Rejoinder (part 2)

L. Mark Berliner

So far as the laws of mathematics refer to reality, they are not certain. And so far as they are certain, they do not refer to reality.

Albert Einstein (1921)

I am grateful to all the discussants for their interesting and thought-provoking comments. I am gratified that Professors Cutler, Geweke, Griffeath, Smith and Tsay generally agree with the propositions that the ideas assoicated with chaos should be of interest to statisticians, and that statisticians can make valuable contributions to this area. Cutler and Smith presented valuable expansions, particularly in the context of the estimation and interpretation of "dimension," on some topics for which my review is only cursory. I strongly agree with Professor Griffeath's view that the real issues discussed in various literatures and under various names really all hinge on "complexity." I also appreciate his discussion of random number generation and cellular automata. Indeed, since I agree with so much of what the discussants said, I will only comment on remaining points of contention or additional suggestions.

RESPONSE TO GEWEKE

I enjoyed Geweke's suggested Bayesian analysis of chaotic models, based on symbolic dynamic data, in the presence no traditional randomness. The analysis is exactly the sort of thing I believe statisticians can contribute to problems of chaotic data analysis. I wish to raise a point involving the specific computational algorithm he used, as outlined in the second paragraph of his Section 4. Specifically, for a fixed parameter value a, he suggested that the twentieth literates of many, equally spaced values may yield an approximation to the natural (Bowen–Ruelle) ergodic distribution for that a. This is not necessarily true. In doing a similar analysis, Steve MacEachern and I noticed that this is not true for a corresponding to an attracting periodic attractor. For finite attractor, the invariant distribution must assign equal probabilities on the attractor. However, uniformly spaced points are not attracted to the limit points uniformly. For example, I ran 5000 equally spaced points in the interval $(0, 0.5)$ for 500 iterates each using the logistic map with $a = 3.4$. This a corresponds to a period 2

attractor, consisting of roughly 0.452 and 0.842. Only 44% of the 5000 points were at 0.842, whereas the remaining 56% were at 0.452. Simply stated, Geweke's original suggestion only guarantees, up to numerical complications, that we approximate the support of the desired ergodic distribution. Therefore, he is right in his concern about the validity of the approximation in the periodic case, although I suspect the problem diminishes as the number of periods increases. For a yielding a continuous ergodic distribution, statistical regularity suggests that Geweke's method would indeed yield a good, again up to numerical complications, approximation.

I think these points are crucial to our interpretation of ergodic probabilities. In the period 2 attractor case above, it is true that almost all (Lebesgue measure) points in the unit interval yield paths which, after a transience period, spend 50% of the time near each of the limit points. However, if the initial condition is randomly and uniformly generated from the unit interval, it is clearly not reasonable to claim that, for every very large N, x_N is equally likely to be near each of the limit points. Further, note that statistical regularity generally requires continuous ergodic distributions. In the case of a finite attractor, if the initial condition is generated according to a distribution that assigns all its mass to the limit points, but not with equal probabilities, the initial distibution never washes out.

RESPONSE TO GRIFFEATH

I have lingering doubts concerning the special role Griffeath appears to ascribe to "truly random" processes. He alludes to the powerful tools of probability theory, namely, the central limit theroem and the law of the iterated logarithm. I suggest that the ergodic theorem may well be included in such a list, and that the relationship between ergodic processes and stationary stochastic processes can be viewed as establishing a link between "truly random" and other complex processes.

Griffeath, in the role of purist, suggests that he debunks my probability statement given in Section 2. Perhaps my Bayesian tint is too strong to appreciate the impact of his argument. In particular, I readily model my uncertainty via probability statements. To clarify, consider the following two points.

