

ACTUARIAL STUDIES AND DEMOGRAPHY

Research Paper Series

Actuarial Applications of Autoregressive Models of Economic Variables

John Pollard

This paper has now appeared in the
Singapore International Insurance and Actuarial Journal,
and is reproduced with the kind permission of the editor.

Research Paper No. 009/97
ISBN No. 1 86408 400 6
November 1997

jpollard@efs.mq.edu.au
School of Economic and Financial Studies
Macquarie University
Sydney NSW 2109 Australia

The Macquarie University *Actuarial and Demographic Studies Research Papers* are written by members or affiliates of the Actuarial and Demographic Studies Department, Macquarie University. Although unrefereed, the papers are under the review and supervision of an editorial board.

Editorial Board:

Nick Parr

Michael Sherris

Additional copies of the papers are available from the World Wide Web site of the School of Economic and Financial Studies at Macquarie University at:

http://www.efs.mq.edu.au/acstdem/research_papers/index.html

Alternatively, requests for hard copies of the papers should be directed to:

The Editorial Board - Nick Parr and Michael Sherris

Actuarial and Demographic Studies Department

School of Economics and Financial Studies

Macquarie University NSW 2109

Fax No: 61 2 9850 9481

Email: lschalch@efs.mq.edu.au

Views expressed in this paper are those of the author and not necessarily those of the Actuarial Studies and Demography Department.

ACTUARIAL APPLICATIONS OF AUTOREGRESSIVE MODELS OF ECONOMIC VARIABLES

J H POLLARD

Macquarie University, Sydney, Australia

Summary

The first autoregressive model of fluctuating interest rates in the actuarial context appeared in 1971. Since then, several alternative but similar models have been proposed. In this paper, we summarize some of the earlier results and provide improved approximations for some of the moment formulae of annuities and assurances certain, and of life annuities and life assurances. It is shown that matrix analysis can be a very useful tool for deriving many of the asymptotic results, which are convenient for “rule-of-thumb” calculations.

The behaviour of the interest process in the steady state is discussed, and methods are developed for analysing accumulations, which are important for maturity guarantee products.

Although exact formulae and accurate approximations are available for the moments of the various interest rate functions, the actual distribution in many cases is unknown. It is suggested that the lognormal distribution can be used as an accurate approximation, and simulation studies confirm this.

The matrix methods developed for the fluctuating interest rate models are generalised to allow the study of simultaneous autoregressive models of inflation, market interest rates and fund investment returns.

1. Introduction

In 1971, a stochastic model of fluctuating interest rates was devised [9], based on the second-order autoregressive equation

$$(\delta_t - \delta) = 2k (\delta_{t-1} - \delta) - k (\delta_{t-2} - \delta) - e_t, \quad (1)$$

where $\delta_t = \ln(1+i_t)$ is the force of interest experienced in the year to time t , and $\delta = \ln(1+i)$ is the assumed long-term average force of interest. Ignoring the factor k and the random disturbance term e_t , equation (1) merely extrapolates the force of interest from year to year in a linear manner; if the interest rate has been rising, the model assumes that it has a tendency to follow that trend. The damping factor k ($0 \leq k < 1$) has the effect of dragging the linear extrapolated value back towards the assumed long-term average force of interest δ . The mutually independent normal random variables $\{e_t\}$ with zero expectations and variances σ^2 represent external influences affecting the damped system.

Models which do not prescribe a tendency to return to an average long term interest rate always allow the possibility of interest rates which become larger and larger, to unrealistic levels, or more and more negative, which is again unrealistic. Simulations using such models can produce extremely erratic results. The model represented by (1) is not of this type.

Equation (1) may also be written conveniently in the form

$$u_t = 2k u_{t-1} - k u_{t-2} + e_t, \quad (2)$$

$$\text{where } u_t = \delta - \delta_t, \quad (3)$$

so that with $v = e^{-\delta}$, the discount function v_t in year t is

$$v_t = \exp(-\delta_t) = v \exp(u_t). \quad (4)$$

The interest rate i corresponding to the long-term average force of interest is given by

$$1 + i = e^\delta. \quad (5)$$

The above autoregressive model allows the derivation of exact formulae for the moments of various compound interest functions including A_n and a_n , when the interest rate is allowed to fluctuate, and the moments of the life contingency functions A_x and a_x , when both age at death varies and the interest rate fluctuates. The exact formulae, however, are somewhat tedious to apply, and in section 2, we propose simple formulae for the moments, which yield accurate approximations to the exact second-order moments, except at very short durations.

A later paper [12] in 1976 showed how a life insurance company might load its non-participating premiums to take account of both variation in age at death and fluctuating investment returns. By far the greater proportion of the variance of each of the various life contingency functions is due to variation in the age at death (covered in the more modern life contingency texts). Paradoxically, however, the small additional variability due to fluctuations in the investment earnings rate (not usually discussed in the textbooks) may be of far greater concern to a life insurance company than the larger mortality variation, since the effects of the latter can be minimized by writing a large number of independent contracts.

Since the original paper of 1971, various alternative and sometimes very similar models have been proposed (e.g. [1], [2], [3], [4], [5], [6], [7], [13], [14], [15]) and the literature continues to grow.

In this paper, the mathematical analysis of the original 1971 model is reformulated in matrix form, which allows generalisation to simultaneous stochastic models of inflation, interest earnings and investment return. The matrix approach is also used to analyse some of the simpler models developed by other authors, and it is interesting to note that, all those which do not allow the force of interest to increase or decline indefinitely exhibit the same asymptotic behaviour.

