Scientific
Research

On Approximating Two Distributions from a Single Complex-Valued Function

William Dana Flanders, George Japaridze

Department of Epidemiology, Rollins School of Public Health, Atlanta, USA
Department of Physics, Clark Atlanta University, Atlanta, USA
E-mail: *flanders@sph.emory.edu*

Abstract

We consider the problem of approximating two, possibly unrelated probability distributions from a single complex-valued function ψ and its Fourier transform. We show that this problem always has a solution within a specified degree of accuracy, provided the distributions satisfy the necessary regularity conditions. We describe the algorithm and construction of ψ and provide examples of approximating several pairs of distributions using the algorithm.

1. Introduction

In this paper we consider the problem of reconstructing two different probability distributions from a single complex-valued function with a specified degree of accuracy. This problem was suggested by the interpretation of the wave function Ψ which is a solution of the Schrödinger equation. As is well known $\Psi^\star(x)\Psi(x)$ is the probability density in a position space and $\hat{\Psi}^\star(p)\hat{\Psi}(p)$ is the corresponding probability density in a momentum space, where $\Psi^\star(x)$ is the complex conjugate of $\Psi(x)$ and $\hat{\Psi}(p)$ is the Fourier transform of $\Psi(x)$ [1]. We raise the related question as to whether *any* two distributions can be approximated within a specified error from a single function by calculating its modulus or that of its Fourier transform, calculations like those used in Quantum Mechanics.

Our answer to this question is affirmative.

We prove this assertion by constructing a complex-valued function ψ which approximates two given distributions within a given error, provided the two distributions satisfy some regularity requirements specified below. In particular, we show that the cumulative distributions can be approximated pointwise using the indefinite integral of the modulus of ψ or of $\hat\psi$.

The paper is organized as follows. In the next section we list assumptions, conventions and notation. In Section 3 we first introduce an expression for evaluating the error in the approximation of the two given distributions from

the modulus of ψ or its Fourier transform. We then show how to construct this function. In the next section we list 4 pairs of distributions approximated using construction described in Section 3. In Section 5 we discuss the method, limitations and possible generalizations. In the appendix we sketch a proof that the function ψ approximates any two given distributions on subintervals as claimed and a proof that the cumulative distributions can be approximated pointwise (uniformly on closed intervals) using the indefinite integral of the modulus of ψ or of $\hat\psi$.

2. Definitions, Conventions and Assumptions

2.1. Assumptions

1) X and P are any two variables; $f_X(x)$ is the probability density function for X with corresponding cumulative distribution $f_X(x)$; $f_P(p)$ is the probability density function for P with corresponding cumulative distribution $F_P(p)$

2) $f_X(x)$ and $f_P(p)$ have finite variances and they (and their square roots) are continuously differentiable with Fourier transforms in L^2 and in L^1 (for definitions of L^1 and L^2 see [2]; throughout, we denote the norm in L^2 as $\| \ \|_2$)

3) for every $\varepsilon > 0$, there exist X_b and P_b such that: a) for all $x \in [-X_b, X_b]$ and for all $p \in [-P_b, P_b]$, $f_X(x) > 0$ and $f_P(p) > 0$; and, b) $F_X(-X_b) < \varepsilon/16$,

$F_P\left(-P_b\right)<\varepsilon/16, F_X\left(X_b\right)>1-\varepsilon/16,$ and $F_P\left(P_b\right)>1-\varepsilon/16.$

2.2. Definitions and Notations

1) Let n be any integer, set $J=2^n$, $\Delta x=X_b/J$, and $\Delta p=P_b/J$, where X_b and P_b are given in assumption 3; if necessary, increase n (and/or P_b) so that $F_P\left(P_b-\Delta p/2\right)>1-\varepsilon/16$; $2J$ is the number of intervals of width Δx and Δp in $\left[-X_b,X_b\right]$ and $\left[-P_b-\Delta p/2,P_b-\Delta p/2\right]$;

2) s is a number defined in the Construction below; it indicates the number of subintervals into which each interval Δx is divided. It is used to control the size of the error in the approximations;

3) $t=s/\left(\Delta x\Delta p\right)$;

4) j and k are integers, between $-sJ$, and $sJ-1$; l, m and m' are integers between $-J$, and $J-1$ or J, unless noted otherwise;

5) $\varphi_s\left(x\right)$ is a function of the form used in the "sinc approximation" [3,4]:

$$\varphi_s\left(x\right)=\sum_{j=-sJ}^{sJ-1}\sqrt{f_X\left(j\Delta x/s\right)}\,sinc\left(\frac{x-j\Delta x/s}{\Delta x/s}\right),$$

where $sinc\left(x\right)\equiv\dfrac{\sin\left(\pi x\right)}{\pi x}$;