First, regarding my original claim in Section 2, I offer Professor Griffeath the following bet. We agree to set x_0 equal a reasonably precise computer representation of π^{-1} and run the logistic map with $a = 4.0$ for a very large number of iterates on a computer in double precision. (The exact number of iterations $N + 1$ will be determined by a neutral party, perhaps the Editor, and announced before Griffeath must decide on the bet.) If $x_N \geq 0.5$ and $x_{N+1} < 0.5$, I pay Griffeath $100. If not, Griffeath pays me $29. Notice that I ask for only $29, rather than $33 as suggested by a fair bet according to my probability statement. This is because my utility function accounts for the fact that I "win" if he even accepts the bet. It is interesting to ask whether or not my probably statement is Bayesian statement. In some sense the answer is no in that a rigorous application of Bayes' Theorem should require that once I know N and x_0, my posterior on x_N and x_{N+1} should be degenerated on the points $f^N(x_0)$ and $f^{N+1}(x_0)$. Thus, I am behaving incoherently in a rigorous sense. On the other hand, if x_0 is gerenated by any random mechanism that is absolutely continuous with respect to Lebesque measure, my probability statement is correct for very large N by statistical regularity. The key to my statement is that I would not change my statement after learning the value of the generated x_0; indeed, I would maintain the probability statement for any transcendental x_0. The only way to implement Bayes' Theorem for such x_0 is to find f^N; there is no shortcut. The complexity implicit in chaotic systems suggests that we may ask, "What does it mean to 'know' f^N?" As a related note, this is essentially the same question asked by Diaconis (1988) in his article on Bayesian numerical analysis.

Second, a "purist's" view of random variables is that they are simply *deterministic*, measurable functions defined on some probability space. Imagine an experiment in which the height, X, of a randomly chosen male listed in the Columbus White Pages 1990–91 Phone Directory is to be observed. A purist is happy to discuss the induced probability distribution for X. Suppose I now tell you that the name selected is on page 314, and indeed the selected person's name is P. Freidli. If we ignore measurement error for the sake of argument, are all traces of purist randomness gone? (A common view of measurement error in the physical sciences is something like, "We can actually learn the height up to the nearest unit of measurement.") Regardless of pure frequentist position, I still think of X as a random variable. (Actually, I did not even pick this person "at random"—314 are the first three

digits of π.) The point is that the knowledge of the person may be uniformative, at least to most of us. (Some readers might actually know Mr. Freidli, or at least something about heights associated with his name. However, this should simply change the distribution for X, although not to a degenerate one.) This parallels the argument in the context of chaos (with continuous ergodic distributions) and the uninformative nature, from a practical viewpoint, of initial conditions, and also suggests that the basic issue of relaxing the notion of true randomness may be quite pervasive in the application of statistical reasoning.

RESPONSE TO GRANGER

While I agree with the skepticism he alludes to concerning the presentation of chaos as the greatest thing since sliced bread, I find little else in this discussion with which to agree. The only exception is Granger's questioning of the value of some of the advertised "tests for chaos." Most of Granger's lessons in logic and the removal of confusion were too subtle for me. His closing paragraph left me bewildered.

Professor Granger's view of the implications of statistical hypothesis testing is so remote from my own, I fear he and I do not have sufficient common ground from which I can proceed in an expeditious fashion. His attempts at teaching me the logic of statistical hypothesis testing, although he may be surprised to know that I have heard it before, failed. The trouble stems from the fact that before studying testing, I learned in probability theory that $P(A \mid B)$ is, alas, not the same as $P(B \mid A)$. Putting debates over the philosophy of statistics aside, I think Granger believes that mathematical models, whether deterministic, random, chaotic or whatever, are truly either right or wrong, and we are able to learn which. I think mathematical models, although subject to constant scrutiny and potential replacement, are simply tools by which we organize and communicate our thoughts and projections about how things work. The quote with which I opened this rejoinder was primarily intended for him.

RESPONSE TO SMITH

As in the case of Professor Cutler, Professor Smith presents an excellent and valuable commentary on recent work in statistics and chaos, as well as some thoughts on the future role of statisticians. The only portion I will respond to is his remarks, toward the end of his discussion, concerning my view of chaos and ergodic theory with regard to philosophy. I do not believe my views are "diametrically

