

AD 714792

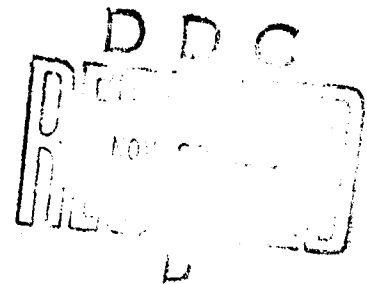
NUSC Report No. NL-3012

**Alternate Forms and Computational Considerations for
Numerical Evaluation of Cumulative Probability
Distributions Directly from Characteristic Functions**

ALBERT H. NUTTALL
*Office of the Associate Technical Director
for Research*



12 August 1970



NAVAL UNDERWATER SYSTEMS CENTER
New London Laboratory

This document has been approved for public release and sale;
its distribution is unlimited.

Reproduced by
**NATIONAL TECHNICAL
INFORMATION SERVICE**
Springfield, Va. 22151

89

ACCESSION FOR		
OFSTI	WHITE SECTION	☒
DC	BLUE SECTION	☐
GR	RED	☐
DATE		
BY		
DESCRIPTION FOR AVAILABILITY CALL		
DIST.	AVAIL. NO. OF SPECIAL	
/		

REVIEWED AND APPROVED: 12 August 1970

W. A. Von Winkle

W. A. Von Winkle
Associate Technical Director
for Research
New London Laboratory

Correspondence concerning this report should be addressed as follows:

Officer in Charge
New London Laboratory
Naval Underwater Systems Center
New London, Connecticut 06320

ABSTRACT

Alternate integral forms for the cumulative probability distribution in terms of the characteristic function are given. In particular, forms that can utilize a fast Fourier transform (FFT) algorithm and special forms for one-sided probability density functions are derived. For a special class of discrete random variables, all integral evaluations are over a finite range. Some computational aspects of utilizing the FFT are discussed.

ADMINISTRATIVE INFORMATION

This study was performed under New London Laboratory Project No. A-041-00-00, "Statistical Communication Through Applications" (U), Principal Investigator, Dr. A. H. Nuttall, Code 2021, and Navy Subproject and Task No. AFXX 112 001, Program Manager, J. H. Huth, NAVSHIPS DPL/MAT 033.

The Technical Reviewer for this report was Dr. E. S. Eby, Research Associate in the Signal Processing Branch of the Information Processing Division.

TABLE OF CONTENTS

	Page
ABSTRACT	i
ADMINISTRATIVE INFORMATION	i
Section	
1. INTRODUCTION	1
2. ANALYSIS	1
2.1 General Distributions	1
2.2 Nonnegative Distributions	4
2.3 Discrete Distributions	8
2.4 Nonnegative Discrete Distributions	10
2.5 Use of FFT for Fourier Transforms	12
3. CONCLUSIONS	16
LIST OF REFERENCES	17
APPENDIX A — BEHAVIOR OF INTEGRAND OF EQ. (3) AT ORIGIN . .	19
APPENDIX B — EVALUATION OF $\sum_{n=1}^{\infty} \exp(\ln x)$	21
APPENDIX C — ALIASED SPECTRUM	25
INITIAL DISTRIBUTION LIST	Inside Back Cover

BLANK PAGE

ALTERNATE FORMS AND COMPUTATIONAL CONSIDERATIONS FOR NUMERICAL EVALUATION OF CUMULATIVE PROBABILITY DISTRIBUTIONS DIRECTLY FROM CHARACTERISTIC FUNCTIONS

1. INTRODUCTION

A recent report [1] on numerical evaluation of cumulative probability distribution functions directly from characteristic functions (CF) gave the cumulative distribution functions (CDF) in terms of a single integral on the CF for both continuous and discrete random variables (RV). In this report some alternate forms for the CDF in terms of the CF will be presented, with an aim toward more accurate, efficient, and expeditious calculations. For the motivation of this study and utility of the results, as well as numerical examples, see Reference 1.

2. ANALYSIS

This section is composed of five subsections. In the first, general distributions are considered; in the second, specialization to a nonnegative random variable is made. In both subsections, forms that utilize a fast Fourier transform (FFT) are derived and their applicability is discussed. In the third and fourth subsections, discrete random variables are considered. The former subsection shows that the distribution function can be evaluated entirely in terms of finite integrals; the latter subsection specializes to nonnegative discrete random variables. The fifth subsection treats some computational aspects of the FFT.

2.1 GENERAL DISTRIBUTIONS

Let RV x have probability density function (PDF) $p(x)$ and CF $f(\xi)$:

$$f(\xi) = \int dx \exp(i\xi x) p(x) , \quad (1)$$

$$p(x) = \frac{1}{2\pi} \int d\xi \exp(-i\xi x) f(\xi) . \quad (2)$$

(Integrals without limits are over the real axis from $-\infty$ to ∞ .) The CDF $\Pr(X)$ is defined as the probability that RV x is less than or equal to X . The modified distribution function (MDF) $P(X)$ is defined equal to $\Pr(X)$ at points of continuity, but it takes a value midway between limit values on either side of a discontinuity.