6) $Q_x=\max f_X\left(x\right),$ for $x\in\left[-X_b,X_b\right]$;

$P_x=\min f_X\left(x\right)>0,$ for $x\in\left[-X_b,X_b\right]$;

$P_p=\min f_P\left(p\right)>0,$ for $p\in\left[-P_b-\Delta p/2,P_b\right]$;

$Q_x'=\max df_X\left(x\right)/dx,$ for $x\in\left[-X_b,X_b\right]$;

7) $a_j=\left[\dfrac{\Delta x}{s}f_X\left(j\Delta x/s\right)\right]^{1/2}$ for $j=-sJ,-sJ+1,\cdots,sJ-1$

8) $pr_{x;j,z}=\sum_{m=j}^{j+z}a_m^2,$ for $0\le z\le s\cdot J-j$;

9) $pr_{p;l}=\int_{l\Delta p}^{\left(l+1\right)\Delta p}f_P\left(z\right)dz$ for $-J\le l\le J-1$;

10) $\hat{f}\left(.\right)$ denotes the Fourier transform of a function $f\left(.\right)$, $f^*\left(.\right)$ denotes the complex conjugate of a function $f\left(.\right)$;

11) M_j is a function, defined recursively in the construction, mapping the integers $j\in\left\{-sJ,-sJ+1,\cdots,sJ-1\right\}$ to the numbers

$$\left\{-sJ/\Delta x,-s\left(J-1\right)/\Delta x,\cdots,s\left(J-1\right)/\Delta x,sJ/\Delta x\right\};$$

12) $m_j=M_j\cdot\Delta x/s$; m_j is an integer;

13) $S_l\equiv\left\{j:j\in\left\{-sJ,-sJ+1,\cdots,sJ-1\right\}\right\}$ and $M_j=\left\{l\cdot s/\Delta x\right\}$ for $l=-J,-J+1,\cdots,J$. Equivalently,

$S_l=f_M^{-1}\left(l\cdot s/\Delta x\right),$ where $f_M^{-1}\left(.\right)$ is the inverse mapping.

3. Construction

Given $f_X\left(x\right),f_P\left(p\right)$, ε and n, and X_b and P_b with the properties described in Assumptions we construct a complex-valued function ψ which satisfies the following two inequalities

$$\sum_{l=-J}^{J-1}\left|\int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p}\left(t\cdot\hat{\psi}\left(qt\right)\cdot\hat{\psi}^*\left(qt\right)-f_P\left(q\right)\right)dq\right|<\varepsilon \qquad (1)$$

and

$$\sum_{l=-J}^{J-1}\left|\int_{l\Delta x}^{\left(l+1\right)\Delta x}\left(\psi\left(x\right)\cdot\psi^*\left(x\right)-f_X\left(x\right)\right)dx\right|<\varepsilon \qquad (2)$$

where t is a scale parameter defined as

$$t=\frac{s}{\Delta x\Delta p} \qquad (3)$$

and s is defined to be the smallest integer so that Inequalities (4)-(8) are satisfied:

$$\frac{\Delta x}{s}Q_x<\frac{\varepsilon}{4J} \qquad (4)$$

$$4\left[\frac{\Delta x}{s}Q_x\right]^{1/2}\left(3+2\cdot log\left(2sJ\right)\right)^3<\frac{\varepsilon}{4\left(2J+1\right)^4} \qquad (5)$$

$$\left\|\varphi_s\left(x\right)-\sqrt{f_X\left(x\right)}\right\|_2<\varepsilon/8\left(2J+1\right)^4 \qquad (6)$$

$$\frac{\left(\Delta x\right)^2}{s}J\,|Q_x'|\,\sqrt{\frac{Q_x}{P_x}}\left(2+2\cdot log\left(2sJ\right)\right)^2\le\frac{\varepsilon}{4\left(2J+1\right)^4} \qquad (7)$$

$$\int_{-X_b}^{X_b}f_X\left(x\right)dx-\sum_{j=-sJ}^{sJ-1}f_X\left(j\Delta x/s\right)\frac{\Delta x}{s}=$$
$$\sum_{j=-sJ}^{sJ-1}\int_{j\Delta x/s}^{\left(j+1\right)\Delta x/s}f_X\left(x\right)dx-\sum_{j=-sJ}^{sJ-1}f_X\left(j\Delta x/s\right)\frac{\Delta x}{s}\le\frac{\varepsilon}{8} \qquad (8)$$

Many of the functions already defined, such as $\varphi_s\left(x\right)$, or to be defined, such as $\psi\left(x\right)$ and M_j, depend on s and n. For brevity we suppress explicit notation of that dependency.