opposite" to those of J. Durbin, as quoted by Smith. Indeed, Smith is right in suggesting that my interpretations of the impact of ergodic theory, as discussed in my Section 2, are quite close to those of Durbin. However, my final opinions, as presented in Section 5, involve the claim that ergodic theory still does not fully specify appropriate distibutions in all problems. For example, Durbin's claim, "... the postulation of objective probability models in physical situations such as games of chance has been strengthened" is certainly reasonable, but does not mean the case has been closed. Ergodic theory generally suggests a natural starting point for analysis. Before I apply ergodic theory and the laws of physics, I must subjectively assess the degree to which my models capture the important aspects of the situation at hand; in the context of games of chance, I must also assess the likelihood of cheating. Furthermore, real problems in which parameters of models are unknown, such as alluded to in my Section 4 and Geweke's comments, are not fully solved by ergodic theory. Even if all pertinent ergodic distributions are completely known, I still need a distribution on the parameters in order to make probabilistic predictions. Similarly, by definition, ergodic theory does not typically provide sufficient information to enable analysts to incorporate observational information in the calculation of short-term predictions of dynamical systems. [I know of, but cannot reference, actual, moderately successful attempts by enterprising applied chaologists to (covertly) electronically monitor spins of roulette wheels to use trial dependent information to decide, just before the croupier closes betting, on bets with improved odds.] To conclude, I find no contradiction in claiming to be a Bayesian, philosophically, who enthusiastically uses scientific reasoning (physics, ergodic theory, etc.) in the construction of models. Indeed, when faced with outrageous computational difficulties such as in the bet I offered to Griffeath, I am ready to abandon fully Bayesian caculations in favor of a readily available, mathematially justifiable solution. I am happy to suggest appropriate ergodic distributions as "objective Bayesian" models. (With no small trepidation, I even suggest that a philosophy based on the sole reliance on ergodic distributions may encompass and lend further credibility to some portions of the maximum entropy school of objective Bayesian statistics.)

REMARKS

Despite the emphasis of this rejoinder and portions of my article, I hope that the discussions here do not suggest that chaos is just another arena for debate between classical and Bayesian statisti-

cians. While these issues first attracted me to this topic and I believe philosophy is both interesting and important, chaos/complexity offers opportunities for contributions from statisticians of all philosophies. Indeed, chaos may suggest the blending of philosophies: to the classical statistician, the notion of "uncertainty modeling via probability" should be strengthened, while the meaning of "truly random" is blurred. On the other hand, ergodic theory suggests that the notion of "true" or "correct" priors deserves more credibility from subjective Bayesians. I think the following advice from Ruelle (1991, page 13) is pertinent: "Don't embark in general abstract discussions as to whether physics is deterministic, or probabilistic, . . . , and so on. The answer depends on the physical theory considered, and how determinism, or chance, . . . , is introduced in this theory."

As to whether chaos is as important an idea as relativity and quantum mechanics, I will not speculate. Except for those statisticians and probabilists who, up to now, exclusively work on problems related either to relativity and/or quantum mechanics, the question does not seem to me to be terribly relevant to our community. It is relevant that chaos offers fascinating and challenging problems for which probabilistic and statistical solutions would be of interest to the larger scientific community.

ADDITIONAL REFERENCES

AMATO, I. (1991). Speculating in precious computronium. *Science* **253** 856–857.

ANOSOV, D. V. (1967). Geodesic flows and closed Riemannian manifolds with negative curvature *Proc. Steklov Inst. Math.* **90**.

ARRATIA, R., GOLDSTEIN, L. and GORDON, L. (1990). Poisson approximation and Chen–Stein method (with discussion). *Statist. Sci.* **5** 403-434.

BARBOUR, A. D. and EAGLESON, G. K. (1984). Poisson convergence for dissociated statistics *J. Roy. Statist. Soc. Ser. B* **46** 397-402.

BASSINGTHWAIGHTE, J. B., KING, R. B. and ROGER, S. A. (1989). Fractal nature of regional myocardial blood flow heterogeneity. *Circulation Research* **65** 578-590.

BENNETT, C. H. and GARDNER, M. (1979). The random number omega bids fair to hold the mysteries of the universe. *Scientific American* **241** 20-34.

BØGSTED HANSEN, M. (1991). The behavior of the correlation integral in the nonlinear time series case. In preparation.

BOWEN, R. (1975). ω-limit sets for Axiom A diffeomorphisms *J. Differential Equations* **18** 333-339.

BOWEN, R. and RUELLE, D. (1975). The ergodic theory of Axiom A flows. *Invent. Math.* **29** 181-202.

BROCK, W. A., DECHERT, W. D. and SCHEINKMAN, J. A. (1987). A Test of Independence Based on the Correlation Dimension. Working Paper 8702, Dept. Economics, Univ. Wisconsin, Madison.

