A BAYESIAN APPROACH TO PERSISTENCY RATES WHEN PROJECTING RETIREMENT COSTS

by

Arnold F. Shapiro
The Pennsylvania State University

For Presentation
at
The 13th Annual Actuarial Research Conference

A Bayesian Approach to Persistency Rates when Projecting Retirement Costs

by

Arnold F. Shapiro

> ...it is almost inconceivable that anybody could be in the position of having no *a priori* knowledge whatever regarding mortality. -- E. T. Whittaker

Introduction

Pension actuaries long have realized that pension cost projections provide valuable insight into the cash flow characteristics of pension plans.[1] Because of this, the more

---

[1]One of the first published accounts of the growth of a pension fund was James J. M'Lauchlan, "The Fundamental Principles of Pension Funds," *TFA*, Vol. IV (1908), pp. 195-227. In that article M'Lauchlan illustrated the necessity of accummulating large investment funds during the early years of a fund's existence so as to provide for the heavy liability which ultimately will be maturing for payment.

elaborate pension plan proposals and valuations typically append a pension cost projection to their report. These range, depending on the size of the plan, from simple projections which assume a closed group with no terminations other than for retirement[2] to more sophisticated models which introduce the full spectrum of pension plan parameters.[3]

While this practice has much to commend it, providing as it does valuable insight into the cash flow of a pension plan, it suffers at least two serious shortcomings. First, it provides no mechanism for incorporating the actuary's "feelings" regarding his confidence in underlying assumptions. Perhaps the most important attribute of an experienced pension actuary is his intuitive notion of what should be, and, ideally, there should be some vehicle for injecting

---

[2]This type of projection is most commonly associated with the valuation of small pension plans. See, for example, Calculating Auxilliary Fund Deposits for the Small Pension Plan. Chicago, Ill.:A. A. Beaven & Company, Inc., 1975.

[3]Articles which discuss this type of projection include Robert J. Myers, "Some Considerations in Pension Fund Valuations," TASA, Vol. XLVI (1945), pp. 51-58, A. M. Niessen, "Projections -- How to Make Them and How to Use Them," TSA, Vol. II (1950), pp. 235-253, Charles L. Trowbridge, "Fundamentals of Pension Funding," TSA, Vol. IV (1952), pp. 17-43, and Frank L. Griffin, Jr., "Concepts of Adequacy in Pension Plans," TSA, Vol. XVIII (1966), pp. 46-63.

this intuition into pension cost projections. Second, since projections invariably are based on expected value models, they provide no mechanism for introducing credibility.* Ideally, attached to any estimate of projected pension costs should be a statement of the actuary's confidence in that estimate.

It might be argued that this degree of refinement of pension cost projections is not warranted. Proponents of this view reason that pension costs are funded sequentially over a number of years and that periodic actuarial valuations will uncover underfunding problems before they can materially affect the solvency of a plan. The implication being that ex ante pension cost projections should be viewed strictly as rough (albeit best) estimates of ultimate pen-

---

*A. Guy Shannon Jr., in Pension Topics, Study Note 71-22-76 of the Society of Actuaries, p. 10, for example, remarked:

> Invariably, projections are based on expected value models and seldom is there a quantified statement of the actuary's confidence in the projection. Ideally, the individual assumptions and the composite results of the valuation should be viewed as the mean of the universe from which the experience of that pension plan will be drawn. The measurement of liabilities would be accompanied by a set of confidence limits based on the combined effect of the entire set of assumptions.

sion plan costs. The fact that such projections may not convey an accurate picture of ultimate cost is regarded as only marginally relevant.

This proposition, however, disregards the question of whether a particular plan or plan liberalization would have been introduced initially had the plan sponsor realized that actual cost might be considerably in excess of the projected costs. Furthermore, this view presumes that the plan sponsor will be able to fund any deficiencies which arise. These considerations have become increasingly important in light of the liability ERISA imposes on plan sponsors.[5] Thus, while ex post reconciliation of pension cost estimates remains an important facet of pension cost funding, there are compelling arguments for developing techniques to measure the variability of ex ante pension cost projections.

These observations suggest the need for a stochastic model for projecting pension costs. A straight forward procedure would be to base such a model on direct or deductive probabilities. One could assume, for example, that the number of participants who succumb to a particular decrement is binomially distributed and based upon a probability of

----

[5]Under ERISA, Section 4062(b), an employer's liability may be as high as thirty percent of its net worth.

decrement which is constant or is given by a degenerate distribution. This assumption of an underlying degenerate distribution, however, is questionable in actual practice. Probabilities of decrement, for example, are obtained either from intercompany experience, which, at best, may only approximate the actual experience of a particular firm, or else it is derived from the firm's own experience, which, for the majority of firms, is not very credible. Thus, what is needed is a model in which underlying parameters also may take on probability distributions.

These additional considerations lead naturally to a Bayesian approach to stochastic pension cost projections. Under this approach, not only are pension cost determinants, such as the number of decrements due to a given cause and the fund accumulation factor, assumed to be stochastically distributed, but the parameters upon which these determinants depend are themselves assumed to be stochastically distributed.

This article uses a Bayesian approach to persistency to explore pension cost projection variability.[6] The analysis

---

[6]Persistency is not the only source of variability in pension cost projections, of course. Deviations resulting from such things as shifts in the distribution of salaries or returns on assets are also extremely important sources of variation. For the purpose of the present study, however,

is divided into two parts. In the first portion the specifications of the model are developed. The second portion shows how the model might be used to generate confidence intervals for pension cost projections.

The article ends with a comment on the use of stochastic models and suggestions for further study.

Stochastic Pension Cost Models

Very few articles have dealt specifically with the development of a stochastic model of pension costs. Stone[7] investigated the impact of mortality fluctuations on pensions paid to pensioners. The main thrust of that study was the use of probability generating functions to develop probabilities, at various durations after employees had begun to retire, that the total actual pension payments would differ from the expected total payments. Taylor[8] investigated the size of the contingency reserve needed to insure, with a given probability, that the funds on hand would be suffi-

---

factors unrelated to persistency are assumed to be invariant.

[7]David G. Stone, "Actuarial Note: Mortality Fluctuations in Small Self-Insured Pension Plans," TASA, Vol. XLIX (1948), pp. 82-91.

[8]Robert H. Taylor, "The Probability Distribution of Life Annuity Reserves and Its Application to a Pension System," PCAPP, Vol. II (1952), pp. 100-150.

cient to pay all promised pensions. Both these studies dealt exclusively with the retired population, under the assumption that the number of retirees was known.

Articles which considered variability in pension cost estimates for active plan participants included the studies of Seal,[9] Knopf,[10] and Shapiro.[11] Seal investigated the impact of death benefits in a trusteed plan using a normal approximation to the binomial distribution to introduce variance minimization into the design of pension plans. Knopf investigated the feasibility of fully trusting small pension plans using a simplified monte carlo approach. Shapiro considered the credibility of projected pension costs using a model based on the direct application of a conditional Bernoulli process.

Of course, to the extent that pensions may be regarded as annuities, there have been a considerable number of other

---

[9]Hilary L. Seal, "The Mathematical Risk of Lump-Sum Death Benefits in a Trusteed Pension Plan," TSA, Vol. V (1953), pp. 135-142.

[10]Myrna Knopf, "A Practical Demonstration of the Risk Run by a Very Small Company with a Trusteed Pension Plan," PCAPP, Vol. 6 (1956-7), pp. 230-43.

[11]Arnold Shapiro, "The Relevance of Expected Persistency Rates when Projecting Pension Costs," JRI, Vol. XLIV, No. 4 (December, 1977), pp. 623-638.

relevant studies. Piper,[12] for example, developed contingency reserves for life annuities based on the mean and variance associated those annuities. Menge,[13] and later Hickman,[14] elaborated on the Piper article: Menge using discrete functions and Hickman using continuous functions. Hickman's article, in addition, extended the development to include loss functions and a probabilistic consideration of multiple decrement theory. The latter, of course, is directly applicable to pension populations.

