appc1, 9/19/2004 09:51

## 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 $f_{1}(x), f_{2}(x), \cdots f_{n}(x)$ defined on the real line and not necessarily nonnegative, suppose that their integrals (zero'th moments) and their first, second, and third moments exist:

$$
\begin{align*}
Z_{i} \equiv \int_{-\infty}^{\infty} f_{i}(x) d x<\infty, & S_{i} \equiv \int_{-\infty}^{\infty} x^{2} f_{i}(x) d x<\infty \\
F_{i} \equiv \int_{-\infty}^{\infty} x f_{i}(x) d x<\infty & T_{i} \equiv \int_{-\infty}^{\infty} x^{3} f_{i}(x) d x<\infty \tag{C-1}
\end{align*}
$$

The convolution of $f_{1}$ and $f_{2}$ is defined by

$$
\begin{equation*}
h(x) \equiv \int_{-\infty}^{\infty} f_{1}(y) f_{2}(x-y) d y \tag{C-2}
\end{equation*}
$$

or in condensed notation, $h=f_{1} * f_{2}$. Convolution is associative: $\left(f_{1} * f_{2}\right) * f_{3}=f_{1} *\left(f_{2} * f_{3}\right)$, so we can write a multiple convolution as ( $h=f_{1} * f_{2} * f_{3} * \cdots * f_{n}$ ) without ambiguity. What happens to the moments under this operation? The zero'th moment of $h(x)$ is

$$
\begin{equation*}
Z_{h}=\int_{-\infty}^{\infty} d x \int_{-\infty}^{\infty} d y f_{1}(y) f_{2}(x-y)=\int d y f_{1}(y) Z_{2}=Z_{1} Z_{2} \tag{C-3}
\end{equation*}
$$

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

$$
\begin{align*}
F_{h} & =\int_{-\infty}^{\infty} d x \int_{-\infty}^{\infty} d y f_{1}(y) x f_{2}(x-y)=\int d y f_{1}(y) \int_{-\infty}^{\infty} d q(y+q) f_{2}(q) \\
& =\int_{-\infty}^{\infty} d y f_{1}(y)\left[y Z_{2}+F_{2}\right]=F_{1} Z_{2}+Z_{1} F_{2} \tag{C-4}
\end{align*}
$$

so the first moments are additive under convolution:

$$
\begin{equation*}
F_{h}=F_{1}+F_{2} \tag{C-5}
\end{equation*}
$$

For the second moment, we have by a similar argument

$$
\begin{equation*}
S_{h}=\int d y f_{1}(y) \int d q\left(y^{2}+2 y q+q^{2}\right) f_{2}(q)=S_{1} Z_{2}+2 F_{1} F_{2}+Z_{1} S_{2} \tag{C-6}
\end{equation*}
$$

or,

$$
\begin{equation*}
S_{h}=S_{1}+2 F_{1} F_{2}+S_{2} \tag{C-7}
\end{equation*}
$$

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

$$
\begin{equation*}
\left[S_{h}-\left(F_{h}\right)^{2}\right]=\left[S_{1}-\left(F_{1}\right)^{2}\right]+\left[S_{2}-\left(F_{2}\right)^{2}\right] \tag{C-8}
\end{equation*}
$$

Proceeding to the third moment, we find

$$
\begin{equation*}
T_{h}=T_{1} Z_{2}+3 S_{1} F_{2}+3 F_{1} S_{2}+Z_{1} T_{2} \tag{C-9}
\end{equation*}
$$

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

$$
\begin{equation*}
T_{h}-3 S_{h} F_{h}+2\left(F_{h}\right)^{3} \tag{C-10}
\end{equation*}
$$

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

$$
\begin{gather*}
F_{h}=\sum_{i=1}^{n} F_{i} \\
S_{h}-F_{h}^{2}=\sum_{i=1}^{n}\left(S_{i}-F_{i}^{2}\right)  \tag{C-11}\\
T_{h}-3 S_{h} F_{h}+2 F_{h}^{3}=\sum_{i=1}^{n}\left(T_{i}-3 S_{i} F_{i}+2 F_{i}^{3}\right)
\end{gather*}
$$

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 $f_{i}(x)$ :

$$
\begin{equation*}
F_{i}(\alpha) \equiv \int_{-\infty}^{\infty} f_{i}(x) e^{i \alpha x} d x \quad f_{i}(x)=\frac{1}{2 \pi} \int_{-\infty}^{\infty} F_{i}(\alpha) e^{-i \alpha x} d \alpha \tag{C-12}
\end{equation*}
$$

Under convolution it behaves very simply:

$$
\begin{align*}
H(\alpha) & =\int_{-\infty}^{\infty} h(x) e^{i \alpha x} d x=\int d y f_{1}(y) \int d x e^{i \alpha x} f_{2}(x-y) \\
& =\int d y f_{1}(y) \int d q e^{i \alpha(y+q)} f_{2}(q)  \tag{C-13}\\
& =F_{1}(\alpha) F_{2}(\alpha)
\end{align*}
$$

