

PROBABILITY DENSITY FUNCTIONS IN EDUCATION USING DYNAMICAL SYSTEMS AND JAVA APPLETS

ABSTRACT

Probability density functions (PDF's) have inherent connections to the field of Dynamical Systems. A methodology for teaching PDF's, using the Logistic equation as a prototype, is suggested. Numerical computation of the PDF that arises in the study of the Logistic equation is carried out using a Java applet. Educational implications of this approach are given.

BY PANKAJ KAMTHAN

Department of Computer Science, Concordia University, Montreal, Canada, H3G 1M8

Key Words: Probability Density Functions; Dynamical Systems; Logistic Equation; Invariant Density; Java Applets; World-Wide Web.

CONTENTS	
1. INTRODUCTION 2. DYNAMICAL SYSTEMS AND PROBABILITY DENSITY FUNCTIONS 2.1 THE LOGISTIC EQUATION AS A DYNAMICAL SYSTEM 3. AN ALGORITHM FOR COMPUTING THE PROBABILITY DENSITY FUNCTIONS 4. JAVA AS THE PROGRAMMING LANGUAGE OF CHOICE 4.1 ADVANTAGES OF JAVA APPLETS IN EDUCATION	5. THE JAVA APPLET FOR THE INVARIANT DENSITY OF THE LOGISTIC EQUATION 5.1 THE JAVA APPLET CONFIGURATION 5.2 THE JAVA APPLET EXPERIMENT 5.3 RELIABILITY OF NUMERICAL RESULTS 5.4 REMARKS 6. EDUCATIONAL IMPLICATIONS 7. CONCLUSION ACKNOWLEDGEMENTS REFERENCES

1. INTRODUCTION

Reform of undergraduate mathematics is the key to revitalizing mathematics education.

- NCTM

Dynamical systems involves the study of change in natural and artificial (human-made) systems with respect to time. It has been emphasized in the past several years [May 1976, Devaney 1990] that the students should be introduced to simple nonlinear dynamical systems early in their mathematical education. There are now various textbooks dedicated to chaotic dynamical systems for pre-university and university level education [Devaney 1990, Peitgen, Jurgens, Saupe 1992], which usually emphasize the geometrical (topological) and numerical approach to dynamical systems. Statistical approach to chaotic dynamical systems has received a lot of attention in the last decade [Lasota and

Mackey 1994, Boyarsky and Gora 1997]. We then ask the following question: *How can traditional topics in basic statistics courses be taught from a dynamical systems viewpoint?*

The last few years have also seen a rapid emergence and broad acceptance of the Web (<http://www.w3.org>) as a global medium for disseminating and processing information that is accessible in multiple formats (multimedia) and at extremely fast speeds (hypermedia). This has opened new vistas in education by taking full advantage of our basic "senses" of learning such as visualizing 3D objects and nonlinear nature of thought processes. Recent years have also seen the inception and development of Java (<http://java.sun.com/>), a powerful programming language from SUN Microsystems. From a pedagogical viewpoint, we ask: *How can the present computational environment of Java and the Web be integrated in education?*

In this paper, we take the topic of PDF's, and suggest how it could be taught from a dynamical systems viewpoint. We also discuss how the environment of Java and the Web can assist in teaching and learning the topic.

2. DYNAMICAL SYSTEMS AND PROBABILITY DENSITY FUNCTIONS

The concept of a PDF is crucial in mathematical statistics courses. There exists an inherent connection between chaotic dynamical systems and PDF's.

Let (I, T) be a dynamical system, where $I = [0, 1]$ and $T: I \rightarrow I$ is a piecewise monotonic transformation. The interest from a dynamical systems viewpoint lies in the study of asymptotic behaviour of the trajectories $\{T^n(x)\}$, $n=0, 1, 2, \dots$, of points x in the phase space I . Instead of studying the evolution $\{T^n(x)\}$ of a single point x in I by the transformation T , we study the evolution of a PDF f of a collection (or ensemble) of points in I :

$$\begin{array}{ccccccc}
 x_1 & x_2 & x_3 & \cdots & x_n & & f \\
 \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & & \downarrow \\
 \tau(x_1) & \tau(x_2) & \tau(x_3) & \cdots & \tau(x_n) & & ?
 \end{array}$$

The evolution of f represents the corresponding change in the distribution of points under the action of T . The question we ask is: *How does f evolve?* Since theoretical study of this question is nontrivial in general, we answer this question for a specific dynamical system using a computational approach.