2. Moments of annuities and assurances certain under the original autoregressive model

If S_t is defined as the cumulative sum of the $\{u_t\}$ as follows

$$S_t = \sum_{r=1}^t u_r, \quad (6)$$

then the discounted value of 1 due at the end of t years is

$$\mathbf{A}_t] = \prod_{r=1}^t v_r = v^t \exp(S_t) . \quad (7)$$

$\mathbf{A}_t]$ has been printed in bold to indicate that it is a random variable (depending on the fluctuating interest rate) to distinguish it from the usual deterministic compound interest function $A_t]$. It is immediately apparent from (2) and (6) that the $\{S_t\}$ have the multivariate normal distribution, and that the discounted value (7) will therefore have the lognormal distribution.

Exact formulae for the moments $\{S_t\}$ were derived in the original paper [7]. Equation(7) and knowledge of the multivariate normal moment generating function allows one to write down explicit formulae for the moments of $\mathbf{A}_t]$. Exact formulae for the moments of the stochastic annuity function $\mathbf{a}_t]$ can then be deduced by summing the relevant $\mathbf{A}_t]$ moment formulae. Most of these formulae are rather tedious for practical application, and we do not reproduce them here. Instead, we develop new approximate formulae, which are easy to apply.

The exact formulae derived in [7] reveal that for large t , the following asymptotic formulae apply (they also emerge directly from the matrix analysis later, in section 4):

$$E(S_t) \cong k(u_0 - u_1)/(1 - k) = \ln C ; \quad (8)$$

$$\text{Var}(S_t) \cong t \sigma^2/(1-k)^2 = t \Delta ; \quad (9)$$

$$\text{Cov}(S_t, S_r) \cong \min(t, r) \sigma^2/(1-k)^2 = \min(t, r) \Delta . \quad (10)$$

We noted earlier that $\mathbf{A}_t]$ has the lognormal distribution. Using (7) and our knowledge of the lognormal moment generating function, we deduce that

$$E(\mathbf{A}_t]) \cong C v^t e^{t\Delta/2} ; \quad (11)$$

$$\text{Var}(\mathbf{A}_t]) \cong [E(\mathbf{A}_t)]^2 (e^{t\Delta} - 1) \quad (12)$$

$$\text{Cov}(\mathbf{A}_t], \mathbf{A}_r]) \cong E(\mathbf{A}_t]) E(\mathbf{A}_r]) (e^{\min(t,r)\Delta} - 1). \quad (13)$$

Since $\mathbf{a}_n]$ is simply the sum of $\mathbf{A}_t]$ for values of t from 1 to n , approximations to the moments of $\mathbf{a}_n]$ can be obtained by summing the geometric progressions implied by inserting the asymptotic formulae (11), (12) and (13). The algebra is somewhat tedious, and the formulae which emerge involve compound interest functions evaluated at the forces of interest $\delta - \Delta/2$, $2\delta - \Delta$, and $2\delta - 2\Delta$. The positive quantity $\Delta = \sigma^2/(1-k)^2$ is usually small. Applying the Taylor series approximations to the standard compound interest functions evaluated at the forces of interest δ and 2δ , we obtain:

$$E(\mathbf{a}_t) \approx C [\delta_{\mathbf{a}_t} + \delta(\text{Ia})_t \Delta/2] ; \quad (14)$$

$$\text{Var}(\mathbf{a}_t) \approx C^2 \{2v(\delta_{\mathbf{a}_t} - 2\delta_{\mathbf{a}_t})/d^2 - (1+v)^{2\delta}(\text{Ia})_t/d\} \Delta ; \quad (15)$$

$$\text{Cov}(\mathbf{a}_t, \mathbf{A}_t) \approx C^2 v^t \delta(\text{Ia})_t \Delta . \quad (16)$$

(The prefixes on the compound interest functions denote the forces of interest to be used in the evaluation of the function.)

The above approximate asymptotic formulae provide reasonably good estimates of the second order moments of \mathbf{A}_t and \mathbf{a}_t at the longer durations, but less satisfactory results at shorter durations (Table 1). The formula for the expected value of \mathbf{A}_t also provides good approximations at all except the shorter durations (Table 2), but the expectation formula in the annuity case is disappointing, even at the longer durations. This is not surprising in view of the fact that the annuity involves the \mathbf{A}_t values at the earlier durations, where the approximate \mathbf{A}_t formula is less satisfactory and these earlier terms are dominant in \mathbf{a}_t . All the formulae give excellent results in the special case $k=0$.

For many purposes (e.g. life annuities and life assurances, dealt with in the following section) the performance of the above formulae at the shorter durations is relatively unimportant, and the formulae can be used to derive useful and reliable results. Improved approximations at the shorter durations will be discussed in a subsequent article.

3. Moments of life assurances and life annuities

The moments of life assurances and life annuities when variation in age at death is taken into account are discussed in all modern life contingency textbooks. These texts, however, usually shy away from the issue of variation in the interest rate at which the assurance or annuity function is evaluated, because of the complexities involved and the lack of a suitable, generally accepted model.

We shall demonstrate later that the approximate asymptotic formulae of the previous section are rather more general than the particular model in which they were developed, and life assurance and annuity functions based on them are therefore of some interest. Approximate formulae for the moments of the life assurance annuity functions follow readily from (11)-(16) using well-known conditional expectation, conditional variance and conditional covariance formulae [8], conditioning on the age at death, and assuming that age at death and the interest process are mutually independent.

The following second order moments follow fairly readily from (12), (15) and (16):

$$\text{Var}(\mathbf{A}_x) \approx C^2 \{ [{}^{2\delta}(\mathbf{IA})_x] \Delta + [{}^{2\delta}A_x - ({}^\delta A_x)^2] \}; \quad (17)$$

$$\text{Cov}(\mathbf{A}_x, \mathbf{a}_x) \approx C^2 \{ [v^t {}^\delta(\mathbf{IA})_x] \Delta + [({}^{2\delta}A_x - ({}^\delta A_x)^2)/d] \}; \quad (18)$$

$$\text{Var}(\mathbf{a}_x) \approx C^2 \{ [2v({}^\delta a_x - {}^{2\delta}a_x)/d^2 - (1+v) {}^{2\delta}(\mathbf{IA})_x/d] \Delta + [({}^{2\delta}A_x - ({}^\delta A_x)^2)/d^2] \}. \quad (19)$$

In each of these formulae, the first term provides the component of the second moment attributable to interest rate variation and the second term is the textbook moment in respect of mortality variation alone. Second moments for an annuity-due are readily deduced from the above.