Although it is clear that the number of lives which persist to a given age from an initial group of lives is generated by a Bernoulli process, the complexity of this process resulted in the development of various approximation methods. Hence, Piper assumed a large group of lives and used a normal distribution, as did Seal; Taylor suggested fitting a Pearson Type III distribution to the total present value of life annuity costs; Boermeester[15] applied a Monte

---

[12]Kenneth B. Piper, "Contingency Reserves for Life Annuities," TASA, Vol. XXXIV (1933), pp. 240-249.

[13]W. C. Menge, "A Statistical Treatment of Actuarial Functions," RAIA, Vol. XXVI (1937), pp. 65-88.

[14]James C. Hickman, "A Statistical Approach to Premiums and Reserves in Multiple Decrement Theory," TSA, Vol. XVI (1964), pp. 1-16.

Carlo approach to the problem as did Knopf; Fretwell and Hickman[16] investigated upper bounds for the cost using the inequalities of Tchebychef and Uspensky; and Bowers[17] investigated the use of the Cornish-Fisher expansion to develop probabilities of sufficient reserves, based on correction factors applied to a standard normal table.

These studies generally relied on distributions whose underlying parameters were given. This study explores the use of a less constrained distribution.

The Probability of a Given Number of Participants at each age

The number of participants at age x in a pension plan may be regarded as a random variable, l_x^{aa} say,[18] that depends on the number of participants at the previous age,

---

[15]J. M. Boermeester, "Frequency Distribution of Mortality Costs," TSA, Vol. VIII (1956), pp. 1-9.

[16]Robert L. Fretwell and James C. Hickman, "Approximate Probability Statements About Life Annuity Costs," TSA, Vol. XVI (1964), pp. 55-60.

[17]N. L. Bowers, "An Approximation to the Distribution of Annuity Costs," TSA, Vol. XIX (1968), pp. 295-309.

[18]Throughout this article standard actuarial notation is adhered to, as far as possible. See the outline of International Actuarial Notation, with revisions, TASA, Vol. 48 (1947), pp. 166- 176.

l_{x-1}^{aa}, which is also a random variable, except at the entry age. Let

$$ {}^{k}\underset{\sim}{l}_x^{aa} = ({}^{k}l_{x-t}^{aa} : t=0,1,\ldots,x-a)' \quad (1) $$

where a is the entry age, denote a vector of l_{x-t}^{aa} values consistent with a final value of l_x^{aa}, and call this vector a feasible l_x^{aa} array. Assuming that there are K distinct feasible l_x^{aa} arrays, the probability that $\tilde{l}_x^{aa}$ takes on some particular value, is given by

$$ Pr\{\tilde{l}_x^{aa}=l_x^{aa}\} = \sum_{k=1}^{K} \prod_{t=0}^{x-a-1} f({}^{k}l_{x-t}^{aa} \mid l_x^{aa}), \quad (2) $$

where f denotes the probability that exactly l_{x-t}^{aa} participants will persist through age x-t-1, consistent with l_x^{aa} participants persisting through age x.

A Conditional Probability Distribution Function for l_x^{aa}

In order to implement equation 2 it is necessary to specify the probability distribution function. Under the assumption that valuations are based only on curtate ages, the number of employees who persist through a given age may be thought of as being generated by a Bernoulli process under which employees either persist as active members or leave the active group.[19] It follows that a conditional distribution of l_{x-t}^{aa} given p_{x-t-1}^{aa} is specified by the binomial

---

[19]Howard Raiffa and Robert Schlaifer, Applied Statistical Decision Theory (Boston: The M.I.T. Press, 1968), Chapter 9.

mass function[20]

$$f_b(l_{x-t}^{aa} \mid p_{x-t-1}^{aa}, l_{x-t-1}^{aa}) \quad (3)$$

$$\propto (p_{x-t-1}^{aa})^{l_{x-t}^{aa}} (1-p_{x-t-1}^{aa})^{l_{x-t-1}^{aa} - l_{x-t}^{aa}}$$

In Figure 1 the binomial distribution is used to project the distribution of the number of plan participants at each age through age 65, assuming there are 100 entrants at age 20. The probabilities of persisting are based on mortality rates from the 1971 Group Annuity Mortality Table;[21] disability rates used in the 1970 Civil Service Pension valuation: and Turnover Table III given by McGinn.[22] This data base, which is used for illustrative purposes, will subsequently be referred to as "the decrement data." The curve to the far right represents the distribution of participants

---

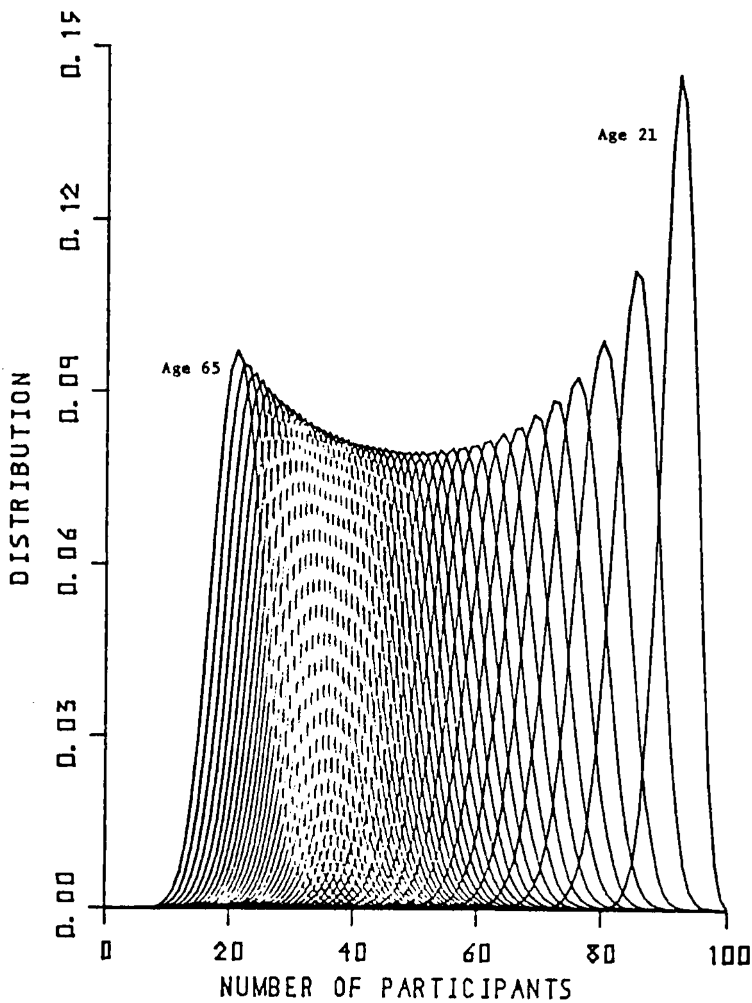
[20]*Ibid.*, p. 213. For an application of this distribution to the problem of projecting pension cost see Shapiro, *op. cit.*

[21]See Harold R. Greenlee, Jr., and Alfonso D. Keh, "The 1971 Group Annuity Mortality Table," *TSA*, Vol. XXIII (1972), pp. 583-584.

[22]Daniel F. McGinn, "Indices to the Cost of Vested Pension Benefits," *TSA*, Vol. XVII (1966), pp. 235-6. If vesting were the topic of this study, bodily shifts in the rates of withdrawal subsequent to a vesting liberalization would be an important additional source of variation. Since this study deals solely with retirement benefits, however, this complication is not introduced.

FIGURE 1

DISTRIBUTION OF NUMBER OF PARTICIPANTS AT AGES 21 THROUGH 65, GIVEN THAT THEY ARE BINOMIALLY DISTRIBUTED



Data base: 1971 GAM, rates of disability from the 1970 Valuation of the Civil Service Retirement System, and McGinn's Turnover Table III.

at age 21. The curve to the far left represents the distribution of the number of participants who will retire at age 65. The intermediate curves are associated with participants at intermediate ages.

