King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 103

March 1988

Generalized Fourier Transform and Applications in Probability and Statistics

M.A. Chaudhry and Munir Ahmad

DHAHRAN 31261 ● SAUDI ARABIA ● www.kfupm.edu.sa/math/ ● E-mail: mathdept@kfupm.edu.sa

GENERALIZED FOURIER TRANSFORM AND APPLICATIONS IN PROBABILITY AND STATISTICS

by

M.A. Chaudhry and Munir Ahmad
Department of Mathematical Sciences
King Fahd University of Petroleum & Minerals
Dhahran, Saudi Arabia

ABSTRACT

In this paper we introduce the theory of generalized functions developed by L-Schwartz (1950-51) and discuss the generalized Fourier Transform and its applications in probability and statistics. The notion of the generalized weight functions for orthogonal polynomials is introduced. The use of the generalized function technique in determining the probability functions corresponding to the given moment functions is discussed.

Subject Classification: 60-02 and 60B15

Keywords: Generalized functions, Fourier Transforms, weight functions and singular functions.

INTRODUCTION

The singular functions have long been used in the fields of Physics and Engineering, although these cannot be properly defined within the framework of classical function theory. The simplest of the singular functions is the delta function. It is common defined as:

$$\delta(s) = \begin{cases} 0 & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases} \tag{1}$$

and

$$\int_{-\infty}^{\infty} \delta(x)dx = 1. \tag{2}$$

According to the classical definition of a function and integrals, the definitions (1) and (2) are inconsistant. There are several extensions and generalizations of the concept of a mathematical function, see [3], [5] and [10]. However, we shall briefly discuss here the theory of generalized functions developed by Schwarts [14] and point out its applications in probability in the subsequent sections.

NOTATIONS

$K_a = \{x \in R^n: |x_2| \leq a,\ 1 \leq i \leq n\}$ is a compact subset of R^n and $||x|| = (x_1^2 + x_2^2 + \ldots + x_n^2)$ is the norm of $x = (x_1, \ldots, x_n)$. Let $k = (k_1, k_2, \ldots, k_n)\ \mathbb{Z}_+^n$. Then, we define

$$|k| = k_1 + k_2 + \ldots + k_n$$

$$x^k = x_1^{k_1}\, x_2^{k_2} \ldots x_n^{k_n}$$

$$D^K = \frac{\partial^{|k|}}{\partial x_1^{k_1}\,\partial x^{k}\,\ldots\,\partial x_n^{k_n}}$$

Supp f = The support of the function $f(x)$

= the closure of the set of all points x such that $f(x) \neq 0$.

1. THE TEST SPACE, $\mathcal{D}(a) \equiv \mathcal{D}(K_a)$

An infinitely differentiable function $\phi(x)$ is said to belong to the space $\mathcal{D}(a)$ if for each $p = 1,2,3,\ldots$

$$\|\phi\|_p = \sup_{|k|\leq p} \sup_{x} |D^K\phi(x)| < \infty$$

$\|\cdot\|_p$; $p = 1,2,3,\ldots$ define a sequence of norms on $\mathcal{D}(a)$ and:

$$\|\phi\|_1 \leq \|\phi\|_2 \leq \ldots \leq \|\phi\|_p \leq \ldots \tag{1}$$

Let $\mathcal{D}_p(a)$ be the completion of $\mathcal{D}(a)$ with respect to the norm $\|\cdot\|_p$. Then, it follows from (1) that:

$$\mathcal{D}_1(a) \supseteq \mathcal{D}_2(a) \supseteq \ldots \supseteq \mathcal{D}_p(a) \supseteq \ldots \supseteq \mathcal{D}(a) \tag{2}$$

As a matter of fact

$$\mathcal{D}(a) = \bigcap_{p=1}^{\infty} \mathcal{D}_p(a) \tag{3}$$

and therefore, it is complete countably normed space [3], [9].

2. THE TEST SPACE $\mathcal{D}$

The space $\mathcal{D}$ consists of infinitely differentiable functions outside a compact set (depending upon the function) vanish identically. It can be seen that:

$$\mathcal{D} = \bigcup_a \mathcal{D}(a)$$

$$= \bigcup_a \left(\bigcap_{p=1}^{\infty} \mathcal{D}_p(a)\right) \tag{4}$$

REMARKS:

1) $\mathcal{D}$ is complete [1], [3], [9]
2) $\mathcal{D}$ is not metrizable [3]
3) $\{\phi_n\}_{n=1} \subset \mathcal{D}$ converges to ϕ iff $D^k\phi_n$ ($\forall$ n = 1,2,3,... and $\forall$ |k| = 0,1,2,3,... .) vanishes outside the same compact set K_a and $D^k\phi_n \to D^k\phi$ for all k = 0,1,2,3, ...