When the conditional expectation formula is applied to (11), conditioning on age at death, the resulting expected value of the life assurance functions turns out to be $C A_x$ evaluated at force of interest $\delta - \Delta/2$, or δ for practical purposes. In the life annuity case, the expected value is simply (14) with the term certain t replaced by the life status x . For reasons given in the preceding section, both these formulae are likely to produce less than satisfactory results, and a simple practical approach would be, at least for the very early durations, to calculate the expectations using the recurrence relation (1) omitting the random error term.

The reliability of the annuity second moment formula is demonstrated in table 3.

4. Matrix analysis

It is clear from equation (7) that, for discounting purposes, S_t is central to the actuarial application of the autoregressive model. The basic model, however, was expressed in terms of the $\{u_t\}$. Replacing u_t by $S_t - S_{t-1}$ in equation (2), we obtain

$$S_t = (1 + 2k) S_{t-1} - 3k S_{t-2} + k S_{t-3} + e_t, \quad (20)$$

so that in matrix notation, we may write

$$\begin{bmatrix} S_t \\ S_{t-1} \\ S_{t-2} \end{bmatrix} = \begin{bmatrix} 1+2k & -3k & k \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} S_{t-1} \\ S_{t-2} \\ S_{t-3} \end{bmatrix} + \begin{bmatrix} e_t \\ 0 \\ 0 \end{bmatrix}. \quad (21)$$

This equation will be valid for all $t > 0$, provided we define

$$\left. \begin{aligned} S_0 &= 0; \\ S_{-1} &= -u_0; \\ S_{-2} &= -(u_0 + u_{-1}). \end{aligned} \right\} \quad (22)$$

Matrix equation (20) may be written more compactly as

$$\mathbf{w}_t = \mathbf{A} \mathbf{w}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (23)$$

which applied recursively yields

$$\mathbf{w}_t = \mathbf{A}^t \mathbf{w}_0 + \mathbf{A}^{t-1} \boldsymbol{\varepsilon}_1 + \mathbf{A}^{t-2} \boldsymbol{\varepsilon}_2 + \dots + \mathbf{A}^0 \boldsymbol{\varepsilon}_t. \quad (24)$$

The dominant eigenvalue of matrix \mathbf{A} is easily shown to be 1, and the corresponding right and left eigenvectors are

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \quad \mathbf{y}^T = [1/(1-k) \quad -2k/(1-k) \quad k/(1-k)]. \quad (25)$$

These eigenvectors have been normalised so that $\mathbf{y}^T \mathbf{x} = 1$, and for large t therefore [11],

$$\mathbf{A}^t \cong \mathbf{x} \mathbf{y}^T. \quad (26)$$

Although this result will not apply to the matrix \mathbf{A} powers attached to the earlier $\boldsymbol{\varepsilon}$ terms in equation (24), as t increases, the later terms will dominate, and we can deduce that

$$\mathbf{w}_t \cong \mathbf{x} \mathbf{y}^T \mathbf{w}_0 + \mathbf{x} \mathbf{y}^T (\boldsymbol{\varepsilon}_1 + \boldsymbol{\varepsilon}_2 + \dots + \boldsymbol{\varepsilon}_t). \quad (27)$$

The random vectors $\{\boldsymbol{\varepsilon}_t\}$ are independent. Each has expectation $\mathbf{0}$, and a covariance matrix $\boldsymbol{\Gamma}$ which has as its top left-hand element σ^2 and zeros elsewhere. It follows from (27) that

$$E(\mathbf{w}_t) \cong (\mathbf{y}^T \mathbf{w}_0) \mathbf{x}; \quad (28)$$

$$\text{Var}(\mathbf{w}_t) \cong t \mathbf{x} \mathbf{y}^T \boldsymbol{\Gamma} \mathbf{y} \mathbf{x}^T; \quad (29)$$

$$\text{Cov}(\mathbf{w}_t, \mathbf{w}_r) \cong \min(t, r) \mathbf{x} \mathbf{y}^T \boldsymbol{\Gamma} \mathbf{y} \mathbf{x}^T. \quad (30)$$

The leading elements of these vectors and matrices provide immediate confirmation of the asymptotic results quoted earlier as (8), (9) and (10).

The advantages of this matrix approach are that it produces the asymptotic results more simply, and it readily allows the analysis of alternative and more complex models. The quoted asymptotic formulae are in fact simply examples of more general results.

5. A simpler model

If one hypothesizes that the force of interest will always tend to move from its current level back towards a given long term value δ , then the following first-order autoregressive model emerges:

$$\delta_t - \delta = k (\delta_{t-1} - \delta) - e_t \quad (0 \leq k < 1). \quad (31)$$

If we then define u_t and S_t in exactly the same way as in the original second-order process (equations (3) and (6) respectively), the process can be written in matrix form as (23). The recurrence vector \mathbf{w}_t is of dimension 2 in this case (listing S_t and S_{t-1}), and the recurrence matrix \mathbf{A} is 2×2 , taking the form:

$$\mathbf{A} = \begin{bmatrix} 1+k & -k \\ 1 & 0 \end{bmatrix}. \quad (32)$$

The dominant eigenvalue of \mathbf{A} is again 1, and the corresponding normalised right and left eigenvectors are easily determined (\mathbf{x} is a 2-dimensional column of ones; the two elements of \mathbf{y} are $1/(1-k)$ and $-k/(1-k)$). With these definitions, all the formulae (23), (24) and (26)-(30) apply to the simplified model (31), and we can deduce that for large t

$$E(S_t) \cong k u_0 / (1-k); \quad (33)$$

$$\text{Var}(S_t) \cong t \sigma^2 / (1-k)^2 = t \Delta; \quad (34)$$

$$\text{Cov}(S_t, S_r) \cong \min(t,r) \sigma^2 / (1-k)^2 = \min(t,r) \Delta. \quad (35)$$

