Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Vol. 3, Issue 1 (June 2008), pp. 42 – 54
(Previously, Vol. 3, No. 1)

Applications and Applied Mathematics:
An International Journal
(AAM)

Some Applications of Dirac's Delta Function in Statistics for More Than One Random Variable

Santanu Chakraborty
Department of Mathematics
University of Texas Pan American,
1201 West University Drive, Edinburg, Texas 78541, USA
schakraborty@utpa.edu

Received July 14, 2006; accepted December 7, 2007

Abstract

In this paper, we discuss some interesting applications of Dirac's delta function in Statistics. We have tried to extend some of the existing results to the more than one variable case. While doing that, we particularly concentrate on the bivariate case.

Keywords: Dirac's Delta function, Random Variables, Distributions, Densities, Taylor's Series Expansions, Moment generating functions.

1. Introduction

Cauchy, in 1816, was the first (and independently, Poisson in 1815) gave a derivation of the Fourier integral theorem by means of an argument involving what we would now recognize as a sampling operation of the type associated with the delta function. And there are similar examples of the use of what are essentially delta functions by Kirchoff, Helmholtz and Heaviside. But Dirac was the first to use the notation δ. The Dirac delta function (δ-function) was introduced by Paul Dirac at the end of the 1920s in an effort to create the mathematical tools for the development of quantum filed theory. He referred to it as an "improper function" in Dirac (1930). Later, in 1947, Laurent Schwartz gave it a more rigorous mathematical definition as a linear functional on the space of test functions D (the set of all real-valued infinitely differentiable functions with compact support) such that for a given function $f(x)$ in D, the value of the functional is given by the property (b) below. This is called the sifting or sampling property of the delta function. Since the delta function is not really a function in the classical sense, one should not consider the "value" of the delta function at x. Hence, the domain of the delta function is D and its value, for $f \in D$ and a given x_0 is $f(x_0)$. Khuri (2004) studied some interesting applications of the delta function in statistics. He mainly studied univariate cases even though he did give some interesting examples for the multivariate case. We shall study some more applications in the multivariate scenario in this work. These might help future researchers in statistics to develop more ideas. In sections 2 and 3, we discuss derivatives of the delta function in both univariate and multivariate case. Then, in section 4, we discuss some applications of the delta function in probability and statistics. In section 5, we discuss calculations of densities in both univariate and multivariate case using transformations of variables. In section 6, we use vector notations for delta functions in the multidimensional case. In section 7, we discuss very briefly the transformations of variables in the discrete case. Then, in section 8, we discuss the moment generating function in the multivariate set up. We conclude with few remarks in section 9.

2. Derivatives of the δ-function in the Univariate Case

In the univariate case, some basic properties satisfied by Dirac's delta function are:

$$\text{(a)} \int_{-\infty}^{\infty} \delta(x)dx = 1,$$

$$\text{(b)} \int_{a}^{b} f(x)\delta(x-x_0)dx = f(x_0) \text{ for all } a < x_0 < b,$$

where $f(x)$ is any function continuous in a neighborhood of the point x_0. In particular, we have,

$$\int_{-\infty}^{\infty} f(x)\delta(x-x_0)dx = f(x_0).$$

This is the sifting property that we mentioned in the previous section. If $f(x)$ is any function with continuous derivatives up to the n^{th} order in some neighborhood of x_0, then

$$\int_{a}^{b} f(x)\delta^{(n)}(x-x_0)dx = (-1)^n f^{(n)}(x_0), \ \ n \geq 0 \text{ for all } a < x_0 < b.$$

In particular, we have,

$$\int_{-\infty}^{\infty} f(x)\delta^{(n)}(x-x_0)dx = (-1)^n f^{(n)}(x_0), \ \ n \geq 0$$

for a given x_0. Here, $\delta^{(n)}$ is the generalized n^{th} order derivative of δ. This derivative defines a linear functional which assigns the value $(-1)^n f^{(n)}(x_0)$ to $f(x)$.