EXAMPLE:

Let

$$\phi(x;a) = \begin{cases} \exp\left(\dfrac{-|a|^2}{|a|^2 - |x|^2}\right) & \text{for } |x| < |a| \\ 0 & \text{for } |x| \geq |a| \end{cases} \tag{5}$$

Then

$$\phi(x;a) \in \mathcal{D}(a) \subset \mathcal{D} \quad \text{and} \quad \text{supp}\ \phi = [-a,a] \tag{6}$$

Let $\psi_n(x) = \frac{1}{n}\,\phi(x,a)$

and $\zeta_n(x) = \frac{1}{n}\,\phi(\frac{x}{n};a)$

Then, $\text{supp}\ \psi_n = \text{supp}\ \phi = [-a,a] = K_a$

and $\text{supp}\ \zeta_n = [-na,na]$

Now $\zeta_n(x) \xrightarrow{\mathcal{D}} 0$. However, $\{\zeta_n(x)\}_{n=1}^{\infty}$ does not converge to zero in the sense of the convergence in $\mathcal{D}$ because all of $\zeta_n(x)$ $(n = 1,2,3,\ldots)$ do not have the same support.

3. THE DISTRIBUTION SPACE $(\mathcal{D})'$

Let $\mathbb{C}$ be the field of complex numbers. Then, $f:\mathcal{D} \to \mathbb{C}$ is said to be a continuous linear function if:

1) $\langle f, \alpha\phi+\beta\psi\rangle = \alpha\langle f,\phi\rangle + \beta\langle f,\psi\rangle$

and

2) Whenever $\phi_n \xrightarrow{\mathcal{D}} \phi$, $\langle f,\phi_n\rangle \xrightarrow{\mathbb{C}} \langle f,\phi\rangle$.

The space of all continuous linear functionals defined on $\mathcal{D}$ is denoted by $(\mathcal{D})'$. The elements of $(\mathcal{D})'$ are called generalized functions. The convergence in $(\mathcal{D})'$ is defined as follows:

A sequence f_n, $n = 1,2,3,\ldots$ of generalized functions is said to converge to generalized function $f \in (\mathcal{D})'$ and we write:

$$f_n \xrightarrow{(\mathcal{D})'} f \quad \text{iff} \quad \langle f_n,\phi\rangle \xrightarrow{\mathbb{C}} \langle f,\phi\rangle \quad \forall\, \phi \in \mathcal{D} \tag{7}$$

4. EXAMPLES OF GENERALIZED FUNCTIONS

1. (Regular generalized functions):

Let $f(x)$ be a locally integrable function, i.e., $\int_{\Omega}|f(x)|dx < \infty$ for every bounded region Ω in $\mathbb{R}^n$. Then, the map $f:\mathcal{D} \to \mathbb{C}$ defined by:

$$\langle f,\phi\rangle = \int_{\mathbb{R}} f(x)\phi(x)\,dx \quad \forall\, \phi \in \mathcal{D} \tag{8}$$

defines a continuous linear functional on $\mathcal{D}$ and hence is an element of $(\mathcal{D})'$. Such type of generalized functions are called regular.

2. (Singular generalized functions):

Let $\delta:\mathcal{D} \to \mathbb{C}$ be defined by:

$$<\delta,\phi> = \phi(o) \quad \forall\ \phi \in \mathcal{D} \tag{9}$$

Then, $\delta \in (\mathcal{D})'$. We cannot find any locally integrable function $f(x)$ for which

$$<\delta,\phi> = <f,\phi> = \int_{\mathbb{R}} f(x)\phi(x)dx \quad \forall\ \phi \in \mathcal{D} \tag{10}$$

Suppose there exists some $f(x)$ (locally integrable) such that (10) is satisfied. Take $\phi(x) = \phi(x;a)$ as defined in (5). Then,

$$\text{L.H.S.} = <\delta,\phi> = <\delta(x);\ \phi(x;a)> = 1/e$$

and

$$\text{R.H.S.} = <f,\phi> = \int_{R} f(x)\phi(x;a)dx \to 0 \quad \text{as} \quad a \to 0$$

which contradicts the assumption.

The type of generalized functions which are not regular (such as δ-distribution) are called singular.

<u>REMARK</u>

The singular generalized functions can be approximated by a sequence (or by a set) of regular generalized functions in the sense of the convergence defined in (7).

EXAMPLES

1. It can be seen in [1], [3], [6], [9] and [10] that among other converging regular generalized functions we also have

i) $\frac{\sin(nx)}{x} \xrightarrow{\mathcal{D}} \delta(x)$ as $n \to \infty$

ii) The normal probability function converges to $\delta(x)$ as $t \to 0^+$, i.e.,

$$\frac{1}{2\sqrt{\pi t}} \exp(- \frac{x^2}{4t}) \xrightarrow{\mathcal{D}} \delta(x) \quad \text{as} \quad t \to 0^+$$

and

iii) $\frac{\eta}{\pi[x^2 + \eta^2]} \xrightarrow{\mathcal{D}} \delta(x)$ as $\eta \to 0^+$.

2. 2. Consider the Heaviside function $H(x)$ defined by:

$$H(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Then, $<H,\phi> = \int_0^\infty \phi(x)dx$

defines a regular generalized function on $\mathcal{D}$.

One can define a probability function on H as follows:

Since $H(x - x_j) = 0$ if $x < x_j$ and $H(x - x_j) = 1$ if $x > x_j$, then $P(H(x - x_j) = 1) = F(x)$ and $P(H(x - x_j) = 0) = 1 - F(x)$.