The MDF $P(X)$ can be obtained from the CF by [1, Eq. (7), or 2, Eq. (4.14)]

$$P(X) = \frac{1}{2} - \frac{1}{\pi} \int_0^{\infty} \frac{d\xi}{\xi} \operatorname{Im} \left\{ f(\xi) \exp(-i\xi X) \right\}, \quad \text{all } X. \quad (3)$$

If we attempt to remove the Imaginary operation from under the integral sign, we obtain an infinite integral since $f(0) = 1$. However, if we express

$$f(\xi) = [f(\xi) - a(\xi)] + a(\xi), \quad (4)$$

where $a(0) = 1$, and split (3) into two integrals, we can move the Imaginary operation out of the first integral in (3). One particularly useful choice for $a(\xi)$, which results in a closed form expression for the second integral in (3), is

$$a_1(\xi) = \exp\left[i\mu\xi - \frac{1}{2}\sigma^2\xi^2\right], \quad \xi > 0, \quad (5)$$

where* μ and σ^2 are the mean and variance of RV x . The mean and variance are available from $f'(0)$ and $f''(0)$, if these quantities can be evaluated; if not, the method to be described is still applicable with arbitrary constants used for μ and σ^2 . When (4) and (5) are substituted into (3), there results [3, Eq. 3.896 4; integrate both sides with respect to b]

$$P(X) = \Phi\left(\frac{X - \mu}{\sigma}\right) - \frac{1}{\pi} \operatorname{Im} \left\{ \int_0^{\infty} d\xi \frac{f(\xi) - a_1(\xi)}{\xi} \exp(-i\xi X) \right\}, \quad \text{all } X. \quad (6)$$

*Actually, μ and σ^2 could be assigned arbitrary values in the form (5); this particular choice gives a second-order fit to $f(\xi)$ at the origin.

where

$$\Phi(y) \equiv \int_{-\infty}^y dt (2\pi)^{-1/2} \exp(-t^2/2) \quad (7)$$

is the Gaussian CDF.

Equation (6) is now in a form where an FFT can be utilized in the integration on ξ (See Subsection 2.5). This equation is exact; we are not making a Gaussian approximation in (6). There is no problem in the integration at $\xi = 0$ because, for the choice of μ and σ^2 as the mean and variance of RV x ,

$$\frac{f(\xi) - a_1(\xi)}{\xi} = O(\xi^2) \text{ as } \xi \rightarrow 0+. \quad (8)$$

Also, since $|a_1(\xi)|$ decays as $\exp(-\sigma^2 \xi^2/2)$, the decay of the left side of (8) for large ξ will often depend on the decay of $f(\xi)/\xi$; this decay will dictate how far the integral in (6) must be carried out for specified accuracy in $P(X)$.

Other choices for $a(\xi)$ are possible and sometimes recommended. For example, if the mean and variance of RV x do not exist (e.g., $p(x) = x^{-1} (1+x^2)^{-1}$, all x), we might choose

$$a_2(\xi) = \exp(-b\xi), \quad \xi > 0. \quad (9)$$

To best match $f(\xi)$ near the origin, we could choose

$$b = -f'(0+) = |f'(0+)| \text{ (assuming } f'(0+) \text{ real)}. \quad (10)$$

Then by substituting (4) and (9) into (3) [3, Eq. 3.941 1], we get

$$P(X) = \frac{1}{2} + \frac{1}{\pi} \arctan(X/b) - \frac{1}{\pi} \operatorname{Im} \left\{ \int_0^\infty d\xi \frac{f(\xi) - a_2(\xi)}{\xi} \exp(-i\xi X) \right\}, \text{ all } X. \quad (11)$$

For the choice of b in (10),

$$\frac{f(\xi) - a_2(\xi)}{\xi} = O(\xi) \text{ as } \xi \rightarrow 0+, \quad (12)$$

so no problem in integration arises at the origin. We must be able to evaluate $f'(0+)$ in this case so that b is known, and it must be real. In cases where $f'(0+)$ is not known or is infinite (See Appendix A, for example), the above methods are inapplicable, and special techniques such as subtracting out the singularity are required.

2.2 NONNEGATIVE DISTRIBUTIONS

When RV x is limited to nonnegative values, some simplifications in the general form (3) occur. (The case of nonpositive RV x can be treated in a similar fashion.) First, if $X < 0$ in (3), then $P(X) = 0$. Letting $X = -a$ yields

$$\frac{1}{2} = \frac{1}{\pi} \int_0^{\infty} \frac{d\xi}{\xi} [f_r(\xi) \sin(a\xi) + f_i(\xi) \cos(a\xi)], \quad a > 0, \quad (13)$$

where subscripts r and i denote real and imaginary parts, respectively. Employing (13) in (3) for $X > 0$, we get

$$P(X) = \frac{2}{\pi} \int_0^{\infty} \frac{d\xi}{\xi} f_r(\xi) \sin(\xi X), \quad X > 0, \quad (14)$$

or

$$P(X) = 1 - \frac{2}{\pi} \int_0^{\infty} \frac{d\xi}{\xi} f_i(\xi) \cos(\xi X), \quad X > 0. \quad (15)$$

Thus, the MDF $P(X)$ can be evaluated from knowledge of either the real part or the imaginary part of the CF $f(\xi)$. For $X = 0$, neither (14) nor (15) is necessarily valid, and we must resort to (3).

There are computational reasons for choosing (14) over (15), or vice versa. The first has to do with ease of calculating $f_r(\xi)$ versus $f_i(\xi)$. For example, in Appendix A, for $p(x) = 2/\pi (1+x^2)^{-1}$ for $x > 0$, we find that $f_r(\xi)$ is a simple exponential, whereas $f_i(\xi)$ is a sum of exponential integrals. Converse examples, where $f_i(\xi)$ is simpler to compute, can also be found.