The asymptotic variance and covariance are the same as those in the more complex second order case (equations (9) and (10)). Only the asymptotic formula for $E(S_t)$ is changed. We conclude therefore that all the asymptotic formulae we derived for the original second order model (equations (9)- (19) are applicable to this simplified model provided we define

$$C = \exp[k u_0 / (1-k)]. \quad (36)$$

6. A more general autoregressive model

The results of the preceding sections suggest that the asymptotic results we have derived are but examples of rather more general results. In this section, we examine therefore the asymptotic behaviour of the r th order model

$$\delta_t - \delta = \sum_{j=1}^r a_j (\delta_{t-j} - \delta) - e_t \quad . \quad (37)$$

To ensure that the force of interest δ_t does not wander off to $+\infty$ or $-\infty$, we shall insist that

$$0 \leq \kappa = \sum_{j=1}^r a_j < 1 \quad . \quad (38)$$

This condition means that, in the absence of the random disturbance term e_t , the process would approach δ as t increased.

Using (3) and (6) to define u_t and S_t , we find that

$$S_t = (1 + a_1) S_{t-1} + \sum_{j=1}^{r-1} (a_{j+1} - a_j) S_{t-j-1} - a_r S_{t-r-1} + e_t \quad , \quad (39)$$

so that in matrix notation, we have

$$\begin{bmatrix} S_t \\ S_{t-1} \\ \vdots \\ S_{t-r} \end{bmatrix} = \begin{bmatrix} 1+a_1 & a_2-a_1 & a_3-a_2 & \dots & -a_r \\ & 1 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & 1 & 0 \end{bmatrix} \begin{bmatrix} S_{t-1} \\ S_{t-2} \\ \vdots \\ S_{t-r-1} \end{bmatrix} + \begin{bmatrix} e_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad , \quad (40)$$

which is valid for all $t > 0$ provided $S_0 = 0$ and

$$S_{-q} = - \sum_{j=0}^{q-1} u_j \quad , \quad (0 < q \leq r) \quad (41)$$

The $(r+1) \times (r+1)$ matrix \mathbf{A} defined by equation (40) has characteristic equation

$$(\lambda - 1) (\lambda^r - a_1 \lambda^{r-1} - a_2 \lambda^{r-2} - \dots - a_r) = 0 \quad . \quad (42)$$

It is immediately apparent that one of the eigenvalues is equal to 1. This is also the dominant eigenvalue, because of the constraint (38). The right eigenvector \mathbf{x} corresponding to the dominant eigenvalue 1 is easily seen to be a column of ones, whilst the corresponding left eigenvalue is

$$\mathbf{y}^T = [1/(1-\kappa) \quad -a_1/(1-\kappa) \quad -a_2/(1-\kappa) \quad \dots \quad -a_r/(1-\kappa)] \quad . \quad (43)$$

We deduce, therefore, using the same argument as in sections 4 and 5 above that

$$E(S_t) \cong \left(\sum_{j=1}^r a_j \sum_{i=0}^{j-1} u_{t-i} \right) / (1-\kappa) = \ln C; \quad (44)$$

$$\text{Var}(S_t) \cong t \Delta ; \quad (45)$$

$$\text{Cov}(S_t, S_u) \cong \min(t, u) \Delta ; \quad (46)$$

with

$$\Delta = \sigma^2 / (1-\kappa) \quad (47)$$

All the previous approximate annuity and assurance results (11)-(19) follow as previously, with κ replacing k and C defined by (44).

7. The interest process in the steady state

It is unlikely that an autoregressive model higher than third order will be used to analyse the interest process. Let us examine therefore the model represented by equation (37) with $r=3$.

In the steady state, the variance of $u_t = \delta - \delta_t$ will be independent of t , and the correlations ρ_1 between u_t and u_{t-1} , and ρ_2 between u_t and u_{t-2} will also be independent of t . One can deduce fairly readily by matrix analysis or otherwise that

$$\rho_1 = (a_1 + a_2 a_3) / [1 - a_2 - a_3(a_1 + a_3)] ; \quad (48)$$

$$\rho_2 = a_2 + (a_1 + a_3) \rho_1 ; \quad (49)$$

$$\text{Var}(\delta_t) \cong \sigma^2 / [(1 - a_1^2 - a_2^2 - a_3^2) - 2(a_1 a_2 + a_2 a_3) \rho_1 - 2a_1 a_3 \rho_2] . \quad (50)$$

For the special case of the autoregressive model (1), we observe that

$$\rho_1 = 2k / (1+k) ; \quad (51)$$

$$\rho_2 = k(3k-1) / (1+k) ; \quad (52)$$

$$\text{Var}(\delta_t) \cong \sigma^2 (1+k) / [(1+3k)(1-k)^2] . \quad (53)$$

The latter result is given in [10]. Using (51)-(53) we can deduce that

$$\text{Var}(\delta_t - \delta_{t-1}) \cong 2\sigma^2 / [(1+3k)(1-k)] . \quad (54)$$

The corresponding formulae for the first order model (31) are

$$\rho_1 = k ; \quad (55)$$

$$\rho_2 = k^2 ; \quad (56)$$

$$\text{Var}(\delta_t) \cong \sigma^2/(1-k^2) ; \quad (57)$$

$$\text{Var}(\delta_t - \delta_{t-1}) \cong 2\sigma^2/(1+k) . \quad (58)$$

Use will be made of some of these results in a later numerical example.

