Actuarial Mathematics and Life-Table Statistics

Eric V. Slud
Mathematics Department
University of Maryland, College Park

Chapter 2

Theory of Interest and Force of Mortality

The parallel development of Interest and Probability Theory topics continues in this Chapter. For application in Insurance, we are preparing to value uncertain payment streams in which times of payment may also be uncertain. The interest theory allows us to express the present values of *certain* payment streams compactly, while the probability material prepares us to find and interpret average or expected values of present values expressed as functions of random lifetime variables.

This installment of the course covers: (a) further formulas and topics in the pure (i.e., non-probabilistic) theory of interest, and (b) more discussion of lifetime random variables, in particular of *force of mortality* or hazard-rates, and theoretical families of life distributions.

2.1 More on Theory of Interest

The objective of this subsection is to define notations and to find compact formulas for present values of some standard payment streams. To this end, newly defined payment streams are systematically expressed in terms of previously considered ones. There are two primary methods of manipulating one payment-stream to give another for the convenient calculation of present values:

- First, if one payment-stream can be obtained from a second one precisely by delaying all payments by the same amount t of time, then the present value of the first one is v^t multiplied by the present value of the second.
- Second, if one payment-stream can be obtained as the superposition of two other payment streams, i.e., can be obtained by paying the total amounts at times indicated by either of the latter two streams, then the present value of the first stream is the sum of the present values of the other two.

The following subsection contains several useful applications of these methods. For another simple illustration, see Worked Example 2 at the end of the Chapter.

2.1.1 Annuities & Actuarial Notation

The general present value formulas above will now be specialized to the case of constant (instantaneous) interest rate $\delta(t) \equiv \ln(1+i) = \delta$ at all times $t \geq 0$, and some very particular streams of payments s_j at times t_j, related to periodic premium and annuity payments. The effective interest rate or APR is always denoted by i, and as before the m-times-per-year equivalent nominal interest rate is denoted by $i^{(m)}$. Also, from now on the standard and convenient notation

$$v \;\equiv\; 1/(1+i) \;=\; 1/\left(1+\frac{i^{(m)}}{m}\right)^m$$

will be used for the present value of a payment of $1 in one year.

(i) If $s_0 = 0$ and $s_1 = \cdots = s_{nm} = 1/m$ in the discrete setting, where m denotes the number of payments per year, and $t_j = j/m$, then the payment-stream is called an **immediate annuity**, and its present value G_n is given the notation $a^{(m)}_{\overline{n}|}$ and is equal, by the geometric-series summation formula, to

$$m^{-1}\sum_{j=1}^{nm}\left(1+\frac{i^{(m)}}{m}\right)^{-j} = \frac{1-(1+i^{(m)}/m)^{-nm}}{m(1+i^{(m)}/m-1)} = \frac{1}{i^{(m)}}\left(1-\left(1+\frac{i^{(m)}}{m}\right)^{-nm}\right)$$

This calculation has shown

$$a^{(m)}_{\overline{n}|} \;=\; \frac{1-v^n}{i^{(m)}} \tag{2.1}$$

All of these immediate annuity values, for fixed v, n but varying m, are roughly comparable because all involve a total payment of 1 per year. Formula (2.1) shows that all of the values $a^{(m)}_{\overline{n}|}$ differ only through the factors $i^{(m)}$, which differ by only a few percent for varying m and fixed i, as shown in Table 2.1. Recall from formula (1.4) that $i^{(m)} = m\{(1+i)^{1/m} - 1\}$.

If instead $s_0 = 1/m$ but $s_{nm} = 0$, then the notation changes to $\ddot{a}^{(m)}_{\overline{n}|}$, the payment-stream is called an **annuity-due**, and the value is given by any of the equivalent formulas

$$\ddot{a}^{(m)}_{\overline{n}|} \;=\; (1+\frac{i^{(m)}}{m})\,a^{(m)}_{\overline{n}|} \;=\; \frac{1-v^n}{m} \,+\, a^{(m)}_{\overline{n}|} \;=\; \frac{1}{m} + a^{(m)}_{\overline{n-1/m}|} \tag{2.2}$$

The first of these formulas recognizes the annuity-due payment-stream as identical to the annuity-immediate payment-stream shifted earlier by the time $1/m$ and therefore worth more by the accumulation-factor $(1+i)^{1/m} = 1+i^{(m)}/m$. The third expression in (2.2) represents the annuity-due stream as being equal to the annuity-immediate stream with the payment of $1/m$ at $t=0$ added and the payment of $1/m$ at $t=n$ removed. The final expression says that if the time-0 payment is removed from the annuity-due, the remaining stream coincides with the annuity-immediate stream consisting of $nm-1$ (instead of nm) payments of $1/m$.