It is apparent from these curves that, even under conditions of perfect information, the actual number of participants at a given age may vary considerably from the best estimate of the number of participants. While this is not surprising, what is interesting is the considerable disparity which is likely to occur. In the graph the locus of the modes of the distribution of participants is convex. The age at which the locus attains a minimum value represents the age at which the distribution of participants is most nearly symmetrical. Below this age the distribution of participants is negatively skewed and above this age the distribution of participants is positively skewed.

It is important to recognize that the binomial mass function is appropriate only under the assumption that the exact probabilities of persisting are known. This assumption, however, is generally not valid. Although it is true that estimates of p_x^{aa} are often available, these estimates may or may not be valid for the particular pension plan under consideration. Furthermore, the binomial mass function provides no mechanism for the actuary to indicate the intensity with which he views the credibility of the estimated value p_x^{aa}. These criticisms suggest the need for a more general probability distribution function. What is needed is a distribution which is not conditional upon a degenerate p_x^{aa}, that is, an unconditional distribution.

An Unconditional Probability Distribution for l_x^{aa}

Bayes' theorem may be used to transform the conditional probability of l_x^{aa} individuals persisting to an unconditional probability.[23] According to Bayes' theorem, if y has the probability density function $f_1(y)$, and the conditional probability distribution function of x, given y, is $h(x|y)$, then the joint distribution of x and y, $f(x,y)$, is given by

$$f(x,y)=h(x|y)f_1(y) \quad (4)$$

---

[23]For discussions of Bayesian analysis with insurance applications see Arthur L. Bailey, "Credibility Procedures," PCAS, Vol. XXXVII (1950), pp. 7-23, Andrew R. Davidson and A. R. Reed, "On the Calculation of Rates of Mortality," TFA, Vol. XI (1927), pp. 183-212, James C. Hickman and Robert B. Miller, "Notes on Bayesian Graduation," TSA, Vol. XXIX (1977), pp. 7-21, Donald A. Jones, "Bayesian Statistics," TSA, Vol. XVII (1965), pp. 33-57, Allen L. Mayerson, "A Bayesian View of Credibility," PCAS, Vol. LI (1964), pp. 85-104, Wilfred Perks, "Some Observations on Inverse Probabilities Including a New Indifference Rule," JIA, Vol. 73 (1947), pp. 285-310, and E. T. Whittaker, "On Some Disputed Questions of Probability," TFA, Vol. VIII (1920), pp. 163-206.

If y has a continuous distribution, it follows that the marginal distribution of x, $f_2(x)$, is

$$f_2(x) = \int f(x,y)\,dy \qquad (5)$$
$$= \int h(x|y)\, f_1(y)\,dy$$

which is independent of y.[24]

From the previous section we know the conditional distribution of exactly l_x^{aa} lives persisting through age x. If we can assume that the probability of persisting, p_x^{aa}, is a random variable from a nondegenerate distribution, the unconditional probability of l_x^{aa} lives persisting can be determined.

Certain properties of p_x^{aa} seem evident. First, $0 \leq p_x^{aa} \leq 1$, so that the distribution from which p_x^{aa} is drawn, generally referred to as the "prior" distribution, must be distributed over this range. Second, the probability of persisting may take any value in this domain, so that p_x^{aa} has a continuous distribution. Finally, for any given age, the probability of persisting may be concentrated at at most one value, so that the distribution of p_x^{aa} has a single

---

[24]A more general formulation would give the marginal of x in terms of a generalized Riemann-Stieltjes integral with respect to y. However, since the distribution of the probability of persisting is continuous this complication need not be introduced.

mode.[25] It is assumed that any probability density function that is chosen to represent the probability of persisting must exhibit these properties.

In addition to the empirical properties mentioned above, another desirable property stems from the fact that it may be impossible to specify the distribution of p_x^{aa} exactly, due to a scarcity of relevant data. The distribution that is used to characterize the probability of persisting should lend itself to updating as more sample information becomes available.[26]

A convenient choice for the prior distribution of p_x^{aa}, from an updating point of view, is the beta distribution in the form

---

[25]While this constraint seems generally appropriate, it has been argued that it may not be a necessary or desirable one. G. E. Lidstone, in his discussion of Whittaker, op. cit., p. 196, for example, suggested the possibility of using U-shaped curves in those instances when high probabilities occur in the upper or lower bounds of the distribution. Such a prior was subsequently developed for the binomial distribution by Perks, op. cit., based on hypothesis that $p_x \cdot dx \propto dx/\sigma_x$, where σ_x is the large sample standard error of x, a parameter in a probability law.

[26]It is important from the point of view of pension plan valuation, i.e., the going concern analysis, to be able to update estimates of pension population parameters as more data becomes available. The development of a stochastic pension valuation model which incorporates this facility is currently under investigation by the author, and will form the basis for a sequel to the present study.

$$f_{\beta}(p|r,n)=p^{r-1}(1-p)^{n-r-1}/B(r,n-r) \quad (6)$$

where

$$B(r,n-r)=\Gamma(r)\Gamma(n-r)/\Gamma(n).$$

This follows since the updated, or "posterior," distribution would also be a beta distribution.[27]

The beta distribution satisfies all the empirical requirements mentioned above except that it is not necessarily unimodal. This requirement is met, however, if r and n are restricted to positive values and max{r,n-r} exceeds unity.[28]

Given that p_x^{aa} has a beta distribution as specified above, the unconditional distribution of the number of employees who persist through age x-1 is given by

$$\int_0^1 f_b(\ell_x^{aa}|p_{x-1}^{aa},\ell_{x-1}^{aa})\, f_\beta(p_{x-1}^{aa}|r_{x-1},n_{x-1})\, dp_{x-1}^{aa} \quad (7)$$

$$=\binom{\ell_{x-1}^{aa}}{\ell_x^{aa}} \frac{B(\ell_x^{aa}+r_{x-1},\ \ell_{x-1}^{aa}-\ell_x^{aa}+n_{x-1}-r_{x-1})}{B(r_{x-1},\ n_{x-1}-r_{x-1})}, \quad \ell_x^{aa}=0,1,\ldots,\ell_{x-1}^{aa}$$

$$n_{x-1}>r_{x-1}>0$$

$$\max\{r_{x-1},n_{x-1}-r_{x-1}\}>1$$

[27]See Raiffa and Schlaifer, op. cit., p. 263. It is interesting to note that Sir G. F. Hardy alluded to this distribution in the form $x^r(1-x)^s$ in a correspondence regarding a Bayesian approach to mortality. Insurance Record Vol. XXVII (October, 1889) ff. 433. It is not clear, however, whether Hardy recommended the method for practical use. See Lidstone's discussion of Davidson and Reed, op. cit., p. 225.

[28]The beta distribution is bimodal if max{r,n-r} is less than one, in which case it has a U-shape. It was this distribution, in the form n=2r=1 that was developed by Perks, op. cit., p. 298.

This distribution is appropriately called the beta-binomial distribution, and its probability distribution function is denoted by $f_{\beta b}(l_x^{aa} | r_{x-1}, n_{x-1}, l_{x-1}^{aa})$.[29]

As a consequence of the foregoing, it follows that an unconditional probability of exactly l_x^{aa} employees persisting to age x is

$$\sum_{k=1}^{K} \prod_{t=0}^{x-a-1} f_{\beta b}({}^{k}l_{x-t}^{aa} | l_x^{aa}, r_{x-t-1}, n_{x-t-1}, {}^{k}l_{x-t-1}^{aa}) . \quad (8)$$

The remainder of this paper assumes the beta distribution adequately describes the distribution of the probability of persisting, and that the beta-binomial distribution appropriately describes the distribution of the number of participants at a given age.

