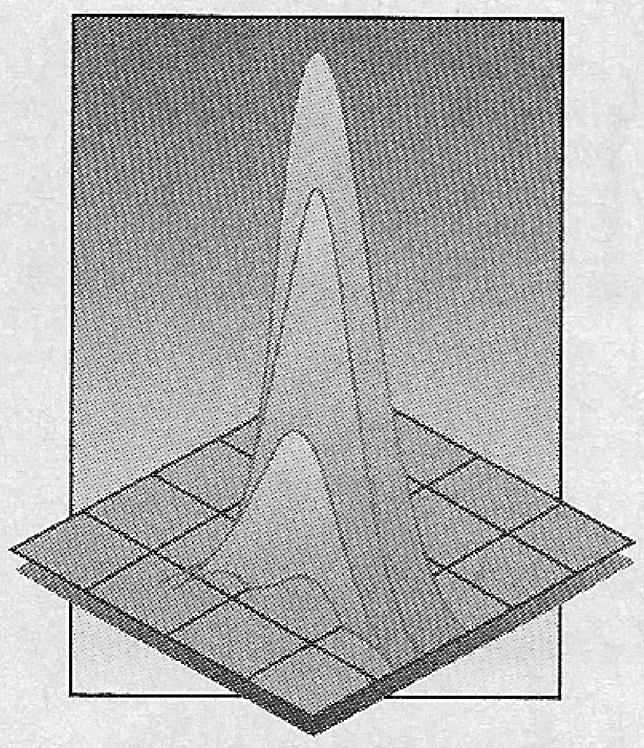

SP Forum

"... most engineers do not understand the need for careful thinking about statistical matters. It is time for the professional statisticians to rise to the challenge of teaching statistics to engineers and scientists."—M.J. Hinich

While *SP Magazine* is not necessarily adverse to the publication of book reviews, for various reasons none has been published for many years. The following articles are included, not so much for their interest as book review material, but for the issues they raise; and also to launch a new department in this Magazine by which readers can express their views on topics of importance in the area of signal processing.

Most of us working in SP would confess to sometimes sloppy or cavalier use of statistical concepts. The reality, is, however, that we rarely have certain justification for any population model assumed, regardless of how mathematically rigorous our analysis based on the model. Professor William Gardner of the University of California at Davis proposes an unorthodox solution to this problem in his recent book in which statistics over ensemble models are replaced by a dual "temporal probability" model over waveforms. In his review of Gardner's book, Professor Melvin Hinich of the University of Texas at Austin takes great exception to this approach in his review written for *SIAM Review*, and reprinted here.

Hinich implies that acceptance of such a model would reveal the engineer's misunderstanding of "the need for careful thinking about statistical matters." Following Hinich's review is Gardner's rejoinder in the form of a humorous piece, "Ensembles in Wonderland," intended to "inform people about both sides of the controversy without trying to discredit either side."
I will let the reader determine whether Professor Gardner has achieved his objective with complete impartiality. Regardless of which side of the debate you find yourself on, the article should be informative.

At a time when several issues centering on statistical signal processing are planned for *SP Magazine*, it is hoped that this set of articles will stimulate debate and interest in the subject. Write to us!
— J. R. Deller

Statistical Spectral Analysis: A Nonprobabilistic Theory. William Gardner. Prentice-Hall, NJ, 1987. xxvi + 566 pages. ISBN 0-13844-572-9.

This book presents a nonstandard treatment of the Fourier analysis approach to the study of deterministic, random continuous, and discrete-time signals. The coverage of the algebra of spectral analysis, and the motivation, examples, and exercises used in his book are basically the same as other modern books in the subfield of electrical engineering, which is usually called "Digital Signal Processing." What is odd about this work is the author's strange "mathematical" approach to stochastic processes.

The subtitle illustrates the bizarre nature of this book, "A Nonprobabilistic Theory" (is nonprobabilistic a word?). In his Preface, Gardner attempts to justify his approach. He writes about the need for an empirical approach to the study of time series (I agree), and then he makes a brief attempt to use past mathematical work on randomness to justify his approach. For example, he uses an out-of-context quote from an obscure book by T. S. Fine (*Theories of Probability: An Examination of Foundations*, Academic Press, New York, 1973).

Gardner does *not* avoid probability and randomness in this exposition. What he does is develop a heuristic treatment of randomness of a realized time series in terms of what he calls the "fraction-of-time probabilistic model." His concept is a confused and confusing attempt to use well-known results from Markov process theory to set up an idiosyncratic theory of probability without the standard discipline from mathematics. Just imagine the confusion that would result in the application of statistical concepts if we had no standard and widely understood theory of probability to rationalize our data analysis and interpretation!

Both the author's Preface and the Forward by Bracewell reveal the confusion that reigns concerning the role of probability models in a statistical analysis of data and experimental design. The science of statistics deals with (1) the design of experiments to control the randomness and uncertainty in any experiment, and (2) the analysis of data as an organized process of scientific inquiry. The basic idea is to take representative samples from some defined population, and then make inferences about the population from the data. In time series analysis, the standard conceptualization is to assume that the time series we observe is a representative sample from an ensemble of series that are mathematically modeled by a joint distribution for the sample times.

One convenient way of thinking about the ensemble (which Gardner abhors) is to think of the observed series as a signal from a machine whose system parameters randomly vary over time in some fashion. The machine that generates the data is one of a set of identical machines that were manufactured at the same plant in the same time. The only difference between the machines is the setting of the varying parameters over time. A probability model is assumed for the random parameters for the set of machines. As long as we believe that the machine that generates the data is a representative randomized selection from the population of identical machines (the ensemble), then it is possible to make inferences about the population from the signal that we observed, provided that the process is stationary and satisfies some probabilistic mixing condition.

The mixing condition is crucial to statistical analysis of time series. It is not enough to assume stationarity and ergodicity. Ergodicity and stationarity are sufficient to develop consistent estimates of system parameters and the spectrum at a set of discrete frequencies. Consistency is convergence in probability. But convergence in probability, or even convergence almost surely to a point, is a very weak condition for an estimation method. In order to make inferences from sample estimates about population parameters, we need to know something about the statistical variability of the sample estimates. This requires the use of probability models and some sort of central limit theorem. Since the interesting cases in time series involve time-dependent random processes, we need to invoke a mixing condition to validate the application of a central limit theorem for asympotic analysis of estimates. If the finite sampling properties of an estimate are known, which may be the case if the process is Gaussian, then we do not need asympotics, but we do need probability models.

The author seems to be aware that he needs to use standard notation in parts of his book, so he writes about expected values using time averages. His approach is very hard to follow because he is "standing on his head" to adhere to this ideology of determinism.