2.1 THE LOGISTIC EQUATION AS A DYNAMICAL SYSTEM

The Logistic equation family is given by

$$T_p(x) = px(1-x),$$

where p is in $(1,4]$. This family has been widely studied for the last several years, particularly in population dynamics [May 1976]. For $p=4$, we obtain the Logistic equation:

$$T_4(x) = 4x(1-x)$$

which is a hallmark of one-dimensional chaotic dynamical systems [May 1976]. See the Figure 1 for a graphical illustration.

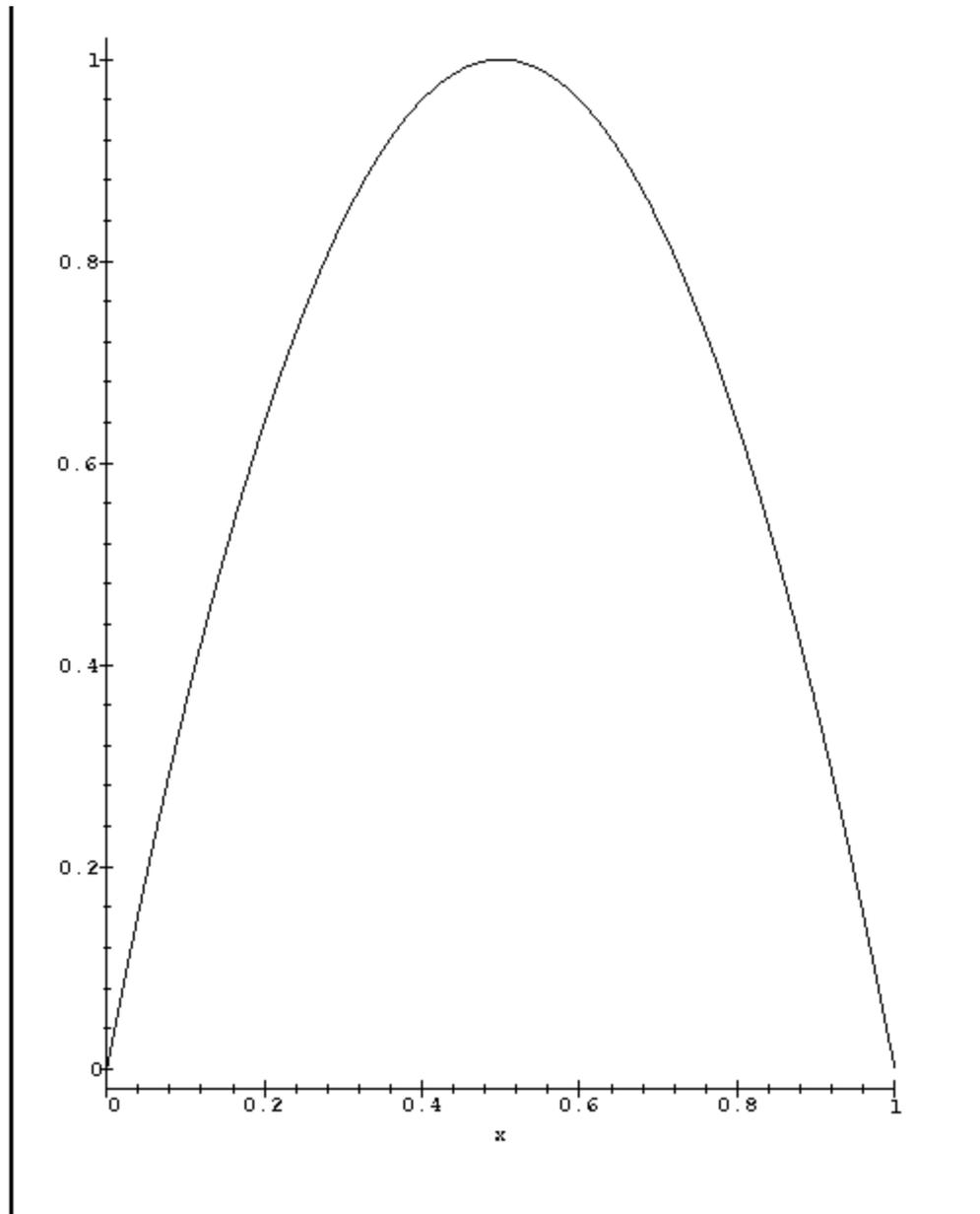


Figure 1: The Logistic equation $T_4(x) = 4x(1-x)$.

3. AN ALGORITHM FOR COMPUTING THE PROBABILITY DENSITY FUNCTIONS

We choose an initial point x_0 and iterate it under the map, say m times. We then observe the parts of the interval $[0,1]$ that are visited by the trajectory x_0, \dots, x_m and their frequency of visitation.

The algorithm consists of the following steps:

1. Start with a 'typical' real number x_0 in I , known as the initial condition. Such a number is usually chosen by a (pseudo) random-number generator.
2. Use this number to generate a sequence of N numbers x_1, x_2, \dots, x_N .
3. Choose a bin size h and divide the interval I into n bins such that $nh = 1$.
4. Count the x values in each bin.
5. Construct a histogram.

The histogram displays the *frequency* with which states along a trajectory $\{T^n(x)\}$ fall into given regions of I . As N tends to infinity, the histogram can be replaced (as a consequence of the Birkhoff Ergodic Theorem) by a probability density function $f(x)$. Then

$$\int_0^1 f(x)dx = 1$$

and $f(x)dx$ is the probability that an arbitrary number in the sequence lies between x and $x+dx$.

This algorithm can be put into practice by implementing it in a programming language, which in our case, is Java. Prior to that, we justify the use of Java as the language of choice in the next section.

4. JAVA AS THE PROGRAMMING LANGUAGE OF CHOICE

There are specific advantages of using Java applets over existing programming environments:

- **Platform Independence and Portability.** Multi-platform versions of educational programs are scarce, and often have a strong dependency on the underlying hardware. On the other hand, users can access a Java applet using different types of architectures and Java-compliant browsers, but still all see the same information in nearly the same format.