Now let us consider the Heaviside function H(x) unit step function defined by

$$H(x) = 0 \text{ for } x < 0$$
$$= 1 \text{ for } x \geq 1.$$

The generalized derivative of $H(x)$ is $\delta(x)$, i.e., $\delta(x) = \frac{dH(x)}{dx}$. As a result, we get a special case of the formula for the n^{th} order derivative mentioned above:

$$\int_{-\infty}^{\infty} x^n \delta^{(i)}(x)dx = 0 \text{ if } i \neq n$$
$$= (-1)^n n! \text{ if } i = n$$

3. Derivatives of the delta function in the bivariate case

Following Saichev and Woyczynski (1997), Khuri (2004) provided the extended definition of delta function to the n-dimensional Euclidean space. But we shall mainly concentrate on the bivariate case. As in the univariate case, we can write down similar properties for the bivariate case as well. In the bivariate case, $\delta(x, y) = \delta(x)\delta(y)$. So, if we assume $f(x, y)$ to be a continuous function in some neighborhood of (x_0, y_0), then we can write

$$\iint_{\Re\times\Re} f(x, y)\delta(x - x_0, y - y_0)dxdy = f(x_0, y_0),$$

where $\Re$ is the real line.

Now, for this function f, if all its partial derivatives up to the n^{th} order are continuous in the abovementioned neighborhood of (x_0, y_0), then,

$$\iint_{\Re\times\Re} f(x, y)\delta^{(n)}(x - x_0, y - y_0)dxdy = (-1)^n \sum_{0\leq k\leq n} {}^nC_k \frac{\partial^n f(x, y)}{\partial x^k \partial y^{n-k}}\Big|_{x=x_0, y=y_0} \;\ldots\ldots \quad (1)$$

where ${}^nC_k$ is the number of combinations of k out of n objects, $\delta^{(n)}(x, y)$ is the generalized n^{th} order derivative of $\delta(x, y)$. In the general p-dimensional case, by using induction on n, it can be shown that

$$\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} f(x_1,\ldots, x_p)\delta^{(n)}(x_1 - x_1^*,\ldots, x_p - x_p^*)dx_1\ldots dx_p$$
$$= (-1)^n \sum_{0}^{n}\ldots\sum_{0}^{n-k_1-\ldots-k_{p-1}} \frac{n!}{k_1!\ldots k_p!} \frac{\partial^{(n)} f}{\partial x_1^{k_1}\ldots\partial x_p^{k_p}} \,|\, \mathbf{x} = \mathbf{x}^*,$$

where $\mathbf{x} = (x_1,...,x_p)'$, $\mathbf{x}^* = (x_1^*,...,x_p^*)'$ and f is a function of p variables, namely, $x_1, x_2, \cdots, x_p$.

4. Use of Delta Function to Obtain Discrete Probability Distributions

If X is a discrete random variable that assumes the values $a_1,...,a_n$ with probabilities $p_1,..., p_n$ respectively such that $\sum_{1\le i\le n} p_i = 1$, then, the probability mass function of X can be represented as $p(x) = \sum_{1\le i\le n} p_i \delta(x - a_i)$.

Now let us consider two discrete random variables X and Y which assume the values $a_1,...,a_n$ and $b_1,...,b_n$, respectively, and the joint probability $P(X = a_i, Y = b_j)$ is given by p_{ij} for $i = 1,...,m$ and $j = 1,...,n$ so that the joint probability mass function $p(x, y)$ is given by $p(x, y) = \sum_{1\le i\le m} \sum_{1\le j\le n} p_{ij} \delta(x - a_i)\delta(y - b_j)$.

Similarly, one can write down the joint probability distribution of any finite number of random variables in terms delta functions as follows:

Suppose we have k random variables $X_1,..., X_k$ with X_i taking values $a_{ij}, j = 1,2,...,n_i$ for $i = 1,2,...,k$ with probability $p_{i_1...i_k}$. Then, the joint probability mass function is

$$P(X_1 = x_1,..., X_k = x_k) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_k=1}^{n_k} p_{i_1\cdots i_k} \delta(x_1 - a_{1i_1}) \cdots \delta(x_k - a_{ki_k})$$