The second reason has to do with the rate of decay of $f_r(\xi)$ versus $f_i(\xi)$. We have

$$f_r(\xi) = \int_0^\infty dx p(x) \cos(\xi x) = \int_0^\infty dx p_e(x) \cos(\xi x) = \int_0^\infty dx p_e(x) \exp(i\xi x), \quad (16)$$

$$f_i(\xi) = \int_0^\infty dx p(x) \sin(\xi x) = \int_0^\infty dx p_o(x) \sin(\xi x) = i^{-1} \int_0^\infty dx p_o(x) \exp(i\xi x), \quad (17)$$

where subscripts e and o denote even and odd parts, respectively. Now, if $p(0+) > 0$, then $p_o(x)$ is discontinuous at the origin, and $f_i(\xi)$ decays only as ξ^{-1} for large ξ . An example is

$$p(x) = e^{-x}, \quad x > 0; \quad f(\xi) = (1 - i\xi)^{-1},$$

$$f_r(\xi) = (1 + \xi^2)^{-1}, \quad f_i(\xi) = \xi(1 + \xi^2)^{-1}. \quad (18)$$

In (18), $f_r(\xi)$ decays as ξ^{-2} for large ξ , giving rise to an integral in (14) that can be terminated earlier than the one in (15). On the other hand, consider that $p(0+) = 0$ and that $p(x)$ and its derivative are continuous except at the origin, but $p'(0+) > 0$. Then $p_o(x)$ and its derivative are continuous, whereas $p_e(x)$ is discontinuous. In this case, $f_r(\xi)$ decays only as ξ^{-2} for large ξ . An example is

$$p(x) = xe^{-x}, \quad x > 0; \quad f(\xi) = (1 - i\xi)^{-2},$$

$$f_r(\xi) = (1 - \xi^2) (1 + \xi^2)^{-2}, \quad f_i(\xi) = 2\xi (1 + \xi^2)^{-2}. \quad (19)$$

Here (15) could be terminated earlier than (14).

The third reason has to do with the region of X of interest. For large X , where $P(X)$ is near unity, Eq. (15), in the form

$$1 - P(X) = \frac{2}{\pi} \int_0^{\infty} \frac{d\xi}{\xi} f_1(\xi) \cos(\xi X), \quad X > 0, \quad (20)$$

is to be recommended, since it is an alternating sum of small quantities and retains significance. Equation (14), for large X , is an alternating sum of large quantities and loses significance. But for small X , Eq. (14) would be recommended.

Equation (15) can be immediately manipulated into a form where an FFT can be utilized. Namely,

$$P(X) = 1 - \frac{2}{\pi} \operatorname{Re} \left\{ \int_0^{\infty} d\xi \frac{f_1(\xi)}{\xi} \exp(-i\xi X) \right\}, \quad X > 0. \quad (21)$$

From Appendix A, we have $f_1(\xi)/\xi \sim \mu_X$ as $\xi \rightarrow 0$, if μ_X exists and is finite.

If we attempt to express (14) in the form

$$\operatorname{Im} \left\{ \frac{2}{\pi} \int_0^{\infty} \frac{d\xi}{\xi} f_r(\xi) \exp(i\xi X) \right\},$$

we obtain an integral that does not converge at the origin. However, if we express

$$f_r(\xi) = [f_r(\xi) - b(\xi)] + b(\xi), \quad (22)$$

where $b(0) = 1$ and $b(\xi)$ is real, then (14) becomes

$$P(X) = \frac{2}{\pi} \operatorname{Im} \left\{ \int_0^{\infty} d\xi \frac{f_r(\xi) - b(\xi)}{\xi} \exp(i\xi X) \right\} + \frac{2}{\pi} \int_0^{\infty} d\xi b(\xi) \frac{\sin(\xi X)}{\xi}, \quad X > 0, \quad (23)$$

and an FFT can be used on the first integral. Preferably, the second integral should be integrable in closed form. A particularly useful choice is

$$b(\xi) = \exp\left(-\frac{1}{2} \mu_2 \xi^2\right), \quad \xi > 0, \quad (24)$$

where* μ_2 is the mean-square value of RV x . This quantity is available from $f_r''(0)$ if it can be evaluated. When (24) is substituted into (23), we get (See (3) through (7))

$$P(X) = 2\Phi\left(\frac{X}{\mu_2^{1/2}}\right) - 1 + \frac{2}{\pi} \operatorname{Im} \left\{ \int_0^{\infty} d\xi \frac{f_r(\xi) - \exp\left(-\frac{1}{2} \mu_2 \xi^2\right)}{\xi} \exp(i\xi X) \right\}, \quad X > 0. \quad (25)$$

The function

$$\frac{f_r(\xi) - \exp\left(-\frac{1}{2} \mu_2 \xi^2\right)}{\xi} \rightarrow 0 \text{ as } \xi \rightarrow 0+$$

in (25) if the mean-square value μ_2 exists. In many cases, it decays as $f_r(\xi)/\xi$ for large ξ .