3. Let us define $F^*(x):\mathcal{D} \to \mathbb{C}$ by:

$$<F^*,\phi> = \sum_{n=1}^{\infty} \phi(n)$$

Then, $F^*(x) = \sum_{n=1}^{\infty} \delta(x - n)$

is a shifted singular generalized function. It is also called sampling distribution because it gives the information about the function $\phi(x)$ at $x = n$.

4. The function $\frac{1}{x}$ does not define a regular generalized function, because $\int_{-\infty}^{\infty} \frac{\phi(x)}{x}\,dx$ is not convergent for all test functions. Let us define $PV(\frac{1}{x})$ as follows:

$$<PV(\frac{1}{x}), \phi> = PV\int_{-\infty}^{\infty} \frac{\phi(x)}{x}\,dx$$

$$= \lim_{\varepsilon \to 0^+} \int_{|x| \geq \varepsilon} \frac{\phi(x)}{x}\,dx$$

$$= \int_{-\infty}^{\infty} \frac{\phi(x) - \phi(0)}{x}\,dx \tag{11}$$

The integral in (11) is convergent because $\phi(x)$ is differentiable at $x = 0$. Moreover $PV(\frac{1}{x})$ is continuous, [1], [3]. Therefore, $PV(\frac{1}{x})$ is a singular generalized function.

5. Let us define

$$\delta_{\pm}(x) = \frac{1}{2}\,\delta(x) \mp (\frac{1}{2\pi})PV(\frac{1}{x}) \tag{12}$$

Then, $\delta_{\pm}$ are also singular generalized functions on $\mathcal{D}$ and are called the Heisenberg distributions. It is shown [1] that:

$$<\delta_{\pm}(x), \phi(x)> = \mp \lim_{\varepsilon \downarrow 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\phi(x)}{x \pm i\varepsilon}\,dx \tag{13}$$

It follows from (12) and (13) that:

$$\frac{1}{x + io} = -\pi i\delta(x) + PV(\frac{1}{x}) \tag{14}$$

$$\frac{1}{x - io} = i\pi\delta(x) + PV(\frac{1}{x}) \tag{15}$$

(14) and (15) are called the Sokhotski-Plemelj relations.

6. OPERATIONS ON GENERALIZED FUNCTIONS

Since the locally integrable functions are examples of generalized functions it is, therefore, natural to define operations on them that will remain valid for integrable functions.

Let $f, g \in (\mathcal{D})'$, $\alpha \in \mathbb{C}$ and $x = Ay - a$ where A is an $n \times n$ matrix with $\det A \neq 0$ and a is a constant vector, be a non-singular linear transformation of the space $\mathbb{R}^n$ onto itself. Then, for every $\phi \in \mathcal{D}$, the following operations are defined:

a) Addition: $\langle f+g, \phi\rangle = \langle f,\phi\rangle + \langle g,\phi\rangle$

b) Linear change of variables:

$$\langle f(Ay - a),\ \phi(y)\rangle = \frac{1}{|\det A|} \langle f(x),\ \phi\{A^{-1}(x + a)\}\rangle \qquad 6.1)$$

For a simple translation, i.e., when $A = I$, the unit matrix, (6.1) yields

$$\langle f(y-a),\ \phi(y)\rangle = \langle f(x),\ \phi(x+a)\rangle \qquad (6.2)$$

For a simple scale expansion $A = cI$, $a = 0$, (6.1) becomes,

$$\langle f(cy),\ \phi(y)\rangle = (\frac{1}{|c|^n} \langle f(x),\ \phi(x/c)\rangle \qquad (6.3)$$

For a simple reflection take $x = -y$, $(c = -1)$

$$\langle f(-y),\ \phi(y)\rangle = \langle f(x),\ \phi(-x)\rangle \qquad (6.4)$$

For a simple rotation, $a = 0$, $A^T = A^{-1}$, we have

$$\langle f(Ay),\ \phi(y)\rangle = \frac{1}{|\det A|}\langle f(x),\ \phi(A^T x)\rangle \tag{6.5}$$

c) Multiplication by $\psi(t)$ $C^\infty(\mathbb{R}^n)$:

$$\langle \psi f,\ \phi\rangle = \langle f,\ \psi\phi\rangle \tag{6.7}$$

e.g. $\langle \psi\delta,\ \phi\rangle = \psi(0)\phi(0)$.

Therefore

$$\psi(x)\delta(x) = \psi(0)\delta(x)$$

d) Differentiation: The generalized derivative $D^k f$ of the generalized function $f \in (\mathcal{D})'$ is defined as follows:

$$\langle D^k f, \phi\rangle = (-1)^{|k|}\langle f,\ D^k\phi\rangle \tag{6.8}$$

e.g.,

$$\langle \frac{d}{dx}H(x),\ \phi(x)\rangle = -\langle H(x),\ \phi'(x)\rangle = -\int_0^\infty \phi'(x)dx = \phi(0)$$

Therefore

$$\frac{d}{dx}H(x) = \delta(x) \tag{6.9}$$

In other words, the derivative of the Heaviside function is Dirac's δ-function which was one of the properties of the δ-function deduced by Dirac.

An important consequence of the definition (6.8) is that generalized functions have derivatives of all orders. It reveals an important fact that continuous and locally integrable functions are infinitely differentiable in the generalized sense which gives us relief from the difficulties that arise with non-differentiable functions.

