Approximating distributions from moments

R. F. Pawula
Random Applications, Inc., 515 South Junction Avenue, Montrose, Colorado 81401
(Received 8 June 1987)

A method based upon Pearson-type approximations from statistics is developed for approximating a symmetric probability density function from its moments. The extended Fokker-Planck equation for non-Markov processes is shown to be the underlying foundation for the approximations. The *approximation* is shown to be *exact* for the beta probability density function. The applicability of the general method is illustrated by numerous pithy examples from linear and nonlinear filtering of both Markov and non-Markov dichotomous noise. New approximations are given for the probability density function in two cases in which exact solutions are unavailable, those of (i) the filter-limiter-filter problem and (ii) second-order Butterworth filtering of the random telegraph signal. The approximate results are compared with previously published Monte Carlo simulations in these two cases.

I. INTRODUCTION

The problem of approximating a probability density function from a finite number of its moments is an old one dating back to the last century.[1–7] Some of the earliest works in this area are the classic investigations of K. Pearson, well known in mathematical statistics. Over the years, Pearson's work has been the subject of criticism partly because of its supposed *ad hoc* nature—this criticism being leveled in spite of the generality of Pearson's results in that his system contained many of the best-known continuous distributions. An attempt to extend Pearson's ideas was made in the voluminous paper by Hansmann in 1934 (Ref. 2) (this work was evidently done under Pearson's tutelage as is evidenced by the acknowledgment at the end of the paper). However, extremely little use has been made of Hansmann's results. In part, this is because Hansmann failed to correctly solve his differential equation in all cases, and his error went undetected for over 50 years.[8,9] In this paper, we provide a more solid foundation for these Pearson-type approximations based upon extended Fokker-Planck equations for non-Markov processes. We also make further extensions along the lines begun by Hansmann. Our *approximate* method is shown to be *exact* for a certain class of probability density functions and some limitations of the method are explored. The method is first validated by some known examples from linear and nonlinear filtering of Markov and non-Markov dichotomous noise, and then applied to further examples in which analytic results are not available.

Many different techniques for approximating a probability density function from its moments have been proposed and some studied in great detail.[3,7] These include quadratures, orthogonal polynomial expansions, Edgeworth and Gram-Charlier–type series, Fisher-Cornish–type expansion, continued fraction expansions, and the use of modified moments. No method appears to work well in all cases and the most common two ways in which they fail are that the approximations are either negative over some domains or that they contain extraneous oscillations. In addition, sometimes knowledge of the moments seems to be necessary to inordinately high degrees of precision. No attempt is made here to give any review of existing techniques except for general comments such as those just made. One attractive feature of the method to be presented here is that it gives relatively simple expressions for the approximating probability densities. Also, it is easy to see by inspection whether or not the approximation ever becomes negative.

Symmetric probability densities are treated in most of the work but indications are given as to how to make extensions to the asymmetric case. The treatments are somewhat simpler in the symmetric case and this was the route taken by Hansmann. We also restrict attention to densities which exist over finite intervals.

The details of our method are presented in Sec. II which begins with the extended Fokker-Planck equation for non-Markov processes. The original Pearson, Hansmann, and later Pearson-type approximations are then shown to follow directly as approximations made in the Fokker-Planck equation itself. The approximations are noted to be exact for certain probability density functions and cases in which the approximations may not work well are discussed. Sec. III contains examples taken from known probability density functions which arise in first-order linear and nonlinear systems driven by Markov and non-Markov dichotomous noise. Further examples from physical problems in which the probability density functions are unknown are considered in Sec. IV, and extensions to the asymmetric case are discussed in Sec. V. The final sections summarize and discuss the results.

All of the probability density functions to be encountered vanish outside their intervals of definition and, as a matter of convenience, this *will not* be explicitly stated each time an expression for a density function is given. $p(y)$ will denote the probability density function of the dynamical variable $y(t)$, $\hat{p}(y)$ will be used to indicate an approximation to $p(y)$, and $f(t)$ will be reserved for the probability density function of the intervals between transitions of the dichotomous noise.

II. GENERAL FORMULATION

A. Extended Fokker-Planck equation

In general, a stationary random process $y(t)$, not necessarily Markov, has a marginal probability density function $p(y)$ which satisfies the *νth-order extended* Fokker-Planck equation[10–13]

$$\frac{1}{2}\frac{d^2}{dy^2}[B(y)p(y)]-\frac{d}{dy}[A(y)p(y)]=0 , \tag{1}$$

in which the conditional moments $A(y)$ and $B(y)$ are defined as

$$A(y)=\lim_{\epsilon\to 0}\frac{1}{\epsilon^{\nu}}E[y(t+\epsilon)-y(t)\,|\,y(t)] , \tag{2}$$

$$B(y)=\lim_{\epsilon\to 0}\frac{1}{\epsilon^{\nu}}E[\{y(t+\epsilon)-y(t)\}^2\,|\,y(t)] . \tag{3}$$

When $\nu=1$, this reduces to the classical Fokker-Planck equation; however, the classical equation degenerates to $0=0$ in some cases [because $A(y)$ and $B(y)$ are each zero] (Ref. 10) and it is sometimes necessary to use $\nu>1$. It can be shown that there exists only one value of ν for which (1) is nondegenerate. In three cases, those of the RC filter and first-order nonlinear systems driven by dichotomous Markov noise[14,15] and the filter-limiter-filter driven by white Gaussian noise,[16] it is found that $\nu=2$ is the necessary choice.

The first integral of (1) is

$$\frac{d}{dy}[B(y)p(y)]-2A(y)p(y)=\kappa , \tag{4}$$

where κ is a constant. Upon integrating this equation over the range of y and noting, from (2), that $E[A(y)]=0$ if the limit and expectation can be interchanged, we find

$$\kappa=\frac{B(y_{\max})p(y_{\max})-B(y_{\min})p(y_{\min})}{y_{\max}-y_{\min}} . \tag{5}$$

It seems plausible to suppose that $\kappa=0$ for any random process; however, we have been unable to prove this in general. We can say that $\kappa=0$ for any statistically symmetric process since the numerator on the right-hand side of (5) is then identically zero. For certain first-order nonlinear systems driven by dichotomous non-Markov noise, it has been shown[15,17] that $B(y)p(y)$ when evaluated at the level $y=l$ is proportional to the average number of crossings of $y(t)$ with the level $y=l$. In these cases, which can be asymmetric, $B(y)p(y)$ must vanish at the extreme values $y_{\min}$ and $y_{\max}$ and we again have $\kappa=0$. We conjecture $\kappa=0$ (without loss of generality in the symmetric case) and write (4) as

$$\frac{d}{dy}[B(y)p(y)]-2A(y)p(y)=0 . \tag{6}$$

B. Pearson-type differential equations

The above equation can further be put into the form

$$\frac{1}{p(y)}\frac{dp(y)}{dy}=\frac{y-\bar{y}}{H(y)} , \tag{7}$$

where $\bar{y}=E[y]$ and

$$H(y)=\frac{(y-\bar{y})B(y)}{2A(y)-B'(y)} . \tag{8}$$

If $H(y)$ were expanded in a power series

$$H(y)=h_0+h_1y+h_2y^2+\cdots , \tag{9}$$

and only a finite number of terms kept, we would expect that the resulting solution to (7) would be an approximation to $p(y)$ which would get better and better as more terms are retained. This is the justification for Pearson's approach (and has been noted as such previously for Markov processes[6]). When powers up to the second are retained, (7) is then of the form assumed by Pearson in his classic work. As these arguments show, the extended Fokker-Planck equation for non-Markov processes provides a solid foundation for Pearson's approach in the general case.