It is too bad that the exposition in his book is harmed by such an odd misdirection of intellectual effort. The book covers a set of topics that are rarely treated in the signal processing literature: cyclostationarity and periodic signals with random components. The treatment of cyclostationarity is very idiosyncratic since Gardner is a pioneer in this field with L. Franks. A cyclostationary process is *not* ergodic, but it may be made so by sampling the process using periodic sampling synchronized with the periodicity of the process.

The simplest such process is a sine function whose amplitude and phase are fixed, but where the phase is drawn at random. Such a process is then stationary but not ergodic. Since it is easy to determine the period of the sinusoid as long as it is not observed with a lot of noise, it is easy to synchronize the sampling. Since the author defines cyclostationarity in a convoluted way using his time averaging formalism, the simplicity of the concept is totally obscured. This is one of many examples in the book where simple concepts are confused by Gardner's compulsion for "home-made" mathematics.

This book should not be used as a single source for a course in statistical spectral analysis because the students will not learn the basic concepts that everyone else learns (or attempts to). There is no reason for students in such a course to learn measure theory, but they should learn the basics of multivariate probability theory. Even social science students are able to learn and properly use these basic mathematical tools. Engineers and scientists have better mathematical background and are constantly drilled in the precise application of mathematical models, so they should be expected to master the simple mathematics required in statistics. The hard part of statistical analysis is the complexities inherent in inference, where there is no *right answer*. It is clear to me from reading this book and its Preface and Forward that most engineers do not understand the need for careful thinking about statistical matters. It is time for the professional statisticians to rise to the challenge of teaching statistics to engineers and scientists.

— Melvin J. Hinich

Author's Comment

Professor Hinich has written a severely critical review [1] of the book [2], and of engineers in scientists in general with regard to their understanding of statistics. As author of this book, I am responding to this criticism, not to try to establish whether or not this is a good book, but rather to counter Professor Hinich's bold claim that the philosophy put forth regarding the utility of a nonstochastic approach to time-series analysis, is misguided and that engineers and scientists who accept this philosophy as a viable conceptual tool do not understand statistics. In Professor Hinich's own words," Both the author's Preface and the Forward by Bracewell reveal the confusion that reigns concerning the role of probability models in a statistical analysis of data...it is clear to me from reading this book that most engineers do not understand the need for careful thinking about statistical matters. It is time for the professional statisticians to rise to the challenge of teaching statistics to engineers and scientists."

Everyone agrees, I think, that stochastic processes have their place— they are indeed useful in some applications. The question before us is, "Must we accept stochastic processes as the only viable approach to dealing with time-series data that is erratic or unpredictable? Or is there a viable alternative that is more useful in some applications?"

The standard approach in statistics is well suited to experimental design, data analysis, and inference for populations: When a population actually exists in the real world, application of the standard conceptual framework of orthodox statistics can be appropriate. However, when no population associated with the available data exists or when no such population can exist (e.g., in astronomy, the concept of a population of universes is not usually considered viable— replicating the "experiment" of creating the universe is rather far-fetched), then the appropriateness of pretending that a population exists should be questioned.

Many statisticians accept and invoke the concept of a population in developing their theories and methods. When one has a single time-series of data to analyze and use as a basis for making inferences about the real-world situation that gave rise to the data, and when one knows that there can be no access to a population of such time-series (because it does not or cannot exist), then one would be remiss in not questioning the appropriateness of the orthodox conceptualization of the time-series as one member of a population, or ensemble, mathematically modeled as a stochastic process.

Many—but by no means all—real-world problems in communications engineering and signal processing involve time-series data for which no population exists; that is, data for which replication of the experiment is impossible or impractical. However, many of these time-series are known to arise from physical phenomena that can be considered to be unchanging in their basic nature of very long periods of time. In such cases, conceptually idealizing this time-invariance by extending the length of the data record without bound enables us to conceive of a model that is derivable from the data in the limit as the amount of data used for measuring the parameters of the model approaches infinity. This leads us to the concept of a fraction-of-time (FOT) probability model that is free from the abstract concept of a population. For example, the FOT probability that a time-series exceeds some specified level is defined to be the fraction of time that this event occurs over the life of the time-series.

Once we have accepted the idea of an infinitely long time-series with an FOT probability model, we can develop a theory of statistical inference and decision that is isomorphic to the theory for stationary stochastic processes.

In summary, I believe Professor Hinich's admitted confusion about the message delivered in [2] results not from any flaw in the philosophy put forth in [2], but rather from his unwillingness to accept this philosophy as a viable alternative to his philosophy—a philosophy to which he clings tightly. Moreover, in defense of engineers and scientists, I mention that the book [3] on stochastic processes, coupled with [2], illustrates that some nonstatisticians are capable of understanding, using, and teaching both the orthodox theory of stochastic processes (for those situations where it is appropriate) and the unorthodox theory of time-series based on FOT probability (for those other situations where it is the more appropriate of the two).

Not all statisticians share Professor Hinich's inflexible position on what he calls "an odd misdirection of intellectual effort." Professor A. M. Yaglom of the Academy of Sciences of Russia, author of well known books on time-series analysis and stochastic processes, states in his review [4] of [2] and [3]:

"It is important, however, that until Gardner's second book was published there was no attempt to present the modern spectral analysis of random processes consistently in language that uses only time-averaging rather than averaging over the statistical ensemble of realizations. Moreover, this book also shows that such a treatment possesses some advantage over the traditional one...Professor Gardner's books are both valuable additions to the available literature on the theory of random processes."

Similarly, Professor Enders A. Robinson of Columbia University, author of thirty books on time-series analysis, states in his review [5]:

"This book can be highly recommended to the engineering profession. Instead of struggling with many unnecessary concepts from abstract probability theory, most engineers would prefer to use methods that are based on the available data. This highly readable book gives a consistent approach for carrying out this task. In this work Professor Gardner has made a significant contribution to statistical spectral analysis, one that would please the early pioneers of spectral theory and especially Norbert Wiener."

Apparently, the distinction to be made is not Professor Hinich's distinction between engineers/scientists and statisticians, but rather it is the distinction between pragmatists (as defined by the American philosophers Charles Sanders Peirce and William James), who can adopt whatever conceptualization best serves the practical purpose at hand, and those others who believe in the sanctity of one particular system of conceptualization, regardless of its practical consequences. The nonpragmatists speak of a controversy over the stochastic and nonstochastic approaches to time-series analysis. But, there is really no basis for controversy. The only real issue is one of judgement—judgement in choosing for each particular time-series analysis problem the most appropriate of two alternative approaches [6].

Making inferences from available data is tricky business for anyone. For example, Professor Hinich makes the following inference about a population numbering in the hundreds of thousands on the basis of data from a single member of this population: "It is clear to me [Hinich] from reading this book that most engineers do not understand the need for careful thinking about statistical matters."