As an example, we may consider the situation of multinomial distributions. Let $X_1, X_2,...X_k$ follow multinomial distribution with parameters $n, p_1, p_2,..., p_k$. Then,

$$P(X_1 = i_1, \cdots, X_k = i_k) = \frac{n!}{i_1!\cdots i_k!} p_1^{i_1} \cdots p_k^{i_k}$$

where $i_1, \cdots, i_k$ add up to n and $p_1, \cdots p_k$ add up to 1. In terms of delta function, the joint probability mass function is

$$P(X_1 = x_1, X_2 = x_2,..., X_k = x_k)$$
$$= \sum_{i_1} \sum_{i_2} ... \sum_{i_k} \frac{n!}{i_1!i_2!...i_k!} p_1^{i_1} p_2^{i_2} ... p_k^{i_k} \delta(x_1 - i_1)\delta(x_2 - i_2)...\delta(x_k - i_k) .$$

We can also consider conditional probabilities and think of expressing them in terms of the δ-function. Let us go back to the example of the two discrete random variables X and Y, where X takes the values $a_1, a_2, \cdots, a_m$ and Y takes the values $b_1, b_2, \cdots, b_n$. Then, the conditional probability of $Y = y$ given $X = x$ is given by

$$p(y \mid x) = P(Y = y \mid X = x) = \frac{P(Y = y, X = x)}{P(X = x)}$$

$$= \frac{p(x, y)}{p(x)} = \frac{\sum_{1 \le i \le m} \sum_{1 \le j \le n} p_{ij} \delta(x - a_i) \delta(y - b_j)}{\sum_{1 \le i \le n} p_i \delta(x - a_i)} .$$

5. Densities of Transformations of Random Variables Using δ-function

If X is a continuous random variable with a density function $f(x)$ and if $Y = g(X)$ is a function of X, then the density function of Y, namely, $h(y)$ is given by

$$h(y) = \int_{-\infty}^{\infty} f(x) \delta(y - g(x)) dx .$$

We can extend this to the two-dimensional case. If X and Y are two continuous random variables with joint density function $f(x, y)$ and if $Z = \phi_1(X, Y)$ and $W = \phi_2(X, Y)$ are two random variables obtained as transformations from (X, Y), then the bivariate density function for Z and W is given by

$$h(z, w) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(z - \phi_1(x, y)) \delta(w - \phi_2(x, y)) dx dy ,$$

where z and w are the variables corresponding to the transformations $\phi_1(X, Y)$ and $\phi_2(X, Y)$. This has obvious extension to the general p-dimensional case.

Khuri (2004) gave an example of two independent Gamma random variables X and Y so that X and Y are gamma random variables with distributions $\Gamma(\lambda, \alpha_1)$ and $\Gamma(\lambda, \alpha_2)$ respectively. If we denote the densities as f_1 and f_2 respectively, then we have,

$$f_1(x) = \frac{\lambda^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1 - 1} e^{-\lambda x}, \quad \text{for} \quad x > 0$$

$$= 0, \qquad \text{for} \quad x \le 0 .$$

$$f_2(y) = \frac{\lambda^{\alpha_2}}{\Gamma(\alpha_2)} y^{\alpha_2 - 1} e^{-\lambda y} \ , \ \text{for} \quad y > 0$$

$$= 0 \, , \qquad\qquad \text{for} \quad y \le 0 \, .$$

In that case, if we define $Z = \dfrac{X}{X+Y}$ and $W = X + Y$, then, Z is distributed as Beta with parameters α_1 , α_2 and W is distributed as Gamma with parameter values 1 and $\alpha_1 + \alpha_2$. From now on, we shall use $\beta(a,b)$ to indicate a Beta random variable with positive parameters a, b and $\Gamma(c,k)$ to indicate a Gamma distribution with positive parameters c and k.