BROCK, W. A. and SAYERS, C. L. (1988). Is the business cycle characterized by deterministic chaos? *Journal of Monetary Economics* **22** 71-90.

CASDAGLI, M. (1992). Chaos and deterministic versus stochastic nonlinear modeling (with discussion). *J. Roy. Statist. Soc. Ser. B.* To appear.

CASTI, J. (1990). *Searching for Certainty.* William Morrow, New York.

CASTI, J. (1991). *Beyond Belief: Randomness, Prediction and Explanation in Science.* CRC Press, Boca Raton, Fla.

CHEN, R. and TSAY, R. S. (1991). On the ergodicity of TAR(1) processes. *Ann. Appl. Probab.* 1 613-634.

CHENG, R. and TONG, H. (1991). On consistent nonparametric order determination and chaos. *J. Roy. Statist. Soc. Ser. B.* To appear.

CORLESS, R. M. (1991). Defect controlled numerical methods and shadowing for chaotic differential equations. Preprint.

CORLESS, R. M., ESSEX, C. and NERENBERG, M. A. H. (1991). Numerical methods can suppress chaos. *Phys. Lett. A* 157 27-36.

COVER, T., GACS, P. and GRAY, R. (1989). Kolmogorov's contributions to information theory and algorithmic complexity. *Ann. Probab.* 17 840-865.

CROWELL, R. E. and PETERS, E. E. (1991). Chaos theory weakens efficient market idea. *Pensions and Investments* 10 14-15.

CUTLER, C. D. (1990a). A dynamical system with integer information dimension and fractal correlation exponent. *Comm. Math. Phys.* 129 621-629.

CUTLER, C. D. (1990b). Connecting ergodicity and dimension in dynamical systems. *Ergodic Theory Dynamical Systems* 10 451-462.

CUTLER, C. D. (1990c). Measure disintegrations with respect to σ-stable monotone indices and the pointwise representation of packing dimension. In *Proceedings of the 1990 Measure Theory Conference at Oberwolfach.* To appear.

CUTLER, C. D. (1991). Some results on the behavior and estimation of the fractal dimensions of distributions on attractors. *J. Statist. Phys.* 62 651-708. (Errata: To appear.)

CUTLER, C. D. and DAWSON, D. A. (1990). Nearest neighbor analysis of a family of fractal distributions. *Ann. Probab.* 18 256-271.

DESTEXHE, A. (1990). Symbolic dynamics from biological time series. *Phys. Lett. A* 143 373-378.

DIACONIS, P. (1988). Bayesian numerical analysis. In *Statistical Decision Theory and Related Topics IV* (S. S. Gupta and J. O. Berger, eds.) 1 163-175. Springer, New York.

DURBIN, J. (1987). Statistics and statistical science (Presidential Address of the Royal Statistical Society). *J. Roy. Statist. Soc. Ser. A* 150 177-191.

DURBIN, J. (1990). Discussion of "Chance or chaos?" by M. S. Bartlett. *J. Roy. Statist. Soc. Ser. A* 153 338.

DURRETT, R. (1988). *Lecture Notes on Particle Systems and Percolation.* Wadsworth, Pacific Grove, Calif.

DURRETT, R. (1991). Some new games for your computer. *Nonlinear Science Today.* To appear.

FAN, J. (1990). Local linear regression smoothers and their minimax efficiency. Technical report, Univ. North Carolina, Chapel Hill.

FARMER, J. D. and SIDOROWICH, J. J. (1990). Optimal shadowing and noise reduction. *Phys. D* 47 373-392.

FISCH, R., GRAVNER, J. and GRIFFEATH, D. (1991). Threshold-range scaling of excitable cellular automata. *Statistics and Computing* 1 23-39.

FISCH, R. and GRIFFEATH, D. (1991). *EXCITE!: A Periodic Wave Modeling Environment* (1 disk for IBM-compatibles with VGA graphics). Freeware. Write D. Griffeath, Mathematics Dept., Univ. Wisconsin, Madison WI 53706.

GRAF, H.-P. (1983). Long range correlations and estimation of self-similarity parameter. Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich.