- **Development and Maintenance.** Usually, the process of writing an educational program is: **main program** (which does the computations) + **graphics** (which plots the results) + **user interface** (to the main program and its graphical component), each of them often in different languages. This can be difficult to extend and maintain. With the support of Java API, these three components can be integrated into one unified environment.

4.1 ADVANTAGES OF JAVA APPLETS IN EDUCATION

The integration of Java applets can help overcome some of obstacles inherent in the current practices of teaching and learning:

- **Open Classroom.** The use of Java applets can encourage asynchronous distance learning and thus help overcome the limitations (involving both time and space) inherent in traditional instructional techniques.
- **Nature of Information.** Java applets can complement a lecture session with information that is difficult to convey in a traditional manner. The use of applets, along with desired interaction and multimedia support, provides a representation that is often better in communicating a concept than a static figure(s) or a written description.
- **Cost.** Commercial application packages and compilers which are used in teaching courses are not readily affordable by all students for personal use. On the other hand, requisite software for Java applets (such as browsers and run-time environments for various architectures) are available free of cost for academic use.
- **Transferability.** Java applets (or even individual Java source/class files) developed for one course could be (re)used by teacher and students in other courses with intersecting syllabi. To realize that, for example, the applets can be made available on the institutional Web site or the intranet.

These advantages have instilled the use of Java applets in education [Kamthan 1999]. Java applets have found various useful applications in statistical education [West and Ogden 1998]. For applications of Internet and the Web in education in general, see the Web site *E-Education on the WWW* at <http://indy.cs.concordia.ca/e-edu/> engineered and maintained by the author. For details on how the Web technology be used on a local intranet, see [Kamthan 1997].