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

$$
\begin{equation*}
F(\alpha)=M_{0}+M_{1}(i \alpha)+M_{2} \frac{(i \alpha)^{2}}{2!}+M_{3} \frac{(i \alpha)^{3}}{3!}+\cdots \tag{C-14}
\end{equation*}
$$

with the coefficients

$$
\begin{equation*}
\left.M_{n}=\frac{1}{i^{n}} \frac{d^{n} F(\alpha)}{d \alpha^{n}}\right]_{\alpha=0}=\int_{-\infty}^{\infty} x^{n} f(x) d x \tag{C-15}
\end{equation*}
$$

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 $\log F(\alpha)$ :

$$
\begin{equation*}
\log F(\alpha)=C_{0}+C_{1}(i \alpha)+C_{2} \frac{(i \alpha)^{2}}{2!}+C_{3} \frac{(i \alpha)^{3}}{3!}+\cdots \tag{C-16}
\end{equation*}
$$

Evidently, all its coefficients

$$
\begin{equation*}
\left.C_{n}=\frac{1}{i^{n}} \frac{d^{n}}{d \alpha^{n}} \log F(\alpha)\right]_{\alpha=0} \tag{C-17}
\end{equation*}
$$

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

$$
\begin{gather*}
C_{0}=\log F(0)=\log \int f(x) d x=\log Z  \tag{C-18}\\
C_{1}=\frac{1}{i} \frac{\int i x f(x) d x}{\int f(x) d x}=\frac{F}{Z}  \tag{C-19}\\
C_{2}=\frac{d^{2}}{d(i \alpha)^{2}} \log F(\alpha)=\frac{d}{d(i \alpha)} \frac{\int x f(x) e^{i \alpha x}}{\int f(x) e^{i \alpha x} d x}=\frac{\int f \int x^{2} f-\left(\int x f\right)^{2}}{\left(\int f\right)^{2}}
\end{gather*}
$$

or,

$$
\begin{equation*}
C_{2}=\frac{S}{Z}-\left(\frac{F}{Z}\right)^{2} \tag{C-20}
\end{equation*}
$$

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 ( $\mathrm{C}-10$ ). 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:

$$
\begin{equation*}
\left.M_{n} \equiv \int_{-\infty}^{\infty} x^{n} f(x) d x=\frac{d^{n}}{d(i \alpha)^{n}} \int f(x) e^{i \alpha x} d x\right]_{\alpha=o}=i^{-n} F^{(n)}(0), \quad n=0,1,2, \ldots \tag{C-21}
\end{equation*}
$$

It is often convenient to use also the notation

$$
\begin{equation*}
M_{n}=\overline{x^{n}} \tag{C-22}
\end{equation*}
$$

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

$$
\begin{equation*}
F(\alpha)=\sum_{n=0}^{\infty} M_{n}(i \alpha)^{n} \tag{C-23}
\end{equation*}
$$

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

$$
\begin{equation*}
C_{1}=M_{1}=\bar{x} \tag{C-24}
\end{equation*}
$$

while for the second cumulant we have $C_{2}=M_{2}-M_{1}^{2}$; but this is

$$
\begin{equation*}
C_{2}=\int\left[x-M_{1}\right]^{2} f(x) d x=\overline{(x-\bar{x})^{2}}=\overline{x^{2}}-\bar{x}^{2}, \tag{C-25}
\end{equation*}
$$

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

$$
\begin{equation*}
\int(x-\bar{x})^{3} f(x) d x=\int\left[x^{3}-3 \bar{x} x^{2}+3 \bar{x}^{2} x-\bar{x}^{3}\right] f(x) d x \tag{C-26}
\end{equation*}
$$

but this is just the third cumulant ( $\mathrm{C}-11$ ):

$$
\begin{equation*}
C_{3}=M_{3}-3 M_{1} M_{2}+2 M_{1}^{3} \tag{C-27}
\end{equation*}
$$

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

$$
\begin{equation*}
\overline{(x-\bar{x})^{4}}=M_{4}-4 M_{3} M_{1}+6 M_{2} M_{1}^{2}-3 M_{1}^{4} \tag{C-28}
\end{equation*}
$$

but the fourth cumulant is

$$
\begin{equation*}
C_{4}=M_{4}-4 M_{3} M_{1}-3 M_{2}^{2}+12 M_{2} M_{1}^{2}-6 M_{1}^{4} . \tag{C-29}
\end{equation*}
$$

So they are related by

$$
\begin{equation*}
\overline{(x-\bar{x})^{4}}=C_{4}+3 C_{2}^{2} . \tag{C-30}
\end{equation*}
$$

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

$$
\begin{equation*}
f(x)=\frac{1}{\sqrt{2 \pi \sigma^{2}}} \exp \left(\frac{(x-\mu)^{2}}{2 \sigma^{2}}\right) \tag{C-31}
\end{equation*}
$$

Then we find the fourier transform

$$
\begin{equation*}
F(\alpha)=\exp \left(i \alpha \mu-\alpha^{2} \sigma^{2} / 2\right) \tag{C-32}
\end{equation*}
$$

so that

$$
\begin{equation*}
\log F(\alpha)=i \alpha \mu-\alpha^{2} \sigma^{2} / 2 \tag{C-33}
\end{equation*}
$$

and so

$$
\begin{equation*}
C_{0}=0, \quad C_{1}=\alpha, \quad C_{2}=\sigma^{2} \tag{C-34}
\end{equation*}
$$

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