Hansmann extended Pearson's work in the symmetric case by approximating $H(y)$ by a fourth-order polynomial; however, he did not correctly solve the ensuing differential equation in all cases.[8] We will further extend Pearson's ideas in the symmetric case by approximating $H(y)$ by higher-order polynomials. In general, we note, the Fokker-Planck conditional moments $A(y)$ and $B(y)$ are unknown [if they were known, we could just solve the Fokker-Planck equation for $p(y)$] and so the function $H(y)$ is not known. Pearson's great insight was not only in formulating his differential equation but also in constructing a scheme for determining the coefficients in the approximating polynomial in terms of the moments $\mu_n=E[y^n]$. The types of approximations with which we shall be concerned in the symmetric case will satisfy equations such as

$$\frac{1}{\hat{p}}\frac{d\hat{p}}{dy}=\frac{y}{c_0+c_2y^2} \tag{10a}$$

(Pearson, 1895),

$$\frac{1}{\hat{p}}\frac{d\hat{p}}{dy}=\frac{y}{c_0+c_2y^2+c_4y^4} \tag{10b}$$

(Hansmann, 1934),

$$\frac{1}{\hat{p}}\frac{d\hat{p}}{dy}=\frac{y}{(y_{\max}^2-y^2)(c_0+c_2y^2)} , \tag{10c}$$

$$\frac{1}{\hat{p}}\frac{d\hat{p}}{dy}=\frac{y}{c_0+c_2y^2+c_4y^4+c_6y^6} , \tag{10d}$$

$$\frac{1}{\hat{p}}\frac{d\hat{p}}{dy}=\frac{y}{(y_{\max}^2-y^2)(c_0+c_2y^2+c_4y^4)} . \tag{10e}$$

As we will presently see, when the domain of the $y(t)$ process is known to be $(-y_{\max}, y_{\max})$ it will be necessary to assume forms like (10c) and (10e) in which the boundary is a root of the denominator polynomial. When the boundary is unknown or infinite, the forms (10a), (10b), and (10d) are more appropriate but may lead to approximations which have probability mass outside the actual ranges of the variables. In the examples to follow in later sections, we will confine most, but not all, of our attention to the sixth-order polynomial approximation of type (10e); however, for some of the theory to be given in the remainder of this section it will be more convenient to use the general form

$$\frac{1}{\hat{p}}\frac{d\hat{p}}{dy}=\frac{y}{(y_{\max}^2-y^2)\sum_{n=0}^{N} c_{2n}y^{2n}} . \tag{11}$$

C. Polynomial coefficients in Pearson-type approximations (Refs. 1 and 2)

Expressions for the unknown coefficients, the c's, can be obtained in the following way. Clearing fractions in (11), multiplying both sides by y^{2m+1}, and integrating over y leads to

$$\int_{-y_{\max}}^{y_{\max}} y^{2m+1}(y_{\max}^2-y^2)\sum_{n=0}^{N} c_{2n}y^{2n}\frac{d\hat{p}}{dy}dy = \int_{-y_{\max}}^{y_{\max}} y^{2m+2}\hat{p}(y)dy . \tag{12}$$

Integrating the left-hand side by parts and assuming that the approximation $\hat{p}(y)$ has the same moments as the true $p(y)$, i.e.,

$$\mu_n=\int_{-y_{\max}}^{y_{\max}} y^n\hat{p}(y)dy , \tag{13}$$

gives

$$y^{2m+1}(y_{\max}^2-y^2)\sum_{n=0}^{N} c_{2n}y^{2n}\hat{p}(y)\Big|_{-y_{\max}}^{y_{\max}} - \sum_{n=0}^{N} c_{2n}\xi_{2n+2m}=\mu_{2m+2} , \tag{14}$$

in which

$$\xi_k=(k+1)y_{\max}^2\mu_k-(k+3)\mu_{k+2} . \tag{15}$$

The integrated terms in (14) vanish at the end points. They would also vanish if $\hat{p}(y)$ were zero at the end points. In either case, it is necessary that these terms vanish for Pearson's method to work. Setting the integrated terms to zero then leads to the system of simultaneous equations in the c's,

$$\sum_{n=0}^{N} \xi_{2n+2m}c_{2n}=-\mu_{2m+2}, \quad m=0,\ldots,N \tag{16}$$

or

$$\Xi\,\mathbf{c}=-\boldsymbol{\mu} , \tag{17}$$

where

$$\mathbf{c}=\begin{bmatrix} c_0\\ c_2\\ \vdots\\ c_{2N}\end{bmatrix}, \quad \boldsymbol{\mu}=\begin{bmatrix} \mu_2\\ \mu_4\\ \vdots\\ \mu_{2N+2}\end{bmatrix}, \quad \Xi=\begin{bmatrix} \xi_0 & \xi_2 & \cdots & \xi_{2N}\\ \xi_2 & \xi_4 & \cdots & \xi_{2N+2}\\ \vdots & \vdots & & \vdots\\ \xi_{2N} & \xi_{2N+2} & \cdots & \xi_{4N}\end{bmatrix} . \tag{18}$$

Consequently,

$$\mathbf{c}=-\Xi^{-1}\boldsymbol{\mu} \tag{19}$$

is the desired expression for the c's. Note that all the moments $\mu_2,\mu_4,\ldots,\mu_{4N+2}$ are required to determine $c_0,c_2,\ldots,c_{2N}$. In the sixth-order polynomial case (10e), this means that all even moments up to μ_{10} are required to determine c_0,c_2, and c_4; *five* moments are necessary to determine *three* coefficients.

For the sixth-order polynomial approximation (10e), the solution to the simultaneous equations for the c's is

$$c_0=\frac{1}{\Delta}[(\xi_6^2-\xi_4\xi_8)\mu_2+(\xi_2\xi_8-\xi_4\xi_6)\mu_4+(\xi_4^2-\xi_2\xi_6)\mu_6] , \tag{20a}$$

$$c_2=\frac{1}{\Delta}[(\xi_2\xi_8-\xi_4\xi_6)\mu_2+(\xi_4^2-\xi_0\xi_8)\mu_4+(\xi_0\xi_6-\xi_2\xi_4)\mu_6] , \tag{20b}$$

$$c_4=\frac{1}{\Delta}[(\xi_4^2-\xi_2\xi_6)\mu_2+(\xi_0\xi_6-\xi_2\xi_4)\mu_4+(\xi_2^2-\xi_0\xi_4)\mu_6] , \tag{20c}$$

where

$$\Delta=\xi_0\xi_4\xi_8+2\xi_2\xi_4\xi_6-\xi_4^3-\xi_0\xi_6^2-\xi_2^2\xi_8 . \tag{20d}$$

D. The Pearson-type approximations

Once the c's are known, it is straightforward to solve the resulting Pearson differential equation. However, the character of the solution is strongly dependent upon the nature of the roots of the denominator polynomial.