GRANGER, C. W. J. (1983). Forecasting white noise. In *Applied Time Series Analysis of Economic Data* (A. Zellner ed.). U.S. Dept. Commerce, Washington D.C.

GRANGER C. W. J. and LIN, J.-L. (1991). Using the mutual information coefficient to identify lags in non-linear models. Working paper, Dept. Economics, Univ. California, San Diego.

GRANGER, C. W. J. and TERÄSVIRTA, T. (1992). *Modelling Nonlinear Economic Relationship.* Oxford Univ. Press.

GREGOBI, C. HAMMEL, S. and YORKE, J. (1987). Do numerical orbits of chaotic dynamical systems represent true orbits? *J. Complexity* 3 136-145.

HALL, P. (1982). On some simple estimates of an exponent of regular variation. *J. Roy. Statist. Soc. Ser. B* 44 37-42.

HALSEY, T. C., JENSEN, M. H., KADANOFF, L. P., PROCACCIA, I. A. and SHRAIMAN, B. I. (1986). Fractal measures and their singularities: The characterization of strange sets. *Phys. Rev. A* 25 3453-3456.

HAMMEL, S. M., YORKE, J. A. and GREBOGI, C. (1988). Numerical orbits of chaotic processes represent true orbits. *Bull. Amer. Math. Soc.* 19 465-469.

HAMPEL, F. (1987). Date analysis and self-similar processes. Invited paper, ISI session, Tokyo, Japan.

HILL, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* 3 1163-1174.

HOCKETT, K. (1990). Chaotic numerics from an integrable Hamiltonian system. *Proc. Amer. Math. Soc.* 108 271-281.

HUNT, F. (1990). Error analysis and convergence of capacity dimension algorithms. *SIAM J. Appl. Math.* 50 307-321.

HUNT, F. Y. and MILLER, W. M. (1990). On the approximation of invariant measures. Preprint, Dept. Mathematics, Howard Univ.

KING, R. B., WEISSMAN, L. J. and BASSINGTHWAIGHTE, J. B. (1990). Fractal descriptions for spatial statistics. *Annals of Biomedical Engineering* 18 111-121.

LAING, J. R. (1991). Efficient chaos. *Barrons* 12 12-13.

LALLEY, S. P. (1988). The packing and covering functions of self-similar fractals. *Indiana Univ. Math. J.* 37 699-709.

LALLEY, S. P. (1989). Renewal theorems in symbolic dynamics with applications to geodesic flows, noneuclidean tessallations and their fractal limits. *Acta Math.* 163 1-55.

LALLEY, S. P. (1990). Travelling salesman with a self-similar itinerary. *Probability in the Engineering and the Informational Sciences* 4 1-18.

LALLEY, S. P. (1991). Probabilistic methods in certain counting problems of ergodic theory. In *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces* (T. Bedford, M. Keane and C. Series, eds.) 223-258. Oxford Univ. Press.

LASOTA, A. and MACKEY, M. (1985). *Probabilistic Properties of Deterministic Systems.* Cambridge Univ. Press.

LEDRAPPIER, F. and MISIUREWICZ, M. (1985). Dimension of invariant measures for maps with exponent zero. *Ergodic Theory Dynamical Systems* 5 595-610.

LIGGETT T. (1985). *Interacting Particle Systems.* Springer, New York.

LIM, K. S. (1992). On the stability of a threshold AR(1) without intercepts. *J. Time Ser. Anal.* To appear.

LIU, T., GRANGER, C. W. J. and HELLER, W. P. (1991). Using the correlation exponent to decide if an economic series is chaotic. Discussion Paper 91-21, Dept. Economics, Univ. California, San Diego.

LORD, M. E. (1979). The method of non-linear variation of constants for difference equations. *J. Inst. Math. Appl.* 23 285-290.

MCCULLAGH, P. and NELDER, J. A. (1986). *Generalized Linear Models,* 2nd ed. Chapman and Hall, London.

NYCHKA, D., ELLNER, S., MCCAFFREY, D. and GALLANT, A. R. (1992). Finding chaos in noisy systems (with discussion). *J. Roy. Statist. Soc. Ser. B.* To appear.

ORNSTEIN, D. S. and WEISS, B. (1991). Statistical properties of chaotic systems. *Bull. Amer. Math. Soc.* **24** 11-117.

