

FRAMEWORK FOR PREDICTING
FREQUENCIES OF EVENTS

by

Donald Rosenfield
Visiting Associate Professor
M.I.T. Sloan School of Management

Senior Consultant
Arthur D. Little, Inc.

Dec. 1984

#1662-85

ABSTRACT

This paper presents a new technique for forecasting frequencies of events over time for individuals. The technique bases future occurrences on an estimation procedure based on historical data. The estimation procedure incorporates two types of uncertainty: population variation and individual propensity to change. An example is presented.

FRAMEWORK FOR PREDICTING FREQUENCIES OF EVENTS

1. Introduction

A new technique and approach for forecasting the frequencies of events over time for individuals is presented in this paper. These events include auto accidents or other loss events for which frequency forecasting is crucial in setting insurance premiums. A second application includes marketing systems (e.g. credit evaluation, catalog distribution) in which individuals are tracked and evaluated at regular intervals. The technique of this paper is designed for use as an alternative to exponential smoothing or present techniques based on Bayesian approaches.

The approach used in developing the technique is to incorporate both types of uncertainty that exist in any system of tracking and estimating individual event frequencies: a) the inability to determine where within the population range of frequencies (i.e. a set of frequencies) an individual is, and b) the incidence of changes over time in individual behavior. These two sources of uncertainty generally require different treatments in updating estimates of individual frequencies. If the only source of uncertainty were due to the population variation, a standard Bayesian estimation scheme [11] could be

applied. In this situation, it is assumed that people do not change, and consequently, prior estimates become heavily weighted with the collection of more observations. As a practical matter, individuals do change. Although changes in individual behavior can be incorporated in a Bayesian framework, the simultaneous treatment of population variation and individual change over a continuous range generally makes a Bayesian approach intractable. In accounting for changes in individual behavior, approaches such as exponential smoothing [2] usually have been implemented. An exponential smoothing system with a fixed smoothing constant uses historical data only to a limited degree and does not directly incorporate either the length of history or differences among groups.

The framework and technique in this paper is also independent of the specific shape of any prior distribution. Updating is based only on moments of the prior.

The development of the general methodology incorporating both aforementioned types of uncertainty is necessary in any system that predicts individual frequencies of events. A simple smoothing system cannot incorporate the additional uncertainty that exists at the beginning of the time horizon. On the other hand, a simple Bayesian update model cannot account for changes

in customer behavior. Classical Bayesian models with Poisson distributions of events for individuals have been proposed within the insurance industry [14]. In that context, it can be shown that estimates of variance between customers are inconsistent when based on time periods of different lengths. The reason for the inconsistency is that the variance within individuals' data (i.e., the variance of the count of events for an individual) is not strictly Poisson, possibly to changes of individual behavior. While there have been attempts to structure Bayesian models using a non-parametric system [11], [13], there is still an assumption that individuals do not change.

The approach presented in this paper is analogous to the Kalman filter approach [9] for estimating the parameters of an unknown normal variate, as well as other adaptive approaches [1], [3]. The Kalman approach assumes both types of uncertainty noted above. Indeed, the update equations are similar, but they are based on a normal distribution rather than a point process. In this paper, the individual data are assumed to be based on a Poisson distribution, and the changes are to the individual means of this process.

The remainder of the paper is divided into the following sections: Section 2 presents the theoretical development of the general smoothing system; and Section 3 presents an example and discussion.

2. Smoothing Framework

The major conceptual feature of the model is a shock process. At any point in time, each individual is characterized by a frequency or an expected number of events per period. This frequency may not be observable. For example if an individual has a frequency of one event per ten periods, it is impossible to measure this, but this "underlying" frequency still governs the individual's behavior. As an example an individual's events may be Poisson, where the Poisson parameter represents the underlying frequency. In each period, the frequency is subject to a random shock or change. The expected value of this change is zero, and the expectation of the frequency is equal to the frequency at the end of the previous period. For example, the frequency of one in ten is subject to a change, but the average of the change is zero. The effect of the shock process is to introduce additional uncertainty in every time period. As a result of this uncertainty, no estimating procedure can precisely determine the underlying frequency. Even after an extended period of time, there must be real smoothing in each time period (i.e. incorporation of the most recent data), since this data provides the only information about the most recent shock.