References

1. M. J. Hinich, *Review of Statistical Spectral Analysis: A Nonprobabilistic Theory*, by W. A. Gardner, Prentice-Hall, 1987, Siam Review (1991), pp. 677-678.

2. W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1987.

3. W. A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems*, Macmillian, New York, 1985 (2nd Edition, McGraw-Hill, New York, 1990).

4. A. M. Yaglom, Review of *Introduction to Random Processes with Applications to Signals and Systems*, by W. A. Gardner, Macmillian, 1985, and *Statistical Spectral Analysis: A Nonprobabilistic Theory,* by W. A. Gardner, Prentice-Hall, 1987; *Theory Probab. Appl.* 35, (1990), pp. 405-407.

5. E. A. Robinson, Review of *Statistical Spectral Analysis: A Nonprobabilistic Theory*, by W. A. Gardner, Prentice-Hall, 1987; *Signal Processing* (EURASIP) **20** (1990), p. 99.

6. W. A. Gardner, "Two alternative philosophies for estimation of the parameters of time-series," *IEEE Transactions on Information Theory,* **37** (1991), pp. 216-218.

Ensembles In Wonderland

by William A. Gardner

Alice was an above-average engineering student at the University of the Queen of Hearts, but this semester she was struggling through a course on statistical signal processing, and things were getting curiouser and curiouser. When she was confronted with the topic of spectral analysis, she was perplexed by the mysterious reasoning given for spectrally smoothing the periodogram of a record of data produced by a receiver in a radio astronomy experiment. The argument Alice was presented with went something like this: Nevermind that you have only one record (which could go on forever) and that you want to use this record to learn something about the sources of radio energy impinging on Earth, pretend that your data is one record from an infinite ensemble of possible records (each of which could go on forever) that might have occurred, but didn't. Also, assume that this hypothetical ensemble is governed by a Gaussian probability law. To reduce the variance over the ensemble of these nonexistent records (from nonexistent universes), you should spectrally smooth the periodogram. Why? Because it is mathematically proven in the textbook that the variance over the make-believe ensemble, which defines a Gaussian stochastic process, is reduced by spectral smoothing. Oh, and by the way, this hypothetical variance also can be reduced by time averaging the periodogram instead of spectrally smoothing it. Why? Same reason: It is mathematically proven ...

Alice whispered to her classmate, "It would be so nice if someone would make sense for a change." Professor M. Hatter overheard Alice and turned to her with a stern look on his face and announced "This does make sense." Holding the textbook up for display, he said, "Just start at the beginning and when you reach the end, stop." Troubled by this, Alice decided she would get a second opinion. Later that day, she followed the Queen's Way to the campus repository of knowledge, the Cabbages and Kings Library. Hoping to find a more intuitively satisfying discussion of statistical spectral analysis, she located a score of books on the subject and proudly carried them off to a quiet corner of the reading room and began her quest for enlightenment. But, before long, her enthusiasm began to fade and in its place resignation set in. Alice found that all the books told the same story, whether authored by Tweedle Dee or Tweedle Dum. Stochastic processes were mathematically defined; the property of stationarity was introduced; the autocorrelation function was defined; the power spectral density function was defined to be the Fourier transform of the autocorrelation function; then the periodogram was introduced as a potential estimator of the power spectral density, but it was shown that even though the expected value over the ensemble could be made to approach the power spectral density by using long enough data records, this did not reduce the large variance over the ensemble. Then spectral smoothing was introduced as a means for reducing variance.

At this point, Alice could only conclude that the apparent absurdity of the make-believe ensemble and corresponding stochastic process was a result of her own failure to grasp something very deep and meaningful. After all, engineers do simulate ensembles on computers, and there was a lot of talk about something called virtual reality. So maybe the concept of reality was itself silly she thought, maybe nothing's impossible. Alice put her nose to the grindstone for the rest of the semester and learned the mechanics of stochastic processes, all the while trying to repress the thought "Why is it I give myself very good advice, but very seldom follow it?"

Time passed and Alice went on to earn a Ph.D. degree in, of all things, statistical signal processing. She considered a research position in industry, but was offered a professorship at the prestigious Stanford Institute of Technology in Berkeley, Massachusetts, which she happily accepted. One day after a lecture in which Alice eloquently spun the yarns about ensembles in wonderland that she had learned at Queen of Hearts, one of her students, with the grin of a cheshire cat, pointed out that a book review on this very topic had just appeared in a mathematics journal, and it seemed to suggest that engineers don't always view this topic the same way statisticians do. Alice got a copy of the book review and was amused to find the reviewer claiming that some foolish professor had written an entire book on an apparently ridiculous way of thinking about statistical spectral analysis that avoided stochastic processes and associated ensembles. The review [1] exposed the "bizarre" book [2] as an "odd misdirection of intellectual effort."

Reaching back in her mind, Alice recalled the difficulties she had experienced as a student at Queen of Hearts trying to grasp what then seemed to her to be a mystery, but now she accepted without question. But did she? There still seemed to be some doubt lurking in the recesses of her mind, in spite of the confidence with which she taught the subject of stochastic processes. Alice thought it would be fun to see first hand what this professor had to say. Upon finding the book in the library, she was surprised to see that the binding was broken and the cover was crumpled and soiled. There was a grimy imprint on the title page that read "Nike." It looked as though someone had thrown the book violently to the floor and stomped on it. She skimmed the table of contents and then turned to Chapter 3, where statistical spectral analysis was introduced. This is what she read:

In order to understand why a time--averaged periodogram can be preferable to one that has not been averaged, we must focus our attention not on the data itself but rather on the source of the data—the physical mechanism that generates the data. Generally speaking, data is nothing more than a partial representation of some physical phenomenon—a numerical representation of some aspects of the phenomenon. The fundamental reason for interest in a time-averaged spectrum of some given data is a belief that interesting aspects of the phenomenon being investigated have spectral influences on the data that are masked by uninteresting (for the purpose at hand) random effects and an additional belief (or, at least, hope) that these spectral influences can be revealed by averaging out the random effects. This second belief (or hope) should be based on the knowledge (or, at least, suspicion) that the spectral *influences of the interesting aspects of the phenomenon are time-invariant, so that the corresponding invariant spectral features (such as peaks or valleys) will be revealed rather than destroyed by time-averaging.*

Hmmm, she thought. This doesn't sound all that foolish. Let me go back and look at the introductory section in Chapter I, entitled "Objectives and Motives." She read:

A premise of this book is that the way engineers and scientists are commonly taught to think about empirical statistical spectral analysis of time-series data is fundamentally inappropriate for many applications. The subject is not really as abstruse as it appears to be from the conventional point of view. The problem is that the subject has been imbedded in the abstract probabilistic framework of stochastic processes, and this abstraction impedes conceptualization of the fundamental principles of empirical statistical spectral analysis. Hence, the probabilistic theory of statistical spectral analysis should be taught to engineers and scientists only after they have learned the fundamental deterministic principles—both qualitative and quantitative. For example, one should first learn 1) when and why sine wave analysis of time-series is appropriate, 2) how and why temporal and spectral resolution interact, 3) why statistical (averaged) spectra are of interest, and 4) what the various methods for measuring and computing statistical spectra are and how they are related. One should also learn how simultaneously to control the spectral and temporal resolution and the degree of randomness (reliability) of a statistical spectrum. All this can be accomplished in a nonsuperficial way without reference to the probabilistic theory of stochastic processes.