The fact that MDF $P(X)$ can be obtained from either the real or imaginary parts of the CF for a nonnegative distribution are manifestations of the fact that $f_r(\xi)$ and $f_i(\xi)$ can be found from each other; in fact, they are related by Hilbert transforms. For $p(x) = 0$ for $x < 0$, and no impulses at the origin, we see that [4, p. 38]

*As in the footnote to Eq. (5), μ_2 could be assigned any convenient value.

$$\begin{aligned}
f(\xi) &= \int dx \, p(x) \exp(i\xi x) = \int dx \, p(x) U(x) \exp(i\xi x) = \mathfrak{F}\{p(x) U(x)\} \\
&= \mathfrak{F}\{p(x)\} \bullet \mathfrak{F}\{U(x)\} = f(\xi) \bullet \left[\frac{1}{2} \delta(\xi) + \frac{1}{2\pi i \xi} \right] = \frac{1}{2} [f(\xi) + i\mathcal{H}\{f(\xi)\}] ,
\end{aligned} \tag{26}$$

where $U(x)$ is the unit step function, $\mathfrak{F}$ denotes a Fourier transform, $\bullet$ denotes convolution, and $\mathcal{H}$ denotes a Hilbert transform. Therefore,

$$f(\xi) = i\mathcal{H}\{f(\xi)\} , \tag{27}$$

or

$$f_1(\xi) = \mathcal{H}\{f_r(\xi)\} , \quad f_r(\xi) = -\mathcal{H}\{f_1(\xi)\} . \tag{28}$$

For the cases when $p(x)$ contains an impulse at the origin of area c_0 , the first part of (28) is still correct, but the second part is incorrect by the additive constant c_0 . However, we can still find $f_r(\xi)$ from $f_1(\xi)$ by utilizing the fact that $f_r(0) = 1$. Thus, either the real or the imaginary part of the CF constitutes complete knowledge about the MDF in the case of a nonnegative distribution.

2.3 DISCRETE DISTRIBUTIONS

In this subsection, the RV x is restricted to take on values that are multiples of some fundamental increment Δ , and can be either positive or negative. Although the equations in Subsection 2.2 are applicable here, it is advantageous to have forms for the distribution function that require finite integrals rather than infinite ones. We have for the PDF

$$p(x) = \sum_k c_k \delta(x - k\Delta) , \tag{29}$$

where the sum ranges from $-\infty$ to ∞ . The CDF $\Pr(M)$ for integer M is given in Reference 1, Eqs. (20) through (26). All the integrals are finite integrals except for the one in (2c) for MDF value $P(0)$:

$$P(0) = \frac{1}{2} - \frac{1}{\pi} \int_0^{\infty} d\xi \frac{f_1(\xi)}{\xi}.$$

We now rectify this situation and obtain a finite integral for $P(0)$ also. From Reference 1, Eq. (15),

$$c_k = \frac{\Delta}{2\pi} \int_{-\pi/\Delta}^{\pi/\Delta} d\xi f(\xi) \exp(-ik\Delta\xi). \quad (30)$$

Therefore, by using Appendix B and $f(-\xi) = f^*(\xi)$, we get

$$\begin{aligned} \sum_{k=-\infty}^{-1} c_k &= \frac{\Delta}{2\pi} \int_{-\pi/\Delta}^{\pi/\Delta} d\xi f(\xi) \sum_{k=-\infty}^{-1} \exp(ik\Delta\xi) \\ &= \frac{\Delta}{2\pi} \int_{-\pi/\Delta}^{\pi/\Delta} d\xi f(\xi) \left[-\frac{1}{2} + \pi \sum_k \delta(\Delta\xi - k2\pi) + i \frac{1}{2} \cot\left(\frac{\Delta\xi}{2}\right) \right] \\ &= -\frac{1}{2} c_0 + \frac{1}{2} + i \frac{1}{2} \frac{\Delta}{2\pi} \int_{-\pi/\Delta}^{\pi/\Delta} d\xi f(\xi) \cot\left(\frac{\Delta\xi}{2}\right) \\ &= -\frac{1}{2} c_0 + \frac{1}{2} - \frac{\Delta}{2\pi} \int_0^{\pi/\Delta} d\xi \frac{f_1(\xi)}{\tan(\Delta\xi/2)}. \end{aligned} \quad (31)$$

Then, we obtain the desired result

$$P(0) = \sum_{k=-\infty}^{-1} c_k + \frac{1}{2} c_0 = \frac{1}{2} - \frac{\Delta}{2\pi} \int_0^{\pi/\Delta} d\xi \frac{f_1(\xi)}{\tan(\Delta\xi/2)}. \quad (32)$$

As $\xi \rightarrow 0+$, the integrand of (32) approaches $2\mu_x/\Delta$ if μ_x exists and is finite. (There is no integral expression for $P(0)$ in terms of $f_r(\xi)$. Since $f_r(\xi)$ is the Fourier transform of $p_e(x)$ (See Eq. (16)), and since

$$\int_{-\infty}^0 dx p_e(x) = \frac{1}{2},$$

irrespective of the form of $p(x)$, $f_r(\xi)$ contains no information about $P(0)$. This is analogous to the general distribution case where $P(0)$ follows from (3) as

$$P(0) = \frac{1}{2} - \frac{1}{\pi} \int_0^{\infty} d\xi f_1(\xi)/\xi.$$

2.4 NONNEGATIVE DISCRETE DISTRIBUTIONS

When RV x is limited to nonnegative values, the CDF $\Pr(M)$ takes on forms requiring either the real or imaginary parts of the CF for its evaluation, just as in Subsection 2.2. To see this, we note that c_k in (28) is zero for $k < 0$. By letting $k = -m$ in (30), we get

$$\int_0^{\pi/\Delta} d\xi \sin(m\Delta\xi) f_1(\xi) = \int_0^{\pi/\Delta} d\xi \cos(m\Delta\xi) f_r(\xi) \text{ for } m > 0. \quad (33)$$