The smoothing (or prediction) problem is thus the determination of the expected value of the frequency in period $n + 1$, given the most recent observation in period n . In order to develop this expectation and initial estimates we use the following notation which can serve as a reference:

Let

- M_n = expected number of events period in period n for an individual
- $f_n()$ prior distribution of M_n
- x_n = random number of events in period n for an individual
- U_{in} = i th moment of $f_n(y)$
- W_{n+1} = $E(M_{n+1}|x_n)$ = current estimate of average
 $M_n = U_{1,n+1}$
- σ_n^2 = variance of $f_n(y) = \text{Var}(M_n|X_{n-1})$
- S_n = skewness of $f_n(y)$
- ϵ_n = change or actual shock at time period n
- V_s = $\text{Var}(\epsilon_n)$
- $\bar{\sigma}_2$ = variance of events among all individuals
- \bar{M} = observed average number of events for all individuals

The key to the model is the behavior of the random process M_n . M_n is the theoretical mean representing the expected number of events per period (i.e. mean frequency). The shock is the random change that occurs in M_n at the start of each period. Since the expected value of the shock is zero, we have the property

$$E(M_{n+1}|M_n) = M_n$$

The purpose of the model is to update the estimate, given observations x_n . At any point, there is prior distribution of M_n based on previous observations of numbers of events and the updating procedure is designed to estimate the mean of M_n .

An important feature of the model is the assumptions on the prior for M_n and the sampling distribution. It is assumed

that most of the mass of the prior and sampling distribution concentrated in small values. In particular, it is assumed that 1) the probability of more than one event is very small, 2) the moments of sampling distributions rapidly converge to zero, and 3) the likelihood of one event is an order of magnitude less than that of zero events.

For infrequent events, such as accidents or consumer purchasing episodes over a short time period such an assumption is valid. Indeed, it is always valid if the time period in question is short enough.

The actual sampling distribution could be, for example, either a Poisson or a Bernoulli distribution. The above assumptions could be satisfied by a Poisson distribution with an unknown mean, or a Bernoulli distribution with an unknown success probability, combined with a prior with small first moment and rapidly declining higher moments. The development in the paper uses a Poisson distribution.

The period-to-period behavior of the process is as follows

$$M_{n+1} = M_n + \epsilon_n$$

and

$$E(\epsilon_n) = 0$$

$$\text{Var } \epsilon_n = V_S$$

The variance of ϵ_n may depend on M_n . Hence V_S is interpreted to be $E(\text{Var } \epsilon_n | M_n)$. It is useful to compare the

M_n process and its estimation with a Kalman filter model. The M_n process is similar to the mean process in a Kalman filter. The mean of the filtered process is similar to the mean process in a Kalman filter. The mean of the filtered process is subject to added shocks superimposed on the original value. The observations of a Kalman filter are normally distributed about its mean. The difference in this process is the Poisson distribution of the observations about its mean.

Under the assumptions noted above, we obtain the following relationships for updating W_n . In any practical updating system, the variable W_n is the basis of evaluating individuals or customers. In an insurance system, for example, premiums would be based on average losses.

$$W_{n+1} = W_n \left(1 - \frac{\sigma_n^2}{W_n}\right) + x_n \frac{\sigma_n^2}{W_n} \quad (1a)$$

$$\sigma_{n+1}^2 = V_s + \sigma_n^2 - S_n + x_n \left(\frac{S_n}{W_n} - \frac{\sigma_n^4}{W_n^2}\right) \quad (1b)$$

Relationships (1) resemble an exponential smoothing update in which the smoothing constant is one minus the ratio of the mean to the variance of the uncertainty to the mean of the process. The widespread use of exponential smoothing is recognition that individuals change according to a shock-type process. The major difference between the smoothing constant in (1a) and an exponential smoothing constant is that the constant in (1a)

generally starts out at a much lower value as uncertainty is greater when an individual is first tracked. The derivation of relationships (1) is presented in the Appendix.