OSBORNE, A. R. and PROVENZALE, A. (1989). Finite correlation dimension for stochastic systems with power-law spectra. *Phys. D* **35** 357-381.

OTT, E. (1981). Strange attractors and chaotic motions of dynamical systems. *Rev. Modern Phys.* **53** 655-671.

PARK, S. and MILLER, K. (1988). Random number generators: Good ones are hard to find. *Comm. ACM* **31** 1192-1201.

PEÑA, D. (1990). Influential observations in time series. *Journal of Business and Economic Statistics* **8** 235-241.

PETERS, E. E. (1991). A chaotic attractor for the S&P 500. *Financial Analysts Journal* **March-April** 55-63.

PREISS, D. (1987). Geometry of measures in $\mathbf{R}^n$: Distribution, rectifiability, and densities. *Ann. Math.* **125** 537-643.

RUCKER, R. (1989). *CA Lab* (software for IBM-compatibles with VGA graphics and 264-page manual). Autodesk, Sausalito, Calif.

RUELLE, D. (1990). Deterministic chaos: The science and the fiction (The Claude Bernard Lecture 1989). *Proc. Roy. Soc. London Ser. A* **427** 241-248.

RUELLE, D. (1991). *Chance and Chaos.* Princeton Univ. Press.

SAUER, T. and YORKE, J. A. (1991). Rigorous verification of trajectories for computer simulation of dynamical systems. *Nonlinearity* **4** 961-979.

SCHROEDER, M. (1990). *Fractals, Chaos, Power Laws.* W. H. Freeman, New York.

SILVERMAN, B. W. and BROWN, T. C. (1978). Short distances, flat triangles and Poisson limits. *J. Appl. Probab.* **15** 815-825.

SMITH, L. A. (1988). Intrinsic limits on dimension calculations. *Phys. Lett. A* **133** 283-288.

SMITH, R. L. (1991). Optimal estimation of fractal dimension. In *Nonlinear Modeling and Forecasting, SFI Studies in the Sciences of Complexity, Proc., Vol. XII* (M. Casdagli and S. Eubank, eds.). Addison-Wesley, Reading, Mass. To appear.

SUGIHARA, G. and MAY, R. M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* **344** 734-741.

TAYLOR, S. J. (1986). The measure theorey of random fractals. *Math. Proc. Cambridge Philos. Soc.* **100** 383-406.

THEILER, J. (1986). Spurious dimension from correlation algorithms applied to limited time-series data. *Phys. Rev. A* **34** 2427-2432.

THEILER, J. (1988). Lacunarity in a best estimator of fractal dimension. *Phys. Lett. A* **133** 195-200.

THEILER, J. (1990). Statistical precision of dimension estimators. *Phys. Rev. A* **41** 3038-3051.

THEILER, J. (1991). Some comments on the correlation dimension of $1/f^\alpha$ noise. *Phys. Lett. A* **155** 480-493.

TILMAN, D. and WEDIN, D. (1991). Oscillations and chaos in the dynamics of a perennial grass. *Nature* **353** 653-655.

TOFFOLI, T. and MARGOLUS, N. (1987). *Cellular Automata Machines.* MIT Press.

USPENSKII, V. A., SEMENOV, A. L. and SHEN, A. K. (1990). Can an individual sequence of zeros and ones be random? *Russian Math. Surveys* **45** 121-1189.

VAN BEEK, J. G. H. M., ROGER, S. A. and BASSINGTHWAIGHTE, J. B. (1989). Regional myocardian flow heterogeneity explained with fractral networks. *American Journal of Physiology* H1670-H1680.

WOLFRAM, S. (1986). Random sequence generation by cellular automata. *Adv. in Appl. Math.* **7** 123-169.

WOLFF, R. C. L. (1992). Local Lyapunov exponents: Looking closely at choas (with discussion). *J. Roy. Statist. Soc. Ser. B.* To appear.

YOUNG, L.-S. (1983). Entropy, Lyapunov exponents, and Hausdorff dimension in differentiable dynamial systems. *IEEE Trans. Circuits and Systems* **30** 599-607.

ZUREK, W. H. (ed.) (1991). *Complexity, Entropy and Physics of Information.* Addison-Wesley, Reading, Mass.