = \alpha, \quad C_2 = \sigma^2 \quad (\text{C-34})$$

and all higher  $C_n$  are zero. A gaussian distribution is characterized by the fact that it has only two nontrivial cumulants, the mean and variance.