To be specific, we consider only (10e), which can be rewritten as

$$\frac{1}{\hat{p}}\frac{d\hat{p}}{dy}=\frac{y}{c_4(a^2-y^2)(b^2-y^2)(y_{\max}^2-y^2)} \;. \tag{21}$$

Formally the solution can be found by partial fraction expansion of the right-hand side and doing so leads to

$$\hat{p}(y)\sim K\,|a^2-y^2|^{\mathcal{A}}\,|b^2-y^2|^{\mathcal{B}}\,|y_{\max}^2-y^2|^{\mathcal{C}}\,, \tag{22}$$

in which

$$\mathcal{A}=\frac{-1}{2c_4}\frac{1}{(b^2-a^2)(y_{\max}^2-a^2)}\,, \tag{23a}$$

$$\mathcal{B}=\frac{-1}{2c_4}\frac{1}{(a^2-b^2)(y_{\max}^2-b^2)}\,, \tag{23b}$$

$$\mathcal{C}=\frac{-1}{2c_4}\frac{1}{(a^2-y_{\max}^2)(b^2-y_{\max}^2)}\;. \tag{23c}$$

It is easily verified that the absolute-value signs are correct in (22) (indeed, this was what Hansmann failed to notice[8]). Now, the roots need not lie in the interval $(0,y_{\max}^2)$, and so, when a^2 and b^2 are real, we have the following three possibilities for the complete solution (without loss of generality, we take $a^2<b^2$):

(i) Neither a^2 nor b^2 in $(0,y_{\max}^2)$,

$$\hat{p}(y)=K\,|a^2-y^2|^{\mathcal{A}}\,|b^2-y^2|^{\mathcal{B}}(y_{\max}^2-y^2)^{\mathcal{C}}\,,\quad |y|\le y_{\max}\;. \tag{24}$$

(ii) One of a^2 or b^2 in $(0,y_{\max}^2)$, say a^2,

$$\hat{p}(y)=\begin{cases}K_1(a^2-y^2)^{\mathcal{A}}(b^2-y^2)^{\mathcal{B}}(y_{\max}^2-y^2)^{\mathcal{C}}\,, & |y|\le a\\ K_2(y^2-a^2)^{\mathcal{A}}(b^2-y^2)^{\mathcal{B}}(y_{\max}^2-y^2)^{\mathcal{C}}\,, & a<|y|\le y_{\max}\;.\end{cases} \tag{25}$$

(iii) Both a^2 and b^2 in $(0,y_{\max}^2)$,

$$\hat{p}(y)=\begin{cases}K_1(a^2-y^2)^{\mathcal{A}}(b^2-y^2)^{\mathcal{B}}(y_{\max}^2-y^2)^{\mathcal{C}}, & |y|\le a\\ K_2(y^2-a^2)^{\mathcal{A}}(b^2-y^2)^{\mathcal{B}}(y_{\max}^2-y^2)^{\mathcal{C}}, & a<|y|\le b\\ K_3(y^2-a^2)^{\mathcal{A}}(y^2-b^2)^{\mathcal{B}}(y_{\max}^2-y^2)^{\mathcal{C}}, & b<|y|\le y_{\max}\;.\end{cases} \tag{26}$$

Finally, if a^2 and b^2 are imaginary, it is readily shown that the solution can be written as follows:

(iv) a^2 and b^2 imaginary,

$$\hat{p}(y)=K\left[\frac{y_{\max}^2-y^2}{(|c_0+c_2y^2+c_4y^4|)^{1/2}}\right]^{\mathcal{E}}\times\exp\left[-\frac{\mathcal{E}(c_2+2c_4y_{\max}^2)}{(4c_0c_4-c_2^2)^{1/2}}\tan^{-1}\frac{c_2+2c_4y^2}{(4c_0c_4-c_2^2)^{1/2}}\right],\quad |y|\le y_{\max} \tag{27}$$

in which

$$\mathcal{E}=\frac{-1}{2(c_0+c_2y_{\max}^2+c_4y_{\max}^4)}\;. \tag{28}$$

The K coefficients in these expressions will usually have to be determined by numerical integration. In the cases of (24) and (27), we need merely normalize the solutions to integrate to *one*. Equations (25) and (26) require the use of μ_2 and μ_2 and μ_4, respectively, to determine the constants.

All of the above forms will be illustrated by the examples to be presented in the following sections.

E. Cases when the approximations are exact

It is somewhat surprising that there exists a family of probability density functions for which the Pearson "approximations" give "exact" results *no matter how many moments are used in the approximations.* This family is the beta distribution given by

$$p(y)=\frac{(1-y^2)^{\alpha-1}}{2^{2\alpha-1}B(\alpha,\alpha)},\quad |y|\le 1 \tag{29}$$

in which B denotes the beta function and $\alpha>0$ is a parameter. The moments of this density are

$$\mu_{2n}=B(n+\tfrac{1}{2},\alpha)\;. \tag{30}$$

It is easy to show from the recursive properties of the gamma and beta functions and the definition (15), that for these moments

$$\xi_{2n}=2(\alpha-1)\mu_{2n+2}\;. \tag{31}$$

Using this to substitute for the μ_n's in the right-hand side of (16) gives

$$\sum_{n=0}^{N}\xi_{2n+2m}c_{2n}=\frac{-\xi_{2m}}{2(\alpha-1)},\quad m=0,\ldots,N\;. \tag{32}$$

By inspection, the solution to this set is

$$c_0=\frac{-1}{2(\alpha-1)},\quad c_2=c_4=\cdots=c_{2N}=0\;. \tag{33}$$

Hence, for all $N\ge 0$, the differential equation (11) reduces to

$$\frac{1}{\hat{p}}\frac{d\hat{p}}{dy}=\frac{-y}{2(\alpha-1)(1-y^2)}\,, \tag{34}$$

which has the solution (29). In the case when $y_{\max}$ is *not* equal to 1, the above argument is easily modified to hold for the normalized random variable $y/y_{\max}$.

This is a wondrous and astonishing result. The implication is that density functions close in shape to one of the beta family can be extremely well approximated by these Pearson-type methods. The more the deviation from the beta density, the more moments we expect to be needed for a good approximation. In addition, for small values of α, the beta density is extremely leptokurtic and it is well known that some expansions have problems in convergence in these cases.

F. Negative features of the approximations

Now that we have seen a case when the approximations are exact, let us look at a case when they fail completely. This is the uniform density

$$p(y)=\frac{1}{2y_{\max}}, \quad |y| \le y_{\max} \tag{35}$$

which has moments

$$\mu_{2n}=\frac{y_{\max}^{2n}}{2n+1} . \tag{36}$$

Now, from (15),

$$\xi_{2n}=0 . \tag{37}$$

Hence, the system of equations (16) can only be satisfied by some of the c's being infinite, which is what should happen. However, the implication is that if the underlying density is very close to uniform, very large and small numbers may enter into the calculations and accuracy lost. Hence, the moments may need to be very accurately known and/or high-precision computing methods employed.