Alice turned back to Chapter 3 and read on to find that one can obtain approximately the same result by frequency smoothing the periodogram rather than time averaging it. There was a proof of this statement that consisted of manipulating the mathematical expression for the time-averaged periodogram into an expression for the frequency-smoothed periodogram. She jumped ahead to Chapter 5 and found a mathematical proof that the variance over time (not over some make-believe ensemble) was indeed reduced by both time averaging and frequency smoothing the periodogram.

In fact, there were many mathematical results that were essentially the same as those in the books that couch everything in terms of stochastic processes; it was the link between the mathematics and the real world that was different. This book used time averaging everywhere the other books used ensemble averaging. This seemed to make a lot of sense when the real world situation involves a single record of data from a physical phenomenon that is not changing with time.

Alice feverishly read on as her original motive for perusing the book turned into a reborn desire to solve the mystery she first encountered as a student at Queen of Hearts many years ago.

When she finally set the book down hours later with a deep feeling of satisfaction, Alice wondered why all the other books mask such naturally simple ideas behind ramblings about ensembles in wonderland. She suddenly remembered the review which led her to this unique book and she returned to her office to reread the review. She went to the concluding line in the review and read:

It is clear to me from reading this book that most engineers do not understand the need for careful thinking about statistical matters. It is time for the professional statistician to rise to the challenge of teaching statistics to engineers and scientists.

She reflected for a moment and then said to herself "How strange for a professional statistician to draw an inference about a population consisting of hundreds of thousands on the basis of observations on a single member of this population; certainly the concept of ergodicity does not apply here. Then something caught her eye. It was the only technical statement in the review. She read, "the simplest such process is a sine function whose amplitude and phase are fixed [over time], but where the phase is drawn at random [from an ensemble]. Such a process is then stationary but not ergodic. "Gee," Alice thought "maybe engineers have something to offer to statisticians: we can explain to them why this simple process is indeed ergodic, even though it does not pass the standard "mixing" test that they are so fond of.

Alice set the review aside and mused for awhile. It was late and she was tired. Little glimmers of curiousities from the history of engineering came to mind and faded—the heated controversy between physicists and mathematicians over Heaviside's operational calculus, now a standard part of engineering curricula—the mathematician's revulsion over the engineer's best friend, the impulse. She wondered if the current protests from some statisticians over this alternative approach to statistical spectral analysis would subside, and it too would become a standard part of engineering. One thing was clear. It would definitely become a standard part of the signal processing curriculum in her department at Stanford Institute of Technology.

The next day, Alice awoke early and went directly to her desk to prepare a new lecture to replace the one she had given yesterday. She wrote:

"The standard approach in statistics is well suited to experimental design, data analysis, and inference for populations: when a population actually exists in the real world, application of the standard conceptual framework of orthodox statistics can be appropriate. However, when no population associated with the available data exists or when no such population can exist, then one should question the appropriateness of pretending that a population does exist.

"Many statisticians accept and invoke the concept of a population in everything they do in developing theory and method. But unquestioning adherence to orthodoxy is not usually considered good science. It is a comforting approach for many but it is not always defensible on scientific grounds. When one has a single time-series of data to analyze and use as the basis for making inferences about the real-world situation that gave rise to the data, and when one knows that there can be no access to a population of such time-series (because none does or can exist), then one would be remiss in not questioning the appropriateness of the orthodox conceptualization of the time-series as one member of a population, or ensemble, mathematically modeled as a stochastic process.

"Many—but by no means all—real-world problems in communications engineering and signal processing involve time-series data for which no population exists, i.e., for which replication of the "experiment" is impossible or impractical. However, many of these time-series arise from physical phenomena that can be considered to be unchanging in their basic nature for a very long time. In such cases, conceptually idealizing this time-invariance by extending the length of time without bound enables us to conceive of a model that is derivable from the data in the limit as the amount of data used for measuring the parameters of the model "approaches infinity." This leads us to the concept of a fraction-of-time probability model that is free from the abstract concept of a population. (For example, the fraction-of-time probability that a time-series exceeds some specified level is defined to be the fraction of time that this event occurs over the life of the time-series.) ..."

Alice stopped writing for a moment to reflect. It occurred to her that the title of the book [2] *Statistical Spectral Analysis: A Nonprobabilistic Theory* could be misleading since it is shown in Chapter 5 of the book that an empirically motivated inquiry into the problem of quantifying the average behavior of spectral measurements leads naturally to a probabilistic theory. Since this probabilistic theory is nonstochastic (it involves only time averages, not ensemble averages), the title could have been *Statistical Spectral Analysis: A Nonstochastic Theory*. Nevertheless, she thought, the majority of the concepts and methods developed in the book are not only nonstochastic, they are indeed nonprobabilistic and a primary goal of the book is to show that, in an empirically motivated development of the fundamental concepts and methods of statistical spectral analysis, probability does not play a seminal role. It does play an important role in the mechanics of quantifying average behavior, but it plays no role in conceptualizing the objectives and methods of statistical spectral analysis of single time-series.

As Alice continued to write her lecture, she was keenly aware of a new level of enthusiasm that she could bring to her course on statistical signal processing. She would have to make a point of thanking the student with the cheshire cat grin for bringing that book review to her attention.

References

1. M. I. Hinich, Review of Statistical Spectral Analysis: A Nonprobabilistic Theory, by W. A. Gardner, Prentice-Hall, 1987, *SIAM Review*, pp. 677-678, December 1991.

2. W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1987.

SP Forum

To the Editor:

Congratulations on your inspired decision to volunteer SP Forum as a field on which the warriors of modern signal processing philosophy and their loyal clans, the Statisticians and the Fraction-of-Timers, may battle. As a partisan spectator, I am pleased to respond to your invitation to cast a thumbs-up/down and to interject my own views. But first, a confession: I am a statistician. It is therefore no surprise that I sympathize more with Professor Hinich's gleefully vicious, no-holds-barred review than with Professor Gardner's carefully worded, slyly mocking replies. (Perhaps my New York upbringing has something to do with this!)

