16.2. Characteristic functions

The characteristic function $\phi(u)$ associated with the p.d.f. $f(x)$ is essentially its (inverse) Fourier transform, or the expectation value of $\exp(iux)$:

$$\phi(u) = E(e^{iux}) = \int_{-\infty}^{\infty} e^{iux} f(x)dx \ . \tag{16.12}$$

It is often useful, and several of its properties follow [1].

It follows from Eqs. (16.3a) and (16.12) that the nth moment of the distribution $f(x)$ is given by

$$i^{-n} \frac{d^n\phi}{du^n}\bigg|_{u=0} = \int_{-\infty}^{\infty} x^n f(x)dx = \alpha_n \ . \tag{16.13}$$

Thus it is often easy to calculate all the moments of a distribution defined by $\phi(u)$, even when $f(x)$ is difficult to obtain.

If $f_1(x)$ and $f_2(y)$ have characteristic functions $\phi_1(u)$ and $\phi_2(u)$, then the characteristic function of the weighted sum $ax + by$ is $\phi_1(au)\phi_2(bu)$. The addition rules for common distributions (*e.g.*, that the sum of two numbers from Gaussian distributions also has a Gaussian distribution) easily follow from this observation.

Let the (partial) characteristic function corresponding to the conditional p.d.f. $f_2(x|z)$ be $\phi_2(u|z)$, and the p.d.f. of z be $f_1(z)$. The characteristic function after integration over the conditional value is

$$\phi(u) = \int \phi_2(u|z)\, f_1(z)dz \ . \tag{16.14}$$

Suppose we can write ϕ_2 in the form

$$\phi_2(u|z) = A(u)e^{ig(u)z} \ . \tag{16.15}$$

Then

$$\phi(u) = A(u)\phi_1(g(u)) \ . \tag{16.16}$$

The semi-invariants κ_n are defined by

$$\phi(u) = \exp\left(\sum_1^{\infty} \frac{\kappa_n}{n!}(iu)^n\right) = \exp\left(i\kappa_1 u - \tfrac{1}{2}\kappa_2 u^2 + \ldots\right) \ . \tag{16.17}$$

The κ_n's are related to the moments α_n and m_n. The first few relations are

$$\kappa_1 = \alpha_1 \ (= \mu, \text{ the mean})$$

$$\kappa_2 = m_2 = \alpha_2 - \alpha_1^2 \ (= \sigma^2, \text{ the variance})$$

$$\kappa_3 = m_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^2 \ . \tag{16.18}$$

16.3. Some probability distributions

Table 16.1 gives a number of common probability density functions and corresponding characteristic functions, means, and variances. Further information may be found in Refs. 1–6; Ref. 6 has particularly detailed tables. Monte Carlo techniques for generating each of them may be found in our Sec. 18.3. We comment below on all except the trivial uniform distribution.

16.3.1. ***Binomial distribution***: A random process with exactly two possible outcomes is called a *Bernoulli* process. If the probability of obtaining a certain outcome (a "success") in each trial is p, then the probability of obtaining exactly r successes ($r = 0, 1, 2, \ldots, n$) in n trials, without regard to the order of the successes and failures, is given by the binomial distribution $f(r; n, p)$ in Table 16.1. If r successes are observed in n_r trials with probability p of a success, and if s successes are observed in n_s similar trials, then $t = r + s$ is also binomial with $n_t = n_r + n_s$.

16.3.2. ***Poisson distribution***: The Poisson distribution $f(r; \mu)$ gives the probability of finding exactly r events in a given interval of x (*e.g.*, space and time) when the events occur independently of one another and of x at an average rate of μ per the given interval. The variance σ^2 equals μ. It is the limiting case $p \to 0$, $n \to \infty$, $np = \mu$ of the binomial distribution. The Poisson distribution approaches the Gaussian distribution for large μ.