In the limit as $m \to \infty$ for fixed n, the notation $\overline{a}_{\overline{n}|}$ denotes the present value of an annuity paid instantaneously at constant unit rate, with the limiting nominal interest-rate which was shown at the end of the previous chapter to be $\lim_m \, i^{(m)} = i^{(\infty)} = \delta$. The limiting behavior of the nominal interest rate can be seen rapidly from the formula

$$i^{(m)} \;=\; m\left((1+i)^{1/m} \,-\, 1\right) \;=\; \delta \,\cdot\, \frac{\exp(\delta/m)-1}{\delta/m}$$

since $(e^z - 1)/z$ converges to 1 as $z \to 0$. Then by (2.1) and (2.2),

$$\overline{a}_{\overline{n}|} \;=\; \lim_{m\to\infty} \ddot{a}^{(m)}_{\overline{n}|} \;=\; \lim_{m\to\infty} a^{(m)}_{\overline{n}|} \;=\; \frac{1-v^n}{\delta} \tag{2.3}$$

Table 2.1: Values of nominal interest rates $i^{(m)}$ (upper number) and $d^{(m)}$ (lower number), for various choices of effective annual interest rate i and number m of compounding periods per year.

$i =$	.02	.03	.05	.07	.10	.15
$m = 2$	.0199	.0298	.0494	.0688	.0976	.145
	.0197	.0293	.0482	.0665	.0931	.135
3	.0199	.0297	.0492	.0684	.0968	.143
	.0197	.0294	.0484	.0669	.0938	.137
4	.0199	.0297	.0491	.0682	.0965	.142
	.0198	.0294	.0485	.0671	.0942	.137
6	.0198	.0296	.0490	.0680	.0961	.141
	.0198	.0295	.0486	.0673	.0946	.138
12	.0198	.0296	.0489	.0678	.0957	.141
	.0198	.0295	.0487	.0675	.0949	.139

A handy formula for annuity-due present values follows easily by recalling that

$$1 - \frac{d^{(m)}}{m} = \Big(1 + \frac{i^{(m)}}{m}\Big)^{-1} \qquad \text{implies} \qquad d^{(m)} = \frac{i^{(m)}}{1 + i^{(m)}/m}$$

Then, by (2.2) and (2.1),

$$\ddot{a}^{(m)}_{\overline{n}|} = (1 - v^n) \cdot \frac{1 + i^{(m)}/m}{i^{(m)}} = \frac{1 - v^n}{d^{(m)}} \tag{2.4}$$

In case m is 1, the superscript $^{(m)}$ is omitted from all of the annuity notations. In the limit where $n \to \infty$, the notations become $a^{(m)}_{\overline{\infty}|}$ and $\ddot{a}^{(m)}_{\overline{\infty}|}$, and the annuities are called **perpetuities** (respectively immediate and due) with present-value formulas obtained from (2.1) and (2.4) as:

$$a^{(m)}_{\overline{\infty}|} = \frac{1}{i^{(m)}} \quad , \quad \ddot{a}^{(m)}_{\overline{\infty}|} = \frac{1}{d^{(m)}} \tag{2.5}$$

Let us now build some more general annuity-related present values out of the standard functions $a^{(m)}_{\overline{n}|}$ and $\ddot{a}^{(m)}_{\overline{n}|}$.

(ii). Consider first the case of the **increasing perpetual annuity-due**, denoted $(I^{(m)}\ddot{a})^{(m)}_{\overline{\infty}|}$, which is defined as the present value of a stream of payments $(k+1)/m^2$ at times k/m, for $k = 0, 1, \ldots$ forever. Clearly the present value is

$$(I^{(m)}\ddot{a})^{(m)}_{\overline{\infty}|} \;=\; \sum_{k=0}^{\infty} m^{-2}\,(k+1)\left(1 + \frac{i^{(m)}}{m}\right)^{-k}$$