When we employ (33) into (30) for $k > 0$, we get

$$c_k = \frac{2\Delta}{\pi} \int_0^{\pi/\Delta} d\xi \cos(k\Delta\xi) f_r(\xi), \quad k > 0, \quad (34)$$

or

$$c_k = \frac{2\Delta}{\pi} \int_0^{\pi/\Delta} d\xi \sin(k\Delta\xi) f_1(\xi), \quad k > 0. \quad (35)$$

Therefore, the CDF $\text{Pr}(M)$ for integer M is given by

$$\begin{aligned} \text{Pr}(M) &= \sum_{k=0}^M c_k = \frac{2\Delta}{\pi} \int_0^{\pi/\Delta} d\xi f_r(\xi) \left[\frac{1}{2} + \sum_{k=1}^M \cos(k\Delta\xi) \right] \\ &= \frac{\Delta}{\pi} \int_0^{\pi/\Delta} d\xi f_r(\xi) \frac{\sin\left[\left(M + \frac{1}{2}\right)\Delta\xi\right]}{\sin\left[\frac{1}{2}\Delta\xi\right]}, \quad M \geq 0, \end{aligned} \quad (36)$$

where we have used (34), (30), the fact that $f(-\xi) = f^*(\xi)$, and Eq. 1.342.2 in Reference 3. Equation (36) enables evaluating the CDF in terms of the real part of the CF alone.

To represent $\text{Pr}(M)$ in terms of $f_1(\xi)$, we first note that for nonnegative RV, the general formula for $P(0)$ in (32) becomes

$$P(0) = \frac{1}{2} c_0 = \frac{1}{2} - \frac{\Delta}{2\pi} \int_0^{\pi/\Delta} d\xi \frac{f_1(\xi)}{\tan(\Delta\xi/2)}. \quad (37)$$

Now,

$$\text{Pr}(M) = \sum_{k=0}^M c_k = c_0 + \frac{2\Delta}{\pi} \int_0^{\pi/\Delta} d\xi f_1(\xi) \sum_{k=1}^M \sin(k\Delta\xi)$$

$$= 1 - \frac{\Delta}{\pi} \int_0^{\pi/\Delta} d\xi f_1(\xi) \frac{\cos \left[\left(M + \frac{1}{2} \right) \Delta \xi \right]}{\sin \left[\frac{1}{2} \Delta \xi \right]}, \quad M \geq 0, \quad (38)$$

where we have employed (35), (37), and Eq. 1.342.1 in Reference 3. Equation (38) is complementary to (36) in the sense that only the imaginary part of the CF is necessary for evaluating the CDF. The reasons given in Subsection 2.2 for selecting (36) or (38) in a particular application are again relevant.

2.5 USE OF FFT FOR FOURIER TRANSFORMS

Many of the integrals in this report take the form

$$\int_0^{\infty} dt g(t) \exp(-i2\pi ft) .$$

Suppose a limit T on the integration can be found such that

$$\left| \int_T^{\infty} dt g(t) \exp(-i2\pi ft) \right| < \epsilon \text{ for all } f, \quad (39)$$

where ϵ is some specified tolerance or error. Then, attention can be focused on evaluating

$$G_T(f) \equiv \int_0^T dt g(t) \exp(-i2\pi ft) . \quad (40)$$

Since the integration in (40) is over an interval of length T , it is seen that $G_T(f)$ will undergo a significant change in value in an interval no smaller than $1/T$ in f . Thus, one might initially anticipate that (40) should be evaluated at values of $f = n/T$, $n = 1, 2, \dots$. However, in many cases, this resolution, $1/T$, may be much too fine, and be the result of satisfying (39) with a very small ϵ . In such cases, values of $G_T(f)$ at some multiple of the fundamental

resolution may be satisfactory, say m/T , where m is an integer. Thus, we might be interested only in evaluating $G_T\left(n\frac{m}{T}\right)$, $n = 1, 2, \dots$. But from (40)

$$\begin{aligned} G_T\left(n\frac{m}{T}\right) &= \int_0^T dt \, g(t) \exp(-i2\pi nmt/T) \\ &= \sum_{k=0}^{m-1} \int_{kT/m}^{(k+1)T/m} dt \, g(t) \exp(-i2\pi nmt/T). \end{aligned} \quad (41)$$

In making the substitution $u = t - kT/m$ in (41) and defining the collapsed function

$$g_0(u) = \begin{cases} \sum_{k=0}^{m-1} g\left(u + k\frac{T}{m}\right), & 0 \leq u \leq T/m \\ 0, & \text{otherwise} \end{cases}, \quad (42)$$

we note that (41) becomes

$$G_T\left(n\frac{m}{T}\right) = \int_0^{T/m} du \, g_0(u) \exp\left(-i2\pi\frac{n}{T/m}u\right). \quad (43)$$

The collapsed function g_0 is obtained from g by "pre-aliasing" g into the interval T/m . If we define the Fourier transform of g_0 as

$$G_0(f) = \int_0^{T/m} du \, g_0(u) \exp(-i2\pi fu), \quad (44)$$

then (43) yields

$$G_T\left(n \frac{m}{T}\right) = G_c\left(n \frac{m}{T}\right). \quad (45)$$

That is, $G_T(f)$ can be evaluated at $f = m/T, 2m/T, \dots$, in terms of the Fourier transform of the collapsed function.