Estimating the Parameters r and n

In order to utilize the beta-binomial distribution the parameters r and n be either known or estimated. In practice, of course, it is unlikely that the exact value of these parameters are known, so it is necessary to estimate them. In this section the method of moments in conjunction with subjective judgement is used to develop an approach for

---

[29] Raiffa and Schlaifer, op. cit., p. 237.

estimating these parameters.[30]

Let the tabular probability of persistency, for some particular age, be denoted by $\hat{p}_x^{aa}$. Since $E(p_x^{aa}) = r_x/n_x$,[31] it follows that if one assumes that the tabular probability of persisting is approximately equal to the probability of persisting, then

$$r_x \doteq n_x \cdot \hat{p}_x^{aa} \quad (9)$$

Furthermore, since the variance of the beta distribution is[32]

$$V(p_x^{aa}) = E(p_x^{aa}) \; E(1-p_x^{aa}) / (n_x+1), \quad (10)$$

it follows that

$$V(p_x^{aa}) \doteq \hat{p}_x^{aa} \; (1-\hat{p}_x^{aa}) / (n_x+1). \quad (11)$$

It is clear that the estimated variance of the prior distribution of the probability of persisting will be inversely proportional to the size of the n_x parameter that is chosen. The greater the confidence in the tabular per-

---

[30]There are, of course, more sophisticated methods for developing sample estimates of r and n, such as the method of maximum likelihood. See, for example, S. W. Dharmadhikari, "A Simple Modification of the Binomial Distribution," JIASS, Vol. 15 (1960), pp. 436-444. However, the simplicity of the approach used in the text is a strong argument in its favor, particularly as a method of forming an initial prior.

[31]See Raiffa and Schlaifer, op. cit.

[32]Ibid., p. 213.

sistency rate the greater the value of n that should be chosen, that is, the smaller should be the estimated variance. Once an appropriate n_x is chosen, r_x is determined by solving equation 9. Denoting by $\hat{r}_x$ and $\hat{n}_x$ the estimated parameters of the prior distribution, the unconditional distribution of l_{x-t}^{aa} becomes

$$f_{\beta b}(l_{x-t}^{aa} \mid \hat{r}_{x-t-1}, \hat{n}_{x-t-1}, l_{x-t-1}^{aa}). \qquad (12)$$

The Implementation of the Foregoing Procedure

Tables 1 and Figures 2 and 3 exemplify the mechanics of the foregoing procedure by showing how one might determine a subjective prior distribution for p_{20}^{aa}. Given the decrement data, the tabular value of $\hat{p}_{20}^{aa}$ is 0.918247. Table 1 which was developed by substituting this value into equation 11, shows the trend of the estimated variance for various choices of $\hat{n}_{20}$. As indicated therein, an actuary who feels extremely confident in the tabular persistency rate might choose an $\hat{n}_{20}$ of 100 or more. This would result in a prior distribution for p_{20}^{aa} which has a variance of .0000075 or less. This distribution would, for all intent and purposes, be degenerate, and a binomial distribution might be used in this case. On the other hand, an actuary may be satisfied that .918247 represents a good estimate of the mean of the prior, but may, at the same time, feel that there is a con-

TABLE 1

IMPACT OF n ON VARIANCE OF THE PRIOR BETA DISTRIBUTION, GIVEN A MEAN OF 0.918

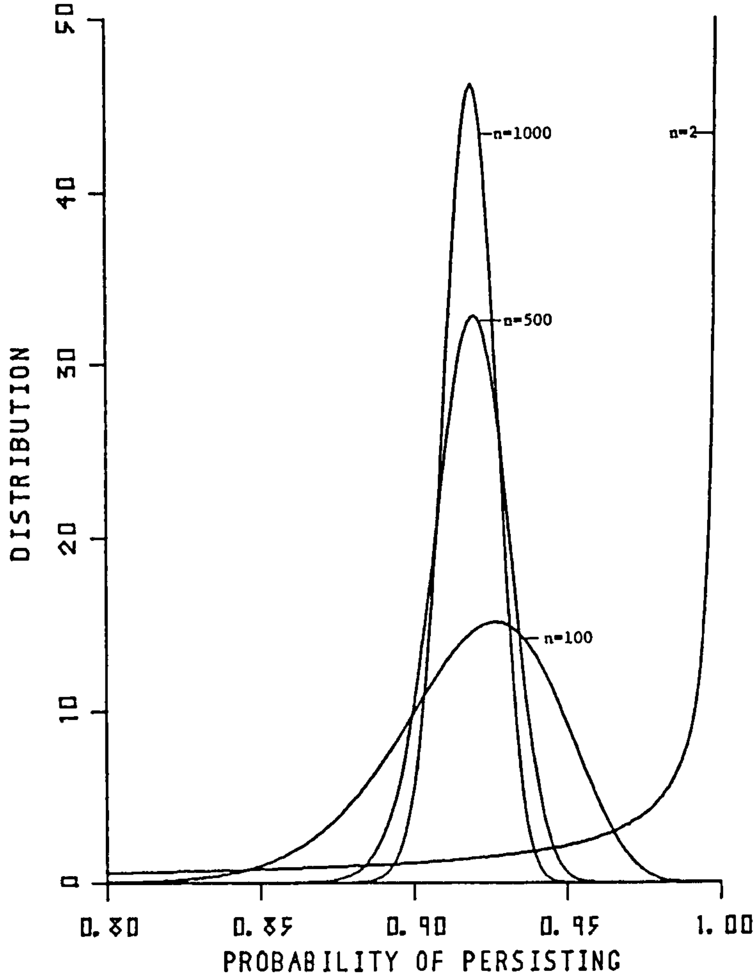
Prior Value of n_{20}	Variance
1+	.0359350*
2	.0250231
100	.0007507
500	.0001501
10000	.0000075
∞	.0000000

*Actual Value of n is 1.089+, which is the smallest value of n which is consistent with a unimodal beta distribution

siderable possibility that $\tilde{p}_{20}^{\wedge\wedge}$ may take on some other value. In an extreme situation of this kind, the actuary may be so uncertain of the outcome that he might choose to introduce considerable variability. This could be done by choosing an $\hat{n}_{20}$ equal to two: a distribution with a mean of .918247 and a variance of approximately .025 would result. Any distribution between these two extremes would also be available.

Figure 2 shows the impact of various choices of $\hat{n}_{20}$ on

FIGURE 2
EFFECT OF THE CHOICE OF n ON THE BETA PROBABILITY DENSITY FUNCTION GIVEN A MEAN OF 0.918



the prior beta distribution of p_{20}^{aa}. A choice of $\hat{n}_{20}$ equal to 1,000 results in an almost symmetric distribution about the mean. At the other extreme, a choice of $\hat{n}_{20}$ equal to two results in a distribution for p_{20}^{aa} which is highly skewed towards the origin and has a maximum value at unity.