In Sec. II C it was noted that five even moments were necessary to determine c_0, c_2, and c_4 for the sixth-order polynomial case (10e). There is no way of going the other way and getting the moments back from the c's with the exception of doing the integrals

$$\hat{\mu}_n = \int_{-y_{\max}}^{y_{\max}} y^n \hat{p}(y)dy . \tag{38}$$

The question is "Is $\hat{\mu}_n=\mu_n$ for $n=2,4,\ldots,N$?" There seems to be nothing in the theory that *guarantees* that these moments be equal in spite of the assumption (13). If $\hat{p}(y)$ is close to $p(y)$ then, of course, the moments will be close. We later give an example that shows unequivocally, however, that the moments are not necessarily the same.

We cannot expect the Pearson-type approximations to reproduce any fine structure in the underlying probability density function. Even coarse structure that is not in keeping with a member of the beta family may be difficult to reproduce without an inordinately high-degree numerator polynomial. We will later see an example of this in which the true density function has multiple lobes. It is to be noted that our approximations need not necessarily be positive. The solutions to (24) and (27) are always positive, but one of the K's in (25) or (26) could conceivably be negative. This does not happen in any of the examples to be presented later.

Finally, we will see in some of the examples that the approximations can contain extraneous zeros or infinities, which manifest themselves as deep chasms or sharp spikes in $\hat{p}(y)$, that are not in the true density function. However, for all of the cases to be examined, these anomalies are such that they have a small effect upon the associated cumulative distribution, and it is the cumulative distribution that has physical significance and not the probability density function.

III. EXAMPLES—KNOWN DENSITIES

All of the examples, both in this section and in Sec. IV, are taken from realistic, nontrivial physical problems involving linear and nonlinear filtering of Markov and non-Markov dichotomous noise. Although the examples are interesting in themselves, they have been chosen to demonstrate various aspects of the Pearson-type approximations. The first example shows that the approximations can be extremely good for a nonbeta probability density function. The second shows that the approximate moments $\hat{\mu}_n$ are not necessarily equal to the exact μ_n, and the final example exhibits a family for which the Pearson-type approximation can work well or not depending upon the value of a parameter in the probability density function. For the case in which the sixth-degree polynomial does not do well, the approximations are compared with those resulting from use of an eighth-order polynomial.

A. Filtering of nonMarkov dichotomous noise with gamma interval PDF

The Langevin equation for this example is

$$\frac{dy(t)}{dt}+\beta y(t)=\beta x(t) ,$$

and $x(t)$ is non-Markov dichotomous noise whose intervals between transitions have the gamma density

$$f(t)=\lambda^2 te^{-\lambda t}, \quad t \ge 0 . \tag{39}$$

It is known[18,19] that the probability density function of the output in the case $\alpha=\lambda/\beta=\frac{1}{2}$ is

$$p(y)=\frac{|\Gamma(1-q)|^2}{\pi^{3/2}}(1-y^2)^{-1/2}\,{}_2F_1(q,q^*;\tfrac{1}{2};y^2) ,$$
$$q=(1+i)/4, \quad |y| \le 1 , \tag{40}$$

in which ${}_2F_1$ is a hypergeometric function. $p(y)$ is U shaped with $(1-y^2)^{-1/2}\ln(1-y^2)$ infinities near $y=\pm 1$. The odd moments are zero and the even moments can be determined recursively from the relation

$$E[y^2(1-y^2)^n]=\frac{3+4n}{8n^2+12n+5}E[(1-y^2)^n] , \tag{41}$$

which follows by multiplying the differential equation [Ref. 18, Eq. (54a); see also Ref. 20] satisfied by $p(y)$ by $(1-y^2)^n$ and integrating over $(-1,1)$. The first five even moments are

TABLE I. Filtering of non-Markov dichotomous noise with gamma interval density. Comparison of sixth-degree polynomial approximation with exact values. ($c_0=0.7228$, $c_2=-0.2731$, $c_4=0.3537$, $K=0.2770$, $\mathcal{E}=-0.6223$.)

y	$p(y)$	$\hat{p}(y)$
0	0.2315	0.2321
0.2	0.2386	0.2388
0.4	0.2636	0.2627
0.6	0.3222	0.3212
0.8	0.4880	0.4910
0.9	0.7638	0.7683
0.95	1.204	1.199
0.99	3.406	3.306
0.999	14.19	13.90

$$\mu_2=\tfrac{3}{5}=0.6\ ,$$
$$\mu_4=\tfrac{61}{125}=0.488\ ,$$
$$\mu_6=\frac{13\times 251}{61\times 125}=0.427\,934\ ,$$
$$\mu_8=\frac{3\times 17\times 6563}{61\times 113\times 125}=0.388\,467\ ,$$
$$\mu_{10}=\frac{11\,220\,991}{25\times 61\times 113\times 181}=0.359\,753\ .$$

For these moments, the sixth-degree polynomial approximation satisfying (10e) comes out to be

$$\hat{p}(y)=K\left[\frac{1-y^2}{\sqrt{c_0+c_2y^2+c_4y^4}}\right]^{\mathcal{E}}\times\exp\left[-\frac{\mathcal{E}(c_2+2c_4)}{(4c_0c_4-c_2^2)^{1/2}}\times\tan^{-1}\left[\frac{c_2+2c_4y^2}{(4c_0c_4-c_2^2)^{1/2}}\right]\right]. \tag{42}$$

This $\hat{p}(y)$ is compared in Table I with the true $p(y)$ given by (40). The agreement is to within about 1% except for $y>0.95$. The approximate moments $\hat{\mu}_n$, $n=2,4,\ldots,10$ agree with the exact values to four decimal places. This example serves to illustrate just how good the Pearson-type approximation can be.

B. Filtering of non-Markov dichotomous noise with McFadden interval PDF

The Langevin equation in this example is the same as that of the previous one but now the intervals of the dichotomous noise have the density

$$f(t)=3e^{-t}(1-e^{-t})^2,\quad t\geq 0 \tag{43}$$

and $\beta=1$. The probability density function of the output has the simple form[18]

$$p(y)=\tfrac{3}{44}(7+y^2),\quad |y|\leq 1 \tag{44}$$

which has moments

$$\mu_{2n}=\frac{3}{11}\frac{8n+11}{(2n+1)(2n+3)}\ . \tag{45}$$

For these moments, the coefficients in the sixth-degree polynomial approximation (10e) come out to have the amazingly simple values

$$c_0=\tfrac{49}{4},\quad c_2=-66,\quad\text{and}\quad c_4=\tfrac{429}{4}\ ,$$

and the Pearson-type approximation is

$$\hat{p}(y)=0.4929\left[\frac{(49-264y^2+429y^4)^{1/2}}{2(1-y^2)}\right]^{1/107}\times\exp\left[\frac{297}{107\sqrt{3597}}\tan^{-1}\frac{429y^2-132}{\sqrt{3597}}\right]. \tag{46}$$

The reason why this example is important is that, because of the simple c's, the accuracy in numerically computing K and $\hat{\mu}_n$ lies only in the integration routines used. Table II(a) compares the approximate and exact probability density functions and Table II(b) compares the corresponding moments.