7. TENSOR PRODUCT AND CONVOLUTION OF TWO GENERALIZED FUNCTIONS

The tensor product of two generalized functions: $f \in (\mathcal{D}(\mathbb{R}^n))'$ and

$g \in (\mathcal{D}(\mathbb{R}^m))'$ is defined by:

$$\langle f(x) \times g(y), \phi(x,y)\rangle = \langle f(x), \langle g(y), \phi(x,y)\rangle\rangle \qquad (7.1)$$
$$\forall\ \phi(x,y) \in \mathcal{D}(\mathbb{R}^{n+m}).$$

The product $f(x) \times g(y)$ belongs to $(\mathcal{D}(\mathbb{R}^{n+m}))'$. It is commutative and associative when extended to any finite number of generalized functions.

Let $f(x)$ and $g(x)$ be two locally integrable functions on $\mathbb{R}^n$. Then, their convolution:

$$f(x) * f(x) = \int_{\mathbb{R}^n} f(t)g(x - t)dt \qquad (7.2)$$

is also a locally integrable function [1], [2], [6] and [9]. Therefore, it defines a functional on $\mathcal{D}$, i.e.,

$$\begin{aligned}\langle f(x) * g(x), \phi(x)\rangle &= \int \phi(x) \int f(t)g(x-t)dt\ dx \\ &= \iint f(x)g(y)\ \phi(x+y)dx\ dy \qquad \text{[By Fubini's Theorem]} \\ &= \langle f(x) \times g(x), \phi(x + y)\rangle \qquad (7.3)\end{aligned}$$

The convolution of two generalized functions may be defined in the same manner. The difficulty is that $\phi(x+y)$ need not have a bounded support even if $\phi(x) \in \mathcal{D}(\mathbb{R}^n)$ [3], [9]. To make the definition meaningful, we can put restrictions on f and g such that the support of $f \times g$ intersects the support of $\phi(x+y)$ in a bounded set.

DEFINITION: A generalized function f is said to vanish on a set $\Omega \subseteq \mathbb{R}^n$ if $\langle f,\phi\rangle = 0$ for all $\phi \in \mathcal{D}$ with supp $\phi \subset \Omega$. The complement of the union of open sets Ω on which f vanishes is a closed set, called the support of the generalized function f.

If g has a compact support, then the convolution $f * g$ is well-defined and is given by:

$$\langle f * g, \phi\rangle = \langle f(x) \times g(y), \eta(y)\phi(x + y)\rangle \tag{7.4}$$

where η is any test function equal to 1 in the neighbourhood of the support of g [2], [6], and [10].

It is easy to verify that $\delta^m * f = f^{(m)}$. Thus, if $P(x)$ is a polynomial, $P(\delta) * y = f$ is an ordinary differential equation. Another property of the convolution is that if f or g has compact support, then

$$D^k(f * g) = (D^k f) * g = f * D^k g \tag{7.5}$$

These properties give a simple proof of the existence theorem for linear partial differential equations with constant coefficients:

$$P(D)y = f \tag{7.6}$$

The existence of the fundamental solution, i.e., the existence of solution of $P(D)E = \delta$ is proved in [3], [6] and [9]. If f has a compact support, then $y = E * f$ is the generalized solution of (7.6). It follows from the fact that:

$$P(D)[E * f] = (P(D)E) * f = \delta * f = f \tag{7.7}$$

8. THE TEST SPACE S

A complex valued function $\phi(x)$ is said to belong to the space S if it has the following properties:

1) $\phi(x)$ is infinitely differentiable, i.e. $\phi(x) \in C^\infty(\mathbb{R}^n)$

2) $\phi(x)$ as well as its derivatives of all orders, vanish at infinity faster than the reciprocal of any polynomial, i.e.,

$$|x^p D^k \phi(x)| \leq C_{pk}, \quad p = 0,1,2,\ldots \tag{8.1}$$

where C_{pk} is a constant depending on p, k and ϕ.

A sequence $\{\phi_m(x)\}_{m=1}$ of test functions is said to converge to $\phi(x)$ in S if for each $|k| = 0,1,2,3,\ldots,$ the sequence $\{D^k\phi_m(x)\}_{m=1}$ converges uniformly to $D^k\phi(x)$ in every bounded region Ω of $\mathbb{R}^n$. This means that constants C_{pk} in (8.1) can be chosen independently of x such that

$$|x^p(D^k\phi_m - D^k\phi)| < C_{pk} \tag{8.2}$$

for all values of m.

The space S is closed and the testing function space $\mathcal{D}$ is dense in S [1], [3], [9] & [12]. The dual space of S is denoted by $(S)'$. The elements of $(S)'$ are called distributions of slow growth or tempered distributions. It follows from the definition of convergence in $\mathcal{D}$ and in S that a sequence $\{\phi_m(x)\}$ converging to a function $\phi(x)$ in the sense of $\mathcal{D}$ also converges to $\phi(x)$ in the sense of S. Accordingly every linear continuous functional on S is also a linear continuous functional on $\mathcal{D}$ and therefore, $(S)' \subset (\mathcal{D})'$. This inclusion is strict because the distributions which grow too rapidly at infinity are not elements of $(S)'$. For example the regular distribution $f = \exp(x^2) \in (\mathcal{D})'$ but is not a member of $(S)'$.