The problem with relationships (1) as they stand is that skewness may not easily be tracked in any actual data collection system. In fact, without a specific assumption on the shape of the distribution, it will follow that S_{n+1} depends on $U_{4,n}$, $U_{4,n}$ depends on $U_{5,n}$, etc. One approach is to update the variance at each step in expectation, that is, without any knowledge of the specific number of observations. While this approach clearly is not exact for any specific stage, it will still capture the appropriate level of smoothing over the course of many periods. That is, since σ_n^2 will decrease in expectation, W_{n+1} will be less sensitive to new observations over time, and smoothing will be greater early on in the process when uncertainty is greatest.

Using the expected value of x_n of W_n in the second relationship (1b), we obtain

$$\sigma_{n+1}^2 = V_S + \sigma_n^2 (1 - \sigma_n^2 / W_n) \quad (2)$$

which together with (1a) represents the final smoothing system.

3. Examples and Discussions

The usefulness of the framework can be illustrated through examples of practical problems. The author has worked with two such practical problems: auto accidents and market response models. Both of the problem areas are characterized by uncertainty between individuals and behavioral changes. An adaptive point process model is required for accurate evaluation of individuals. In this section, we present an example from the auto insurance industry. The classic approach is a Bayesian model with no individual change. Since behavioral changes often occur, the type of framework proposed in this report might be more appropriate.

When examining variations across a population, using a Poisson distribution

$$\begin{aligned}\overline{\sigma^2} &= E(\text{Var } x_n | M_n) + \text{Var}(E x_n | M_n) \\ &= \bar{M} + \text{Var } M_n\end{aligned}$$

Thus initial variances can be estimated from aggregate data. For example, the following statistics are based on aggregate three-year automobile accident data for the state of North Carolina:

$$\begin{aligned}M &= .243 \text{ for three years, } .081 \text{ for 1 year} \\ \sigma^2 &= .297 \text{ for three years, } .667 \text{ for six years.}\end{aligned}$$

With a Bayesian system, a gamma prior and Poisson sampling distribution, we obtain a prior variance $\text{Var } M_n$ of .051 for three years and $.051/9 = .006$ for 1 year. The initial parameters of a gamma distribution for one year averages are

$$k = 1.094$$

$$a = 13.5$$

where

$$f_0(y) = \frac{a^k}{\Gamma(k)} y^{k-1} e^{-ay}$$

and $\Gamma(\cdot)$ is the gamma function. After each (Yearly) observation, the posterior mean is updated as follows

$$W_n = \frac{a}{a+1} W_{n-1} + \frac{1}{a+1} x_n$$

and a is increased by one. $a/(a+1)$ is initially .931 but approaches one asymptotically, and there is little accounting of new information.

Under the alternative approach of this paper, the initial smoothing constant is

$$1 - \frac{\text{Var } M_n}{\bar{M}} = 1 - \frac{.006}{.081} = .925$$

Depending on the shock variance, however, there is always emphasis on experience. As the Bayesian approach ultimately discounts experience, it is possible that initial group statistics from insurance classifications are overemphasized.

The asymptotic value for σ_n^2 using relationship 1b and assuming that W_n reaches a limit and that V_s is constant is

$$\sqrt{V_s \lim_{n \rightarrow \infty} W_n}$$

With a shock variance of .0001 and a limit of W_n of .081, for example, the limiting value of σ_n^2 is .003 and the limiting smoothing constant is

$$1 - \frac{.003}{.081} = .963$$

and the system still weighs experience.

A second application area is the evaluation of customers in marketing operations. These include, for example, catalog or credit card operations where customers are evaluated for the purpose of preferential treatment. Such customer evaluations in terms of purchasing history are presently being used for one of the nation's largest retailing operations.