As the tables show, both the approximate probability density and the approximate moments are quite close to the true values. *However,* the computations were done using double precision and the moments do not agree in the third decimal place. It can only be concluded that the theory does not guarantee that the approximate and exact moments be the same. In some examples later, we shall see even larger differences.

TABLE II. Filtering of non-Markov dichotomous noise with McFadden interval PDF. (a) Comparison of sixth-degree polynomial approximation with exact values. ($c_0=49/4$, $c_2=-66$, $c_4=429/4$, $K=0.4929$.) (b) Comparison of approximate and exact moments.

(a)

y	$p(y)$	$\hat{p}(y)$
0	0.4773	0.4730
0.2	0.4880	0.4738
0.4	0.4882	0.4788
0.6	0.5018	0.5053
0.8	0.5209	0.5319
0.9	0.5325	0.5401
0.95	0.5388	0.5453
0.99	0.5441	0.5547
0.999	0.5453	0.5670
1	0.5455	∞

(b)

n	μ_n	$\hat{\mu}_n$
2	0.3455	0.3494
4	0.2104	0.2136
6	0.1515	0.1540
8	0.1185	0.1204
10	0.0973	0.0989

C. Nonlinear filtering of Markov dichotomous noise

The Langevin equation for this case is

$$\frac{dy(t)}{dt}+\beta\sin y(t)=\beta x(t) \tag{47}$$

and $x(t)$ is the random telegraph signal, i.e., Markov dichotomous noise with interval density

$$f(t)=\lambda e^{-\lambda t}, \quad t\geq 0 . \tag{48}$$

The output probability density is known to be[15]

$$p(y)=\frac{\sec^2 y}{2K_1(2\alpha)}e^{-2\alpha\sec y}, \quad |y|\leq\frac{\pi}{2}, \quad \alpha=\frac{\lambda}{\beta}, \tag{49}$$

where K_1 is a modified Bessel function of the second kind. The nature of $p(y)$ depends strongly upon the parameter α. When α is large, $p(y)$ is a single bell-shaped lobe at the origin; for $\frac{1}{5}<\alpha<\frac{1}{2}$ it becomes a pair of lobes; and for very small α it approaches a pair of impulse functions at $\pm\pi/2$. Evidently, the moments μ_n cannot be determined in closed form and so they were calculated from (49) by numerical integration.

For $\alpha=1$, 2, and 5 the sixth-order polynomial Pearson-type approximations all come out to have the form

$$\hat{p}(y)=K\left[\frac{y_{\max}^2-y^2}{|c_0+c_2y^2+c_4y^4|^{1/2}}\right]^{\mathcal{E}} \times\exp\left[-\frac{\mathcal{E}(c_2+2c_4y_{\max}^2)}{(4c_0c_4-c_2^2)^{1/2}}\tan^{-1}\frac{c_2+2c_4y^2}{(4c_0c_4-c_2^2)^{1/2}}\right] \tag{50}$$

and the approximate and exact densities and moments are within percentage points of one another. Consequently, we do not bother to give the results in these cases since the agreement is so good.

For $\alpha=\frac{1}{5}$ and $\frac{1}{2}$, the approximating densities are each of the form

$$p(y)=\begin{cases} K_1(a^2-y^2)^{\mathcal{A}}(b^2-y^2)^{\mathcal{B}}(y_{\max}^2-y^2)^{\mathcal{C}}, \\ |y|\leq a \\ K_2(y^2-a^2)^{\mathcal{A}}(b^2-y^2)^{\mathcal{B}}(y_{\max}^2-y^2)^{\mathcal{C}}, \\ a<|y|\leq b \\ K_3(y^2-a^2)^{\mathcal{A}}(y^2-b^2)^{\mathcal{B}}(y_{\max}^2-y^2)^{\mathcal{C}}, \\ b<|y|\leq y_{\max} . \end{cases} \tag{51}$$

and the parameter values for these are listed in Table III along with the moments μ_n. This sixth-order Pearson-type approximation is compared with the actual density in Fig. 1(a). The approximation does not do well in reproducing the two-lobe structure of the density function, but the approximate moments agreed with the true values to four decimal places.

In order to improve the approximation, an eighth-order polynomial was used and this resulted in a density of the form

$$\hat{p}(y)\sim K\frac{|a^2-y^2|^{\mathcal{A}}|c^2-y^2|^{\mathcal{C}}}{|b^2-y^2|^{-\mathcal{B}}(y_{\max}^2-y^2)^{-\mathcal{D}}}, \tag{52}$$

which has spikes at $y=b$ and $y=y_{\max}$ and zeros at $y=a$

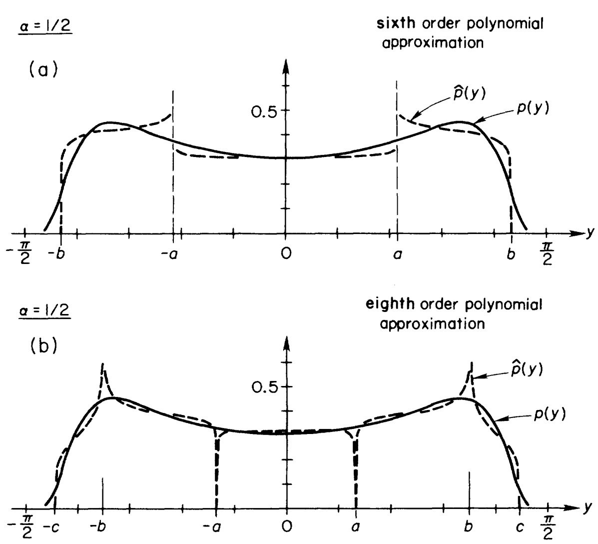

FIG. 1. Probability-density function for nonlinearly filtered Markov dichotomous noise. (a) Sixth-order polynomial approximation, and (b) eighth-order polynomial approximation. $\alpha=\frac{1}{2}$.

TABLE III. Nonlinear filtering of Markov dichotomous noise. Parameters for sixth- and eighth-order Pearson-type approximations.

(a) Moments

n	μ_n, $\alpha=\frac{1}{5}$	μ_n, $\alpha=\frac{1}{2}$
2	1.113 66	0.659 168
4	1.702 17	0.728 577
6	2.868 99	0.952 481
8	5.099 18	1.364 10
10	9.384 29	2.073 71
12	a	3.291 12
14	a	5.398 54

(b) Parameter values

	$\alpha=\frac{1}{5}$ Sixth order	$\alpha=\frac{1}{2}$ Sixth order	$\alpha=\frac{1}{2}$ Eighth order
a^2	0.3657	0.4495	0.1789
b^2	1.898	1.837	1.243
c^2	a	a	1.195
$\mathcal{A}$	0.1276	−0.027 49	0.028 58
$\mathcal{B}$	−0.4712	0.088 01	−0.1337
$\mathcal{C}$	0.3436	−0.060 52	0.1905
$\mathcal{D}$	a	a	−0.085 42
K_1	0.1903	0.3060	0.3226
K_2	0.3319	0.4341	0.3799
K_3	0.1441	0.004 361	0.3105
K_4	a	a	0.001 798

[a]Not applicable.

and $y=c$. There are four multiplicative constants, one for each of the intervals $(0,a)$, (a,b), (b,c), and $(c,\pi/2)$ and these are listed in Table III(b) along with the other parameter values. The resulting approximation is shown in Fig. 1(b) and there is much better agreement with the true density than with the sixth-order polynomial approximation. The moments in the eighth-order case agreed extremely well with the actual values.