Here are two methods to sum this series, the first purely mathematical, the second with actuarial intuition. First, without worrying about the strict justification for differentiating an infinite series term-by-term,

$$\sum_{k=0}^{\infty} (k+1)\,x^k \;=\; \frac{d}{dx}\sum_{k=0}^{\infty} x^{k+1} \;=\; \frac{d}{dx}\,\frac{x}{1-x} \;=\; (1-x)^{-2}$$

for $0 < x < 1$, where the geometric-series formula has been used to sum the second expression. Therefore, with $x = (1 + i^{(m)}/m)^{-1}$ and $1 - x = (i^{(m)}/m)/(1 + i^{(m)}/m)$,

$$(I^{(m)}\ddot{a})^{(m)}_{\overline{\infty}|} = m^{-2}\Big(\frac{i^{(m)}/m}{1 + i^{(m)}/m}\Big)^{-2} \;=\; \Big(\frac{1}{d^{(m)}}\Big)^2 \;=\; \Big(\ddot{a}^{(m)}_{\overline{\infty}|}\Big)^2$$

and (2.5) has been used in the last step. Another way to reach the same result is to recognize the increasing perpetual annuity-due as $1/m$ multiplied by the superposition of perpetuities-due $\ddot{a}^{(m)}_{\overline{\infty}|}$ paid at times $0,\ 1/m,\ 2/m,\ \ldots$, and therefore its present value must be $\ddot{a}^{(m)}_{\overline{\infty}|} \cdot \ddot{a}^{(m)}_{\overline{\infty}|}$. As an aid in recognizing this equivalence, consider each annuity-due $\ddot{a}^{(m)}_{\overline{\infty}|}$ paid at a time j/m as being equivalent to a stream of payments $1/m$ at time j/m, $1/m$ at $(j+1)/m$, etc. Putting together all of these payment streams gives a total of $(k+1)/m$ paid at time k/m, of which $1/m$ comes from the annuity-due starting at time 0, $1/m$ from the annuity-due starting at time $1/m$, up to the payment of $1/m$ from the annuity-due starting at time k/m.

(iii). The **increasing perpetual annuity-immediate** $(I^{(m)}a)^{(m)}_{\overline{\infty}|}$ — the same payment stream as in the increasing annuity-due, but deferred by a time $1/m$ — is related to the perpetual annuity-due in the obvious way

$$(I^{(m)}a)^{(m)}_{\overline{\infty}|} \;=\; v^{1/m}\,(I^{(m)}\ddot{a})^{(m)}_{\overline{\infty}|} \;=\; (I^{(m)}\ddot{a})^{(m)}_{\overline{\infty}|}\Big/(1 + i^{(m)}/m) \;=\; \frac{1}{i^{(m)}\,d^{(m)}}$$

(iv). Now consider the **increasing annuity-due of finite duration** n years. This is the present value $(I^{(m)}\ddot{a})^{(m)}_{\overline{n}|}$ of the payment-stream of $(k+1)/m^2$ at time k/m, for $k = 0, \ldots, nm-1$. Evidently, this payment-stream is equivalent to $(I^{(m)}\ddot{a})^{(m)}_{\overline{\infty}|}$ minus the sum of n multiplied by an annuity-due $\ddot{a}^{(m)}_{\overline{\infty}|}$ starting at time n together with an increasing annuity-due $(I^{(m)}\ddot{a})^{(m)}_{\overline{\infty}|}$ starting at time n. (To see this clearly, equate the payments $0 = (k+1)/m^2 - n\cdot\frac{1}{m} - (k-nm+1)/m^2$ received at times k/m for $k \geq nm$.) Thus

$$\begin{aligned}
(I^{(m)}\ddot{a})^{(m)}_{\overline{n}|} &= (I^{(m)}\ddot{a})^{(m)}_{\overline{\infty}|}\left(1-(1+i^{(m)}/m)^{-nm}\right) - n\ddot{a}^{(m)}_{\overline{\infty}|}(1+i^{(m)}/m)^{-nm} \\
&= \ddot{a}^{(m)}_{\overline{\infty}|}\left(\ddot{a}^{(m)}_{\overline{\infty}|} - (1+i^{(m)}/m)^{-nm}\left[\ddot{a}^{(m)}_{\overline{\infty}|} + n\right]\right) \\
&= \ddot{a}^{(m)}_{\overline{\infty}|}\left(\ddot{a}^{(m)}_{\overline{n}|} - n\,v^n\right)
\end{aligned}$$

where in the last line recall that $v = (1+i)^{-1} = (1+i^{(m)}/m)^{-m}$ and that $\ddot{a}^{(m)}_{\overline{n}|} = \ddot{a}^{(m)}_{\overline{\infty}|}(1-v^n)$. The latter identity is easy to justify either by the formulas (2.4) and (2.5) or by regarding the annuity-due payment stream as a superposition of the payment-stream up to time $n-1/m$ and the payment-stream starting at time n. As an exercise, fill in details of a second, intuitive verification, analogous to the second verification in pargraph (ii) above.