APPENDIX

DERIVATION OF RELATIONSHIPS (1)

We derive the first relationship using a Bayes expression for M_{n+1} . $E(M_{n+1}|x_n)$ = Posterior mean before the shock

$$= \frac{\int y f_n(y) P(x_n | y) dy}{\int f_n(y) P(x_n | y) dy}$$

The assumption is that the distribution of an individual's number of events in any given time period is Poisson. The change process simply stipulates that the mean of this Poisson distribution changes from period to period. In a particular time period a Poisson assumption can still hold. Thus, we have.

$$P(x_n | y) = \frac{y^{x_n} e^{-y}}{x_n!}$$

so

$$E(M_{n+1} | x_n) = \frac{\int e^{-y} y^{x_n+1} f_n(y) dy}{\int e^{-y} y^{x_n} f_n(y) dy} \quad (A1)$$

The issue is the evaluation of (A1). While Bayesian expressions can be developed for a particular form of distribution without shock processes (such as a gamma), such treatment becomes intractible after the introduction of a shock process. Consequently, while the process expression is Bayesian, our approach differs from the usual Bayesian approach that the shock factor is introduced. Furthermore, it is desirable to develop an expression that holds for any form of prior distributions

as long as most of the mass is located at small values, as previously discussed.

To calculate the integrals in (A1) as well as additional integrals in evaluating the variance of the uncertainty process, a Taylor series expansion is used for the exponential function:

$$\begin{aligned}
 & \int e^{-y} y^{x_n+k} f_n(y) dy \\
 &= \int \sum_{i=0}^{\infty} y^{x_n+k} \frac{(-y)^i}{i!} f_n(y) dy \\
 &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \int y^{i+x_n+k} f_n(y) dy \\
 &= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} U_{i+x_n+k} \tag{A2}
 \end{aligned}$$

Assuming that the moments of $f_n(y)$ rapidly converge to zero (which follows, for example, if the range of f is $(0, \epsilon)$ for small ϵ), we specifically can assume that the i^{th} moment can be bounded by c for some c less than one. Thus the absolute summation in (A2) is bounded by $c^{x+k}e^c$ and the interchange of summation and integral is valid.

By (A1),

$$\begin{aligned}
 E(M_{n+1} | x_n) &= \frac{U_{x_n+1,n} - U_{x_n+2,n} + \frac{U_{x_n+3,n}}{2} \cdot \dots}{U_{x_n,n} - U_{x_n+1,n} + \frac{U_{x_n+2,n}}{2} + \dots} \\
 &= \frac{U_{x_n+1,n}}{U_{x_n,n}} \frac{1 - \frac{U_{x_n+2,n}}{U_{x_n+1,n}} - \frac{U_{x_n+3,n}}{2U_{x_n+1,n}} - \dots}{1 - \frac{U_{x_n+1,n}}{U_{x_n,n}} + \frac{U_{x_n+2,n}}{2U_{x_n,n}} + \dots}
 \end{aligned}$$

Here, we again invoke the assumption that most of the mass of $f(y)$ is located near zero and that the moments of y drop off rapidly. Thus, insignificant terms can be eliminated, and we can approximate all expressions by dropping terms whose orders of magnitude are two or more degrees higher than the dominant terms. Thus, for example, if the largest term is $U_{x_n+1,n}$, we can drop all terms $U_{x_n+k,n}$ where $k \geq 3$, or ratios $U_{x_n+k,n}/U_{x_n+m,n}$ where $k-m \geq 3$. In general, we drop the ratio if $k-m \geq 2+j-x_n$ if the highest term is $U_{x_n+j,n}$. For example, suppose a prior is uniform between zero and .02, then $U_{1,n} = 0.1$, $U_{2,n} = .0002$, $U_{3,n} = 2.7 \times 10^{-6}$, and $U_{4,n} = 4 \times 10^{-8}$. $U_{1,n}U_{2,n}$, $U_{3,n}$, $U_{4,n}$, $U_{1,n}$ and $U_{4,n}/U_{1,n}$ are all very small compared to $U_{1,n}$. It also follows that central moments such as σ_n^2 will have the same order of magnitude as $U_{2,n}$ and that powers of moments such as $U_{1,n}^2$ will have an order of magnitude determined by the product of that power and moment.