The next step in the construction of ψ is to define subintervals. First we split the interval $\left[-X_b,X_b\right]$ into $2J$ intervals of length $\Delta x:\left[l\Delta x,\left(l+1\right)\Delta x\right]$, $l=-J,-J+1,\cdots,J-1$. Next we further divide each interval $\left[l\Delta x,\left(l+1\right)\Delta x\right]$ into s subintervals and

number them from $j = -sJ, -sJ+1, \cdots,$ to $sJ-1$. Define $(j = -sJ, -sJ+1, \cdots, sJ-1)$

$$a_j = \left[\frac{\Delta x}{s} f_X \left(\frac{j\Delta x}{s}\right)\right]^{1/2} \quad (9)$$

We continue the construction by introducing the function M_j which maps the integers $j \in \{-sJ, -sJ+1, \cdots, sJ-1\}$ to the numbers $\{-sJ/\Delta x, -s(J-1)/\Delta x, \cdots, s(J-1)/\Delta x, sJ/\Delta x\}$, using the following recursive algorithm: initialize j to $-sJ$ and l to $-J$. Then proceed with the following steps:

1) Set $pr_{x;j,z} = 0$ for $z = -1$ and set $pr_{x;j,z} = \sum_{m=j}^{j+z} a_m^2$ for each integer $z \in \{0, 1, \cdots, sJ - j\}$, and set $pr_{p;l} = \int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p} f_P(z) dz$ for $-J \le l \le J - 1$

2) If an integer z exists between -1 and $sJ - j$ such that the inequalities

$$0 \le pr_{p;l} - pr_{x;j,z} \le \varepsilon / (2(2J+1)),$$

and $pr_{x;j,z+1} - pr_{p;l} > 0$ hold, then set $k = z$; otherwise, set $k = sJ - j$

3) If $k \ge 0$ then set all $M_j, \cdots, M_{j+k}$ equal to $ls/\Delta x$

4) Set $j = j + k + 1$ and $l = l + 1$; if $j = sJ$, stop, otherwise go to step 5)

5) If $l + 1 = J$, set $M_j, M_j + 1, \cdots, M_{sJ-1}$ all to $sJ/\Delta x$ and stop; otherwise return to step 1).

Step 2) always returns a value of k. The iterative algorithm continues until stopping, either at step 4) or 5).

Properties of M_j that will be used subsequently include: for each $j = -sJ, \cdots, sJ - 1$, the mapping M_j is a non-decreasing function of j; each $M_j \cdot s/\Delta x$ is an integer (*i.e.*, $M_j \cdot \Delta x/s = m_j$ for some integer $m_j \in -J, \cdots, J$).

Using this definition of M_j, we can now define $\psi(x)$ appearing in Equations (1)-(2) as:

$$\begin{aligned}\psi(x) &= \sqrt{\frac{s}{\Delta x}} \sum_{j=-sJ}^{sJ-1} a_j \, sinc\left(\frac{x - j\Delta x/s}{\Delta x/s}\right) e^{2i\pi M_j x} \\ &= \sum_{j=-sJ}^{sJ-1} \sqrt{f_X\left(\frac{j\Delta x}{s}\right)} \, sinc\left(\frac{x - j\Delta x/s}{\Delta x/s}\right) e^{2i\pi M_j x}\end{aligned} \quad (10)$$

Apart from $\exp(2i\pi M_j x)$ and the sJ th term in the series (10), the function $\psi(x)$ represents the Whittaker-Shannon interpolation formula [5,6] for $\sqrt{f_X(x)}$.

In addition to approximations given in (1)-(2) we also

prove uniform approximations to the cumulative distributions: let us define

$$F_P(p) = \int_{-P_b - \Delta p/2}^{p} f_P(q) dq$$

$$\tilde{F}_P(p) = \int_{-P_b - \Delta p/2}^{p} t \cdot \hat{\psi}(qt) \cdot \hat{\psi}^*(qt) dq \quad (11)$$

and $F_X(x) = \int_{-X_b}^{x} f_X(z) dz$

$$\tilde{F}_X(x) = \int_{-X_b}^{x} \psi(z) \cdot \psi^*(z) dz \quad (12)$$

Then, it can be shown that with $f_X(x), f_P(p)$, and ε^* given and with X_b and P_b having the properties described in Assumptions, one can choose n large enough that the complex-valued function ψ given by the construction can be used to uniformly approximate the cumulative distributions $F_X(x)$ and $F_P(p)$ for $x \in [-X_b, X_b]$, and for $p \in [-P_b - \Delta p/2, P_b - \Delta p/2]$:

$$\left|F_X(x) - \tilde{F}_X(x)\right| \le \varepsilon^* \quad (13)$$

and

$$\left|F_P(p) - \tilde{F}_P(p)\right| \le \varepsilon^* \quad (14)$$

For derivation of (1) and (13) see Appendix. Approximations (2) and (14) are derived in a similar fashion.