5. THE JAVA APPLLET FOR THE INVARIANT DENSITY OF THE LOGISTIC EQUATION

This section describes the use of a Java applet for numerically computing a PDF f based on the approach discussed in Section 3.

5.1 THE JAVA APPLLET CONFIGURATION

An implementation of the algorithm discussed in Section 3 is available in form of the Java applet at

<http://avalon.epm.ornl.gov/~re7/applets/Quad/QuadDistApp.html>.

The issue of performance of the applet depends (among others), on the following factors:

1. **Speed of access.** The access can be slow in case of low bandwidth or over a crowded network.

2. **Speed of the CPU.** When accessed, the applet runs on the local machine and the performance depends on the speed of the CPU.
3. **Nature of the task being carried out.** The computation is numerically intensive in our case.

Taking into consideration these factors, running the applet remotely can be time consuming. Therefore, it is preferable to run the applet locally. To make it simultaneously accessible to many users, it could, for example, be installed on an local network. This can be done by carrying out the following steps:

1. Download the source code and compile the file `QuadDistApp.java` locally using a Java compiler such as JDK (Java Development Kit) available at <http://java.sun.com/>. This will yield (among others) the file `QuadDistApp.class`.
2. Include the class file in an HTML document, say `QuadDistApp.html` using the applet tag: `<applet code="QuadDistApp.class" width=600 height=480></applet>`. This assumes that the HTML and Java class files are in the same directory.
3. Open the file `QuadDistApp.html` in a Java-compliant Web browser to carry out the desired computation. Note that the Logistic equation in the applet is scaled slightly differently, which in our case translates to $p=4(\lambda)$.

5.2 THE JAVA APPLET EXPERIMENT

The applet illustrates the real-time evolution of the PDF f for T_4 . The computation begins with a uniform density (the points are uniformly distributed). As time goes by (i.e., under the action of T_4), the histogram changes. After changing shape for some time, the histogram seems to settle down, and further computation does not yield any noticeable change. In other words, f evolves to a PDF that "stabilizes." This stable PDF f^* is known as the *invariant density* since it is a fixed point of the Frobenius-Perron operator [Boyarsky and Gora 1997] $P: L^1 \rightarrow L^1$, i.e., $Pf^* = f^*$.

A snapshot of the invariant density obtained by using the Java applet is shown in Figure 2. The data used was 10,000 points, 1000 bins and 500 iterations.