We also can assume that the likelihood of $x_n = 1$ on an actual sampling distribution is an order of magnitude less than the likelihood of $x_n = 0$, and that values of 2 or larger have likelihoods that are at least an order of magnitude less than that of $x_n = 1$. In the example above, $P(x_n = 0) = .98$, $P(x_n = 1) = .02$ and $P(x_n \geq 2) < .0004$ for the largest possible prior parameter of .02. Here, we can drop all terms involving $x_n = 1$. On the expectation basis, this procedure preserves the two highest orders of magnitude.

Hence,

$$\begin{aligned}
 E(M_{n+1} | x_n) &\approx \frac{U_{x_n+1,n}}{U_{x_n,n}} \left[\frac{1 - \frac{U_{x_n+2,n}}{U_{x_n+1,n}}}{1 - \frac{U_{x_n+1,n}}{U_{x_n,n}}} \right] \\
 &\approx \frac{U_{x_n+1,n}}{U_{x_n,n}} \left[1 - \frac{U_{x_n+2,n}}{U_{x_n+1,n}} + \frac{U_{x_n+1,n}}{U_{x_n,n}} \right]
 \end{aligned}$$

by binominal expansion of the denominator and truncating higher order terms.

Thus for $x_n = 0$,

$$\begin{aligned}
 E(M_{n+1} | x_n=0) &\approx U_{1,n} \left[1 - \frac{U_{2,n}}{U_{1,n}} + U_{1,n} \right] \\
 &= U_{1,n} - U_{2,n} + U_{1,n}^2 \\
 &= U_{1,n} - (\sigma_n^2 + U_{1,n}^2) + U_{1,n}^2 \\
 &= U_{1,n} - \sigma_{1,n}^2
 \end{aligned}$$

(A3)

For $x_n=1$

$$\begin{aligned}
E(M_{n+1}|x_n=1) &\approx \frac{U_{2,n}}{U_{1,n}} - \frac{U_{3,n}}{U_{2,n}} + \frac{U_{2,n}}{U_{1,n}} \\
&= \frac{U_{2,n}}{U_{1,n}} - \frac{U_{3,n}}{U_{1,n}} + \frac{U_{2,n}^2}{U_{1,n}^2} \\
&= \frac{\sigma_n^2 + U_{1,n}^2}{U_{1,n}} - \frac{S_n + 3\sigma_n^2 U_{1,n} + U_{1,n}^3}{U_{1,n}^2} + \frac{(\sigma_n^2 + U_{1,n}^2)^2}{U_{1,n}^2} \\
&= \frac{\sigma_n^2}{U_{1,n}} + U_{1,n} - \frac{S_n}{U_{1,n}} - 3\sigma_n^2 - U_{1,n}^2 + \frac{\sigma_n^4}{U_{1,n}^2} + 2\sigma_n^2 + U_{1,n}^2 \\
&= \frac{\sigma_n^2}{U_{1,n}} + U_{1,n} - \frac{S_n}{U_{1,n}} - \sigma_n^2 + \frac{\sigma_n^4}{U_{1,n}^2} \\
&\approx U_{1,n} + \frac{\sigma_n^2}{U_{1,n}} \tag{A4}
\end{aligned}$$

Note that terms multiplied by x_n will lower their values in expectation by another order of magnitude and we need only preserve one order of magnitude above. Hence since $x_n \geq 2$ is two orders of magnitude less likely than $x_n = 0$,

$$\begin{aligned}
E(M_{n+1}|x_n) &\approx (1 - x_n)E(M_{n+1}|x_n=0) + x_n E(M_{n+1}|x_n=1) \\
&= (U_{1,n} - \sigma_n^2)(1 - x_n) + x_n \left(U_{1,n} + \frac{\sigma_n^2}{U_{1,n}} \right) \\
&= U_{1,n} - \sigma_n^2 + x_n \left(\sigma_n^2 + \frac{\sigma_n^2}{U_{1,n}} \right) \\
&\approx U_{1,n} - \sigma_n^2 + x_n \frac{\sigma_n^2}{U_{1,n}} \\
&= U_{1,n} \left(1 - \frac{\sigma_n^2}{U_{1,n}} \right) + \frac{\sigma_n^2}{U_{1,n}} x_n
\end{aligned}$$

which is the first of the two relationships (1).