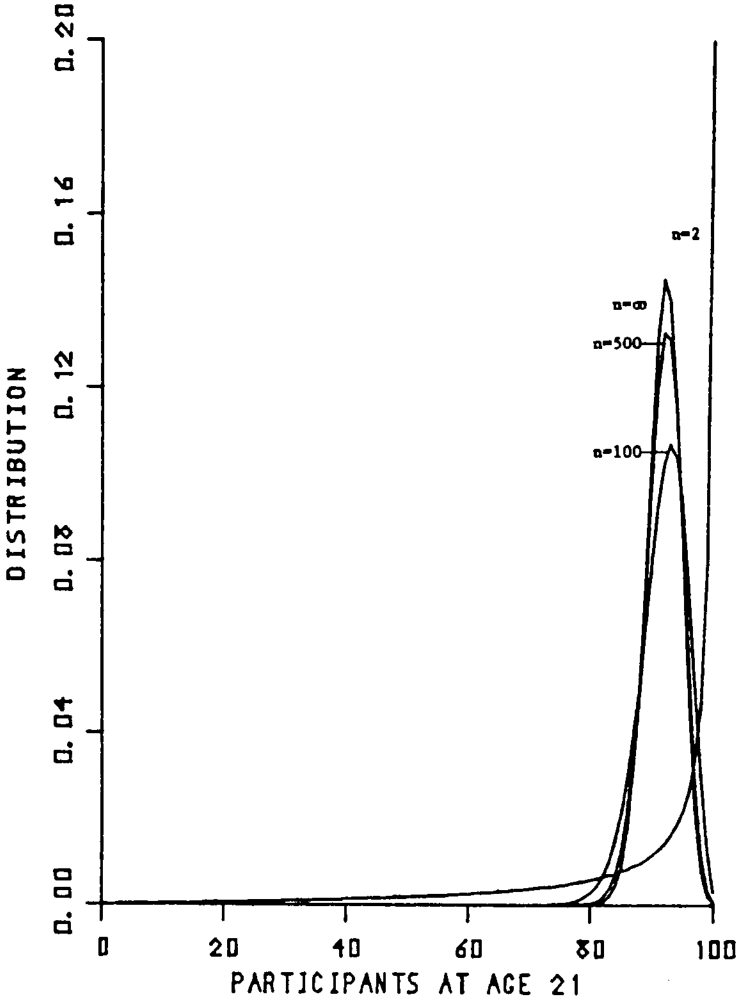
Figure 3 shows the probability of a given number of participants persisting to age twenty-one given 100 entrants at age twenty, based on some of the distributions given in Figure 2. The expected number of participants at age twenty-one is 91.82. It is apparent that as the variance of the distribution of the probability of persisting tends zero, the distribution of the number of employees tends to its limiting distribution, the curve labeled '$n=\infty$', which is based on a binomial mass function. This is as would be expected since, in the limit, the beta-binomial distribution tends to the binomial distribution.[33] Once again, if $\hat{n}_{20}$ is equal to two a hyperbolic curve results.

---

[33]Intuitively, the fact that, in the limit, the beta-binomial distribution tends to the binomial distribution follows from the observation that the binomial distribution obtains when the prior becomes degenerate.

FIGURE 3

EFFECT OF THE CHOICE OF n ON THE DISTRIBUTION OF EMPLOYEES AT AGE 21, GIVEN A PROBABILITY OF 0.918 OF PERSISTING



Data base: 1971 GAM, rates of disability from the 1970 Valuation of the Civil Service Retirement System, and McGinn's Turnover Table III.

The Implications of the Choice of the Parameters r and n

Before proceeding it is appropriate to mention the implications of a particular choice of the parameters r and n. The choice of a small prior n is tantamount to the assumption that the estimated probability of persisting at a given age, although the best available estimate, is questionable. There is a reasonable chance, based on the subjective judgement of the actuary, that the probability of persisting will take some value more or less than the best available estimate. On the other hand, the choice of a large n is tantamount to the assumption that the best available estimate of the probability of persisting is representative of the actuary's subjective evaluation of what that probability should be.

It also should be noted that the variance associated with the distribution of p_x^{aa} need not be the same for each age. The variance may, for example, be somewhat larger for the ages in the vicinity of the initial or full vesting ages, where an actuary might be unsure of his best estimate of p_x^{aa}. For other ages, where the impact of vesting might be slight, an actuary may have considerable confidence in his estimate, and he might chose a somewhat smaller variance for the distribution of p_x^{aa}.

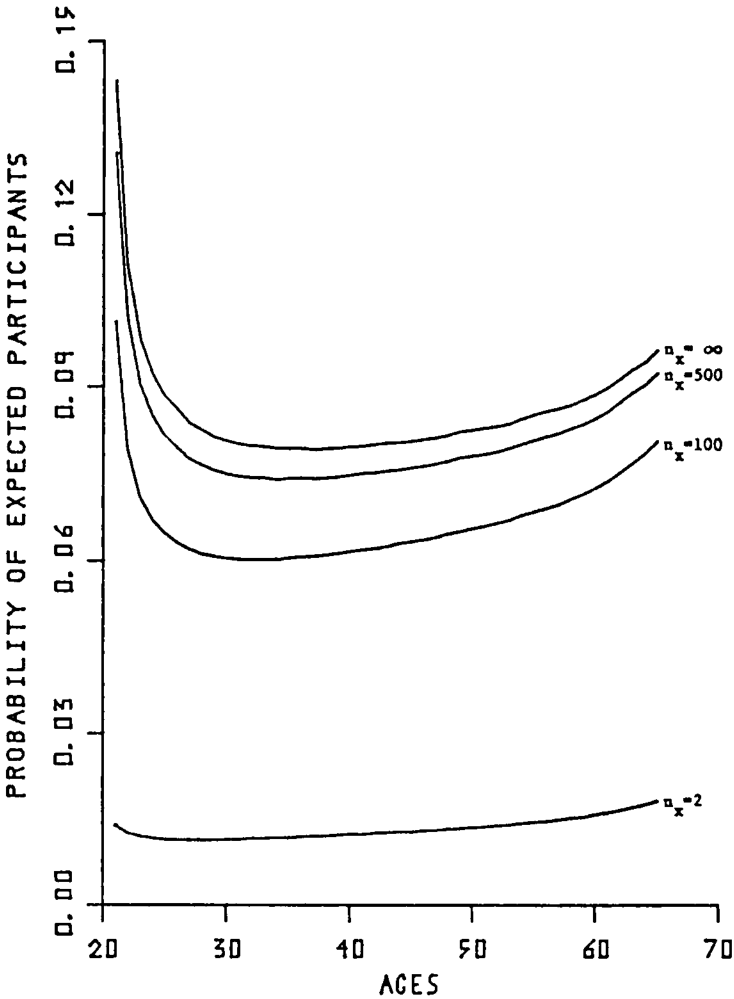
The Probability of the Projected Number of Participants

Although the expected number of participants at a given age is the same regardless of the assumption concerning the prior distribution of the probability of persisting, this is not the case for the probability that the expected number of participants will obtain. In fact, the probability that the actual number of participants at a given age is equal to the projected number of participants is very much a function of the distribution of p_x^{aa}. Figure 4, which was derived using equation 8, exemplifies this characteristic for 100 entrants aged twenty. The values given are, of course, interpolated values, since in most cases the projected number of participants is not integral.

As the prior distribution tends towards degeneracy the probability of the expected number of participants obtaining increases. For example, if the prior distribution of p_x^{aa} is degenerate at its mean, that is, if n_x is infinitely large, the probability that the projected number of retirees at age sixty-five is 0.0962404. This is more than 5 times the probability obtained using an n_x equal to two, 0.0183497. Note, however, that even with a degenerate distribution, that is, with perfect information, the probability that the

FIGURE 4

THE PROBABILITY OF THE EXPECTED NUMBER OF PARTICIPANTS GIVEN 100 ENTRANTS AT AGE 20 AND VARIOUS VALUES OF n



Data base: 1971 GAM, rates of disability from the 1970 Valuation of the Civil Service Retirement System, and McGinn's Turnover Table III.