Now let us consider 3 random variables X_1, X_2, X_3 distributed independently so that X_i is distributed as gamma, i.e., $\Gamma(\frac{1}{2}, \alpha_i)$ for i = 1, 2, 3, and if we define

$$Y_1 = \frac{X_1}{X_1 + X_2} \, ,$$
$$Y_2 = \frac{X_1 + X_2}{X_1 + X_2 + X_3} \, ,$$
$$Y_3 = X_1 + X_2 + X_3 \, .$$

Then, we have,

Y_1 distributed as $\beta(\alpha_1, \alpha_2)$,

Y_2 distributed as $\beta(\alpha_1 + \alpha_2, \alpha_3)$,

Y_3 distributed as $\Gamma(\frac{1}{2}, \alpha_1 + \alpha_2 + \alpha_3)$.

This can be shown following exactly the same technique used in Khuri (2004) which is a one step generalization of the result proved by him. So, we define

$$\delta^*(x_1, x_2, x_3, y_1, y_2, y_3) = \delta(\frac{x_1}{x_1 + x_2} - y_1)\delta(\frac{x_1 + x_2}{x_1 + x_2 + x_3} - y_2)\delta(x_1 + x_2 + x_3 - y_3) \, .$$

The joint density of Y_1, Y_2 and Y_3 is given by

$$g(y_1, y_2, y_3) = \int_0^\infty \int_0^\infty \int_0^\infty f(x_1, x_2, x_3) \delta^*(x_1, x_2, x_3, y_1, y_2, y_3) dx_1 dx_2 dx_2$$
$$= \frac{1}{2^{\alpha_1 + \alpha_2 + \alpha_3} \Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \int_0^\infty \int_0^\infty \int_0^\infty x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} x_3^{\alpha_3 - 1} e^{-\frac{x_1 + x_2 + x_3}{2}} \delta^*(x_1, x_2, x_3, y_1, y_2, y_3) dx_1 dx_2 dx_3$$

Using properties of the delta function,

the innermost integral (integral with respect to x_1)

$$=\int_0^\infty x_1^{\alpha_1-1}e^{-\frac{x_1+x_2+x_3}{2}}\delta^*(x_1,x_2,x_3,y_1,y_2,y_3)dx_1$$

$$=\int_0^\infty x_1^{\alpha_1-1}e^{-\frac{x_1+x_2+x_3}{2}}\delta\Big(\frac{x_1}{x_1+x_2}-y_1\Big)\delta\Big(\frac{x_1+x_2}{x_1+x_2+x_3}-y_2\Big)\delta(x_1-(y_3-x_2-x_3))dx_1$$

$$=(y_3-x_2-x_3)^{\alpha_1-1}e^{-\frac{y_3}{2}}\delta\Big(\frac{y_3-x_2-x_3}{y_3-x_3}-y_1\Big)\delta\Big(\frac{y_3-x_3}{y_3}-y_2\Big).$$

Next, we use delta function properties to deal with the second integral (the integral with respect to x_2) as

$$\int_0^\infty x_2^{\alpha_2-1}(y_3-x_2-x_3)^{\alpha_1-1}e^{-\frac{y_3}{2}}\delta\Big(\frac{y_3-x_2-x_3}{y_3-x_3}-y_1\Big)dx_2$$

$$=\int_0^\infty x_2^{\alpha_2-1}\frac{(y_3-x_2-x_3)^{\alpha_1-1}}{y_3-x_3}e^{-\frac{y_3}{2}}\delta(x_2-(y_3-x_3)(1-y_1))dx_2$$

$$=(y_3-x_3)^{\alpha_1+\alpha_2-1}y_1^{\alpha_1-1}(1-y_1)^{\alpha_2-1}e^{-\frac{y_3}{2}}$$

Then, we use delta function properties for the outermost integral (without the constant terms) is

$$\int_0^\infty x_3^{\alpha_3-1}(y_3-x_3)^{\alpha_1+\alpha_2-1}y_1^{\alpha_1-1}(1-y_1)^{\alpha_2-1}e^{-\frac{y_3}{2}}\delta(x_3-y_3(1-y_2))\frac{1}{y_3}dx_3$$

$$=y_1^{\alpha_1-1}(1-y_1)^{\alpha_2-1}y_2^{\alpha_1+\alpha_2-1}(1-y_2)^{\alpha_3-1}y_3^{\alpha_1+\alpha_2+\alpha_3-1}e^{-\frac{y_3}{2}}$$

Finally, putting the constant terms together, we get

$$g(y_1,y_2,y_3)=\frac{1}{2^{\alpha_1+\alpha_2+\alpha_3}\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)}y_1^{\alpha_1-1}(1-y_1)^{\alpha_2-1}y_2^{\alpha_1+\alpha_2-1}(1-y_2)^{\alpha_3-1}y_3^{\alpha_1+\alpha_2+\alpha_3-1}e^{-\frac{y_3}{2}}.$$

This completes the proof.