In order to develop a usable system, it is necessary to develop the second relationship for the updated σ_n^2 , the variance for the prior distribution. The variance update for the period following the period just observed is:

$$\text{Var} (M_{n+1}|x_n) = E(\text{Var} (M_{n+1}|M_n)|x_n) + \text{Var} (E(M_{n+1}|M_n)|x_n)$$

The first term is the shock variance $\text{Var} (\epsilon_n)$. The second term, since $E (M_{n+1}|M_n) = M_n$, as the shock is a zero mean process, is simply $\text{Var} (M_n|x_n)$, which is the posterior variance of M_n . Thus,

$$\sigma_{n+1}^2 = V_s + E(M_n^2|x_n) - [E(M_n|x_n)]^2$$

Now, the posterior expectation for M_n^2 is

$$\begin{aligned} E(M_n^2|x_n) &= \frac{\int y^2 f_n(y) P(x_n|y) dy}{\int f_n(y) P(x_n|y) dy} \\ &= \frac{\int y^{x_n+2} e^{-y} f_n(y) dy}{\int y^{x_n} e^{-y} f_n(y) dy} \quad \text{[by (A2)]} \\ &= \frac{U_{x_n+2,n} - U_{x_n+3,n} + \frac{U_{x_n+4,n}}{2} + \dots}{U_{x_n,n} - U_{x_n+1,n} + \frac{U_{x_n+2,n}}{2} \dots} \\ &= \frac{U_{x_n+2,n}}{U_{x_n,n}} \left(1 - \frac{U_{x_n+3,n}}{U_{x_n+2,n}} + \frac{U_{x_n+1,n}}{U_{x_n,n}} \right) \end{aligned}$$

again by binomial expansion and assumption.

So for $x_n = 0$,

$$\begin{aligned}
E(M_n^2 | x_n) &= U_{2,n} \left[1 - \frac{U_{3,n}}{U_{2,n}} + U_{1,n} \right] \\
&= U_{2,n} - U_{3,n} + U_{1,n} U_{2,n} \\
&= \sigma_n^2 + U_{1,n}^2 - S_n - 3\sigma_n^2 U_{1,n} - U_{1,n}^3 + U_{1,n}^3 + U_{1,n} \sigma_n^2 \\
&= \sigma_n^2 + U_{1,n}^2 - S_n - 2\sigma_n^2 U_{1,n}
\end{aligned}$$

and thus

$$\begin{aligned}
\text{Var}(M_n | x) &\approx \sigma_n^2 + U_{1,n}^2 - S_n - 2\sigma_n^2 U_{1,n} - (U_{1,n} - \sigma_n^2)^2 && \text{[by (A3)]} \\
&\approx \sigma_n^2 - S_n
\end{aligned}$$

For $x_n = 1$

$$\begin{aligned}
E(M_n^2 | x_n) &= \frac{U_{3,n}}{U_{1,n}} \left[1 - \frac{U_{4,n}}{U_{3,n}} + \frac{U_{2,n}}{U_{1,n}} \right] \\
&\approx \frac{U_{3,n}}{U_{1,n}} \quad (\text{Where we again preserve one order of} \\
&\quad \text{magnitude for a term involving } x_n = 1) \\
&= \frac{S_n}{U_{1,n}} + 3\sigma_n^2 + U_{1,n}^2
\end{aligned}$$

and

$$\begin{aligned}
\text{Var}(M_n^2 | x_n) &\approx \frac{S_n}{U_{1,n}} + 3\sigma_n^2 + U_{1,n}^2 - \left(U_{1,n} + \frac{\sigma_n^2}{U_{1,n}} \right)^2 && \text{[by A4)]} \\
&= \sigma_n^2 + \frac{S_n}{U_{1,n}} - \frac{\sigma_n^4}{U_{1,n}^2}
\end{aligned}$$