9. THE FOURIER TRANSFORM

An essential part of the theory of generalized function and its application rests on the concept of the Fourier Transform. If $\phi(x)$ is an absolutely integrable on the real line then, its Fourier Transform is defined by:

$$\hat{\phi}(u) = \int_{-\infty}^{\infty} e^{iux}\phi(x)\,dx \tag{9.1}$$

the integral in (8.1) exists, since by assumption $|\hat{\phi}(u)| \leq \int_{-\infty}^{\infty} dx|\phi(x)| < \infty$. If, moreover, $\hat{\phi}(u)$ is absolutely integrable, the inverse Fourier Transform is given by:

$$\phi(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iux}\,\hat{\phi}(u)\,du \tag{9.2}$$

It follows that

$$\hat{\hat{\phi}}(x) = 2\pi\phi(-x) \tag{9.3}$$

Now we state a theorem which reveals the characteristics feature of the space S [1], [3] and [12].

THEOREM 9.1

The Fourier Transform as defined in (9.1) and its inverse are continuous linear, and one-to-one mapping of S onto itself.

DEF (Fourier Transform of Tempered Distribution):

The Fourier Transform $\hat{\tau}(u)$ of a tempered distribution $\tau(x) \in (S)'$ is defined by:

$$\langle\hat{\tau}(u), \phi(u)\rangle = \langle\tau(u), \hat{\phi}(u)\rangle, \quad \phi \in S \tag{9.4}$$

The functional on the right hand side of (9.4) is well defined because $\hat{\phi}(u) \in S$. It is clearly linear and continuous. Hence $\hat{\tau}(u) \in (S)'$. As a matter of fact we have the following theorem [3], [6], [9] and [12].

THEOREM 9.2

The generalized Fourier Transform as defined in (9.4) and its inverse as defined by:

$$\langle F^{-1}(\tau), \phi\rangle = \langle\tau, F^{-1}(\phi)\rangle \tag{9.5}$$

are continuous linear and one-to-one mapping of $(S)'$ onto itself.

The definitions (9.4) and (9.5) are consistent with the classical definitions (9.1) and (9.2) whenever the latter are applicable.

The extensions to n-dimensional space of the definitions and results are straight forward. We shall mention the n-dimensional generalization of the results when it is necessary.

10. EXAMPLES

(a) The delta function

$$\langle\hat{\delta}(x),\phi\rangle = \langle\delta(x),\hat{\phi}\rangle = \langle\delta(x),\int_{-\infty}^{\infty} e^{ixy}\phi(y)\,dy = \int_{-\infty}^{\infty}\phi(y)\,dy = \langle 1,\phi\rangle$$

Thus, $\hat{\delta}(x) = 1$ (10.1)

According to (9.3) we have

$$[1]^{\wedge} = [\hat{\delta}]^{\wedge} = (2\pi)^n \delta(-x) = (2\pi)^n \delta(x)$$

or

$$F^{-1}[\delta(x)] = \frac{1}{(2\pi)^n} \tag{10.2}$$

For $n = 1$ this gives the well-known integral representation formula for the delta function:

$$\delta(x) = \frac{1}{2\pi} F[1] = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixy}\, dy \tag{10.3}$$

(b) If $\tau \in (S)'$, then,

i) $D^k F[\tau] = F[(ix)^k \tau]$

ii) $F[P(\frac{d}{dx})\tau] = P(-iu)F[\tau]$

iii) $F[x^k \tau] = (-i \frac{d}{du})^k F[\tau]$

iv) $F[\tau(x - a)] = e^{iau} F[\tau]$

v) $F[P(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n})\tau] = P(-iu_1, -u_2, \ldots, -u_n)F[\tau]$

vi) $F[P(x_1, x_2, \ldots, x_n)\tau] = P(-i \frac{\partial}{\partial u_1}, -i \frac{\partial}{\partial u_2}, \ldots, -i \frac{\partial}{\partial u_n})F[\tau]$

vii) $F[\tau(Ax)] = |\det A|^{-1} F[\tau((A^T)^{-1}(x))]$

For the proof see [1], [9], [11] and [14].

(c) The Heaviside function, $n = 1$

Since $(x-\xi)\delta(x-\xi) = 0$, it follows that $y_c = \delta(x-\xi)$ is a solution to the homogeneous differential equation

$$(x - \xi)\delta(x - \xi) = 0$$

Moreover, $y_p = a(x)\ PV(\frac{1}{x - \xi})$ is a generalized solution of the inhomogeneous differential equation:

$$(x - \xi)\tau(x - \xi) = a(x) \tag{10.4}$$

where $a(x) \in C^{\infty}(\mathbb{R})$.

Therefore,

$$y = \delta(x - \xi) + a(x)\ PV(\frac{1}{x - \xi}) \tag{10.5}$$

is the solution of (10.4).