4. Examples

To illustrate the approximation, we compare the integral of the probability density function over a series of intervals with the correpsonding approximation generated by $\psi$ from (1), (2) and (10). Specifically, in P space we plot $\int_{p-\Delta p/2}^{p+\Delta p/2} f_P(q) dq$ and the corresponding approximation $\int_{p-\Delta p/2}^{p+\Delta p/2} t \cdot \hat{\psi}(qt) \hat{\psi}^*(qt) dq$. Similarly, in X space we plot $\int_{x-\Delta x/2}^{x+\Delta x/2} f_X(q) dq$ and the corresponding approximation $\int_{x-\Delta x/2}^{x+\Delta x/2} \psi(q) \cdot \psi^*(q) dq$; we approximate the last integral by $\psi(x)\psi^*(x)\Delta x$, as Δx is small.

We applied the method to four pairs of distributions; for each pair, we use a single ψ (and its Fourier transform) to approximate them with $n = 7$, $s = 64$ and $\Delta x = \Delta p = 5/2^n \approx 0.039$:

1) **Figure 1** illustrates the approximation using Equation (10), for the distributions

$$f_X(x) = \frac{4}{\sqrt{2\pi}}\left[0.75 e^{-8\left(x+\frac{1}{2}\right)^2} + 0.25 e^{-8\left(x-\frac{1}{2}\right)^2}\right]$$
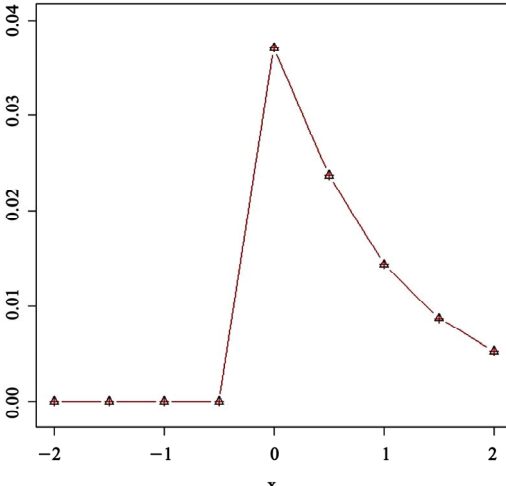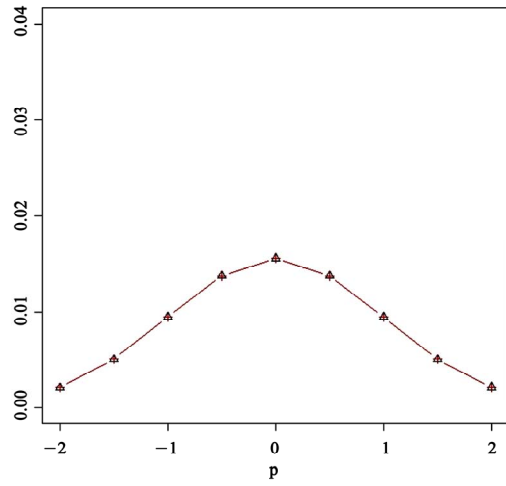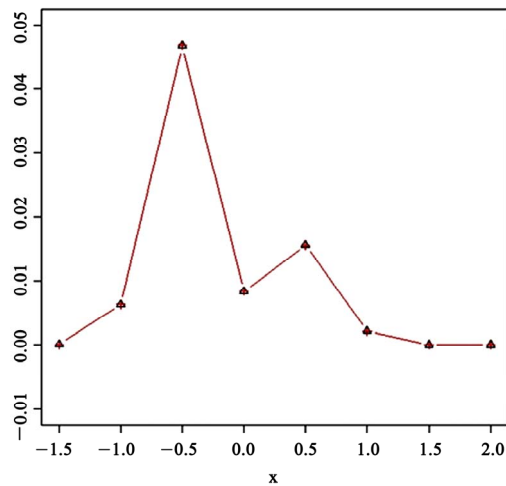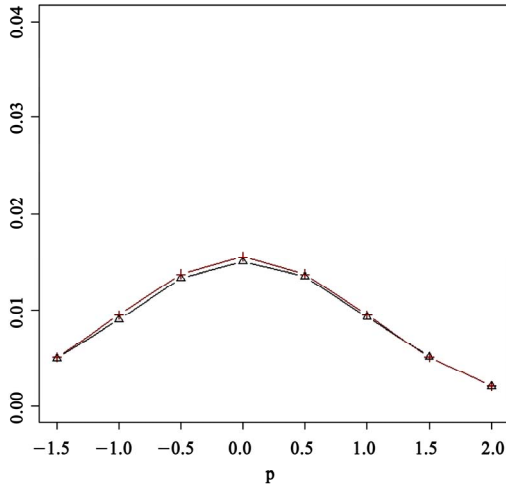
Figure 1. Approximation to the integrals $\int_{p-\Delta p/2}^{p+\Delta p/2} f_P(q)dq$ **(top) and** $\int_{x-\Delta x/2}^{x+\Delta x/2} f_X(q)dq$ **(bottom) based on Equation (10) for the pair of distributions** $f_P(p)=\exp(-p^2)\big/\sqrt{2\pi}$ **, and**