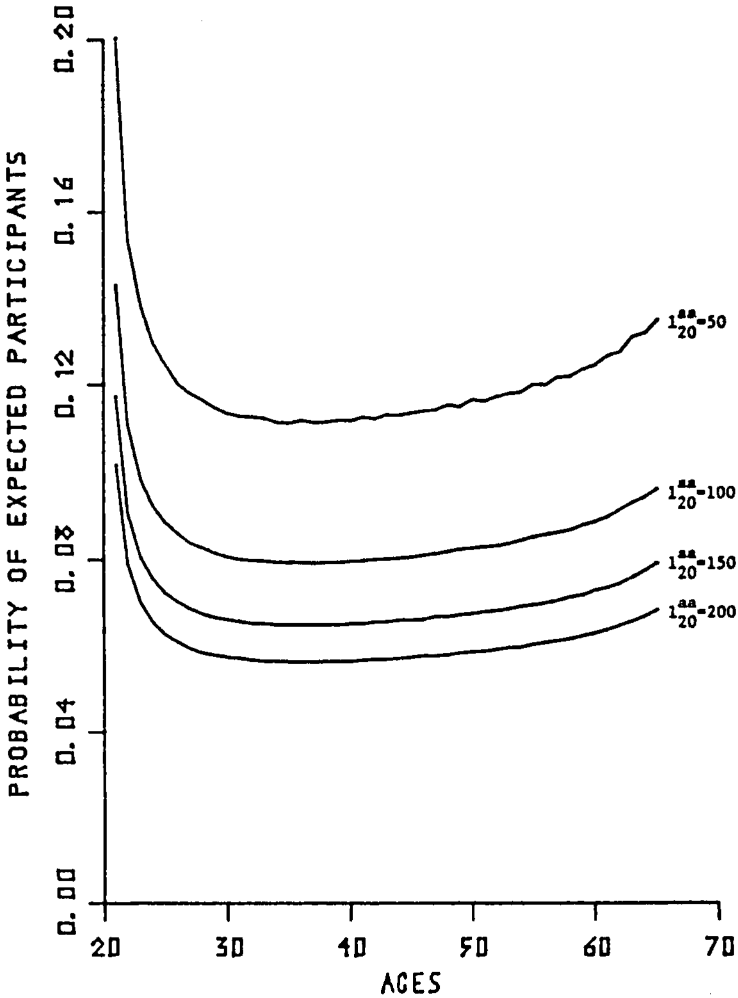
expected number of participants will obtain at any age is relatively small.

While the probability of the projected number of participants at age x varies directly as the size of n parameter of the prior distribution, it varies inversely as the size of the active population. Other things being equal, the larger the size of the population, the smaller the likelihood that the actual number of participants at a given age will equal the projected number of participants at that age. This follows directly from considering the impact on equation 2 as the number of entrants tends to zero.

Figure 5 exemplifies the foregoing observation by showing the probability that the projected number of participants obtains at each age, given various numbers of entrants at age twenty and a degenerate prior. Of particular note is the result that, for 50 entrants at age twenty, the probability that the number that retire will be equal to the projected number, 0.1348059, is approximately twice the probability that the projected number will retire given 200 entrants at age twenty, 0.0664788. Thus, although a large data base acts to increase the credibility associated with probabilities of decrement, one must not make the mistake of attributing a higher confidence to a large exposure estimate of the "expected number" of participants at a given age.

FIGURE 5

THE PROBABILITY OF THE EXPECTED NUMBER OF PARTICIPANTS, GIVEN A DEGENERATE PRIOR AND FROM 50 TO 200 ENTRANTS AT AGE 20



Data base: 1971 GAM, rates of disability from the 1970 Valuation of the Civil Service Retirement System, and McGinn's Turnover Table III.

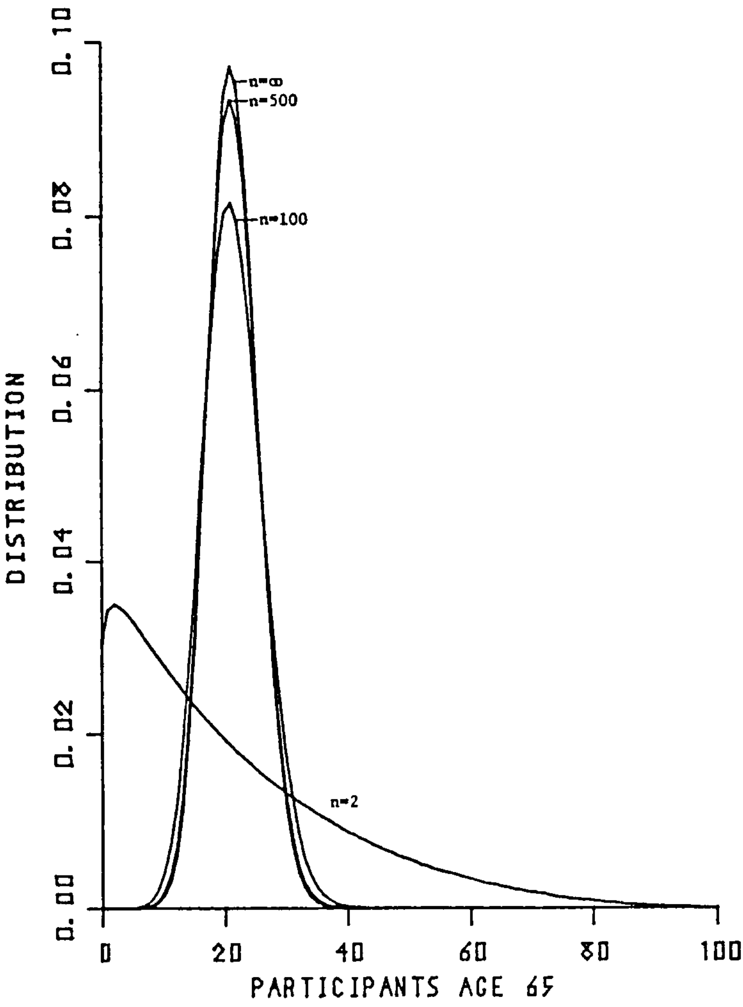
The Projected Number of Retirees

Consider now the application of the beta-binomial mass function to the problem of projecting the distribution of retirees. Figure 6 shows the distribution participants at age sixty-five resulting from 100 entrants at age twenty, given the decrement data and various precision parameters. The projected number of retirees is 21.38. It is apparent that as the probability of persisting at each age tends to degeneracy the distribution of retirees tends to its limiting distribution. It should also be noted that the less credible the prior distribution of the probability of persisting, the greater the probability that the projected number of retirees will exceed the actual number of retirees.

Under a condition of considerable uncertainty, a precision parameter equal to two for all ages, the probability that the actual number of retirements will be exceeded by the projected number is 59.88 percent. This is due to the extreme skewed nature of the distribution of retirements under a condition of high uncertainty. Attributing a high uncertainty to the estimated value of the probability of persisting is tantamount to assuming that the probabilities of decrement may be higher than their best estimates indicate. Thus, there is considerable likelihood that the actual number of retirees will be exceeded by the estimated

FIGURE 6

EFFECT OF THE CHOICE OF n ON THE DISTRIBUTION OF RETIREES AT AGE 65



Data base: 1971 GAM, rates of disability from the 1970 Valuation of the Civil Service Retirement System, and McGinn's Turnover Table III.

number of retirees. Under a condition of high certainty, on the other hand, a precision parameter approaching infinity for all ages, the probability that the actual number of retirements will be exceeded by the projected number is 52.09 percent.

Projected Retirement Costs

Turn now to the development of projected retirement costs.[34] This development proceeds in three stages. First, the concept of a select group is extended both to generalize the model and to simplify notation. Next, the probability that projected retirement costs exceed actual retirement costs is developed. Finally, a contingency charge is introduced.

Total Attribute Groups

To facilitate the development of the model it is convenient to segregate pension populations by qualification ages. To accomplish this, the types of qualification ages first are isolated. Hence, all possible entry ages are

---

[34]Ultimately, studies of the stochastic nature of pension costs will encompass all elements of such costs, including such things as vesting and early retirement. These refinements are beyond the scope of the present paper.

grouped, all possible initial vesting ages are grouped, and so on. Then, each type of qualification age is partitioned by age. For example, there may be ten different entry ages, ten different initial vesting ages, and so on. Given this classification scheme, a plan participant may be assigned to a unique group based on the ages when he or she qualifies under each plan provision. Let each such group of participants be defined as a "total attribute" group, that is, a group having all qualification ages in common, and let "C" denote the set of all total attribute groups.[35]

An example of a total attribute group would be those active participants with an entry age of 25, an initial vesting age and initial disability qualification age of 30, an early retirement age of 55, a normal retirement age of 65, and a mandatory retirement age of 70. This particular total attribute group would be denoted by the six-tuple (25,30,30,55,65,70).