Now suppose $g_c(u)$ is sampled at increments of h in (44), where $h = \frac{T/m}{M}$, and weighting $\{w_k\}$ applied to the samples in an effort to approximate* (44). That is,

$$G_c(f) \approx h \sum_{k=0}^M w_k g_c(kh) \exp(-i2\pi fkh) = \hat{G}_c(f). \quad (46)$$

The approximation in (46) will be good if g_c and the exponential are sampled frequently enough. Thus, if the exponential is not to vary by more than a radian between samples, we require

$$|f| < \frac{1}{2\pi h}. \quad (47)$$

When (45) through (47) are combined, the desired values are given by

$$G_T\left(n \frac{m}{T}\right) = G_c\left(n \frac{m}{T}\right) \approx \hat{G}_c\left(n \frac{m}{T}\right) = h \sum_{k=0}^M w_k g_c(kh) \exp(-i2\pi kn/M) \quad (48)$$

if

$$\left|n\right| \frac{m}{T} < \frac{1}{2\pi h}, \text{ or } |n| < \frac{1}{\pi} \frac{M}{2}. \quad (49)$$

By defining

*For example, Simpson's rule has $w_0 = 1/3$, $w_{2n+1} = 4/3$, $w_{2n+2} = 2/3$, $w_M = 1/3$.

$$d_k = \begin{cases} hw_k g_c(kh), & 1 \leq k \leq M-1 \\ hw_0 g_c(0) + hw_M g_c(T/m), & k = 0 \end{cases}, \quad (50)$$

we can express (48) as

$$G_T\left(n \frac{m}{T}\right) = G_c\left(n \frac{m}{T}\right) \cong \hat{G}_c\left(n \frac{m}{T}\right) = \sum_{k=0}^{M-1} d_k \exp(-j2\pi kn/M), \quad (51)$$

which is an M -point discrete Fourier transform (DFT) of the sequence $\{d_k\}$. The factor $1/r$ in the upper bound on $|n|$ in (49) is due to the aliasing in the frequency domain that takes place at $|n| = M/2$. In fact, letting $\{w_k\}$ be the samples of waveform $w(t)$ at $t = kh$ and $W(f)$ its Fourier transform, it can be shown that (Appendix C)

$$\hat{G}_c(f) = W(f) \bullet \sum_k G_c\left(f - kM \frac{m}{T}\right). \quad (52)$$

Thus, the value $\hat{G}_c\left(\frac{M}{2} \frac{m}{T}\right)$ is composed of at least two overlapping tails of $G_c(f)$. In order to avoid this aliasing, we must observe (49).

To summarize, the values of $G_T(f)$ at $f = n \frac{m}{T}$ are given approximately as an M -point DFT in (51) of the sequence $\{d_k\}$ in (50). When (42) is substituted into (50), this sequence can be expressed as

$$d_k = \begin{cases} hw_k \sum_{j=0}^{m-1} g\left(kh + j \frac{T}{m}\right), & 1 \leq k \leq M-1 \\ hw_0 \sum_{j=0}^{m-1} g\left(j \frac{T}{m}\right) + hw_M \sum_{j=1}^m g\left(j \frac{T}{m}\right), & k = 0 \end{cases}. \quad (53)$$

As is obvious from (53), $g(t)$ must still be evaluated from 0 to T in increments of h , that is, at $mM + 1$ values. However, collapsing reduces the

size of the FFT from mM to M , with an attendant reduction in computation time and round-off error. This method is related to one given in Reference 5, p. 81.

In applying this technique to numerical integration of CF's, since the exponentials take the form $\exp(i\xi X)$, we note that the increment in X at which values are obtained by employing an FFT are $2\pi/T$, or $2\pi m/T$ for coarser resolution as above.

3. CONCLUSIONS

Several alternate forms for direct numerical evaluation of the CDF or MDF from the CF have been presented that have utility in different situations, including ease of calculation, rate of decay of the integrands, and the probability region of interest. Also, the speed of the FFT and the large number of values of the distribution functions that are quickly available make the formulas presented attractive in a large number of practical applications.

In the case of discrete distributions with RV that can take on positive as well as negative values, all integrals for the CDF are finite and over a half-period of the CF. Reevaluation of the sines or cosines, as in (36) or (38) for different values of M , can be avoided if one notes that

$$\sin\left[\left(M + \frac{1}{2}\right)a\right] = \sin\left[\left(M - \frac{1}{2}\right)a + a\right] = \sin\left[\left(M - \frac{1}{2}\right)a\right] \cos[a] + \cos\left[\left(M - \frac{1}{2}\right)a\right] \sin[a],$$

with a similar result for cosine. Thus, if a table of $\sin(a)$ and $\cos(a)$ for the values of a ($\Delta\xi$) is constructed, this recurrence relation can be used to obtain the higher order M -dependence required in (36) and (38) without reevaluating sines and cosines.