Similar computations were performed for the case $\alpha=\frac{1}{5}$ and similar behavior obtained. The eighth-order polynomial approximation may not be practical in a situation when the moments are computed from experimental data, but there is no reason why it cannot be implemented in a situation when the moments are known analytically, but the underlying density is unknown. Two such cases arising in physical problems are considered in Sec. IV.

IV. EXAMPLES—UNKNOWN DENSITIES

All of the examples in Sec. III were tests or benchmarks of the Pearson-type approximations since the probability density functions were known *a priori*. In the examples to be considered in this section, the probability density functions are unknown. Indeed, the results presented here are, to our knowledge, the first times good approximate solutions have been given to these problems. We first consider the filter-limiter-filter problem and then a second-order oscillator driven by Markov dichotomous noise.

A. Filter-limiter-filter

The Langevin equations for the filter-limiter-filter system are

$$\frac{dy(t)}{dt}+\beta y(t)=\beta\,\mathrm{sgn}[\eta(t)]\ , \tag{53}$$

$$\frac{d\eta(t)}{dt}+\frac{1}{\tau_c}\eta(t)=w(t)\ , \tag{54}$$

in which $w(t)$ is white Gaussian noise with autocorrelation function $R_w(\tau)=2D\delta(\tau)$. The signum function in (53) models the operation of a *limiter* in an electrical network and, because of its presence, the intensity D of the white noise is unimportant. It is known that $p(y)$ has a beta density when $\alpha=\beta\tau_c=2$, but there are no exact solutions for $p(y)$ for general α.[16,21,22] However, in the general case, the moments μ_n can be calculated successively from the recursion relation [Ref. 22, Eq. (49) *et seq.*]

$$E[(1+y)(1-y)^n]=\frac{1}{\pi}B\left[\frac{n\alpha+1}{2},\frac{1}{2}\right]E[(1-y_0)^n]\ . \tag{55}$$

In this equation n is an integer, y_0 is an auxiliary random variable, and the odd moments of both y and y_0 are all zero. The moments calculated from this are given in Table IV for $\alpha=\frac{1}{5}$, $\frac{1}{2}$, 1, 2, and 5.

The forms of the approximations depend upon α, and they are, for $\alpha=\frac{1}{5}$ and $\frac{1}{2}$,

TABLE IV. Filter-limiter-filter.

(a) Output moments

α	μ_2	μ_4	μ_6	μ_8	μ_{10}
$\frac{1}{5}$	0.116 847	0.032 664 9	0.012 693 9	0.005 933 46	0.003 128 92
$\frac{1}{2}$	0.237 240	0.112 836	0.067 399 0	0.045 287 5	0.032 725 8
1	0.363 380	0.227 687	0.167 151	0.132 551	0.110 053
2	0.5	0.375	0.3125	0.273 438	0.246 094
5	0.660 469	0.566 609	0.516 323	0.483 184	0.458 969

(b) Parameter values

	$\alpha=\frac{1}{5}$	$\alpha=\frac{1}{2}$	$\alpha=5$
a^2	−7.648	−24.91	−1.277
b^2	2.809	2.034	2.045
$\mathcal{A}$	−0.6137	−0.033 62	0.2438
$\mathcal{B}$	−2.935	−0.8427	0.5311
$\mathcal{C}$	3.548	0.8763	−0.7749
K	75.71	1.319	0.1288

	$\alpha=1$
c_0	4.059
c_2	−9.962
c_4	15.13
$\mathcal{E}$	−0.054 15
K	0.4661

$$\hat{p}(y)=\frac{K(1-y^2)^{\mathcal{C}}}{(-a^2+y^2)^{-\mathcal{A}}(b^2-y^2)^{-\mathcal{B}}}\ , \tag{56}$$

for $\alpha=1$,

$$\hat{p}(y)=K\left[\frac{1-y^2}{(c_0+c_2y^2+c_4y^4)^{1/2}}\right]^{\mathcal{E}} \times\exp\left[-\frac{\mathcal{E}(c_2+2c_4)}{(4c_0c_4-c_2^2)^{1/2}}\tan^{-1}\frac{c_2+2c_4y^2}{(4c_0c_4-c_2^2)^{1/2}}\right]\ , \tag{57}$$

for $\alpha=2$,

$$\hat{p}(y)=p(y)=\frac{1}{\pi(1-y^2)^{1/2}}\ , \tag{58}$$

for $\alpha=5$,

$$p(y)=\frac{K(-a^2+y^2)^{\mathcal{A}}(b^2-y^2)^{\mathcal{B}}}{(1-y^2)^{-\mathcal{C}}}\ , \tag{59}$$

in which $|y|\le 1$ in each case. The parameter values for these are listed in Table IV(b). For $\alpha=\frac{1}{5}$, $\frac{1}{2}$, and 2, the roots a^2 and b^2 both lie outside the interval (0,1).

An approximation that makes use only of μ_2 was developed in 1969 by solving an approximate Fokker-Planck equation[16] and is

$$\hat{p}_{\rm FP}(y)=\frac{(1-y^2)^{\gamma-1}}{2^{2\gamma-1}B(\gamma,\gamma)},\quad \gamma=\frac{1}{2}\left[\frac{B\left[\frac{\alpha+1}{2},\frac{1}{2}\right]}{\pi-B\left[\frac{\alpha+1}{2},\frac{1}{2}\right]}\right]\ . \tag{60}$$

This approximation was shown to give a cumulative distribution that agreed to within 5% with Monte Carlo simulations.[16] Since the true density is close to being beta distributed, we expect our Pearson-type approximations to work extremely well. Unquestionably, the above approximations $\hat{p}(y)$ will be better than this $\hat{p}_{\rm FP}(y)$ since five even moments are used in the former and only one in the latter. It is nevertheless interesting to compare $\hat{p}(y)$ and $\hat{p}_{\rm FP}(y)$ and this is done in Table V for $\alpha=\frac{1}{5}$ and $\frac{1}{2}$. The agreement is so close that it follows that $\hat{p}(y)$ is within a few percent of the Monte Carlo [the actual simulation values in Ref. 16 are no longer available and the simulations were not repeated—this is why we are comparing $\hat{p}(y)$ to $\hat{p}_{\rm FP}(y)$]. The moments $\hat{\mu}_n$ agreed with μ_n to within 1% for all values of α and there is little to be gained by tabulating them.

TABLE V. Filter-limiter-filter. Comparison of $\hat{p}(y)$ and $\hat{p}_{\rm FP}(y)$ for $\alpha=\frac{1}{5}$ and $\alpha=5$.