----

[35]It should be mentioned that the "uniqueness" of the qualification ages is to be interpreted in a computational sense. Hence, for example, two employees whose entry age nearest birthday is twenty might both be given an entry age of twenty for computational purposes, even though they were not, in fact, the exact same age. Another computational convenience that is often used is to classify entry ages into quinquennial age groupings. Under this procedure each quinquennial age would constitute a unique entry age.

The concept of a total attribute group has seen generous implementation in the pension literature, albeit in a somewhat disguised form. Many articles, for example, have used "select" groups, where the common attribute has been the entry age. However, in most of the articles the other qualification ages were the same for each member of a select group. This meant that each select group had all qualification ages in common and was, in fact, a total attribute group.

The Probability that Projected Retirement Costs Exceed Actual Retirement Costs

The retirement cost associated with any particular total attribute group is

$$ {}^{c}\tilde{l}_r^{aa} \cdot {}^{cB}\bar{a}_r^{rr}, \; c \in C, \qquad (14) $$

where ${}^{cB}\bar{a}_r^{rr}$ represents the present value, at the age of retirement, age r, of the pension benefits. ${}^{cB}\bar{a}_r^{rr}$ is assumed to be given.[36] The probability that the projected retirement cost for this group exceeds the actual pension cost is equal to

$$ Pr\{ {}^{c}\bar{l}_r^{aa} \cdot {}^{cB}\bar{a}_r^{rr} \geq {}^{c}\tilde{l}_r^{aa} \cdot {}^{cB}\bar{a}_r^{rr} \}, \qquad (15) $$

----

[36]The retired life annuity could also could also be regarded as a random variable. See Piper, op. cit. For the purpose of this study, however, annuities are assumed to be purchased at an annuity purchase rate of $\bar{a}_r^{rr}$.

which reduces to[37]

$$\Pr\{{}^{c}\bar{l}_r^{aa} \geq {}^{c}\tilde{l}_r^{aa}\}. \quad (16)$$

In view of the integral properties of the function ${}^{c}l_r^{aa}$, this latter probability becomes[38]

$$\sum_{{}^{c}\ell_r^{aa}=0}^{\left[{}^{c}\bar{\ell}_r^{aa}\right]} \Pr\{{}^{c}l_r^{aa}\} \quad (17)$$

Most pension plans, of course, have entrants at more than one age, so it is appropriate to extend the foregoing analysis to recognize this situation.

In general, the probability that the total projected retirement cost exceeds the total actual retirement cost is

$$\Pr\{\sum_{c} {}^{c}\bar{l}_r^{aa} \cdot {}^{cB}\bar{a}_r^{aa} \geq \sum_{c} {}^{c}\tilde{l}_r^{aa} \cdot {}^{cB}\bar{a}_r^{rr}\}. \quad (18)$$

The solution to equation 18 is facilitated by defining two arrays: a retirement benefit array and a feasible retirement array. Let

$$\underset{\sim}{{}^{c}\bar{a}_r^{rr}} = ({}^{cB}\bar{a}_r^{rr} \mid c \in C)' \quad (19)$$

be defined as the retirement benefit array associated with

---

[37]Note that for a given total attribute group, the probability that the expected pension cost will be adequate is independent of the benefit function defined by the plan.

[38]The function $[m]$ represents the largest integer in m.

the pension plan under consideration, that is, the array whose elements are the present value of the retirement benefits, at retirement, associated with each total attribute group. Additionally, let

$$ {}^{n}\underset{\sim}{l}_{r}^{aa} = ({}^{nc}l_{r}^{aa} \mid c \in C)' \quad (20) $$

be defined as a feasible retirement array, that is, an array whose elements are composed of possible numbers of participants who reach their normal retirement age from each total attribute group, and which satisfy the condition

$$ ({}^{n}\underset{\sim}{l}_{r}^{aa})^{T} \; {}^{B}\underset{\sim}{\bar{a}}_{r}^{rr} \leq \sum_{c} {}^{c}\bar{l}_{r}^{aa} \cdot {}^{cB}\bar{a}_{r}^{rr} . \quad (21) $$

Assuming that there are N distinct feasible retirement arrays, it follows that the probability that the total expected retirement costs will exceed the total actual retirement cost is

$$ \sum_{n=1}^{N} \Pr\{{}^{n}\underset{\sim}{l}_{r}^{aa}\} . \quad (22) $$

However, since the probability of a given feasible retirement array is simply the product of the probabilities of the joint occurrence of each element of the array, the probability that the total projected cost will exceed the total actual cost becomes

$$ \sum_{n=1}^{N} \prod_{c} \Pr\{{}^{nc}l_{r}^{aa}\} . \quad (23) $$

The Contingency Charge

The determination of the contingency charge needed to increase the probability of adequate funds to a given level follows immediately from the foregoing analysis. The only change being that instead of defining a feasible array in terms of projected cost, one would define a contingent feasible array in terms of some multiple of the projected cost. For example, one might define a contingent feasible retirement array as a feasible retirement array which satisfies the condition

$$({}^{n}\underset{\sim}{1}_{r}^{aa})^{T}\ {}^{B}\underset{\sim}{\bar{a}}_{r}^{rr} \leqslant (1+m) \sum_{c} {}^{c}\bar{1}_{r}^{aa} \cdot {}^{cB}\bar{a}_{r}^{rr}, \qquad (24)$$

where the factor (1+m) defines the multiple of the projected cost which is to be funded, and where the product of m and the projected cost represents the contingency charge.

In practice the factor (1+m) would be determined such that the probability of adequate funds obtains some desirable level.

Working Formulae for the Determination of the Probable Adequacy of Projected Retirement Costs

Given that the number of participants at a given age has a specified distribution, it is a simple matter to set down a working formula for the probable adequacy of the projected retirement cost. For a specific total attribute group, the probability that the projected retirement cost exceeds the actual retirement cost is

$$\sum_{{}^{c}\ell_r^{aa}=0}^{\left[{}^{c}\bar{\ell}_r^{aa}\right]} \sum_{k=1}^{K} \prod_{t=0}^{x-a-1} f({}^{k}l_{r-t}^{aa} \mid {}^{c}l_r^{aa}), \quad (25)$$

obtained by substituting equation 2 into equation 17. To incorporate a beta-binomial distribution, $f({}^{k}l_{r-t}^{aa} \mid {}^{c}l_r^{aa})$ is replaced by equation 7. On the other hand, the probability that the total projected retirement cost exceeds the total actual retirement cost is

$$\sum_{n=1}^{N} \prod_{c} \sum_{k=1}^{K(c)} \prod_{t=0}^{{}^{c}r-{}^{c}a-1} f({}^{k}l_{r-t}^{aa} \mid {}^{nc}l_r^{aa}), \quad (26)$$

obtained by substituting equation 2 into equation 23. A working formula for the probable adequacy of some multiple of the projected cost is similarly defined.