Explicit formulae for processes of higher-order than three are cumbersome. The asymptotic variance of δ_t however can be written fairly simply as

$$\text{Var}(\delta_t) \cong \sigma^2 \sum_{i=1}^{\infty} b_i^2 , \quad (59)$$

where, with $b_0 = 1$ and $b_u=0$ for $u < 0$,

$$b_i = \sum_{j=1}^r a_j b_{i-j} . \quad (60)$$

8. Accumulations

So far, we have considered only discounted values. For some purposes (e.g. maturity guarantees), accumulations are more relevant. We start with the relationship

$$(1+i_{t+1})(1+i_{t+2})\dots(1+i_n) = (1+i)^{n-t} \exp(S_t - S_n) , \quad (61)$$

in which i_t is the interest rate experienced in year t and i is the interest rate corresponding to the assumed long-term central force of interest δ . In the general case of section 6, we can deduce for large t and n that

$$E[(1+i_{t+1})\dots(1+i_n)] \cong (1+i)^{n-t} \exp[(n-t)\Delta/2] ; \quad (62)$$

$$E[(1+i_{t+1})\dots(1+i_n)]^2 \cong (1+i)^{2(n-t)} \exp[(n-t)\Delta] ; \quad (63)$$

$$\begin{aligned} E[(1+i_{t+1})\dots(1+i_n)][(1+i_{u+1})\dots(1+i_n)] \\ \cong (1+i)^{n-t}(1+i)^{n-u} \exp[(n-t)\Delta/2 + (n-u)\Delta/2 + \min(n-t, n-u)\Delta] . \end{aligned} \quad (64)$$

Summing these results in the same manner as we did for the annuities in section 2, we deduce the following approximate formulae:

$$E(\mathbf{s}_n) \cong \delta \mathbf{s}_n + \delta (\mathbf{Ds})_n \Delta / 2 ; \quad (65)$$

$$\text{Var}(\mathbf{s}_n) \cong \{2(1+i)(\delta \mathbf{s}_n - 2\delta \mathbf{s}_n) / i^2 + (2+i) 2\delta (\mathbf{Ds})_n / i\} \Delta . \quad (66)$$

The similarities between these formulae and (14) and (15) should be noted. The approximate \mathbf{s}_n moment formulae, however, do not involve the initial conditions factor C . The reason for this is that only the earlier payments in the accumulation are affected to any extent by the initial conditions; all the terms of an annuity, on the other hand, are affected by the interest rates in the early years. Formulae (65) and (65) are in fact identical to (14) and (15) respectively if the sign of the force of interest δ is reversed and the initial condition constant C is omitted.

Formula (66) provides an estimate of $\text{Var}(\mathbf{s}_n)$ which tends to be somewhat on the low side. A far better approximation, which is close to the exact variance can be obtained using the following expectation and second moment formulae (from which (66) was actually derived):

$$E(\mathbf{s}_n) \cong e^{\delta+0.5\Delta} \mathbf{s}_n ; \quad (67)$$

$$E(\mathbf{s}_n)^2 \cong 2 [e^{\delta+1.5\Delta} / (e^{\delta+1.5\Delta} - 1)] (e^{2\delta+2\Delta} \mathbf{s}_n - e^{\delta+0.5\Delta} \mathbf{s}_n) - e^{2\delta+2\Delta} \mathbf{s}_n . \quad (68)$$

We shall use these latter formulae in our subsequent calculations.

9. The shape of the \mathbf{s}_n distribution

As Wilkie (1976) [15] has noted, in the case of certain maturity guarantees, we may require the the accumulation random variable \mathbf{s}_n to obey the relationship

$$P(\mathbf{s}_n > n) = 1 - \varepsilon , \quad (69)$$

where ε is small and positive. We are able to compute exact values for the expectation and variance of \mathbf{s}_n (via the exact moments of the $\{S_t\}$) and may also calculate reasonable approximations using (67) and (68). To compute the probability (69), we also need to know the distribution of \mathbf{s}_n , or at least have an accurate approximation to it.

In all the autoregressive model we have discussed, the random error terms $\{e_t\}$ are independent normal random variables. The $\{S_t\}$ which we calculate are therefore correlated normal random variables. It follows from (7) and (61) that the discount functions $\{\mathbf{A}_t\}$ and the accumulation factors $\{(1+i_{t+1})(1+i_{t+2})\dots(1+i_n)\}$ have multivariate

lognormal distributions. Knowing the expectations and second-order moments of these variables, we can therefore make exact probabilistic statements about them.

The random variables a_t and $s_{n,t}$ are more problematic, as each comprises the sum of a series of random variables which have the multivariate lognormal distribution. Given their structure, both these random variables can be expected to retain some of the skewness inherent in their individual correlated lognormal terms. The use of the lognormal distribution would seem a plausible approximation for the distribution of each.

As a check on this proposed approximation, 1,000 simulations of the second-order autoregressive process (1) were performed, with parameters $k=0.5$, $\sigma=0.05$, $\delta=\ln(1.15)$, and $\delta_{-1}=\delta_0=\delta$. The resulting cumulative distribution of the s_{10} distribution is shown as column (2) in table 4. The simulated distribution has mean 24.412 and variance 36.775.

If exact theoretical calculations of the s_{10} moments are performed, a mean of 24.051 and variance of 33.686 emerge. The lognormal distribution having these moments is shown in column (3) of table 4. A Kolmogorov-Smirnov test indicates no significant difference between the simulated distribution and the theoretical lognormal distribution.

The interest process used in this simulation study has parameters which allow quite large and rapid changes in the interest rate (figure 1). Yet the lognormal distribution provided an accurate approximation to the unknown distribution of s_{10} . It would seem appropriate therefore to use either the lognormal distribution with exactly calculated moments, or the lognormal distribution with the approximate moments (67) and (68) to approximate the unknown theoretical distribution of s_n in the more general situation.

We would also expect the lognormal to provide an accurate approximation to the distribution of a_t in the same way.