For me, the statistical approach to signal analysis begins with a probabilistic model (e.g., ARMA) for the signal. The signal time series is viewed as a single realization and as data arising from the model. The time series data is used in conjunction with statistical techniques (e.g., maximum likelihood) to infer parameters, order, appropriateness, etc. of the model. The abstract notion of an infinite population plays no role.

When fitting a deterministic (e.g., chaotic) time series model, it may not be possible to employ a statistical approach in the manner described above. In this case, if the time series is ergodic (and many chaotic time series are), the fraction-of-time approach may be required, though not necessarily: in [1], it is shown that statistical model-fitting techniques developed for stochastic time series models can also be useful in fitting chaotic time series models.

Conversely, when the time series is not ergodic, as would be the case when the signal is transient, the fraction-of-time approach is not appropriate. It can be useful to model stochastic transient time series as harmonizable [2]. To estimate the generalized spectrum of such a time series, one would smooth the raw biperiodogram computed from each successive fixed-length block of data (cf [3]), and not average such estimates across blocks (as most blocks will contain no transient energy) nor let the block size increase without limit.

—Neil L. Gerr

References

1. Stochastic versions of chaotic time series: generalized logistic and Henon time series models, N. L. Gerr and J. C. Allen, *Physica* D, 68, pp. 232-249, 1993.

2. Time-delay estimation for harmonizable signals, N. L. Gerr and J. C. Allen, *Digital Signal Processing*, 4, pp. 49-62, 1994.

3. Spectral coherence of nonstationary and transient stochastic processes, H. L. Hurd, *Proceedings Fourth IEEE ASSP Workshop on Spectrum Estimation and Modeling*, pp. 387-390, 1988.

SP Forum

To the Editor:

It is hard for me to decide whether or not Mr. Gerr's letter in the Forum section of the October 1994 issue of this magazine deserves a response. He does not seem to address the basic issue of whether or not fraction-of-time probability is a useful concept. This is the issue being debated, isn't it? In fact, I cannot find one technical point in his letter that is both valid and clearly stated. But, because Mr. Gerr has clearly stated in his letter that, regarding philosophical issues in science and engineering, he prefers "New York" style vicious attacks like Hinich's to carefully worded slyly mocking replies, like mine, it has occurred to me that I might get through a little better to the Mr. Gerrs out there if I tried my hand at being just a little vicious. I hope the readers will understand that I am new at this; I give them my apologies now in case I fail to overcome my propensity for writing carefully and, when appropriate, slyly. But, before proceeding, I would like to set the record straight regarding what is and what is not slyly mocking. My reply to Hinich's review is nothing but plain fact without embellishment of any sort. It is my fictional article "Ensembles in Wonderland" which was obviously written to be humorous, that one could classify as slyly mocking.

Mr. Gerr's letter reveals a lot of misunderstanding and this provides us with some insight into what may motivate vicious attacks on attempts to educate people about alternative ways to conceptualize problem solving. It is hard for me to imagine how Mr. Gerr could have missed the main point of my response to Hinich's review. This point, which is clearly stated in both the book [1] under attack and the unappreciated response to this attack, is that, and I quote from my response,

> "There is really no basis for controversy. The only real issue is one of judgement—judgement in choosing for each particular time-series analysis problem the most appropriate of two alternative approaches."

To argue against this point is to be a zealot in the truest sense of the word, fanatically fighting for the One True Religion in statistics.

Sociologists and psychologists tell us that vicious behavior is often the result of paranoia born out of ignorance. In the example before us, both Hinich and Gerr demonstrate substantial ignorance regarding nonstochastic statistical concepts and methods, including fraction-of-time (FOT) probability. This case has already been made for Hinich in the Forum section of the April 1994 issue of this magazine. So let us consider Gerr's letter. First off, Gerr admits to the kind of behavior that is supposed to have no place in science and engineering, by identifying himself as a "partisan spectator." Webster's Ninth New Collegiate Dictionary defines partisan as "a firm adherent to a party, faction, or cause, or person, esp: one exhibiting blind, prejudiced, and unreasoning allegiance." On the basis of this admission alone one has to wonder whether to continue reading Gerr's letter or flip the page. (It's interesting that Gerr is into partisanship and Hinich's university appointment is in the Government Department.) But what the heck, let's see if we can find some technical content in his letter.

Mr. Gerr's first of three technical remarks is quoted here:

> "For me, the statistical approach to signal analysis begins with a probabilistic model (e.g., ARMA) for the signal. The signal time series is viewed as a single realization and as data arising from the model. The time series data is used in conjunction with statistical techniques (e.g., maximum likelihood) to infer parameters, order, appropriateness, etc. of the model. The abstract notion of an infinite population plays no role."

Not too surprisingly, it is difficult to tell what point Mr. Gerr is trying to make here. He starts with a probabilistic model and ends with a denial of the notion of a population. Would Mr. Gerr care to tell us how he interprets "probability" in "probabilistic model" if he denies the notion of population? My guess is that his thinking does not go this deep. But let's try to extract some meaning by reading between the lines. In spite of his sympathy with Hinich, Mr. Gerr seems to be agreeing that the problem-solving machinery of probability theory (e.g., ARMA modeling and maximum likelihood estimation) can be used regardless of whether one conceptualizes its use in terms of stochastic probability (with its associated ensembles or populations) or in terms of fraction-of-time (FOT) probability. This is the point that is made by the book [1] under attack: This book does include ARMA models and the maximum likelihood method as parts of the nonstochastic theory. True to the "blind allegiance" definition of partisanship, Mr. Gerr is apparently agreeing with the book while sympathizing with the attack on the book. Either Mr. Gerr has not read the book at all, or he may simply not have thought hard enough and long enough about these things. This is important to point out because I suspect it is the primary reason that there is any controversy at all.

Mr. Gerr then goes on to admit that the FOT approach may be required for chaotic time series. But again, true to form, he then makes a remark that is difficult to interpret:

> "The fraction-of-time approach may be required, though not necessarily: in [1], it is shown that statistical model-fitting techniques developed for stochastic time series models can also be useful in fitting chaotic time series models."

This sounds like Mr. Gerr is again confused about the fact that many probabilistic models can be interpreted or conceptualized in terms of either stochastic probability or FOT probability. Thus, regardless of the fact that a model was originally derived in the stochastic probability framework, it can—depending on the particular model—still be used (and/or rederived) in the FOT framework. In fact, AR models were originally derived within the FOT framework, not the stochastic framework [2] -[3]. This will probably surprise Mr. Gerr. And if he is not confused about this, then he is again agreeing with the book [1] whose attack he supports.