(v). The **decreasing annuity** $(D^{(m)}\ddot{a})^{(m)}_{\overline{n}|}$ is defined as (the present value of) a stream of payments starting with n/m at time 0 and decreasing by $1/m^2$ every time-period of $1/m$, with no further payments at or after time n. The easiest way to obtain the present value is through the identity

$$(I^{(m)}\ddot{a})^{(m)}_{\overline{n}|} + (D^{(m)}\ddot{a})^{(m)}_{\overline{n}|} = (n + \frac{1}{m})\,\ddot{a}^{(m)}_{\overline{n}|}$$

Again, as usual, the method of proving this is to observe that in the payment-stream whose present value is given on the left-hand side, the payment amount at each of the times j/m, for $j = 0, 1, \ldots, nm-1$, is

$$\frac{j+1}{m^2} + (\frac{n}{m} - \frac{j}{m^2}) = \frac{1}{m}(n + \frac{1}{m})$$

2.1.2 Loan Amortization & Mortgage Refinancing

The only remaining theory-of-interest topic to cover in this unit is the breakdown between principal and interest payments in repaying a loan such as a mortgage. Recall that the present value of a payment stream of amount c per year, with c/m paid at times $1/m, 2/m, \ldots, n-1/m, n/m$, is $c\, a^{(m)}_{\overline{n}|}$. Thus, if an amount *Loan-Amt* has been borrowed for a term of n years, to be repaid by equal installments at the end of every period $1/m$, at fixed nominal interest rate $i^{(m)}$, then the installment amount is

$$\text{Mortgage Payment} \;=\; \frac{\textit{Loan-Amt}}{m\, a^{(m)}_{\overline{n}|}} \;=\; \textit{Loan-Amt}\; \frac{i^{(m)}}{m\,(1-v^n)}$$

where $v = 1/(1+i) = (1+i^{(m)}/m)^{-m}$. Of the payment made at time $(k+1)/m$, how much can be attributed to interest and how much to principal ? Consider the present value at 0 of the debt per unit of *Loan-Amt* less accumulated amounts paid up to and including time k/m :

$$1 - m\, a^{(m)}_{\overline{k/m}|}\, \frac{1}{m\, a^{(m)}_{\overline{n}|}} \;=\; 1 - \frac{1-v^{k/m}}{1-v^n} \;=\; \frac{v^{k/m}-v^n}{1-v^n}$$

The remaining debt, per unit of *Loan-Amt*, valued just after time k/m, is denoted from now on by $B_{n,\,k/m}$. It is greater than the displayed present value at 0 by a factor $(1+i)^{k/m}$, so is equal to

$$B_{n,\,k/m} \;=\; (1+i)^{k/m}\, \frac{v^{k/m}-v^n}{1-v^n} = \frac{1-v^{n-k/m}}{1-v^n} \tag{2.6}$$

The amount of interest for a Loan Amount of 1 after time $1/m$ is $(1+i)^{1/m} - 1 \;=\; i^{(m)}/m$. Therefore the interest included in the payment at $(k+1)/m$ is $i^{(m)}/m$ multiplied by the value $B_{n,\,k/m}$ of outstanding debt just after k/m. Thus the next total payment of $i^{(m)}/(m(1-v^n))$ consists of the two parts

$$\text{Amount of interest} \;=\; m^{-1}\, i^{(m)}\,(1-v^{n-k/m})/(1-v^n)$$

$$\text{Amount of principal} \;=\; m^{-1} i^{(m)} v^{n-k/m}/(1-v^n)$$

By definition, the principal included in each payment is the amount of the payment minus the interest included in it. These formulas show in particular that the amount of principal repaid in each successive payment increases geometrically in the payment number, which at first seems surprising. Note as a check on the displayed formulas that the outstanding balance $B_{n,(k+1)/m}$ immediately after time $(k+1)/m$ is re-computed as $B_{n,k/m}$ minus the interest paid at $(k+1)/m$, or