LIST OF REFERENCES

1. A. H. Nuttall, Numerical Evaluation of Cumulative Probability Distribution Functions Directly from Characteristic Functions, NUSL Report No. 1032, 11 August 1969. For an abbreviated version of this report, see IEEE Proceedings (Letters), vol. 57, no. 11, November 1969, pp. 2071-2072.
2. M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, vol. 1, Hafner Publishing Company, New York, 1958.
3. I. S. Gradshteyn and I. M. Ryzhik, Tables of Integrals, Series, and Products, Academic Press, New York, 1965.
4. A. Papoulis, The Fourier Integral and Its Applications, McGraw-Hill Book Co., New York, 1962.
5. J. W. Cooley, P. A. W. Lewis, and P. D. Welch, "Application of the Fast Fourier Transform to Computation of Fourier Integrals, Fourier Series, and Convolution Integrals," IEEE Transactions on Audio and Electroacoustics, vol. AU-15, no. 2, June 1967, pp. 79-84.

*Reports prepared by the New London Laboratory prior to 1 July 1970 bear the Laboratory's earlier acronym NUSL.

BLANK PAGE

Appendix A

BEHAVIOR OF INTEGRAND OF EQ. (3) AT ORIGIN

The integrand of (3) is given by

$$\frac{1}{\xi} \operatorname{Im} \{ f(\xi) \exp(-i\xi X) \} = \frac{f_i(\xi) \cos(\xi X)}{\xi} - \frac{f_r(\xi) \sin(\xi X)}{\xi},$$

where subscripts r and i denote real and imaginary parts, respectively. Now,

$$\frac{f_r(\xi) \sin(\xi X)}{\xi} \rightarrow X \text{ as } \xi \rightarrow 0+.$$

And

$$\frac{f_i(\xi) \cos(\xi X)}{\xi} = \int dx \frac{\sin(\xi x)}{\xi} p(x) \cos(\xi X).$$

$$\sim \int dx x p(x) \equiv \mu_x \text{ as } \xi \rightarrow 0+ \text{ if } \mu_x \text{ exists and is finite.}$$

Here μ_x is the mean of RV x . Therefore, the integrand of (3) approaches

$$\mu_x - X \text{ as } \xi \rightarrow 0+ \text{ if } \mu_x \text{ exists and is finite.}$$

An example where μ_x is infinite is given by

$$p(x) = \begin{cases} 0, & x < 0 \\ \frac{2/\pi}{1+x^2}, & x > 0 \end{cases}.$$

Then,^{A1}

$$f_r(\xi) = \exp(-|\xi|),$$

$$f_i(\xi) = \operatorname{sgn}(\xi) \frac{1}{\pi} \{ \exp(-|\xi|) \operatorname{Ei}(|\xi|) - \exp(|\xi|) \operatorname{Ei}(-|\xi|) \}.$$

Since^{A2}

$$\operatorname{Ei}(-|\xi|) = \ln |\xi| + C - |\xi| + \frac{1}{4} |\xi|^2 + O(|\xi|^3) \text{ as } |\xi| \rightarrow 0,$$

$$\operatorname{Ei}(|\xi|) = \ln |\xi| + C + |\xi| + \frac{1}{4} |\xi|^2 + O(|\xi|^3) \text{ as } |\xi| \rightarrow 0,$$

there follows

$$\begin{aligned} f_i(\xi) &= \operatorname{sgn}(\xi) \left\{ \frac{2}{\pi} |\xi| (-\ln |\xi| + 1 - C + \frac{1}{4} |\xi|^2) + O(|\xi|^3) \right\} \\ &\sim \frac{2}{\pi} \xi \ln \left(\frac{1}{|\xi|} \right) \text{ as } |\xi| \rightarrow 0. \end{aligned}$$

Therefore,

$$\frac{f_i(\xi) \cos(\xi X)}{\xi} \sim \frac{2}{\pi} \ln \left(\frac{1}{|\xi|} \right) \text{ as } |\xi| \rightarrow 0,$$

which is unbounded, but integrable. So, in those cases where μ_x is infinite, the behavior of $f_i(\xi)/\xi$ at the origin must be handled carefully in order to accurately evaluate the integral. One possibility is to subtract out the singularity and integrate it analytically.

^{A1}I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series and Products, Academic Press, New York, 1965, Eq. (3.723 1).

^{A2}Ibid., Eqs. (8.214 1) and (8.214 2).

Appendix B

EVALUATION OF $\sum_{n=1}^{\infty} \exp(inx)$

Consider the ordinary function

$$f(x) = \ln \left| \sin \frac{x}{2} \right|, \quad x \neq 0, \pm 2\pi, \pm 4\pi, \dots$$

Since $(1+x^2)^{-1} f(x)$ is absolutely integrable from $-\infty$ to ∞ , the generalized function $f(x)$ corresponding to ordinary function $f(x)$ can be defined.^{B1} In fact, the generalized function $f(x)$ equals the ordinary function $f(x)$ (See definition 8 by Lighthill^{B2}). Furthermore, the generalized function $f(x)$ is periodic, with period 2π ,^{B3} and, therefore, can be expressed as^{B4}

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}.$$

The generalized function $f(x)$ is absolutely integrable over a period, since the ordinary function $f(x)$ is absolutely integrable over a period.^{B5} Therefore, the coefficients $\{c_n\}$ in the expansion of the generalized function $f(x)$ are given by^{B6}

$$\begin{aligned} c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx f(x) e^{-inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \ln \left| \sin \frac{x}{2} \right| e^{-inx} \\ &= \frac{1}{\pi} \int_0^{\pi} dx \ln \left(\sin \frac{x}{2} \right) \cos(nx) = 2 \int_0^{1/2} dt \ln(\sin \pi t) \cos(2n\pi t) \end{aligned}$$

^{B1}M. J. Lighthill, An Introduction to Fourier Analysis and Generalized Functions, Cambridge University Press, New York, 1959, p. 21, definition 7.