Since $\frac{dH}{dx}(x) = \delta(x)$ we find that

$$F[\frac{dH}{dx}(x)] = (-iu)\hat{H}(u) = 1$$

Therefore, by using (10.4) and (10.5) we get:

$$F[H(x)](u) = C\delta(u) + i\ PV(\frac{1}{u}) \tag{10.6}$$

Changing x to $-x$ in this formula we get:

$$F[H(-x)](u) = C\delta(u) - iPV(\frac{1}{u}) \tag{10.7}$$

$$\therefore \qquad F[H(x)] + F[H(-x)] = 2C\delta(u) \tag{10.8}$$

$$\Longrightarrow \qquad F[H(x)] + F[H(-x)] = F[1] = 2\pi\delta(x) \tag{10.9}$$

From the uniqueness of the Fourier Transform and by equating (10.8) & (10.9) we get $C = \pi$.

$$\therefore \qquad F[H(x)](u) = \pi\delta(u) + iPV(\tfrac{1}{u}) \qquad (10.10)$$

If we write (10.10):

$$\int_{-\infty}^{\infty} H(x)\ e^{ixu}\, dx = \pi\delta(u) + iPV(\tfrac{1}{u})$$

and separate real and imaginary parts we get:

$$\int_0^{\infty} \cos(xu)\,dx = \pi\delta(u) \qquad (10.11)$$

$$\int_0^{\infty} \sin(xu)\,dx = PV(\tfrac{1}{u}). \qquad (10.2)$$

(d) <u>The signum function</u>:

Since Sgn $x = H(x) - H(-x)$

$$\Longrightarrow \quad F[\operatorname{sgn}(x)](u) = F[H(x)](u) - F[H(-x)](u)$$

$$= 2iPV(\tfrac{1}{u}) \qquad (10.13)$$

Therefore, by inversion formula:

$$F[PV(\tfrac{1}{u})](x) = \pi i \operatorname{sgn} x \qquad (10.14)$$

$$\Longrightarrow \quad F[PV(\tfrac{1}{u-a})](x) = \pi i\ e^{iax} \operatorname{sgn} x \qquad (10.15)$$

(e) Since $\frac{1}{x^m} = \frac{(-1)^{m-1}}{(m-1)!} \frac{d^{m-1}}{dx^{m-1}} (\frac{1}{x})$

$$\Longrightarrow \quad F[\tfrac{1}{x^m}](u) = \frac{(-1)^{m-1}}{(m-1)!} (-iu)^{m-1}\ F[\tfrac{1}{x}](u)$$

$$= i^m \pi \frac{u^{m-1}}{(m-1)!} \operatorname{sgn} u \qquad (10.16)$$

By the Translation property:

$$F\left[\frac{1}{(x-a)^m}\right](u) = i^m \frac{u^{m-1}}{(m-1)!} e^{iau} \operatorname{sgn} u \tag{10.17}$$

11. DISTRIBUTIONAL WEIGHT FUNCTIONS

Let $P_n(x)$ be a polynomial of degree n such that

$$\int_a^b P_n(x)P_m(x)w(x)dx = 0, \quad m \neq n \tag{11.1}$$

i.e., $\{P_n(x)\}_{n=1}$ is an orthogonal sequence of polynomials with respect to the weight function $w(x)$. The numbers μ_n $(n=0,1,2,...)$ defined by:

$$\mu_n = \int_a^b x^n w(x)dx \tag{11.2}$$

are called moments. The moments play an important role in the theory of orthogonal polynomials. As a matter of fact, every polynomial can be expressed in terms of its moments [13].

It follows from (11.2) that if we know the weight function of an orthonormal polynomial sequence, then we can calculate the moments. The theory of generalized functions helps us in solving the inverse problem. That is, given the moments, we can determine the corresponding weight functions. For this purpose (11.2) can be written in the functional form as:

$$\mu_n = \langle w(x), x^n \rangle \qquad \text{for all} \quad n = 0,1,2,3,... \tag{11.3}$$

Let $\psi(x)$ be a real analytic function whose Taylor series converges to for all x. Then,

$$\langle w,\phi\rangle = \langle w, \sum_{n=0}^{\infty} \frac{\psi^{(n)}(0)}{n!} x^n\rangle$$

$$= \sum_{n=0}^{\infty} \frac{\psi^{(n)}(0)}{n!} \quad w,\ x^n\rangle \qquad (11.4)$$

Since $(-1)^n\psi^{(n)}(0) = \langle\delta^{(n)}(x),\ \psi(x)\rangle$ [By definition of $\delta^{(n)}(x)$] (11.5)

Therefore,

$$\langle w,\psi\rangle = \langle \sum_{n=0}^{\infty} (-1)^n \frac{\mu_n}{n!} \delta^{(n)}(x),\ \psi\rangle$$

which implies

$$w(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \mu_n\ \delta^{(n)}(x) \qquad (11.6)$$

in the sense of generalized function. It is important to note that when the moments $\{U_i\}_{i=0}^{\infty}$ are those associated with the classical orthogonal polynomials - the Legendre polynomials, the Laguerre polynomials, or the Hermite polynomials - the weight function $w(x)$ yields the same results as the classical weight functions concerning orthogonality and norms. However, when the moments $\{U_i\}_{i=0}^{\infty}$ are those associated with the Jacobi polynomials or the generalized Laguerre polynomials, then w remains a suitable generalized weight function belonging to certain space of generalized functions [6], [12] and [13].