	$\alpha=\frac{1}{5}$		$\alpha=5$	
y	$\hat{p}(y)$	$\hat{p}_{\rm FP}(y)$	$\hat{p}(y)$	$\hat{p}_{\rm FP}(y)$
0	1.049	1.061	0.2000	0.1949
0.2	0.9436	0.9474	0.2058	0.2009
0.4	0.6628	0.6537	0.2256	0.2218
0.6	0.3131	0.3070	0.2709	0.2715
0.8	0.0568	0.0621	0.3992	0.4163
0.9	0.0074	0.0105	0.6245	0.6693
0.95			1.016	1.099
0.99			3.381	3.578
0.999			19.92	19.73

Equations (56), (57), and (59) are *new* approximation solutions to the filter-limiter-filter problem.

B. Dichotomous-noise-driven oscillator

Our final example is that of a second-order Butterworth filter excited by Markov dichotomous noise with Langevin equation[23]

$$\frac{d^2y(t)}{dt^2}+2\beta\frac{dy(t)}{dt}+2\beta^2y(t)=2\beta^2x(t) \tag{61}$$

and $x(t)$ has intervals with the probability density function

$$f(t)=\lambda e^{-\lambda t},\quad t\ge 0\ . \tag{62}$$

In this case, the output probability density function is symmetric and lies in the interval $|y|\le y_{\rm max}$ where $y_{\rm max}=\coth(\pi/2)=1.090\,33$. The output moments are given by

$$\mu_{2n}=\sum_{\text{all } q_j=\pm 1}\prod_{k=1}^{2n}\frac{ikq_k}{[1-(-1)^k]\alpha+k+iS_k}\ , \tag{63}$$

in which

$$\alpha=\frac{\lambda}{\beta}\ , \tag{64a}$$

$$S_k=\sum_{m=1}^{k}q_m\ . \tag{64b}$$

Again, the forms of the sixth-order polynomial Pearson-type approximations depend upon the parameter α. We find, for $\alpha=\frac{1}{5}$ and $\frac{1}{2}$,

$$\hat{p}(y)=\begin{cases} K_1\dfrac{(a^2-y^2)^{\mathcal{A}}(y_{\rm max}^2-y^2)^{\mathcal{C}}}{(b^2-y^2)^{-\mathcal{B}}}, & |y|\le a\\[2ex] K_2\dfrac{(y^2-a^2)^{\mathcal{A}}(y_{\rm max}^2-y^2)^{\mathcal{C}}}{(b^2-y^2)^{-\mathcal{B}}}, & a<|y|\le b\\[2ex] K_3\dfrac{(y^2-a^2)^{\mathcal{A}}(y_{\rm max}^2-y^2)^{\mathcal{C}}}{(y^2-b^2)^{-\mathcal{B}}}, & b<|y|\le y_{\rm max}\ ,\end{cases} \tag{65}$$

for $\alpha=1$,

$$\hat{p}(y)=\begin{cases} K_1\dfrac{(a^2-y^2)^{\mathcal{A}}(b^2-y^2)^{\mathcal{B}}}{(y_{\rm max}^2-y^2)^{-\mathcal{C}}}, & |y|\le a\\[2ex] K_2\dfrac{(y^2-a^2)^{\mathcal{A}}(b^2-y^2)^{\mathcal{B}}}{(y_{\rm max}^2-y^2)^{-\mathcal{C}}}, & a<|y|\le y_{\rm max}\ ,\end{cases} \tag{66}$$

and for $\alpha=2$ and 5,

$$\hat{p}(y)=K\left[\frac{y_{\max}^2-y^2}{|c_0+c_2y^2+c_4y^4|^{1/2}}\right]^{\mathcal{E}}\times\exp\left[-\frac{\mathcal{E}(c_2+2c_4y_{\max}^2)}{(4c_0c_4-c_2^2)^{1/2}}\tan^{-1}\frac{c_2+2c_4y^2}{(4c_0c_4-c_2^2)^{1/2}}\right],\quad |y|\le y_{\max}\,. \tag{67}$$

The parameter values for the various cases are given in Table VI along with the output moments.

The approximation $\hat{p}(y)$ is shown in Fig. 2 for $\alpha=\frac{1}{5}$. The deep chasms at $y=\pm a$ have little effect upon the cumulative distribution. The infinities near $y=\pm b$ can be explained by the physics of the problem; i.e., the output $y(t)$ overshoots and oscillates about the steady states at $y=\pm 1$, and, for this case, b is close to 1.

The cumulative distributions $\hat{P}(y\le Y)=\int_{-y_{\max}}^{Y}\hat{p}(y)dy$ are tabulated in Table VII along with results of Monte Carlo simulations reproduced from Ref. 23. There is excellent, almost uncanny, agreement between the two sets of numbers—especially since the simulations were done without prior knowledge of the true or approximate values.

The approximate moments $\hat{\mu}_n$ were computed in each case and found to be within 1% of the true values except in the case $\alpha=1$ where the difference was of the order of 10% for μ_{10}.

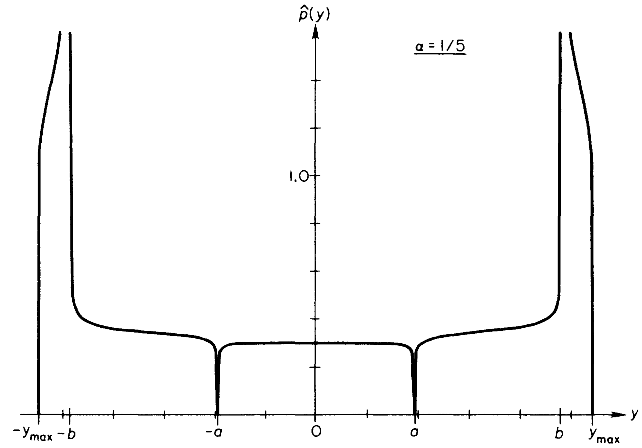

FIG. 2. Probability-density function for output of second-order Butterworth filter driven by Markov dichotomous noise, $\alpha=\frac{1}{5}$.

V. THE ASYMMETRIC CASE

Much of what we have done can be extended to the asymmetric case. The asymmetric analog of the sixth-degree polynomial differential equation (10e) is

$$\frac{1}{\hat{p}}\frac{d\hat{p}}{dy}=\frac{y-\bar{y}}{(y-y_{\min})(y_{\max}-y)(c_0+c_1y+c_2y^2)}\,. \tag{68}$$

TABLE VI. Second-order Butterworth filter driven by Markov dichotomous noise.