6. Vector notations for delta functions in the multidimensional case

In the multidimensional case, if the transformation is linear, i.e., $\mathbf{Y} = \mathbf{AX}$ where $\mathbf{Y}$ and $\mathbf{X}$ are $m \times 1$ and $n \times 1$ vectors respectively and $\mathbf{A}$ is an $m \times n$ matrix, then we can express $g(\mathbf{y})$, the density of $\mathbf{Y}$**,** in vector notation in terms of $f(\mathbf{x})$, the density of $\mathbf{X}$ as follows

$$g(\mathbf{y}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(\mathbf{x}) \delta(\mathbf{y} - \mathbf{Ax}) d\mathbf{x}, \tag{2}$$

where

$$\delta(\mathbf{y}) = \delta(y_1) \dots \delta(y_m).$$

So,

$$\delta(\mathbf{y} - \mathbf{Ax}) = \delta(y_1 - \mathbf{a_1}^{\mathbf{T}} \mathbf{x}) \dots \delta(y_m - \mathbf{a_m}^{\mathbf{T}} \mathbf{x}),$$

where $\mathbf{a_1}^{\mathbf{T}}, \dots, \mathbf{a_m}^{\mathbf{T}}$ are the rows of the matrix $\mathbf{A}$. Now, in the one-dimensional set up, if a is a scalar, then $\delta(ax)$ is given by $\delta(ax) = \dfrac{\delta(x)}{|a|}$. Similarly, in the multidimensional set up, if $\mathbf{Y} = \mathbf{AX}$ as above and $\mathbf{A}$ is a nonsingular matrix so that $m = n$, then, we must have

$$\delta(\mathbf{y} - \mathbf{Ax}) = \frac{\delta(\mathbf{x} - \mathbf{A}^{-1}\mathbf{y})}{|\mathbf{A}|}. \tag{3}$$

This is because of the following: since the transformation is nonsingular, we have

$$g(\mathbf{y}) = \frac{1}{|\mathbf{A}|} f(\mathbf{A}^{-1}\mathbf{y})$$

and therefore, from (2)

$$f(\mathbf{A}^{-1}\mathbf{y}) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(\mathbf{x}) |\mathbf{A}| \delta(\mathbf{y} - \mathbf{Ax}) d\mathbf{x}.$$

But we know that

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(\mathbf{x}) \delta(\mathbf{x} - \mathbf{A}^{-1}\mathbf{y}) d\mathbf{x} = f(\mathbf{A}^{-1}\mathbf{y}).$$

Therefore, (3) follows. Similarly, if $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ where $\mathbf{A}$ is a nonsingular matrix and $\mathbf{b}$ is a vector, then, we have

$$\delta(\mathbf{y}-(\mathbf{AX}+\mathbf{b}))=\frac{\delta(\mathbf{x}-\mathbf{A}^{-1}(\mathbf{y}-\mathbf{b}))}{|\mathbf{A}|}.$$

Using this, one can conclude that if $\mathbf{X}$ is multivariate normal with $\boldsymbol{\mu}$ as the mean vector and $\sum$ as the covariance matrix, then, for a nonsingular transformation $\mathbf{A}$ and a constant vector $\mathbf{b}$, the transformed vector $\mathbf{Y}=\mathbf{AX}+\mathbf{b}$ follows multivariate normal with $\mathbf{A}\boldsymbol{\mu}+\mathbf{b}$ as the mean vector and $\mathbf{A}\sum\mathbf{A}^T$ as the variance-covariance matrix.