$$f_X(x)=\frac{4}{\sqrt{2\pi}}\left(0.75e^{-8\left(x+\frac{1}{2}\right)^2}+0.25e^{-8\left(x-\frac{1}{2}\right)^2}\right).$$

and $f_P(p)=\dfrac{1}{\sqrt{2\pi}}e^{-p^2/2}$

2) **Figure 2** illustrates the approximation using Equation (10), for the distributions

$$f_X(x)=e^{-x},\ x\ge 0\ \text{ and }\ f_P(p)=1\sqrt{2\pi}e^{-p^2/2}$$

3) **Figure 3** illustrates the approximation using Equation (10), for the distributions

$$f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\ \text{ and }\ f_P(p)=e^{-p},\ p\ge 0$$
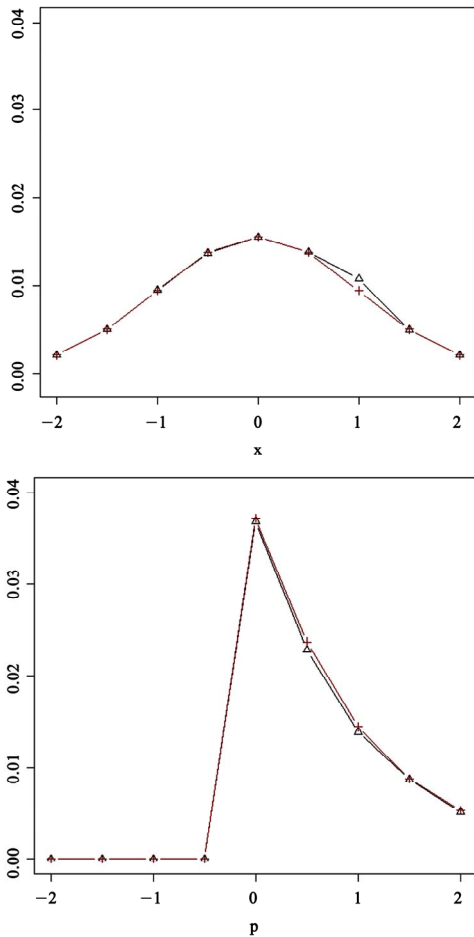
4) **Figure 4** illustrates the approximation using

Figure 2. Approximation to the integrals $\int_{p-\Delta p/2}^{p+\Delta p/2} f_P(q)dq$ **(top) and** $\int_{x-\Delta x/2}^{x+\Delta x/2} f_X(q)dq$ **(bottom) based on Equation (10) for the pair of distributions** $f_P(p)=\dfrac{1}{\sqrt{2\pi}}e^{-p^2/2}$ **and** $f_X(x)=e^{-x}$ **for** $x\ge 0$ **and** $f_X(x)=0$ **for** $x<0$ **.**
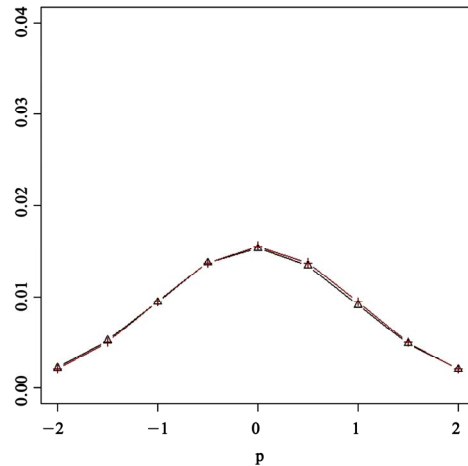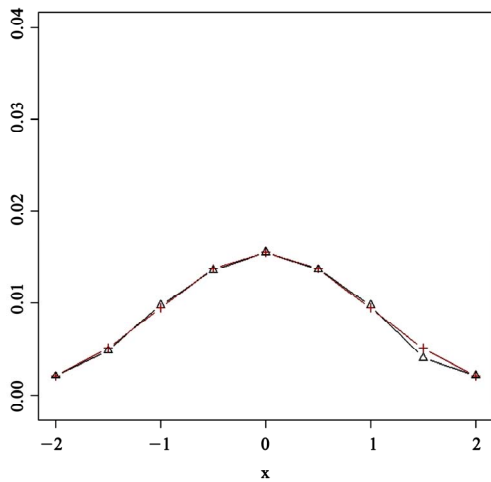
Equation (10), for the distributions

$$f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\ \text{ and }\ f_P(p)=\frac{1}{\sqrt{2\pi}}e^{-p^2/2}$$