(a) Output moments

α	μ_2	μ_4	μ_6	μ_8	μ_{10}
$\frac{1}{5}$	0.810 811	0.787 772	0.794 920	0.816 050	0.846 848
$\frac{1}{2}$	0.6	0.529 412	0.511 658	0.514 680	0.529 669
1	0.4	0.291 584	0.251 469	0.234 333	0.228 184
2	0.230 769	0.118 981	0.080 569 0	0.062 666 7	0.052 962 0
5	0.098 360 7	0.025 646 6	0.009 936 8	0.004 846 0	0.002 752 9

(b) Parameter values

	$\alpha=\frac{1}{5}$	$\alpha=\frac{1}{2}$	$\alpha=1$
a^2	0.058 81	0.1523	0.3517
b^2	0.9479	0.9365	1.252
$\mathcal{A}$	0.063 22	0.031 09	0.030 26
$\mathcal{B}$	−0.2965	−0.1277	0.4020
$\mathcal{C}$	0.2333	0.096 64	−0.4322
K_1	0.1657	0.3088	0.4916
K_2	0.1671	0.3338	0.3892
K_3	2.102	1.343	a

	$\alpha=2$	$\alpha=5$
c_0	−0.3257	−0.1016
c_2	−0.2524	−0.023 58
c_4	−0.080 00	−0.027 36
$\mathcal{E}$	0.6767	2.970
K	1.788	0.043 26

[a]Not applicable.

TABLE VII. Second-order Butterworth filter driven by Markov dichotomous noise. Comparison of sixth-degree polynomial Pearson-type approximations with Monte Carlo results.

	$\alpha=\frac{1}{5}$		$\alpha=\frac{1}{2}$		$\alpha=1$		$\alpha=2$		$\alpha=5$	
Y	$\hat{P}(y\le Y)$	Monte Carlo	$\hat{P}(y\le Y)$	Monte Carlo	$\hat{P}(y\le Y)$	Monte Carlo	$\hat{P}(y\le Y)$	Monte Carlo	$\hat{P}(y\le Y)$	Monte Carlo
0	0.500	0.500	0.500	0.496	0.500	0.497	0.500	0.501	0.500	0.499
0.1	0.515	0.513	0.529	0.526	0.547	0.543	0.573	0.575	0.619	0.622
0.2	0.529	0.530	0.560	0.555	0.594	0.587	0.643	0.641	0.729	0.704
0.3	0.542	0.543	0.589	0.585	0.641	0.633	0.711	0.699	0.822	0.799
0.4	0.557	0.557	0.618	0.615	0.688	0.679	0.773	0.757	0.894	0.881
0.5	0.573	0.575	0.649	0.646	0.734	0.723	0.829	0.812	0.943	0.937
0.6	0.590	0.589	0.682	0.678	0.778	0.769	0.878	0.866	0.973	0.972
0.7	0.608	0.608	0.716	0.711	0.819	0.816	0.919	0.912	0.990	0.989
0.8	0.627	0.628	0.751	0.748	0.861	0.858	0.952	0.950	0.997	0.997
0.9	0.648	0.656	0.788	0.789	0.906	0.900	0.977	0.981	1.000	1.000
0.95	0.661	0.670	0.809	0.826	0.929	0.927	0.987	0.988	1.000	1.000
1.00	0.804	0.798	0.879	0.871	0.953	0.944	0.994	0.995	1.000	1.000
1.05	0.936	0.937	0.952	0.953	0.979	0.974	0.998	1.000	1.000	1.000

Two cases need to be distinguished—those in which the mean $\bar{y}$ is known and those in which it is not. In the latter, the mean would be determined as an unknown along with c_0, c_1, and c_2. Similar comments apply if the extreme values $y_{\min}$ and $y_{\max}$ are unknown.

When $y_{\min}=-1$ and $y_{\max}=1$, the symmetric form of the beta probability density function is[19]

$$p(y)=\frac{(1+y)^{u-1}(1-y)^{v-1}}{2^{u+v-1}B(u,v)}, \quad |y|\le 1 . \tag{69}$$

It would be expected that some form of the general result that the approximations are exact for beta densities in the symmetric case would hold for this asymmetric form also. The approximations would have the same negative features that they have in the symmetric case.

VI. SUMMARY AND CONCLUSIONS

A method based upon Pearson-type approximations from statistics was developed for approximating a symmetric probability density function from its moments. The method was shown to be exact for beta probability density functions and to work extraordinarily well in numerous examples. The sixth-degree polynomial approximation with known end points received the most attention primarily because this is the highest-degree polynomial case that can be handled conveniently analytically. Some comparisons were given with the eighth-degree polynomial approximations.

Computations were performed using a fourth-degree polynomial with unknown end points, the Hansmann case, for the second-order Butterworth filter example. They were found to work well only for values $\alpha\ge 1$. In addition, the number of places of accuracy in the input moments μ_n was varied and it was found that as little as three-place accuracy in the input moments was sufficient to give three-place accuracy in the approximations in most cases. The moments in all of our examples were determined analytically; however, in many applications they result from sample data. In these cases, one might well question the significance and reliability of the higher moments. Nevertheless, there are many cases, like the last two examples, in which all of the moments are precisely known but the probability density not.

Although we have given some general comments about the asymmetric case, much more work needs to be done in this area. Also the case when one or both of the extreme values is infinite merits further study.

ACKNOWLEDGMENTS

This work was supported by the Air Force Office of Scientific Research (AFSC) under Contract No. F49620-85-C-0093 to Random Applications, Inc. The author would like to express his gratitude to S. D. Rice, who provided much insight and encouragement before his untimely death.

[1]K. Pearson, Philos. Trans. R. Soc. London, Ser. A **185**, 719 (1894); **186**, 343 (1895).
[2]G. H. Hansmann, Biometrika **26**, 129 (1934).
[3]J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, revised ed. (American Mathematical Society, 1950).
[4]E. S. Pearson, Biometrika **50**, 95 (1963).
[5]J. K. Ord, *Families of Frequency Distributions* (Griffin, London, 1972).
[6]J. K. Ord, in Vol. 6 of *Encyclopedia of Statistical Sciences*, edited by S. Katz and N. L. Johnson (Wiley, New York, 1981), pp. 653–659.
[7]J. C. Wheeler, J. Chem. Phys. **80**, 472 (1984).
[8]R. F. Pawula and S. O. Rice (unpublished).
[9]J. K. Ord (private communication).

[10]R. F. Pawula, IEEE Trans. Inf. Theory **13**, 33 (1967).
[11]R. F. Pawula, Phys. Rev. **162**, 186 (1967).
[12]W. Horsthemke and R. Lefever, *Noise Induced Transitions* (Springer-Verlag, New York, 1984).
[13]H. Risken, *The Fokker-Planck Equation* (Springer-Verlag, New York, 1984).
[14]R. F. Pawula, Int. J. Control **12**, 25 (1970).
[15]R. F. Pawula, Int. J. Control **25**, 283 (1977).
[16]R. F. Pawula and A. Y. Tsai, IEEE Trans. Inf. Theory **IT-15**, 532 (1969).
[17]R. F. Pawula (unpublished).
[18]R. F. Pawula and S. O. Rice, IEEE Trans. Inf. Theory **IT-32**, 63 (1986).
[19]R. F. Pawula and S. O. Rice, IEEE Trans. Inf. Theory (to be published).
[20]A. Munford, IEEE Trans. Inf. Theory **IT-32**, 824 (1986).
[21]W. Doyle, J. A. McFadden, and I. Marx, J. Soc. Ind. Appl. Math. **10**, 381 (1962).
[22]J. E. Mazo, R. F. Pawula, and S. O. Rice, IEEE Trans. Inf. Theory **IT-19**, 404 (1973).
[23]R. F. Pawula, Phys. Rev. A **35**, 3102 (1987).