12. APPLICATIONS TO PROBABILITY AND STATISTICS

The theory of generalized functions developed by Schwartz has advantage over the measure theory and Stieltjes integral treatments used to explain singular integrals occuring in probability and statistics.

12.1. Probability Distributions

Let X be a random variable taking real values in $(-\infty,\infty)$ and $\Phi(t)$ be the probability distribution function. The probability distribution $\Phi(t)$ is called discrete or continuous according to whether t takes on discrete or continuous values. For example,

$$\Phi(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{t} e^{-(t-\mu)^2/2\sigma^2}\, dt \tag{12.1}$$

is well known Gaussian distribution which is continuous, and in this case:

$$\phi(t) = \frac{d\Phi}{dt} = \frac{1}{\sigma\sqrt{2\pi}} e^{-(t-\mu)^2/2\sigma^2} \tag{12.2}$$

It should be noted that $\Phi(t)$ is a function, not a functional. Therefore, the probability distributions and generalized functions refer to different mathematical objects.

Suppose that it is certain that the random variable X takes the value x_0. Then:

$$\Phi(t) = 0 \quad \text{for} \quad t < x_0$$

$$\Phi(t) = 1 \quad \text{for} \quad t > x_0$$

Thus, $\Phi(t) = H(t - x_0)$ is the Heaviside step function. In this case the probability density $\phi(t)$ does not exist in the ordinary sense. However, in the sense of generalized functions we have:

$$\phi(t) = \delta(t - x_0) \tag{12.3}$$

Similarly, if the random variable X takes the values $x_1,x_2,\ldots,x_n$ with the probabilities $p_1,p_2,\ldots,p_n$, respectively, such that $\Sigma_{i=1}^{n} p_i = 1$, then, the probability distribution $\Phi(t)$ is given by:

$$\Phi(t) = \sum_{i=1}^{n} p_i H(t - x_i) \tag{12.4}$$

and the probability density function $\phi(t)$ is the generalized function given by:

$$\phi(t) = \sum_{i=1}^{n} p_i \delta(t - x_i) \tag{12.5}$$

EXAMPLE

The binomial probability distribution function $\Phi(t)$ is defined by:

$$\Phi(t) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} H(t - k) \tag{12.6}$$

The probability density function is the generalized function

$$\phi(t) = \frac{d}{dt} \Phi(t) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \delta(t - k) \tag{12.7}$$

12.2. Characteristic Functions

Given a probability density $\phi(t)$, the characteriztic function $\chi(u)$ is defined as:

$$\chi(u) = \int_{-\infty}^{\infty} e^{iut} \phi(t)\, dt \tag{12.8}$$

i.e., $\chi(u)$ is the Fourier transform of $\phi(t)$. Since $\phi(t)$ is the derivative of the bounded function $\Phi(t)$, the characteristic function in (12.8) exists in the generalized sense.

Conversely, given a characteristic function $\chi(u)$ it follows from (12.8) that the probability density would be the inverse Fourier transform of $\chi(u)$, i.e.,

$$\phi(t) = F^{-1}\{\chi(u)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-iut}\chi(u)du \qquad (12.9)$$

Distributional Fourier transform permits us to treat the discontinuous distributions and the casual distributions alongside of the continuous distributions.

EXAMPLE 1

Let us take $\chi(u) = e^{i\lambda u}$. Then, from (12.9)

$$\phi(t) = F^{-1}\{e^{i\lambda u}\} = \delta(t - \lambda) \qquad (12.10)$$

EXAMPLE 2

For the Gaussian distribution the probability density function $\phi(t)$ is given by:

$$\phi(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-(t-\mu)^2/2\sigma^2}$$

Therefore,

$$\chi(u) = F^{-1}\{\phi(t)\} = \exp(i\mu u - \frac{1}{2}\,\sigma^2u^2)$$

For the special case of μ taken to be zero, we have

$$\chi(u) = \exp(-\frac{1}{2}\,\sigma^2u^2)$$

and

$$\phi(t) = F^{-1}\{\chi(u)\} = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{1}{2}(t^2/\sigma^2)}, \quad -\infty < t < \infty.$$

12.3. Probability Fields

Let Ω be the set of elementary events and $\mathbb{B}$ be the class of the subsets of Ω such that

(i) The family $\mathbb{B}$ contains the empty set ϕ and the total set Ω;

(ii) If $A \in \mathbb{B}$, and α is a real number. Then $\alpha A \in \mathbb{B}$; and

(iii) If the sets $A_1, A_2, \ldots, A_n$ belong to $\mathbb{B}$, then their sum belongs to $\mathbb{B}$.

The probability measure P or $\mathbb{B}$ has the following properties:

(i) $P(\Omega) = 1$

(ii) If the sets of $A_1, A_2, \ldots, A_n, \ldots$ are mutually disjoint, that is, if $A_i \cap A_j = \phi$ for $i \neq j$, then

$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i).$$

The system $(\Omega, \mathbb{B}, P)$ is called the probability field.