^{B2}Ibid, p. 25.

^{B3}Ibid, p. 60, definition 22.

^{B4}Ibid, p. 66, Theorem 26.

^{B5}Ibid, p. 48, definition 19.

^{B6}Ibid, p. 66, Theorem 26, Note.

$$= \begin{cases} -\ln 2, & n = 0 \\ -\frac{1}{2|n|}, & n \neq 0 \end{cases},$$

where we have used Eq. (4.384 3) by Gradshteyn and Ryzhik.^{B7} Therefore, the generalized function $\ln |\sin(x/2)|$ can be expressed as

$$\ln \left| \sin \frac{x}{2} \right| = -\ln 2 - \frac{1}{2} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{\operatorname{sgn}(n)}{n} e^{inx}.$$

If we define the derivative of the generalized function $\ln |\sin(x/2)|$ as the generalized function $\cot(x/2)/2$, differentiation of the last equation yields the expression of the generalized function $\cot(x/2)/2$ as^{B8}

$$\frac{1}{2} \cot\left(\frac{x}{2}\right) = -i \frac{1}{2} \sum_{n=-\infty}^{\infty} \operatorname{sgn}(n) e^{inx}.$$

(This equation says that the spectrum of the generalized function $\cot(x/2)$ is the odd impulse train.^{B9}) And since^{B10}

$$\pi \sum_{n=-\infty}^{\infty} \delta(x - n2\pi) = \frac{1}{2} \sum_{n=-\infty}^{\infty} e^{inx},$$

we obtain

$$\pi \sum_{n=-\infty}^{\infty} \delta(x - n2\pi) + i \frac{1}{2} \cot\left(\frac{x}{2}\right) = \frac{1}{2} + \sum_{n=1}^{\infty} e^{inx},$$

^{B7}I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series and Products, Academic Press, New York, 1965.

^{B8}Op. cit., M. J. Lighthill, p. 28, Theorem 15.

^{B9}Ibid., p. 66, Theorem 26, Eq. (36).

^{B10}Ibid., p. 67, Example 38.

or

$$\sum_{n=1}^{\infty} e^{inx} = -\frac{1}{2} + \pi \sum_{n=-\infty}^{\infty} \delta(x - n2\pi) + i \frac{1}{2} \cot\left(\frac{x}{2}\right)$$

in the sense of generalized functions.

Appendix C

ALIASED SPECTRUM

If we define the infinite impulse train

$$\delta_h(t) = \sum_n \delta(t - nh)$$

and use the time-limited character of g_c , as given in (42), it is possible to manipulate (46) as follows:

$$\hat{G}_c(f) = h \sum_{k=0}^M w_k g_c(kh) \exp(-i2\pi fkh)$$

$$= \int_0^T dt w(t) g_c(t) h \delta_h(t) \exp(-i2\pi ft)$$

$$= \int dt w(t) g_c(t) h \delta_h(t) \exp(-i2\pi ft)$$

$$= \mathcal{F} \{ w(t) g_c(t) h \delta_h(t) \} = W(f) \bullet G_c(f) \bullet \delta_{1/h}(f)$$

$$= W(f) \bullet \sum_k G_c\left(f - k \frac{1}{h}\right).$$

Using $h = \frac{T/m}{M}$, we see that (52) results.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D		
<i>(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)</i>		
1. ORIGINATING ACTIVITY (Corporate author) New London Laboratory Naval Underwater Systems Center New London, Connecticut 06320		2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
		2b. GROUP
3. REPORT TITLE ALTERNATE FORMS AND COMPUTATIONAL CONSIDERATIONS FOR NUMERICAL EVALUATION OF CUMULATIVE PROBABILITY DISTRIBUTIONS DIRECTLY FROM CHARACTERISTIC FUNCTIONS		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Research Report		
5. AUTHOR(S) (First name, middle initial, last name) Albert H. Nuttall		
6. REPORT DATE	7a. TOTAL NO. OF PAGES 30	7b. NO. OF REFS 17
8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S) NL-3012	
b. PROJECT NO. A-041-00-00		
c. AF XX 112 001	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.		
10. DISTRIBUTION STATEMENT This document has been approved for public release and sale; its distribution is unlimited.		
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY Department of the Navy
13. ABSTRACT Alternate integral forms for the cumulative probability distribution in terms of the characteristic function are given. In particular, forms that can utilize a fast Fourier transform (FFT) algorithm and special forms for one-sided probability density functions are derived. For a special class of discrete random variables, all integral evaluations are over a finite range. Some computational aspects of utilizing the FFT are discussed.		

DD FORM 1473 (PAGE 1)
NOV 61
S/N 0102-014-6600

UNCLASSIFIED

Security Classification

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Cumulative Probability Distributions from Characteristic Functions Discrete Random Variables Fast Fourier Transform (FFT) One-sided Probability Density Functions						

UNCLASSIFIED

Security Classification