5. Discussion

Based on the algorithm described in Section 3 we have shown how to approximate two given distributions from a single complex-valued function with a specified accuracy. The main result of our paper is given in expressions (1-2), (10) and (13-14). Expression (10) can be viewed as an extension of well-known numerical methods based on the sinc function [3], providing a

AM

Figure 4. Approximation to the integrals $\int_{x-\Delta x/2}^{x+\Delta x/2} f_X(q)dq$ **(bottom) and** $\int_{x-\Delta x/2}^{x+\Delta x/2} f_X(q)dq$ **(top) based on Equation (10) for the pair of distributions** $f_P(p) = \dfrac{1}{\sqrt{2\pi}}e^{-p^2/2}$ **, and** $f_X(x) = \dfrac{1}{\sqrt{2\pi}}e^{-x^2/2}$ **.**

Figure 3. Approximation to the integrals $\int_{p-\Delta p/2}^{p+\Delta p/2} f_P(q)dq$ **(bottom) and** $\int_{x-\Delta x/2}^{x+\Delta x/2} f_X(q)dq$ **(top) based on Equation (10) for the pair of distributions** $f_X(x) = \dfrac{1}{\sqrt{2\pi}}e^{-x^2/2}$ **, and** $f_P(p) = \exp(-p)$ **for** $p \geq 0$ **and** $f_P(p) = 0$ **for** $p < 0$ **.**

method for approximating two distributions simultaneously.

The construction and arguments given here show that there always exists a function of the form given by (10) which approximates two given distributions with aspecified degree of accuracy provided our assumptions (1-3) are satisfied. For successively higher degrees of accuracy, the method allows construction of a sequence of approximating functions but this sequence may not necessarily converge, say in L^2, to a limiting function.

A modification of our approach or a similar method based not on the sinc function but on some other basis or wavelet series [7] might be used to approximate the two distributions and converge to a limiting function. For example, one might attempt to use the Hermite polynomials multiplied by the appropriate Gaussian function as a basis.

It remains speculative, however, whether our approach can be modified so as to approximate arbitrary pairs of distributions with a basis other than the sinc functions. Further speculation about this potential generalization is beyond the scope of this paper. We would like to stress that the main goal of the present work is not to find the best numeric approximation (though we can always meet an increasing demand in accuracy) but to establish the existence of an algorithm allowing uniform approximation of the cumulative distributions.

6. References

[1]   J. von Neumann, "Mathematical Foundations of Quantum

Mechanics," Princeton University Press, Princeton, 1955.

[2] W. Rudin, "Functional Analysis," McGraw-Hill, New York, 1991.

[3] F. Stenger, "Summary of Sinc Numerical Methods," *Journal of Computational and Applied Mathematics*, Vol. 121, No. 1-2, September 2000, pp. 379-420.

[4] P. L. Butzer, J. R. Higgins and R. L. Stens, "Classical and Approximate Sampling Theorems; Studies in the $L^2(\mathrm{R})$ and the Uniform Norm," *Journal of Approximation Theory*, Vol. 137, No. 2, December 2005, pp. 250-263.

[5] J. M. Whittaker, "Interpolation Function Theory," Cambridge Tracts in Mathematics and Mathematical Physics, No. 33, Cambridge University Press, Cambridge, 1935.

[6] C. E. Shannon, "Communication in the Presence of Noise," *Proceedings of Institute of Radio Engineers*, Vol. 37, No. 1, January 1949, pp. 10-21.

[7] I. Doubechis, "Ten Lectures on Wavelets," CBMS-NSF Regional Conference Series for Applied Mathematics, 1992.

Appendix

We sketch the derivation of (1) and (13). First we write down the Fourier transform of $\psi(x)$:

$$\hat{\psi}(q) = \int \psi(x) e^{-i \cdot 2\pi \cdot x \cdot q} dx$$

$$= \sqrt{\frac{s}{\Delta x}} \int \sum_{j=-sJ}^{sJ-1} a_j e^{-i \cdot 2\pi \cdot M_j \cdot x} sinc\left(\frac{x - j\Delta x/s}{\Delta x/s}\right) e^{-i \cdot 2\pi \cdot x \cdot q} dx \quad (1)$$

$$= \sqrt{\frac{\Delta x}{s}} \sum_{j=-sJ}^{sJ-1} a_j R\left(\frac{\Delta x}{s}(q - M_j)\right) e^{-i \cdot 2\pi \cdot j \frac{\Delta x}{s}(q - M_j)}$$