A random variable X is a mapping $x: \Omega \to \mathbb{R}$ such that

$$X^{-1}((-\infty,t)) = \{w \in \Omega: \chi(w) \in (-\infty,t)\} \in \mathbb{B}.$$

We associated with the random variable X a probability distribution function $\Phi(t)$ as follows:

$$\Phi(t) = P(x^{-1}(-\infty,t)) = P(x(w) < t) \tag{12.12}$$

The probability distribution function $\Phi(t)$ is locally integrable and hence generates a <u>regular distribution</u> Φ defined by:

$$\langle\Phi,\psi\rangle = \int_{-\infty}^{\infty} \Phi(t)\psi(t)dt; \quad \forall\ \psi(t) \quad \mathcal{D} \tag{12.13}$$

Therefore, $\Phi \quad (\mathcal{D})'$.

On the other hand,

$$\langle \Phi', \psi \rangle = -\langle \Phi, \psi' \rangle = -\int_{-\infty}^{\infty} \Phi(t)\psi'(t)dt$$

$$= -[\Phi(t)\psi(t)]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \psi(t) \frac{d\Phi(t)}{dt} dt$$

$$= 0 + \int_{-\infty}^{\infty} \psi(t) d\Phi(t)$$

$$= \langle \phi, \psi \rangle$$

So the probability density $\phi(t)$ is the distributional derivative of the probability distribution $\Phi(t)$.

Now we can define the classical quantities in the following way:

(1) The expectation value is:

$$E(x) = \langle t, \phi \rangle = \int_{-\infty}^{\infty} t\,\phi(t)dt = \int_{-\infty}^{\infty} t\, d\Phi(t) = \int_{\Omega} x(w)dP(w) \tag{12.14}$$

(2) The variance is:

$$\sigma^2 = \langle (t - E(x))^2, \phi \rangle = \int_{-\infty}^{\infty} \{t - E(x)\}^2 d\Phi(t)$$

$$= \int_{\Omega} \{x(w) - E(x)\}^2 dP(w) \tag{12.15}$$

(3) The non-central mth moment is:

$$\langle t^m, \phi \rangle = \int_{-\infty}^{\infty} t^m d\Phi(t) = \int_{\Omega} [x(w)]^m dP(w).$$

(4) The central mth moment is:

$$\langle (t - E(x))^m, \phi \rangle = \int_{-\infty}^{\infty} \{t - E(x)\}^m d\Phi(t) = \int_{\Omega} \{x(w) - E(x)\}^m dP(w) \tag{12.16}$$

In this notation we can define the characteristic function $\chi(u)$ by:

$$\chi(u) = E(e^{itu}) \tag{12.17}$$

$$= \langle e^{itu}, \phi \rangle$$

and the Inverse Fourier Transform by

$$\phi = \langle e^{-itu}, \chi(u) \rangle$$

The foregoing concepts can be extended to a finite system of random variables $X_1, X_2, \ldots, X_n$. This sytem may be considered as a mapping from the set Ω into the n-dimensional space $\mathbb{R}^n$. Such a mapping is called an n-dimensional random variable, the probability distribution is now:

$$\Phi(t_1, t_2, \ldots, t_n) = P(x_1(w) < t_1, x_2(w) < t_2, \ldots, x_n(t) < t_n) \tag{12.18}$$

The moments are given by the formula:

$$\mu_{|k|} = \int_{\mathbb{R}^n} t^k d\Phi = \int_{\mathbb{R}^n} t_1^{k_1} t_2^{k_2} \ldots t_n^{k_n} \, d\Phi$$

$$= \int_{\Omega} \{(x_1(w))^{k_1} (x_2(w))^{k_2} \ldots (x_n(w))^{k_n} \, dP(w) \tag{12.19}$$

where $|k| = k_1 + k_2 + \ldots + k_n$.

Acknowledgement

The authros are thankful to the KFUPM for the excellent research and computing facilities.

REFERENCES

[1] H.J. Bremmerman, Distributions, Complex Variables and Fourier Transforms, Addison Wesley, 1968.

[2] P.L. Butzer and R.L. Nessel, Fourier Analysis and Approximation, Vol.I, Academic Press, 1971.

[3] I.M. Gel'fand and G.E. Shilov, Generalized Functions, Vol.2, Academic Press.

[4] I.M. Gel'fand and Vilenkin, M. Ya; Generalized Functions, Vol.4, Academic Press, New York, 1964.

[5] A. Erdel'yi, Operational Calculus and Generalized Functions, New York: Holt, Rinehart Co; 1966.

[6] R.P. Kanwal, Generalized Functions: Theory and Technique, Academic Press, 1983.

[7] Gistescu, R., and Marinescu, G; "Applications of Theory of Distributions", Editura Academici, Bucharest, 1973.

[8] Roose, B.W., "Analytic Functions and Distributions in Physics and Engineering", Wiley, New York, 1969.

[9] A.H. Zemanian, Distribution Theory and Transform Analysis, New York: McGraw Hill (1965).

[10] A.H. Zemanian, Generalized Integral Transformations, New York: Interscience, 1968.

[11] V.S. Vladimirov, Equations of Mathematical Physics, Marcel Dekker, Inc., New York, 1971.

[12] Chihara, T.S., "An introduction to orthogonal polynomials," Gordon & Breach, New York, (1979).

[13] Morton, R.D., and Krall, A.M., SIAM J. Math. Anal, 9, 626 (1978).