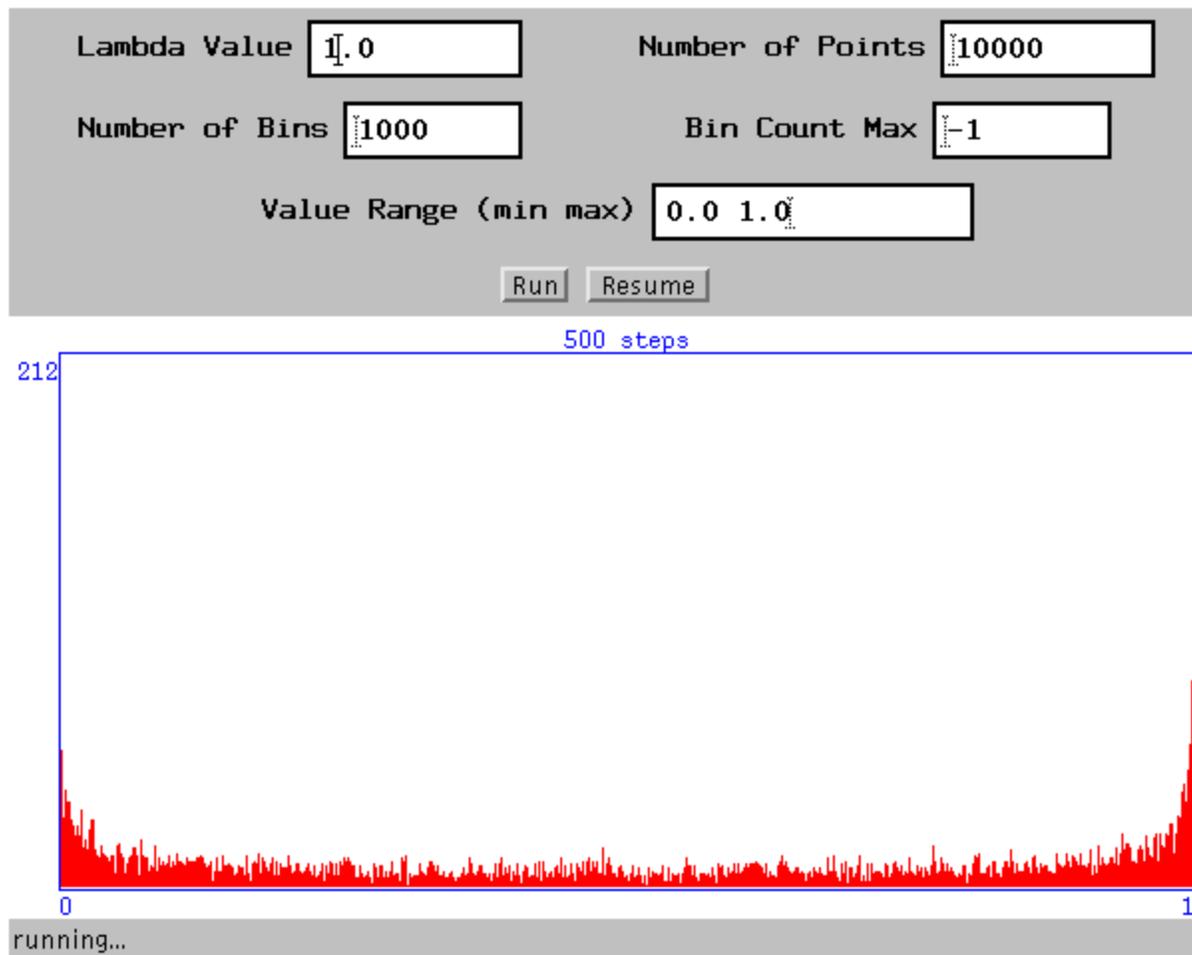


Figure 2: The Invariant Density f^* of the Logistic equation $T_4(x) = 4x(1-x)$ obtained by Computation of the Histogram of a Numerical Trajectory.

Numerically computed invariant density is in a good agreement with the theoretical invariant density given by [Lasota and Mackey 1994]

$$f^*(x) = \frac{1}{\pi\sqrt{x(1-x)}}$$

and shown in (the Maple generated) Figure 3.