Two or more Poisson processes (*e.g.*, *signal* + *background*, with parameters μ_s and μ_b) that independently contribute amounts n_s and n_b to a given measurement will produce an observed number $n = n_s + n_b$, which is distributed according to a new Poisson distribution with parameter $\mu = \mu_s + \mu_b$.

16.3.3. ***Normal or Gaussian distribution***: The normal (or Gaussian) probability density function $f(x; \mu, \sigma^2)$ given in Table 16.1 has mean $\overline{x} = \mu$ and variance σ^2. Comparison of the characteristic function $\phi(u)$ given in Table 16.1 with Eq. (16.17) shows that all semi-invariants κ_n beyond κ_2 vanish; this is a unique property of the Gaussian distribution. Some properties of the distribution are:

- rms deviation $= \sigma$
- probability x in the range $\mu \pm \sigma = 0.6827$
- probability x in the range $\mu \pm 0.6745\sigma = 0.5$
- expection value of $|x - \mu|$, $E(|x - \mu|) = (2/\pi)^{1/2}\sigma = 0.7979\sigma$
- half-width at half maximum $= (2\ln 2)^{1/2}\sigma = 1.177\sigma$

The cumulative distribution, Eq. (16.1), for a Gaussian with $\mu = 0$ and $\sigma^2 = 1$ is related to the error function $\mathrm{erf}(y)$ by

$$F(x; 0, 1) = \tfrac{1}{2}\left[1 + \mathrm{erf}(x/\sqrt{2})\right] \ . \tag{16.19}$$

The error function is tabulated in Ref. 6 and is available in computer math libraries and personel computer spreadsheets. For a mean μ and variance σ^2, replace x by $(x - \mu)/\sigma$. The probability of x in a given range can be calculated with Eq. (17.29).

For x and y independent and normally distributed, $z = ax + by$ obeys $f(z;\, a\mu_x + b\mu_y,\, a^2\sigma_x^2 + b^2\sigma_y^2)$; that is, the weighted means and variances add.

The Gaussian gets its importance in large part from the *central limit theorem*: If a continuous random variable x is distributed according to *any* p.d.f. with finite mean and variance, then the sample mean, $\overline{x}_n$, of n observations of x will have a p.d.f. that approaches a Gaussian as n increases. Therefore the end result $\sum^n x_i \equiv n\overline{x}_n$ of a large number of small fluctuations x_i will be distributed as a Gaussian, even if the x_i themselves are not.

For $\boldsymbol{x}$ a set of n (not necessarily independent) Gaussian random variables x_i arranged into a column vector, the joint p.d.f. is the *multivariate Gaussian*:

$$f(\boldsymbol{x}; \boldsymbol{\mu}, V) = \frac{1}{(2\pi)^{n/2}}\ |V|^{-1/2} \tag{16.20}$$

$$\times\ \exp\left[-\tfrac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T\, V^{-1}\, (\boldsymbol{x} - \boldsymbol{\mu})\right] \ , \ |V| \neq 0 \ .$$

Here V is the *covariance matrix* of the x's, with $V_{ii} = \mathrm{Var}(x_i)$ and $V_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)] \equiv \rho_{ij}\, \sigma_i\, \sigma_j$, and $|V|$ is the determinant of V. (If V is singular, there is a linear relation among some of the variables, and so one should eliminate dependent variables and work with an independent set.) The quantity ρ_{ij} is the correlation coefficient for x_i and x_j, and $|\rho_{ij}|^2 \leq 1$. For $n = 2$, $f(\boldsymbol{x}; \boldsymbol{\mu}, V)$ is

$$f(x_1, x_2;\ \mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1 - \rho^2}}$$

$$\times\ \exp\left\{\frac{-1}{2(1 - \rho^2)}\left[\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2}\right.\right.$$

$$\left.\left. + \frac{(x_2 - \mu_2)^2}{\sigma_2^2}\right]\right\} \ . \tag{16.21}$$