In (1) R is a rectangular function:

$$R(q) = \begin{cases} 1 & \text{when } -\frac{1}{2} < q < \frac{1}{2}, \\ \frac{1}{2} & \text{when } q = \pm\frac{1}{2}, \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

We consider

$$I(l) \equiv \int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p} t \cdot \hat{\psi}(qt) \cdot \hat{\psi}^*(qt) dq, \quad (3)$$

the integral of $t \cdot \hat{\psi}(qt) \cdot \hat{\psi}^*(qt)$ over the interval $\left[\left(l-\frac{1}{2}\right)\Delta p, \left(l+\frac{1}{2}\right)\Delta p\right]$, for each $l \in -J, -J+1, \cdots, J-1$. Using expression (1), rescaling the integration variable, and using definition of integer $m_j : m_j = \Delta x M_j / s$, for this integral we obtain:

$$I(l) = \int_{\left(l-\frac{1}{2}\right)}^{\left(l+\frac{1}{2}\right)} dq' \sum_{j,k=-sJ}^{sJ-1} a_j a_k R(q' - m_j)$$

$$R(q' - m_k) e^{-2i\pi\left((j-k)q' + \left(j \cdot m_j - k \cdot m_k\right)\right)}$$

where $a_j = \left[\Delta x f_X(j\Delta x/s)/s\right]^{1/2}$.

Property (2) of the rectangular function R implies that each j, k term in the integrand is either 0 for the full range of integration or vanishes for all but a subinterval of length 1. The latter holds only if $m_j = m_k$ and we have: $l \le m_j \pm \frac{1}{2} \le (l+1)$ if and only if $l = m_j$, for some integer m_j. After interchanging the order of integration and summation, each integral in the sum is 0 due to the factor $e^{i \cdot 2\pi \cdot (j-k) \cdot q}$, unless $j = k$. On the other hand, if $j = k$ the integral is just a_j^2. Thus, we can write $(l = -J, -J+1, \cdots, J-1)$:

$$I(l) = \int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p} t \cdot \hat{\psi}(qt) \cdot \hat{\psi}^*(qt) dq = \sum_{j \in S_l} a_j^2, \quad (4)$$

where we recall the definition of S_l: S_l is the set of all $j \in -sJ, -sJ+1, \cdots, sJ$ for which $M_j \cdot (\Delta x/s) = l$.

There are $2J$ terms $I(l)$ of the form given in (4); the approximation given by Inequality (1) will be proven if we can show that for $j = -J, -J+1, \cdots, J-1$ the following relation holds:

$$\left|\sum_{j \in S_l} a_j^2 - pr_{p;l}\right| = \sum_{j \in S_l}\left|a_j^2 - \int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p} f_P(z) dz\right| \le \frac{\varepsilon}{2J} \quad (5)$$

Either Inequality (5) holds for $l = -J, -J+1, \cdots, J-1$, and we are done, or there is a smallest l violating (5), say l_1, such that $\sum_{j \in S_{l_1}} |a_j^2 - \int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p} f_p(z) dx| > \varepsilon/(2J)$.

This implies that $j_1 \ge sJ - 1$, where $j_1 = \max\{j : j \in S_l\}$. Indeed, if this were not the case (i.e., $j_1 < sJ - 1$), than M_{j_1+1} would have been defined by the algorithm described in section 3 as l_1 since $a_{j_1+1} \le \varepsilon/(2J+1)$ and $j_1 + 1$ would also have been in S_{l_1}. This would be a contradiction since j_1 was supposed to be the largest element of S_{l_1}. Therefore, $j_1 \ge sJ - 1$.

By choice of s in section 3, requirement (8): $\left| \sum_{l=-J}^{J-1} \sum_{j \in S_l} a_j^2 - \int_{j\Delta x/s}^{(j+1)\Delta x/s} f_X(x) dx \right|$ is less or equal $\varepsilon/8$, and because $f_X(x)$ integrates over the interval $\left[-X_b, X_b \right]$ to at least $1 - \varepsilon/16$ (Assumption 3), we have:

$$\sum_{j=-sJ}^{sJ-1} a_j^2 \geq 1 - \varepsilon/16 - \varepsilon/8 > 1 - \varepsilon/4 \qquad (6)$$

Since $f_P(p)$ is a probability density function and by Assumption 3 and the algorithm described in section 3:

$$1 \geq \int_{-P_b - 12\Delta p}^{\left(l_1 + \frac{1}{2}\right)\Delta p} f_P(z) dz =$$
$$\sum_{l=-J}^{l_1} \int_{\left(l + \frac{1}{2}\right)\Delta p}^{\left(l + \frac{1}{2}\right)\Delta p} f_P(z) dz \geq \sum_{l=-J}^{l_1} \sum_{j \in S_l} a_j^2 = \sum_{j=-sJ}^{sJ-1} a_j^2 \geq 1 - \varepsilon/4. \qquad (7)$$