7. Transformation of Variables in the Discrete Case

Transformation of variables can be applied to the discrete case as well. If X is a discrete random variable taking the values $a_1,\cdots,a_n$ with probabilities $p_1,\cdots,p_n$ and if $Y=g(X)$ is a transformed variable, then the corresponding probability mass function for Y is given by

$$q(y)=\int_{-\infty}^{\infty}p(x)\delta(y-g(x))dx=\sum_{1}^{n}p_i\delta(y-g(a_i)).$$

In the two-dimensional case, if $p(x,y)$ is the joint probability mass function for the two variables X and Y, then $q(z,w)$, the joint probability mass function for the transformed pair $Z=\phi_1(X,Y)$ and $W=\phi_2(X,Y)$ is given by

$$\begin{aligned}q(z,w)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}p(x,y)\delta(z-\phi_1(x,y))\delta(w-\phi_2(x,y))dxdy\\&=\sum_{1}^{m}\sum_{1}^{n}p_{ij}\delta(z-\phi_1(a_i,b_j))\delta(w-\phi_2(a_i,b_j))\end{aligned}$$

Where the pair (X,Y) is discrete having the values (a_i,b_j) with probabilities p_{ij} for $i=1,2,...,n$ and $j=1,2,...,m$.

8. Moments and Moment Generating Functions

In the univariate set up, the k^{th} non-central moment of X is written as

$$\int_{-\infty}^{\infty}x^k p(x)dx=\int_{-\infty}^{\infty}x^k\sum_{1\le i\le n}p_i\delta(x-a_i)dx=\sum_{1\le i\le n}p_i\int_{-\infty}^{\infty}x^k\delta(x-a_i)dx=\sum_{1\le i\le n}p_i a_i^{\,k}.$$

In the bivariate set up, the non-central moment of order (k,l) for (X,Y) is given by

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^k y^l p(x,y)dxdy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^k y^l\sum_{1}^{m}\sum_{1}^{n}p_{ij}\delta(x-a_i)\delta(y-b_j)dxdy$$

$$=\sum_{i=1}^{m}\sum_{j=1}^{n}p_{ij}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^k y^l \delta(x-a_i, y-b_j)dxdy=\sum_{i=1}^{m}\sum_{j=1}^{n}a_i^{\,k}b_j^{\,l}p_{ij}\ .$$

For example, if (X,Y) follow trinomial distribution, then

$$p(x,y)=\sum_{k=0}^{n}\sum_{l=0}^{n-k}\frac{n!}{k!l!(n-k-l)!}p_1^{\,k}p_2^{\,l}(1-p_1-p_2)^{n-k-l}\delta(x-k)\delta(y-l)\ .$$

As a result, the corresponding non-central moment of order (r,s) is given by

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}x^r y^s p(x,y)dxdy=\sum_{k=0}^{n}\sum_{l=0}^{n-k}\frac{n!}{k!l!(n-k-l)!}p_1^{\,k}p_2^{\,l}(1-p_1-p_2)^{n-k-l}a_k^{\,r}b_l^{\,s}\ .$$

The most interesting part in Khuri's article is the representation of the density function of a continuous random variable X in terms of its non-central moments. Thus, if $f(x)$ is the density function for X, then $f(x)$ is represented as

$$f(x)=\sum_{0\le m<\infty}\frac{(-1)^m}{m!}\mu_m\delta^{(m)}(x),$$

where $\delta^{(m)}(x)$ is the generalized m^{th} derivative of $\delta(x)$ and μ_m is the m^{th} order non-central moment for the random variable X. One can see a proof of this result in Kanwal (1998). Let us briefly mention the technique used by him to derive the above expression. He showed that, for any real function ψ defined and differentiable of all orders in a neighborhood of zero,

$$\langle f,\psi\rangle=\langle\sum\nolimits_{0\le n<\infty}\frac{(-1)^n}{n!}\mu_n\delta^{(n)}(x),\psi\rangle,$$

where $\langle f,\psi\rangle$, the inner product between the functions f and ψ, is defined as

$$\langle f,\psi\rangle=\int_{-\infty}^{\infty}f(x)\psi(x)dx\,.$$

This leads us to conclude

$$f(x)=\sum\nolimits_{0\le n<\infty}\frac{(-1)^n}{n!}\mu_n\delta^{(n)}(x)....\qquad(4)$$