On the assumption that people working with stochastic processes would have enough of an understanding of the subject to compare it with the nonstochastic theory presented in [1], this comparison was not made very explicit in [1]. Responses to [1], such as those of Mssrs. Hinich and Gerr, suggest that this assumption is false more often than it is true. To make up for this, an explicit comparison and contrast between the theories of stochastic processes and nonstochastic time-series is made in Chapter 1 of [4].

Mr. Gerr concludes his letter by considering transient time-series and erroneously concluding that time averaging a biperiodogram over successive blocks of data (which he identifies with FOT methodology) is inappropriate, whereas spectrally smoothing a biperiodogram is appropriate. Obviously, he does not realize that the infamous book [1] that proposes FOT concepts and methods shows that when the data block, over which spectral smoothing of the biperiodogram is performed, is partitioned into subblocks over which time averaging of the biperiodogram is performed instead, the iesults from these two methods can closely approximate each other if the subblock length and window shape are chosen properly. In other words, it is very clearly explained in [1] that the FOT framework for spectral analysis includes frequency smoothing as well as time-averaging methods. This again brings up the question, did Mr. Gerr read the book [1], and if so, did he comprehend anything?

It is my recommendation to Mr. Gerr, and others who would entertain joining this discussion of the merit of considering alternatives to stochastic thinking, that the book [1] that started the furor so nicely exemplified by Hinich's review, and Chapter 1 of [4], be read carefully, the way they were written. This should be a prerequisite to criticism, vicious or otherwise.

Before closing this letter, I should point out that the so-called controversy that statisticians like Hinich and Gerr are promoting is about as productive as the statisticians' endless debate between the "Bayesians" and the "frequentists" over whether or not prior probabilities ("prior" meaning "before data collection") should be included in the One True Religion of statistics [5]. The debate is endless, because it is based on the faulty premise that there is One True Religion. In fact, the subject of our "controversy" is not unrelated to the Bayesian/frequentist debate. This debate dates back to the 1920s, and involves many well-known statisticians, some 40 of whom are referenced in [5] for their contributions to this debate. The conclusion in [5], published just last month, is, I am happy to report:

> "The Bayesians have been right all along! And so have the frequentists! Both schools are correct (and better than the other) under specific (and complementary) circumstances ... Neither approach will uniformly dominate the other... knowing when to [use] one or the other remains a tricky question. It is nonetheless helpful to know that neither approach can be ignored."

This is very encouraging! These pragmatic statisticians are attempting to dispell belief in the concept of One True Religion.

I conclude this reply with a little dialogue that I find both amusing and supportive of my response to vicious attacks:

> Can old dogs be taught new tricks?
>
> Maybe, but the teacher might get barked at for trying.
>
> Should the teacher accept the barking graciously?
>
> Maybe, but if the old dogs band together into a pack, the teacher better bark back.

— William A. Gardner

References

1. *Statistical Spectral Analysis: A Nonprobabilistic Theory*. W. A. Gardner, Prentice-Hall, 1987.

2. "On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers," G. U. Yule, *Phil. Trans. Royal Soc. London*, A, Vol. 226, pp. 267-298, 1927.

3. "On periodicity in series of related terms," G. Walker, *Proc. Royal Soc.*, Vol. 131, pp. 518- 532, 1931.

4. *Cyclostationariry in Communications and Signal Processing*, edited by W. A. Gardner, IEEE Press, 1994.

5. "Toward a reconciliation of the Bayesian and frequentist approaches to point estimation," F.J. Samaniego and D.M. Reneau, *Journal of the American Statistical Association*, Vol. 89, pp. 947-957, 1994.

SP Forum

To the editor:

I suppose when one ventures onto the field where the Fraction-of-Timers and Statisticians do battle, one must expect a few bumps and bruises. Still, I was surprised by the vehemence of Professor William Gardner's January response to my note in the October 1994 SP Forum. Even by New York standards it seemed a bit much!

Needless to say, there are many points on which Professor Gardner and I disagree, but only two that are worthy of further discussion. The first disagreement I would like to touch on regards the utility of the fraction-of-time approach for spectrum analysis of transient signals in particular, and nonstationary signals in general. Professor Gardner states:

> "...that when the data block, over which spectral smoothing of the biperiodogram is performed, is partitioned into subblocks over which time averaging of the biperiodogram is performed instead, the results from these two methods can closely approximate each other if the subblock length and window shape are chosen properly."

Why not smooth the biperiodogram computed on each subblock? The resulting generalized spectrum estimates can be used in their own right or in a time averaging scheme. The fact is, there is value in spectral smoothing when estimating the generalized spectrum of a harmonizable (i.e., nonstationary) signal, and these techniques do not require a careful choice of subblock lengths and windows (cf. [1], which contains basic theory, simulations, and an application to real data). Thus, spectral smoothing of the biperiodogram is to be preferred when little is known of the signal *a priori*. Conversely, when analyzing a cyclostationary signal whose cycle frequency is known, time averaging of the raw (i.e., unsmoothed) biperiodogram computed from disjoint data blocks that span individual cycle periods can, as Professor Gardner has shown in many of his writings, be of great utility.

The second disagreement I would like to address relates to what I have referred to as a battle of philosophies: fraction-of-time versus probability/statistics. Few would argue with the assertion that a tremendous amount of progress in the development of signal processing methodology has resulted from work based on the statistical/probabilistic paradigm, and that this is a paradigm with which most of the signal processing community is familiar. I find little to suggest that this paradigm has run out of gas. As evidence for this claim, I note the current high levels of interest and creativity worldwide in high order statistics/cumulant polyspectra and the recent extension and refinement of an important concept like circularity in the (probabilistic) context of harmonizable processes [2]. Professor Gardner has chosen to work within the context of an alternative paradigm, in many respects equivalent to the old but with no obvious advantages, and is surprised to find that there are those who are skeptical. (Some may even be hostile, though I do not count myself among that group!) Professor Gardner errs when he likens such skeptics to religious fanatics. The fact is, Professor Gardner has failed to make a compelling case for his paradigm. From my perspective, developing signal processing results using the fraction-of-time approach (and not probability/statistics) is like building a house without using power tools: it can certainly be done, but to what end?

—Neil L. Gerr

References

1. The Generalized Spectrum and Spectral Coherence of a Harmonizable Time Series, N. L. Gerr and J. C. Allen, *Digital Signal Processing*, vol. 4, pp. 222-238, 1994.