$$\frac{1-v^{n-k/m}}{1-v^n} - \frac{i^{(m)}}{m}\frac{v^{n-k/m}}{1-v^n} = \frac{1-v^{n-k/m}(1+i^{(m)}/m)}{1-v^n}$$

$$= \frac{1-v^{n-(k+1)/m}}{1-v^n} = \left(1-\frac{a^{(m)}_{\overline{(k+1)/m}|}}{a^{(m)}_{\overline{n}|}}\right) v^{-(k+1)/m} \tag{2.7}$$

as was derived above by considering the accumulated value of amounts paid. The general definition of the principal repaid in each payment is the excess of the payment over the interest since the past payment on the total balance due immediately following that previous payment.

2.1.3 Illustration on Mortgage Refinancing

Suppose that a 30–year, nominal-rate 8%, \$100,000 mortgage payable monthly is to be refinanced at the end of 8 years for an additional 15 years (instead of the 22 which would otherwise have been remaining to pay it off) at 6%, with a refinancing closing-cost amount of \$1500 and *2 points*. (The points are each 1% of the refinanced balance including closing costs, and costs plus points are then extra amounts added to the initial balance of the refinanced mortgage.) Suppose that the **new** pattern of payments is to be valued at each of the nominal interest rates 6%, 7%, or 8%, due to uncertainty about what the interest rate will be in the future, and that these valuations will be taken into account in deciding whether to take out the new loan.

The monthly payment amount of the initial loan in this example was $\$100,000(.08/12)/(1-(1+.08/12)^{-360}) = \$733.76$, and the present value as of time 0 (the beginning of the old loan) of the payments made through the end of the 8th year is $(\$733.76)\cdot(12a^{(12)}_{\overline{8}|}) = \$51,904.69$. Thus the present value, *as of the end of 8 years*, of the payments still to be made under the *old* mortgage, is $\$(100,000-51,904.69)(1+.08/12)^{96} = \$91,018.31$. Thus, if the loan were to be refinanced, the new refinanced loan amount would be

$\$91,018.31 + 1,500.00 = \$92,518.31$. If *2 points* must be paid in order to lock in the rate of 6% for the refinanced 15-year loan, then this amount is $(.02)92518.31 = \$1850.37$. The new principal balance of the refinanced loan is $92518.31 + 1850.37 = \$94,368.68$, and this is the present value at a nominal rate of 6% of the future loan payments, no matter what the term of the refinanced loan is. The new monthly payment (for a 15-year duration) of the refinanced loan is $\$94,368.68(.06/12)/(1 - (1 + .06/12)^{-180}) = \$796.34$.

For purposes of comparison, what is the present value at the current going rate of 6% (nominal) of the continuing stream of payments under the old loan ? That is a 22-year stream of monthly payments of \$733.76, as calculated above, so the present value at 6% is $\$733.76 \cdot (12a^{(12)}_{\overline{22}|}) = \$107,420.21$. Thus, if the new rate of 6% were really to be the correct one for the next 22 years, and each loan would be paid to the end of its term, then it would be a financial disaster *not* to refinance. Next, suppose instead that right after re-financing, the economic rate of interest would be a nominal 7% for the next 22 years. In that case *both* streams of payments would have to be re-valued — the one before refinancing, continuing another 22 years into the future, and the one after refinancing, continuing 15 years into the future. The respective present values (as of the end of the 8^{th} year) at nominal rate of 7% of these two streams are:

Old loan: $733.76\,(12a^{(12)}_{\overline{22}|}) = \$98,700.06$

New loan: $796.34\,(12a^{(12)}_{\overline{15}|}) = \$88,597.57$

Even with these different assumptions, and despite closing-costs and points, it is well worth re-financing.

Exercise: Suppose that you can forecast that you will in fact sell your house in precisely 5 more years after the time when you are re-financing. At the time of sale, you would pay off the cash principal balance, whatever it is. Calculate and compare the present values (at each of 6%, 7%, and 8% nominal interest rates) of your payment streams to the bank, (a) if you continue the old loan without refinancing, and (b) if you re-finance to get a 15-year 6% loan including closing costs and points, as described above.

2.1.4 Computational illustration in Splus or R

All of the calculations described above are very easy to program in any language from Fortran to Mathematica, and also on a programmable calculator; but they are also very handily organized within a spreadsheet, which seems to be the way that MBA's, bank-officials, and actuaries will learn to do them from now on.