10. A numerical example

An insurance company markets an investment product under which annual premiums of \$1 are payable for 10 years. At the end of the 10-year period, the insured is entitled to 97.5% of the accumulated premiums. There is a guarantee that the amount paid to the policyholder will be no less than the full premiums accumulated at 7.5% per annum.

In recent years, the earnings rate of the fund has been in the neighbourhood of 15% per annum, and the company believes that future earnings should vary around this level. The earnings of the fund in any given future year are expected to lie in the range 0% to 30% with a high probability, and a change in investment return from one year to the next of more than 12 percentage points is considered most unlikely.

What is the probability that the insurer will receive less than 2.5% of the accumulated fund? Find also the insurer's expected income from one such policy.

The accumulation of ten \$1 annual premiums at 7.5% is ${}^{.075}s_{10}|=15.2081$. If 97.5% of the accumulated premiums is less than this amount, the insurer will have to meet the maturity guarantee that it has given, and its income from the policy will be less than 2.5% of the accumulated value. We require therefore

$$P(0.975 s_{10}| < 15.2081)$$

i.e. the probability that the random accumulation function $s_{10}|$ is less than $15.2081/0.975=15.598$.

Let us adopt the second-order autoregressive model (1) for demonstration purposes, and assume a long-term central investment force of return $\delta=\ln(1.15)=0.13976$. If the return in any given future year is to lie in the range 0% to 30% with a high probability, then (with a 95% confidence level in mind) the standard error of any δ_t in the future should be about 0.08. Using similar reasoning, if a change in investment return from one year to the next of more than 12 percentage points is unlikely, then the standard deviation of $\delta_t-\delta_{t-1}$ must be about 0.065.

To obtain the parameters σ^2 and k of the autoregressive process, we equate (53) to the square of 0.08 and (54) to the square of 0.065, yielding $\sigma=0.0513$ and $k=0.5037$. Our calculations below will be based on $\sigma=0.05$ and $k=0.5$, which correspond to standard deviations of 0.0775 and 0.0632 for δ_t and $\delta_t-\delta_{t-1}$ respectively for a given year well in the future. With these parameters, (67) and (68) indicate an expected value of 24.122 for $s_{10}|$ and a variance of 30.775. Using a lognormal approximation for the distribution of $s_{10}|$, we deduce that the probability that the insurer will receive less than 2.5% of the accumulation is about 0.035 and the expected income to the insurer from such a policy is \$0.57.

11. Inflation and investment earnings

The actuary advising a defined benefit pension scheme will be very interested in future inflation rates, market rates of investment return and the return that will be achieved on the investments of the fund. These three rates are not unrelated. Higher rates of inflation, for example will tend to induce higher market rates of return, which in turn will impinge on the return obtained by the fund.

The matrix methods of sections 4, 5 and 6 are readily adapted to the analysis of simultaneous autoregressive models in which the fluctuating forces of inflation, market investment return and fund return are related, providing asymptotic formulae for the moments and useful approximations. The model we describe below is solely for demonstration purposes.

Let us define x_t to be the force of inflation during the year to time t , and postulate a central force of inflation ξ about which the inflation rate will vary, and assume for demonstration purposes that

$$x_t - \xi = 2k(x_{t-1} - \xi) - k(x_{t-2} - \xi) + e_t. \quad (70)$$

As before, we assume that $0 \leq k < 1$. The random error term e_t is normally distributed with zero mean and variance σ_x^2 , and the $\{e_t\}$ are mutually independent.

If we argue that the force of market return y_t in year t is influenced partly by its own recent history and by inflationary trends, a possible model for the deviation of the market rate of return from its assumed long term central value η is a combination of its previous level and the expected deviation of inflation from its long term central value:

$$y_t - \eta = b(y_{t-1} - \eta) + c[2k(x_{t-1} - \xi) - k(x_{t-2} - \xi)] + f_t, \quad (71)$$

where b and c are constants ($0 \leq |c| < 1$), and the $\{f_t\}$ are mutually independent random error terms with zero means and constant variances σ_y^2 , which are also independent of the $\{e_t\}$.

The force of return on the fund will be influenced by its existing investments as well as the market, and for demonstration purposes we shall assume that the deviation of the fund return z_t in year t from its postulated long term central level ζ is influenced by both its previous level and the expected level of the market:

$$z_t - \zeta = d(z_{t-1} - \zeta) + (1-d)[c(y_{t-1} - \eta) + 2kb(x_{t-1} - \xi) - kb(x_{t-2} - \xi)] + g_t. \quad (72)$$

The constant d is constrained so that $0 \leq d < 1$, and the random error terms $\{g_t\}$ are independent normal random variables with zero means and variances σ_z^2 , independent of the $\{e_t\}$ and $\{f_t\}$.

In section 4, we replaced $u_t = \delta - \delta_t$ by the difference $S_t - S_{t-1}$ of the cumulative sum S_t , and obtained a matrix equation for the $\{S_t\}$. Let us now define X_t , Y_t and Z_t to be the cumulative sums of $x_t - \xi$, $y_t - \eta$ and $z_t - \zeta$ respectively, and proceed in the same manner. We find that the vector of cumulative sums w_t , defined by

$$\mathbf{w}_t^T = (X_t \ X_{t-1} \ X_{t-2} \ Y_t \ Y_{t-1} \ Z_t \ Z_{t-1}) \quad (73)$$

obeys the matrix recurrence equation

$$\mathbf{w}_t = \mathbf{A} \mathbf{w}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (74)$$

where $\boldsymbol{\varepsilon}_t$ is a column vector with e_t as its first element, f_t as its fourth element, g_t as its sixth element, and zeros elsewhere, and

$$\mathbf{A} = \left[\begin{array}{ccc|cc|cc} 1+2k & -3k & k & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 2kb & -3kb & kb & 1+c & -c & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 2k(1-d)b & -3k(1-d)b & k(1-d)b & (1-d)c & -(1-d)c & 1+d & -d \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array} \right]. \quad (75)$$