The last inequality above is simply Inequality (6). By the Construction, $\int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p} f_P(z) dz \geq \sum_{j \in S_l} a_j^2$ for every l. This fact and (7) imply:

$$\int_{-P_b - \frac{1}{2}\Delta p}^{\left(l_1 + \frac{1}{2}\right)\Delta p} f_P(z) dz - \sum_{l=-J}^{l_1} \sum_{j \in S_l} a_j^2 =$$
$$\sum_{l=-J}^{l_1} \left| \int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p} f_P(z) dz - \sum_{j \in S_l} a_j^2 \right| \leq \frac{1}{4}\varepsilon \qquad (8)$$

Further, the upper bound of 1 in expression (6) implies that:

$$\int_{\left(l_1 + \frac{1}{2}\right)\Delta p}^{P_b - \frac{1}{2}\Delta p} f_P(z) dz = \sum_{l=l_1+1}^{J-1} \int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p} f_P(z) dz \leq \frac{3}{4}\varepsilon \qquad (9)$$

Now S_l is empty for $l > l_1$ since $M_{sJ-1} = l_1$ and M_j is non-decreasing. In other words $\sum_{j \in S_l} a_j^2 = 0$ for $l > l_1$. This fact and the inequality in Expression (9) yield:

$$\left| \sum_{l=l_1+1}^{J-1} \int_{l\Delta p}^{(l+1)\Delta p} f_P(z) dz - \sum_{j \in S_l} a_j^2 \right| = \left| \sum_{l=l_1+1}^{J-1} \int_{\left(l-\frac{1}{2}\right)\Delta p}^{\left(l+\frac{1}{2}\right)\Delta p} f_P(z) dz - 0 \right| \leq \frac{3}{4}\varepsilon \qquad (10)$$

Finally, Inequalities (8) and (10) lead to the desired result, proving Inequality (1).

Proof of Inequality (2) requires no new qualitative features but straightforward tedious calculations; we do not cite it here.

To prove the uniform approximation for the cumulative distributions, Inequality (13), choose n large enough that $Q_x \cdot \Delta x = Q_x \cdot X_b / 2^n < \varepsilon^* / 2$, construct $\psi(x)$ so that Inequalities (1) and (2) hold with $\varepsilon = \varepsilon^* / 2$ and let $\delta(x) = \left| F_X(x) - \tilde{F}_X(x) \right|$. We now show that $\delta(x) \leq \varepsilon^*$ for all $x \in \left[j\Delta x, (j+1)\Delta x \right]$ and for $j = -J, -J+1, \cdots, J-1$. Suppose the maximum of $\left| F_X(x) - \tilde{F}_X(x) \right|$ occurs at $X = x_0$, for $x_0 \in \left[j\Delta x, (j+1)\Delta x \right]$. Let M_{F_j} denote $\max\left\{ F_X(x), x \in \left[j\Delta x, (j+1)\Delta x \right] \right\}$. If $F_X(x_0) > \tilde{F}_X(x_0)$, then

$$\delta(x) \leq M_{F_j} - \tilde{F}_X(j\Delta x) \qquad (11)$$

since $\tilde{F}_X(x)$ is non-decreasing. Further, the mean value theorem implies $M_{F_j} \leq F_X(j\Delta x) + Q_x \cdot \Delta x$ since Q_x is the maximum of $f_X(x)$. Using this result in (11) gives

$$\delta(x) \leq F_X(j\Delta x) + Q_x \cdot \Delta x - \tilde{F}_X(j\Delta x) \leq \varepsilon^*/2 + \varepsilon^*/2 = \varepsilon^* \qquad (12)$$

since $\left| F_X(j\Delta x) - \tilde{F}_X(j\Delta x) \right| \leq \varepsilon^*/2$ follows from (1) and $| Q_x \cdot \Delta x | \leq \varepsilon^*/2$ by choice of n.

On the other hand, if $F_X(x_0) < \tilde{F}_X(x_0)$, then

$$\delta(x) \leq \tilde{F}_X\left((j+1)\Delta x\right) - F_X(j\Delta x) \leq F_X\left((j+1)\Delta x\right)$$
$$+ \varepsilon^*/2 - F_X(j\Delta x) \leq F_X(j\Delta x) + Q_x \cdot \Delta x + \varepsilon$$
$$- F_X(j\Delta x) \leq \varepsilon^*/2 + \varepsilon^*/2 = \varepsilon^* \qquad (13)$$

The chain of estimates above follows from the fact that $\tilde{F}_X(x)$ is non-decreasing, $Q_x n$ is the maximum of $f_X(x)$, and by choice of n.

The proof for $\left| F_P(p) - \tilde{F}_P(p) \right|$ is similar.

AM