The crucial step in the proof was to use the Taylor's expansion in one-dimension for the function ψ. Thus, we have,

$$\psi(x)=\sum\nolimits_{0\le n<\infty}\frac{\psi^{(n)}(0)x^n}{n!}$$

Then,

$$\psi^{(n)}(0) = (-1)^n \int_{-\infty}^{\infty} \psi(x)\delta^{(n)}(x)dx = (-1)^n \langle \psi, \delta^{(n)} \rangle .$$

These two steps give us the relation (4).

When we move to the two-dimensional scenario, we shall have to use Taylor's series expansion for a two-dimensional analytic function ψ about the point (0,0) which is given by

$$\psi(x,y) = \sum_{0\le n<\infty} \frac{1}{n!}[x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}]^n \psi(x,y)|_{x=0,y=0}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{\infty} {}^nC_k \frac{\partial^n}{\partial x^k \partial y^{n-k}} \psi(x,y)|_{x=0,y=0} \; x^k y^{n-k} \tag{5}$$

Now, here also, we follow the same technique and so we compute $\langle f, \psi \rangle$ which is defined as

$$\langle f, \psi \rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\psi(x,y)dxdy$$

Now using Taylor’s series expansion of $\psi(x,y)$ about (x_0, y_0) from (5) and assuming that interchange of integrals with summations permissible, we get

$$\langle f, \psi \rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)[\sum_{0\le n<\infty} \frac{1}{n!} \sum_{0\le k\le n} {}^nC_k \frac{\partial^n}{\partial x^k \partial y^{n-k}} \psi(x,y)|_{x=0,y=0} \; x^k y^{n-k}]dxdy$$

$$= \sum_{0\le n<\infty} \sum_{0\le k\le n} \frac{1}{k!(n-k)!} \frac{\partial^n}{\partial x^k \partial y^{n-k}} \psi(x,y)|_{x=0,y=0} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) x^k y^{n-k} dxdy$$

$$= \sum_{0\le n<\infty} \sum_{0\le k\le n} \frac{1}{k!(n-k)!} \frac{\partial^n}{\partial x^k \partial y^{n-k}} \psi(x,y)|_{x=0,y=0} \langle f, x^k y^{n-k} \rangle .$$

Now we define the $(r,s)^{\text{th}}$ order non-central moment for the pair (X,Y) as

$$\mu_{r,s} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) x^r y^s dxdy = \langle f, x^r y^s \rangle .$$

Then from above,

$$\langle f, \psi \rangle = \sum_{0\le n<\infty} \sum_{0\le k\le n} \frac{1}{k!(n-k)!} \frac{\partial^n}{\partial x^k \partial y^{n-k}} \psi(x,y)|_{x=0,y=0} \; \mu_{k,n-k}$$

$$= \langle \sum_{0\le n<\infty} \sum_{0\le k\le n} \frac{1}{k!(n-k)!} \mu_{k,n-k} (-1)^n \delta^{(k)}(x)\delta^{(n-k)}(y), \psi \rangle .$$

The last equality follows because of the following relation

$$\langle \psi, \delta^{(k)}(x)\delta^{(n-k)}(y)\rangle = (-1)^n \frac{\partial^n}{\partial x^k \partial y^{n-k}} \psi(x,y)|_{x=0,y=0} .$$

Therefore, we have

$$f(x,y) = \sum_{0\le n<\infty} \sum_{0\le k<n} \frac{1}{k!(n-k)!} \mu_{k,n-k} (-1)^n \delta^{(k)}(x)\delta^{(n-k)}(y) .$$

Now, the non-central moment of order (r,s) is given by

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) x^r y^s dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^r y^s \sum_{0\le n<\infty} \sum_{0\le k\le n} \frac{1}{k!(n-k)!} \mu_{k,n-k} (-1)^n \delta^{(k)}(x)\delta^{(n-k)}(y) dxdy .$$

But, we also have

$$\int\int x^r y^s \delta^{(i)}(x)\delta^{(j)}(y) dxdy = 0, \text{ if } i \ne r \text{ or } j \ne s$$

$$= (-1)^{r+s} r! s!, \text{ if } i = r \text{ and } j = s .$$

Therefore, the non-central moment of order (r,s) reduces to $\mu_{r,s}$.