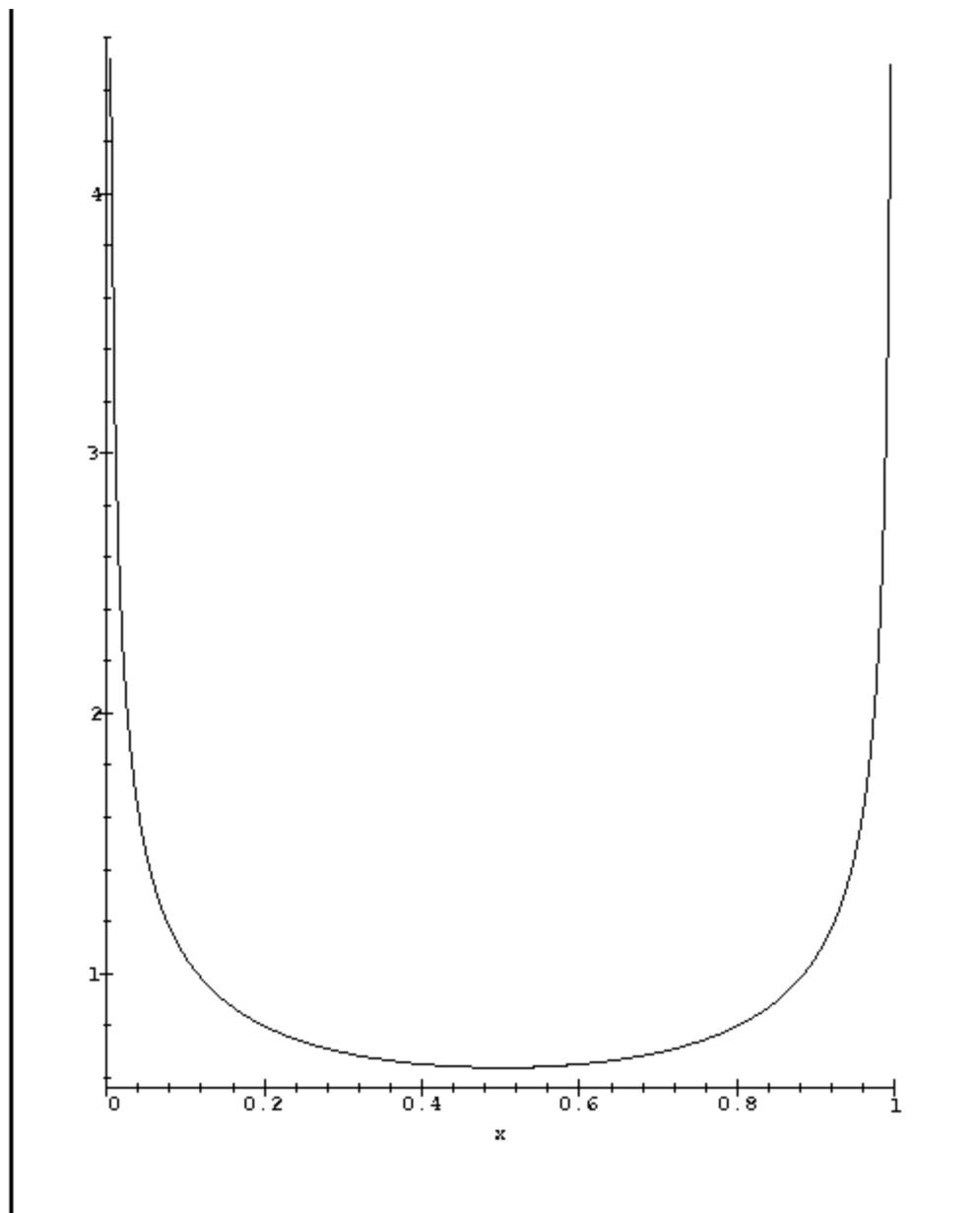


Figure 3: The invariant density f^* for the Logistic Map $T_4(x) = 4x(1-x)$.

5.3 RELIABILITY OF NUMERICAL RESULTS

A word of caution is in order and can be brought to students' attention. Care is to be taken in using the histogram to compute the invariant density. Calculation of an invariant density of a chaotic system on a computer by computation of the histogram of a sufficiently long numerical trajectory can depend strongly on various factors:

- **Arithmetic Precision.** This is the *precision* of the arithmetic due to the presence of roundoff/truncation errors (which may disappear as the precision is increased). It has been shown [Peitgen, Jurgens, Saupe 1992]

that the results obtained by using single precision arithmetic (which can be spurious) can be drastically different from that obtained from the use of double precision.

- **Number System.** The inherent nature of computers, that they do arithmetic in binary, can once again yield misleading results. As an example, consider the shift dynamical system $S_2(x) = 2x \pmod{1}$. It is well known [Boyarsky and Gora 1997] that S_2 has a unique uniform invariant density, i.e., $f^* = 1$. However, during computation, it is reflected that any subset of $[0,1]$ containing 0 has measure 1 and others have measure 0!

It can be shown [Gora and Boyarsky 1988], however, that the method is reliable in our case. It also gives a theoretical justification for the method of calculating the invariant density of a chaotic system on a computer by computation of the histogram of a sufficiently long numerical trajectory.

5.4 REMARKS

- The availability of an exact form of the invariant density f^* of T_4 presented in Section 5.2, is an exception rather than a rule. In general, obtaining a symbolic expression for invariant densities is very difficult, and one has to resort to numerical approximations.
- Using the histogram of a sufficiently long numerical trajectory is only one, and perhaps the simplest, among the methods available to compute an invariant density of a chaotic system. There are more general, more efficient and more accurate ways of computing invariant densities [Kamthan 1996]. However, they also require substantially more work to implement.