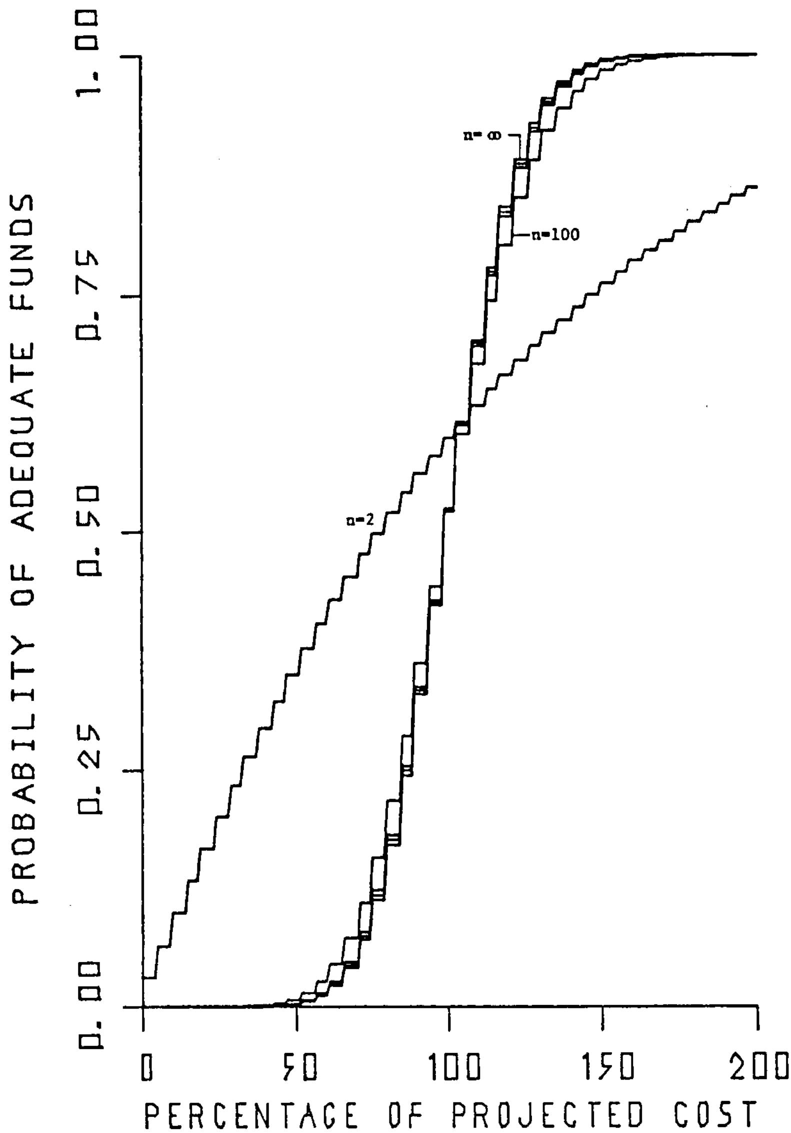
The final section of this paper deals with the application of these formulae.

Estimating the Adequacy of Projected Pension Costs

Turning now to the development of estimates of the adequacy of projected retirement costs, consider first the probability that the projected cost will exceed the actual cost, given a specific total attribute group. Figure 7

FIGURE 7

THE PROBABILITY OF ADEQUATE FUNDS -- SINGLE ENTRY AGE CASE



Data base: 1971 GAM, rates of disability from the 1970 Valuation of the Civil Service Retirement System, and McGinn's Turnover Table III.

shows this probability for 100 entrants at age twenty and various degrees of confidence in the decrement data. Two observations are apparent. First, if the accumulated funds are less than the projected cost, the greater the variability assumed for the probability of persisting, the greater the probability that the projected funds will be adequate. The opposite is true if the funds are greater than the projected cost. Once again, this is caused by the skewness of the beta-binomial distribution under a condition of uncertainty.

As a second example of the implementation of the beta binomial distribution, consider the determination of the contingency charge for a plan as a whole. For the purpose of example, assume that entry takes place quinquennially from ages twenty through forty, inclusive, with the proportion of entrants at each age being .28, .24, .18, .12, and .08, respectively. Assume, also, that the total number of entrants is chosen so that, if entry were to take place annually, an ultimate population of approximately two thousand employees would result. In addition, the benefit function is based on two percent of final salary for each year of service, using Salary Scale 3 of the <u>Actuary's</u> <u>Pension</u>

Handbook.

Figure 8 shows the results of this analysis. Once again, the probability that the projected cost will be adequate is greatest under a condition of high uncertainty regarding the probability of persisting. This is shown by the curve labeled "n=2." However, as a contingency charge is added, its impact is directly proportional to the confidence in the prior distribution of the persistency rate. The greater the degeneracy of the prior distribution the smaller the contingency charge needed to obtain a given probability of adequate funds. This is shown by the curve labeled "n=∞."

Note that, on the basis of the decrement data, even with perfect information a contingency charge of 40 percent of projected cost would be required to attain a 99 percent probability of adequate funds.[39] Note, also, under a condition of high uncertainty a contingency charge of 100 percent

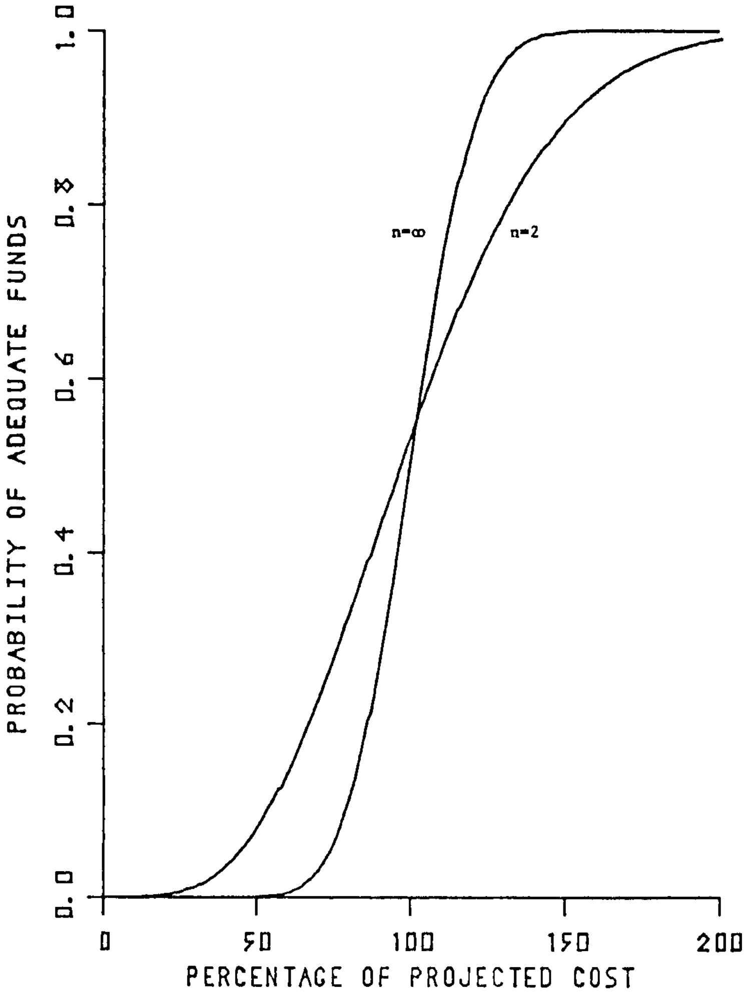
---

[39]It might be argued that a ninety-nine percent confidence interval is too high to be realistic. However, a comment by Piper in his discussion of Menge, *op. cit.*, p. 609, bears repeating. Piper, in discussing a 99.9 percent confidence interval, observed that

> It would be possible for fluctuations during the early years of observation to exhaust the contingency reserve and compel borrowing from some undefined R.F.C. which is assumed to be ready to lend its funds at 4 percent interest. A satisfactory

FIGURE 8

THE PROBABILITY OF ADEQUATE FUNDS
-- MULTIPLE ENTRY AGE CASE



Data base: 1971 GAM, rates of disability from the 1970 Valuation of the Civil Service Retirement System, McGinn's Turnover Table III, and Salary Scale S-3 of the Actuary's Pension Handbook.

of projected cost would be needed to raise the probability of adequate funds to 99 percent.

Comment

It is hoped that the model which was developed in this paper will be instrumental in stimulating both theoretical and empirical research into the stochastic nature of pension costs. As regards the former, there are many refinements that might be incorporated into the model, including such things as a stochastic accumulation of funds and a stochastic retirement annuity. As regards the latter, although this paper did attempt to obtain certain specific results, these results were intended primarily as examples of the implementation of the model, and, as such, were far from exhaustive. Future researchers should find the empirical study of the stochastic nature of pension costs a fruitful area for exploration, particularly if they have at their disposal an accommodating computer facility.

---

answer in statistical terms to the retention question would require, it seems to me, a calculation of the contingency reserve, which would at no time exhaust the available funds.

It is to be hoped that a further investigation can be made along this line, for a rational solution to the problem of retention limits is extremely important to companies of small to medium size.