When we talk about the moment generating function in the one variable case, we have

$$\begin{aligned}
\phi(t) = E(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx} \sum_{0\le n<\infty} \frac{(-1)^n}{n!} \mu_n \delta^{(n)}(x) dx \\
&= \sum_{0\le n<\infty} \frac{(-1)^n}{n!} \mu_n \int_{-\infty}^{\infty} e^{tx} \delta^{(n)}(x) dx \\
&= \sum_{0\le n<\infty} \frac{(-1)^n}{n!} \mu_n (-1)^n \frac{d^n}{dx^n} e^{tx} |_{x=0} \\
&= \sum_{0\le n<\infty} \frac{\mu_n}{n!} t^n .
\end{aligned}$$

In the two-variable case, the moment generating function is given by

$$\begin{aligned}
\phi(s,t) = E(e^{sX+tY}) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{sx+ty} f(x,y) dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{sx+ty} \sum_{0\le n<\infty} \sum_{0\le k<n} \frac{1}{k!(n-k)!} \mu_{k,n-k} (-1)^n \delta^{(k)}(x)\delta^{(n-k)}(y) dxdy \\
&= \sum_{0\le n<\infty} \sum_{0\le k<n} \frac{1}{k!(n-k)!} \mu_{k,n-k} (-1)^n \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{sx+ty} \delta^{(k)}(x)\delta^{(n-k)}(y) dxdy
\end{aligned}$$

$$= \sum_{0\le n<\infty} \sum_{0\le k<n} \frac{1}{k!(n-k)!} \mu_{k,n-k} s^k t^{n-k}$$

$$= \sum_{0\le n<\infty} \sum_{0\le k<n} \frac{s^k}{k!} \frac{t^{n-k}}{(n-k)!} \mu_{k,n-k} .$$

In the general p-dimensional case, the moment generating function could be obtained exactly in the similar fashion so that, it is given by

$$\phi(t_1, t_2, \cdots, t_p) = \sum_{0\le n<\infty} \sum_{0\le k_1<n} \cdots \sum_{0\le k_p<n-k_1-\cdots-k_{p-1}} \frac{t_1^{k_1}}{k_1!} \frac{t_2^{k_2}}{k_2!} \cdots \frac{t_p^{k_p}}{k_p!} \mu_{k_1,k_2,\cdots,k_p} .$$

9. Concluding Remarks

The study of generalized functions is now widely used in applied mathematics and engineering sciences. The δ-function approach provides us with a unified approach in treating discrete and continuous distributions. This approach has the potential to facilitate new ways of examining some classical concepts in mathematical statistics. However, some interesting applications can be found in the paper by Pazman and Pronzato (1996). In this paper, the authors use delta function approach for densities of nonlinear statistics and for marginal densities in nonlinear regression. We are also looking forward to obtain some interesting applications of the delta function in statistics.

Acknowledgement:

I am sincerely thankful to Professor Lokenath Debnath in the University of Texas-Pan American for bringing this problem to my notice.

REFERENCES

Dirac, P.A.M. (1930). The Principles of Quantum Mechanics, Oxford University Press.

Hoskins, R.F. (1998). Generalized Functions, Ellis Horwood Limited, Chichester, Sussex, England.

Kanwal, R.P. (1998). Function Theory and Technique (2nd Edition), Boston, MA, Birkhauser.

Khuri, A.I. (2004). Applications of Dirac's delta function in statistics, International Journal of Mathematical Education in Science and Technology, 35, no. 2, 185-195.

Pazman, A and Pronzato, L. (1996). A Dirac-function method for densities of nonlinear statistics and for marginal densities in nonlinear regression, Statistics & Probability Letters, 26, 159-167.

Saichev, A.I. and Woyczynski, W.A. (1997). Distributions in the Physical and Engineering Sciences, Boston, MA, Birkhauser.