Because of the block structure of \mathbf{A} , its eigenvalues are those of the three principal submatrices, each of which has a dominant eigenvalue of 1. The seven eigenvalues are in fact 1, 1, 1, $k_{\pm}\sqrt{[k(1-k)]}$, c and d . Denoting the dominant right and left eigenvectors of the three principal submatrices of \mathbf{A} by α_x and β_x^T , α_y and β_y^T , α_z and β_z^T respectively, we note that \mathbf{A}^t approaches

$$\mathbf{A}^{\infty} = \left[\begin{array}{cc|cc} \alpha_x \beta_x^T & \mathbf{0} & \mathbf{0} & \\ \mathbf{B}_{yx} & \alpha_y \beta_y^T & \mathbf{0} & \\ \mathbf{B}_{zx} & \mathbf{B}_{zy} & \alpha_z \beta_z^T & \end{array} \right]. \quad (76)$$

A little algebra reveals that the submatrices \mathbf{B}_{yx} , \mathbf{B}_{zx} and \mathbf{B}_{zy} take the forms $\alpha_y \gamma_x^T$, $\alpha_z \gamma_x^T$ and $\alpha_z \gamma_y^T$ respectively, where

$$\gamma_x^T = (L \quad -2L \quad L); \quad (77)$$

$$\gamma_y^T = (M \quad -M); \quad (78)$$

with

$$L = kb / [(1-c)(1-k)]; \quad (79)$$

$$M = c / (1-c). \quad (80)$$

Using these results and the same methods as in section 4, we deduce that for large t

$$E(X_t) \cong k(x_0 - x_{-1}) / (1-k); \quad (81)$$

$$E(Y_t) \cong L(x_0 - x_{-1}) + c(y_0 - \eta) / (1-c); \quad (82)$$

$$E(Z_t) \cong L(x_0 - x_{-1}) + M(y_0 - \eta) + d(z_0 - \zeta) / (1-d); \quad (83)$$

$$\text{Var}(X_t) \cong t \sigma_x^2 / (1-k)^2; \quad (84)$$

$$\text{Var}(Y_t) \cong t [\sigma_y^2 / (1-c)^2 + L^2 \sigma_x^2]; \quad (85)$$

$$\text{Var}(Z_t) \cong t[\sigma_z^2/(1-d)^2 + L^2\sigma_x^2 + M^2\sigma_y^2] ; \quad (86)$$

$$\text{Cov}(X_t, Y_t) \cong t L \sigma_x^2/(1-k) ; \quad (87)$$

$$\text{Cov}(X_t, Z_t) \cong t L \sigma_x^2/(1-k) ; \quad (88)$$

$$\text{Cov}(Y_t, Z_t) \cong t [L^2 \sigma_x^2 + M \sigma_y^2/(1-c)] . \quad (89)$$

The forms of the variances and covariances are not without interest. If, for example, $\sigma_y^2 = \sigma_z^2 = 0$, the random variables X_t and Z_t are asymptotically fully correlated. Other relationships are also evident.

Under this model, the present value at the earnings rate of the fund of a lump sum, currently 1, but adjusted for inflation, is the lognormal random variable $\exp(X_t - Z_t)$. The moments are readily approximated using the asymptotic moments for X_t and Z_t and the approximate formulae we derived for the earlier models.

12. Concluding remarks

Varying levels of inflation and fluctuating investment returns are problems with which the actuary must contend on an almost daily basis. Unlike mortality and other decrements or movements, for which deterministic and stochastic models are readily available, these economic factors are rather more difficult to model. Autoregressive methods are attractive for this purpose and have been proposed by a variety of authors. Many of the formulae which emerge from the autoregressive modelling of these processes are rather complicated. A major objective of this paper has been to show that practical approximate formulae are available in many instances, that these formulae do provide adequate approximations, and that day-to-day problems can be tackled using them.

13. References

- [1] Boyle, P.P. (1976) Rates of return as random variables. **J. Risk and Insurance**, **43**, 693-713.
- [2] Dufresne, D. (1988) Moments of pension fund contributions and fund levels when rates of return are random. **J Institute of Actuaries**, **115**, 535-544.
- [3] Giaccotto, C. (1986) Stochastic modelling of interest rates: actuarial vs. equilibrium approach. **J. Risk and Insurance**, **53**, 435-453.
- [4] Haberman, S. (1993) Pension funding with time delays and autoregressive rates of return. **Insurance: Mathematics and Economics**, **13**, 45-56.
- [5] Haberman, S. (1994) Autoregressive rates of return and the variability of pension contributions and funds levels for a defined benefit pension scheme. **Insurance: Mathematics and Economics**, **14**, 219-240.
- [6] Panjer, H.H. and Bellhouse, D.R. (1980) Stochastic modelling of interest rates with applications to life contingencies. **J. Risk and Insurance**, **47**, 91-110.

- [7] Panjer, H.H. and Bellhouse, D.R. (1981) Stochastic modelling of interest rates with applications to life contingencies - Part II. **J. Risk and Insurance**, **48**, 628-637.
- [8] Parzen, E. (1962) **Stochastic Processes**. Holden-Day, San Francisco.
- [9] Pollard, A.H. and Pollard, J.H. (1969) A stochastic approach to actuarial functions. **J. Institute of Actuaries**, **95**, 79-113.
- [10] Pollard, J.H. (1971) On fluctuating interest rates. **Bulletin de l'Association Royale des Actuaire Belges**, **66**, 68-97.
- [11] Pollard, J.H. (1973) **Mathematical Models for the Growth of Human Populations**, 44. Cambridge University Press.
- [12] Pollard, J.H. (1976) Premium loadings for non-participating business. **J. Institute of Actuaries**, **103**, 205-212.
- [13] Scott, W.F. (1977) A reserve basis for maturity guarantees in unit-linked life assurance. **Transactions of the Faculty of Actuaries**, **35**, 365-385.
- [14] Westcott, D.A. (1981) Moments of compound interest functions under fluctuating interest rates. **Scandinavian Actuarial Journal**, (1981 volume), 237-244.
- [15] Wilkie, A.D. (1976) The rate of interest as a stochastic process. Theory and applications. **Proceedings of the 20th International Congress of Actuaries**, **1**, 325-338.