6. EDUCATIONAL IMPLICATIONS

There are various educational implications of such an endeavour:

- **Curriculum Saturation.** One problem that is faced during curriculum development to incorporate new topics in courses is *saturation*. Undergraduate curriculums at universities are often already crowded. Instead of replacing the traditional topics, dynamical systems suggests *new* approaches in their study, thus strengthening the traditional curriculum while enriching it.
- **Prerequisites.** Such an approach can be introduced at an elementary level as there are not many prerequisites. A introductory course in Calculus is all that is required.
- **Connections.** This approach makes *connections* between fields: Probability Density Functions, Dynamical Systems, Computation, Java applets. This can help the students appreciate the applicability of the subject under study and make them realize that it does not exist in isolation.
- **Further Questions.** This approach gives rise to further questions that can be explored. This gives a sense that the subject is alive and well, and evolving. For example:

1. *How does f^* change as the system parameters are changed?*

The answer to this question can turn out to be nontrivial in general. For example, for the Rényi equation,

$$S_r(x) = rx \pmod{1}$$

the invariant density f^* is uniform on $[0,1]$ for p any *integer* and f^* has a very complicated structure for r *not* an integer [Matthews and Strogatz 1991].

2. *What can be said about f^* for other systems?*

The answer to this question in context of general piecewise monotonic transformations on an interval has been an area of active study in the last decade [Boyarsky and Gora 1997].

- **Computer Experimentation.** To explore the question 1 above computationally for the Logistic equation T_p , students can use the Java applet of Section 5. They can manipulate the parameter p , and subsequently compute and observe the corresponding invariant density. This can encourage the spirit of computer experimentation among students, and induce "learning by doing." This can also make them appreciate that trial-and-error plays an important role in learning. Furthermore, since this trial-and-error process need not be time consuming, it can serve as a useful complement to the student's course work.

To explore question 2 computationally using a Java applet such as of Section 5, would require (a) writing a new Java procedure that programs the system under consideration and, (b) rewriting part of the Java code for the interface which deals with the user input.

- **Higher Education.** This approach sets a background for higher education in the areas of current active research such as Ergodic Theory of Dynamical Systems, in particular, Linear Operator Theory of the Frobenius-Perron operator.

7. CONCLUSION

Teaching PDF from a dynamical systems viewpoint provides a novel approach to a traditional topic in statistics. This approach, equipped with Java applets, has various useful pedagogical and epistemological implications.

ACKNOWLEDGEMENTS

The author is grateful to Dr. Roger Peck, California State University, Bakersfield, for his careful reading of the manuscript and various suggestions for improvement.

REFERENCES

R. L. Devaney (1990), *A First Course in Chaotic Dynamical Systems*, Addison-Wesley.

H.-O. Peitgen, H. Jürgens and D. Saupe (1992), *Fractals for the Classroom: Complex Systems and Mandelbrot Set*, Volume 2, Springer-Verlag.

A. Lasota and M. C. Mackey (1994), *Chaos and Fractals*, Springer-Verlag.

A. Boyarsky and P. Gora (1997), *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension*, Birkhauser.

P. K. Kamthan (1999), *Java Applets in Education*, irt.org.

R. M. May (1976), *Simple mathematical models with very complicated dynamics*, Nature, V.261, 459-467.

R. W. West and R. T. Ogden (1998), *Interactive Demonstrations for Statistics Education on the World Wide Web*, Journal of Statistics Education, V.6, N.3.

P. K. Kamthan (1997), *Intranets in Education*, irt.org.

P. Gora and A. Boyarsky (1988), *Why Computers like Lebesgue Measure?*, Comput. Math. Applic., V.16, N.4, 321-329.

P. K. Kamthan (1996), *Computation of Invariant Densities of Chaotic Dynamical Systems*, Preprint, 1996.

P. C. Matthews and S. H. Strogatz (1991), *Chaotic mappings and probability distributions*, College Mathematics Journal, V.22, N.1, 45-47.