So $\text{Var} (M_n^2 | x_n)$

$$= (\sigma_n^2 - S_n)(1 - x_n) + x_n \left(\frac{S_n}{U_{1,n}} + \sigma_n^2 - \frac{\sigma_{1,n}^4}{U_{1,n}^2} \right)$$

which directly yields (1b) after we drop the term $x_n S_n$, which is of lower order of magnitude in expectation than other terms in (2b).

We also note that the same relationships hold under the assumption of a Bernoulli distribution. In this case, Poisson probabilities for 2,3,... events are small and

$$P(x_n = 1) = U_{1,n}$$

$$P(x_n = 0) = 1 - U_{1,n}$$

It then follows that

$$U_{i,n+1} = \frac{U_{i,n} - U_{i+1,n}}{1 - U_{1,n}} (1 - x_n) + \frac{U_{i+1,n}}{U_{1,n}} x_n$$

and relationships (1) follow after 2nd order terms are dropped.

The only other issue is the establishment of the relationships is to show inductively that

$$U_{1,n} \ll U_{2,n} \dots$$

for successive values of n . This follows as long as the shock process has small enough moments. For example since $E(M_{n+1} | M_n) = 0$, $U_{1,n}$ remains small from stage to stage. For the second moment, we note from (1) that as long, as V_S is small, σ_n^2 does not grow substantially in expectation as n increases.

REFERENCES

1. Bretschneider, S., and Gorr, W., "On the Relationship of Adaptive Filtering Forecasting to Simple Brown Smoothing," Management Science, 27, No. 8, P. 965-969, August, 1981.
2. Brown, Robert G., Smoothing, Forecasting and Prediction of Discrete Time Series, Prentice-Hall, Englewood Cliffs, N.J., 1963.
3. Carbone, Robert and Longini, Richard L., "A Feedback Model for Automated Real Estate Assessment," Management Science, Vol., 24, No. 1, 1977.
4. Duncan, G.T., and Job, B.L., "Probability Forecasting in International Affairs," Expect-Generated Data: Applications in International Affairs, G.W. Hopple, and J.A. Kuhlman, ed. Boulder: Westview Press, 1981.
5. Gelb, Arthur, (ed.) Applied Optimal Estimation, M.I.T. Press, Cambridge, Mass., 1974.
6. Feller, W., An Introduction to Probability Theory and its Applications, Vol. II, 2nd edition, New York, Wiley, 1971.
7. Ferreira, Jr., J., "Identifying Equitable Insurance Premiums for Risk Classes: An Alternative to the Classical Approach," in Automobile Risk Classification Equity and Assuracy, 1978, Massachusetts Division of Insurance.
8. Herniter, J., "A Probabalistic Market Model of Purchase Timing and Brand Selection," Management Science, 18, No. 4, Part II, p. 102-113, December, 1971.
9. Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," J. Basic Engineering (ASME), Vol. 82D, 1960.
10. Morrison, D., and Schmittlein, D. "Predicting Future Random Events Based on Past Performance," Management Science, 27, No. 9, p. 1006-1023, September, 1981.
11. Raiffa, H. and Schalifer, R., Applied Statistical Decision Theory, Harvard University Press, Cambridge, Mass., 1961.
12. Robbins, H. and Murone, S., "A Stochastic Approximation, Method," Ann. Math. Statisti. Vol. 22, No. 1, 1951.
13. Robbins, H., "Prediction and Estimation for the compound Poisson Distribution," Proc. National Academy of Science, U.S.A., 74, p. 2670-2671, 1977.
14. SRI International, "The Role of Risk Classification in Property and Casualty Insurance: A Study of the Risk Assessment Process," 1976.