Table 2
Assurance and annuity certain expectations: exact and approximate values for the
second-order model (1) with $\delta=\ln(1.05)$, $\sigma=0.001$ and $k=0.9$.

Term n	Assurance		Annuity	
	Exact	Approx	Exact	Approx
<i>Case $i_1=0.05, i_0=0.05$</i>				
20	0.3773	0.3773	12.469	12.473
30	0.2318	0.2317	15.383	15.391
40	0.1424	0.1423	17.173	17.184
60	0.0537	0.0537	18.948	18.963
80	0.0203	0.0203	19.618	19.634
<i>Case $i_1=0.06, i_0=0.07$</i>				
20	0.3559	0.3467	11.374	11.463
30	0.2100	0.2130	14.025	14.144
40	0.1318	0.1308	15.681	15.791
60	0.0495	0.0493	17.309	17.426
80	0.0186	0.0186	17.924	18.043
<i>Case $i_1=0.07, i_0=0.06$</i>				
20	0.3982	0.4105	13.470	13.573
30	0.2569	0.2521	16.627	16.748
40	0.1532	0.1549	18.579	18.699
60	0.0582	0.0584	20.510	20.635
80	0.0220	0.0220	21.239	21.366

Table 1
Assurances and annuities certain: exact and approximate second moments for the
second order model (1) with $\delta=\ln(1.05)$, $\sigma=0.001$ and $k=0.9$.

	Term Assurance Variance $\times 10^4$		Assurance/annuity covariance $\times 10^3$		Annuity variance	
	Exact	Approx.	Exact	Approx.	Exact	Approx.
<i>Case $i_{-1}=0.05, i_0=0.05$</i>						
20	3.358	2.844	3.8579	4.1816	0.0848	0.0882
30	1.856	1.608	4.2115	4.2572	0.1619	0.1727
40	0.896	0.809	3.3802	3.5043	0.2378	0.2491
60	0.186	0.172	1.7456	1.7842	0.3369	0.3508
80	0.035	0.033	0.7534	0.7652	0.3834	0.3982
<i>Case $i_{-1}=0.06, i_0=0.07$</i>						
20	2.988	2.402	3.3686	3.5314	0.0706	0.0745
30	1.523	1.359	3.4805	3.5952	0.1349	0.1458
40	0.768	0.683	2.8774	2.9594	0.1993	0.2104
60	0.158	0.146	1.4760	1.5067	0.2827	0.2962
80	0.029	0.028	0.6358	0.6462	0.3220	0.3362
<i>Case $i_{-1}=0.7, i_0=0.06$</i>						
20	3.740	3.367	4.4614	4.9516	0.1034	0.1045
30	2.279	1.905	5.1140	5.0412	0.1948	0.2045
40	1.038	0.958	3.9720	4.1496	0.2854	0.2950
60	0.219	0.204	2.0649	2.1127	0.4031	0.4154
80	0.041	0.039	0.8932	0.9061	0.4584	0.4715

Table 3
Life annuity variance, and components due to interest rate variation
and variation in age at death^a

Age x	Exact variance of \mathbf{a}_x (age at death and interest both varying)	Approx. variance of \mathbf{a}_x (equation (19))	Age at death component of (3)	Varying interest component of (3)
(1)	(2)	(3)	(4)	(5)
20	4.779	4.851	4.472	0.379
30	6.015	6.074	5.753	0.321
40	9.398	9.438	9.190	0.248
50	14.122	14.142	13.974	0.167
60	16.392	16.398	16.304	0.095

^aIn this example, $\sigma=0.01$, $k=0.9$, $i=0.05$, $i_1=i_0=0.05$.

Table 4
Cumulative distribution of s_{10}

P(s_{10}] < x)			P(s_{10}] < x)			P(s_{10}] < x)		
x	Simul- ation	Log- normal	x	Simul- ation	Log- normal	x	Simul- ation	Log- normal
(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
10.5	.000	.001	27.5	.778	.776	44.5	.990	.997
11.5	.002	.003	28.5	.815	.818	45.5	.994	.998
12.5	.004	.007	29.5	.840	.853	46.5	.995	.999
13.5	.013	.016	30.5	.866	.882	47.5	.998	.999
14.5	.023	.031	31.5	.890	.906	48.5	.998	.999
15.5	.037	.055	32.5	.915	.926	49.5	.998	.999
16.5	.071	.090	33.5	.930	.942	50.5	.998	1.000
17.5	.109	.136	34.5	.947	.955	51.5	.999	1.000
18.5	.167	.192	35.5	.958	.965	52.5	.999	1.000
19.5	.246	.256	36.5	.963	.973	53.5	.999	1.000
20.5	.304	.326	37.5	.970	.979	54.5	.999	1.000
21.5	.387	.399	38.5	.976	.984	55.5	.999	1.000
22.5	.461	.473	39.5	.980	.988	56.5	.999	1.000
23.5	.539	.544	40.5	.982	.991	57.5	.999	1.000
24.5	.601	.611	41.5	.983	.993	58.5	1.000	1.000
25.5	.664	.673	42.5	.987	.995	59.5	1.000	1.000
26.5	.719	.728	43.5	.989	.996	60.5	1.000	1.000

Figure 1
Three independent simulations of the interest rate trajectory under the second-order autoregressive model (1) with $i=0.15$, $\sigma=0.05$, $k=0.5$, and $i_1=i_0=0.15$.