2. On Circularity, B. Picinbono, *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3473-3482, December 1994.

References

1. J. Lynn Smith, "Breaking the Nyquist Barrier," *IEEE Signal Processing Magazine*, July 1995, pp 41-43.

2. C. Shannon, "Communications in the presence of noise," *Proceedings of the IRE*, January, 1949, pp 10-21.

3. Oppenheim and Wilsky, *Signals and Systems*, Prentice-Hall, 1983, pp 514-526.

Questioning 'Distinction without a Difference' Debate

While I was entertained by the exchange of letters concerning fraction-of-time and probabilistic models, I was startled by the reference (*Signal Processing Magazine*, March 1995, SP Forum) to [1] as a "recent extension and refinement of an important concept..." This contention is too much for me, and I feel compelled to take [1] together with the fraction-of-time versus probablistic models debate as examples to reemphasize a broader point. Engineering is applied science, has a product, and we do not need to carry on "distinction without a difference" debates, either for time series or complex data models. Ideally, authors submitting papers to the IEEE journals have something to contribute to engineering. Contributions include reporting upon techniques that proved effective for achieving an engineering goal, comparing techniques, or explaining in a general context why techniques are effective, and identifying limitations. I see a growing disconnect between scientists and engineers, and the publications of academics, which are too often dubious mathematics.

The reference paper supplies only abstract results, and if better examples were used, it would be apparent to nearly all readers that the introduced terms are unnecessary and overblown. The paper invents jargon to compensate for a lack of a consistent definition for the common concept of "complex random number." A complex random number is nothing more than notation for pairs of random numbers, a notation which simplifies many algebraic manipulations. Measurable physical quantities, including those modeled as random variables, take on values that are real numbers. Imaginary numbers are, well, imaginary. Even in the abstract, there is no useful distinction between a complex random variable, and pairs of real random variables. There is an isomorphism between the two-dimensional Euclidean plane, in which probability measures are naturally defined, and the complex numbers. The probability distribution functions of "complex" random variables are not even functions solely of the complex variables. The probability distribution functions are functions of real combinations of the complex variables and their complex conjugates. In other words, the probability distribution functions of "complex" random variables are real functions of the real and imaginary components of the complex random variables. (And conversely. Every distribution of pairs of random variables, x, y, can be written in terms of $z = x + iy$ and $z^* = x - iy$ using $2x = z + z^*$ and $2iy = z - z^*$). The complex shorthand is convenient for algebraic manipulation of, for example, analytic signals, but this convenience does not make the results of any measurement complex. Complex random numbers are a well accepted, often convenient, notation for pairs of real random numbers. The author suggests no definition for a complex random number that requires an estimation or detection theory distinct from the well established formalisms. Even [1] notes that, "the standard procedure is to use the real and imaginary parts that are the components of a 2-D real random vector. However, doing so results in considering that complex numbers are nothing else but pairs of real numbers, and the complex theory loses most of its interest." Indeed.

The author's "complex LMSE" solutions are constrained to be complex constructs put together for notational convenience. Only by misplacing physical significance to the algebraic construct does there become any need to distinguish to types of linear estimates. A real formulation is the correct formulation for engineering problems, and consequently is the only linear estimate that need be considered. Estimates with independent coefficients for each measured physical quantity are the formulations of interest to engineering, albeit, there are common cases for which the optimal solution for the coefficients has symmetries that permit the optimal linear estimates to be written in a "complex" form. But, many solutions have distinguishable properties, and each distinguishable property of particular solutions does not require a reformulation for the general problem. Taking note of conditions that imply solutions of a particular form does lead to efficiency for solution finding. Nevertheless, alternative algorithms selected for computational efficiency are not reformulations of the problem.

For models exhibiting circularity, the constrained "complex" estimates coincide with the optimal estimates. But, circularity is only a sufficient, and not a necessary condition for the constrained solution to coincide with the optimal solution.

I am not aware of a single system that is suboptimal because a design engineer mistakingly insisted upon solving the overly constrained, "complex" optimization problem. None of the machinations introduced in the cited reference are required if there is no insistence on making too much of the complex notation for pairs of random variables. If there are notable applications of circularity, [1] chooses not to enlighten us. (And I do not count putting already solved problems into new notation.) As a young mathematician once stated, "we cannot continue to forget at the present rate. Total ignorance provides a convenient lower bound."

—Glenn Johnson
TASC, Reston, VA

Reference

1. Bernard Picobono, "On Circularity," IEEE Transactions on Signal Processing, Vol. 42, pp. 3473-3482, Dec. 1994.

Gardner Rests His Case

This is my final letter to *SP Forum* in the debate initiated by Mr. Melvin Hinich's challenge to the resolution made in the book [1], and carried on by Mr. Neil Gerr through his letters to SP *Forum*.

In this letter, I supplement my previous remarks aimed at clarifying the precariousness of Hinich's and Gerr's position by explaining the link between my argument in favor of the utility of fraction-of-time (FOT) probability, and the subject of a plenary lecture delivered at ICASSP '94. In the process of discussing this link, I hope to continue the progress made in my previous two letters in discrediting the naysayers and thereby moving toward broader acceptance of the resolution that was made and argued for in [1], and is currently being challenged. My continuing approach is to show that the position taken by the opposition—that the fraction-of-time probability concept and the corresponding time-average framework for statistical signal processing theory and method have nothing to offer in addition to the concept of probability associated with ensembles and the corresponding stochastic process framework—simply cannot be defended if argument is to be based on fact and logic.

Thomson's Waveguide Problem

To illustrate that the stochastic-process conceptual framework is often applied to physical situations where the time-average framework is a more natural choice, I have chosen an example from D. J. Thomson's recent plenary lecture on the project that gave birth to the multiple-window method of spectral analysis [2]. The project that was initiated back in the mid-1960s was to study the feasibility of a transcontinental waveguide for a telecommunications transmission system potentially targeted for introduction in the mid-1980s. It was found that accumulated attenuation of a signal propagating along a circular waveguide was directly dependent on the spectrum of the series, indexed by distance, of the erratic diameters of the waveguide. So, the problem that Thomson tackled was that of estimating the spectrum for the more than 4,000 mile long distance-series using a relatively small segment of this series that was broken into a number of 30-foot long subsegments. (It would take more than 700,000 such 30-foot sections to span 4,000 miles). The spectrum had a dynamic range of more than 100 dB and contained many periodic components, indicating the unusual challenge faced by Thomson.

When a signal travels down a waveguide (at the speed of light), it encounters the distance-series of erratic waveguide-diameters. Because of the constant velocity, the distance-series is equivalent to a time-series. Similarly, the series of diameters that is measured for purposes of analysis is— due to the constant effective velocity of the measurement device—equivalent to a time series. So, here we have a problem where there is one and only one long time-series of interest (which is equivalent to a distance-series)—there is no ensemble of long series over which average characteristics are of interest and, therefore, there is no obvious reason to introduce the concept of a stochastic process. That is, in the physical problem investigated, there was no desire to build an ensemble of transcontinental waveguides. Only one (if any at all) was to be built, and it was the spectral density of distance-averaged (time-averaged) power of the single long distance-series (time-series) that was to be estimated, using a relatively short segment, not the spectral density of ensemble-averaged power.

Similarly, if one wanted to analytically categorize the average behavior of the spectral density estimate (the estimator mean), it was the average of a sliding estimator over distance (time), not the average over some hypothetical ensemble, that was of interest. Likewise, to characterize the variability of the estimator, it was the distance-average squared deviation of the sliding estimator about its distance-average value (the estimator variance) that was of interest, not the variance over an ensemble. The only apparent reason for introducing a stochastic process model with its associated ensemble, instead of a time-series model, is that one might have been trained to think about spectral analysis of erratic data *only* in terms of such a conceptual artifice and might, therefore, have been unaware of the fact that one could think in terms of a more suitable alternative that is based entirely on the concept of time averaging over the single time-series. (Although it is true that the time-series segments obtained from multiple 30 ft. sections of waveguide could be thought of as independent random samples from a population, this still does not motivate the concept of an ensemble of infinitely long time-series—a stationary stochastic process. The fact remains that, physically, the 30-foot sections represent subsegments of one long time-series in the communications system concept that was being studied.)

It is obvious in this example that there is no advantage to introducing the irrelevant abstraction of a stochastic process (the model adopted by Thomson) except to accommodate unfamiliarity with alternatives. Yet Gerr turns this around and says there is no obvious advantage to using the time-average framework. Somehow, he does not recognize the mental gyrations required to force this and other physical problems into the stochastic process framework.

Gerr's Letter

Having explained the link between my argument in favor of the utility of FOT probability and Thomson's work, let us return to Gerr's letter. Mr. Gerr, in discussing what he refers to as "a battle of philosophies," states that I have erred in likening skeptics to religious fanatics. But in the same paragraph, we find him defensively trying to convince his readers that the "statistical/probabilistic paradigm" has not "run out of gas," when no one has even suggested that it has. No one, to my knowledge, is trying to make blanket negative statements about the value of what is obviously a conceptual tool of tremendous importance (probability) and no one is trying to denigrate statistical concepts and methods. It is only being explained that interpreting probability in terms of the fraction-of-time of occurrence of an event is a useful concept in some applications. To argue, as Mr. Gerr does, again in the same paragraph, that in general this concept "has no obvious advantages" and using it is "like building a house without power tools: it can certainly be done, but to what end?" is, as I stated in my previous letter, to behave like a religious fanatic—one who believes there can be only One True Religion. This is a very untenable position in scientific research.

As I have also pointed out in my

previous letter, Mr. Gerr is not at all careful in his thinking. To illustrate his lack of care, I point out that Gerr's statement "Professor Gardner has chosen to work within the context of an alternative paradigm (fraction-of-time probability)," and the implications of this statement in Gerr's following remarks completely ignore the facts that I have written entire books and many papers within the stochastic process framework, that I teach this subject to my students, and that I have always extolled its benefits where appropriate. If Mr. Gerr believes in set theory and logic, then he would see that I cannot be "within" paradigm A and also within paradigm B, unless A and B are not mutually exclusive. But he insists on making them mutually exclusive, as illustrated in the statement "From my perspective, developing signal processing results using the fraction-of-time approach (and not probability/statistics)..." (The parenthetical remark in this quotation is part of Mr. Gerr's statement.) Why does Gerr continue to deny that the fraction-of-time approach involves both probability and statistics?

Another example of the lack of care in Mr. Gerr's thinking is the convoluted logic that leads him to conclude "Thus, spectral smoothing of the biperiodogram is to be preferred when little is known of the signal *a priori*." As I stated in my previous letter, it is mathematically proven in [1] that the frequency smoothing and time averaging methods yield approximately the same result (a more detailed and tutorial proof of this fundamental equivalence is given in the article "The history and the equivalence of two methods of spectral analysis," in review for this publication). Gerr has given us no basis for arguing that one is superior to the other and yet he continues to try to make such an argument. And what does this have to do with the utility of the fraction-of-time concept anyway? These are data processing methods; they do not belong to one or another conceptual framework.

To further demonstrate the indefensibility of Gerr's claim that the fraction-of-time probability concept has "no obvious advantages," I cite two more examples to supplement the advantage of avoiding "unnecessary mental gyrations" that was illustrated using Thomson's waveguide problem. The first example stems from the fact that the fundamental equivalence between time averaging and frequency smoothing referred to above was first derived by using the fraction-of-time conceptual framework [1]. If there is no conceptual advantage to this framework, why wasn't such a fundamental result derived during the half century of research based on stochastic processes that preceded [1]?

The second example is taken from the first attempt to develop a theory of higher-order cyclostationarity for the conceptualization and solution of problems in communication system design. In [3], it is shown that a fundamental inquiry into the nature of communication signals subjected to nonlinear transformations led naturally to the fraction-of-time probability concept, and to a derivation of the cumulant as the solution to a practically motivated problem. This is, to my knowledge, the first *derivation* of the cumulant. All other work, which is based on stochastic processes (or non-fraction-of-time probability) and which dates back to the turn of the century, cumulants are defined, by analogy with moments, to be coefficients in an infinite series expansion of a transformation of the probability density function (the characteristic function), which has some useful properties. If there is no conceptual advantage to the fraction-of-time framework, why wasn't the cumulant derived as the solution to the above-mentioned practical problem or some other practical problem using the orthodox stochastic-probability framework?

Conclusion

Since no one in the preceding year has entered the debate to indicate that they have new arguments for or against the philosophy and corresponding theory and methodology presented in [1], it seems fair to proclaim the debate closed. The readers may decide for themselves whether the resolution put forth in [1] was defeated or was upheld. But regarding the skeptics, I sign off with a humorous anecdote:

When Mr. Fulton first showed off his new invention, the steamboat, skeptics were crowded on the bank, yelling "It'll never start, it'll never start."

It did. It got going with a lot of clanking and groaning and, as it made its way down the river, the skeptics were quiet.

For one minute.

Then they started shouting, "It'll never stop, it'll never stop."

—William A. Gardner
University of California, Davis

References

1. W. A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory*. Prentice-Hall, Englewood Cliffs, NJ, 1987.

2. D. J. Thomson. "An overview of multiple-window and quadratic-inverse spectrum estimation methods," Plenary Lecture. In *Proceedings of the 1994 International Conference on Acoustics, Speech, and Signal Processing*, pp. VI-185 - VI-194.

3. W. A. Gardner and C. M. Spooner, "The cumulant theory of cyclostationary time-series, Part I: Foundation," *IEEE Transactions on Signal Processing*, Vol. 42, December 1994, pp. 3387-3408.