In this section, an **Splus** or **R** function (*cf.* Venables & Ripley 2002) is provided to do some comparative refinancing calculations. Concerning the syntax of **Splus** or **R**, the only explanation necessary at this point is that the symbol $<-$ denotes assignment of an expression to a variable: $A <\!\!- B$ means that the variable A is assigned the value of expression B. Other syntactic elements used here are common to many other computer languages: * denotes multiplication, and ^ denotes exponentiation.

The function *RefExmp* given below calculates mortgage payments, balances for purposes of refinancing both before and after application of administrative costs and points, and the present value under any interest rate (not necessarily the ones at which either the original or refinanced loans are taken out) of the stream of repayments to the bank up to and including the lump-sum payoff which would be made, for example, at the time of selling the house on which the mortgage loan was negotiated. The output of the function is a list which, in each numerical example below, is displayed in 'unlisted' form, horizontally as a vector. Lines beginning with the symbol # are comment-lines.

The outputs of the function are as follows. *Oldpayment* is the monthly payment on the original loan of face-amount *Loan* at nominal interest $i^{(12)} =$ *OldInt* for a term of *OldTerm* years. *NewBal* is the balance $B_{n,\,k/m}$ of formula (2.6) for $n =$ *OldTerm*, $m = 12$, and $k/m =$ *RefTim*, and the refinanced loan amount is a multiple $1+$ *Points* of *NewBal*, which is equal to *RefBal* + *Costs*. The new loan, at nominal interest rate *NewInt*, has monthly payments *Newpaymt* for a term of *NewTerm* years. The loan is to be paid off *PayoffTim* years after *RefTim* when the new loan commences, and the final output of the function is the present value at the start of the refinanced loan with nominal interest rate *ValInt* of the stream of payments made under the refinanced loan up to and including the lump sum payoff.

```
 Splus or R FUNCTION CALCULATING REFINANCE PAYMENTS & VALUES

RefExmp
function(Loan, OldTerm, RefTim, NewTerm, Costs, Points,
       PayoffTim, OldInt, NewInt, ValInt)
{
# Function calculates present value of future payment stream
#      underrefinanced loan.
# Loan = original loan amount;
# OldTerm = term of initial loan in years;
# RefTim = time in years after which to refinance;
# NewTerm = term of refinanced loan;
# Costs = fixed closing costs for refinancing;
# Points = fraction of new balance as additional costs;
# PayoffTim (no bigger than NewTerm) = time (from refinancing-
#      time at which new loan balance is to be paid off in
#      cash (eg at house sale);
# The three interest rates OldInt, NewInt, ValInt are
#      nominal 12-times-per-year, and monthly payments
#      are calculated.
   vold <- (1 + OldInt/12)^(-12)
   Oldpaymt <- ((Loan * OldInt)/12)/(1 - vold^OldTerm)
   NewBal <- (Loan * (1 - vold^(OldTerm - RefTim)))/
       (1 - vold^OldTerm)
   RefBal <- (NewBal + Costs) * (1 + Points)
   vnew <- (1 + NewInt/12)^(-12)
   Newpaymt <- ((RefBal * NewInt)/12)/(1 - vnew^NewTerm)
   vval <- (1 + ValInt/12)^(-12)
   Value <- (Newpaymt * 12 * (1 - vval^PayoffTim))/ValInt +
        (RefBal *  vval^PayoffTim * (1 - vnew^(NewTerm -
         PayoffTim)))/(1 -  vnew^NewTerm)
   list(Oldpaymt = Oldpaymt, NewBal = NewBal,
       RefBal = RefBal, Newpaymt = Newpaymt, Value = Value)
}
```

We begin our illustration by reproducing the quantities calculated in the previous subsection:

```
> unlist(RefExmp(100000, 30, 8, 15, 1500, 0.02, 15,
         0.08, 0.06, 0.06))
 Oldpaymt NewBal RefBal Newpaymt Value
   733.76  91018  94368   796.33 94368
```

Note that, since the payments under the new (refinanced) loan are here valued at the same interest rate as the loan itself, the present value Value of all payments made under the loan must be equal to the the refinanced loan amount *RefBal*.

The comparisons of the previous Section between the original and refinanced loans, at (nominal) interest rates of 6, 7, and 8 %, are all recapitulated easily using this function. To use it, for example, in valuing the old loan at 7%, the arguments must reflect a 'refinance' with no costs or points for a period of 22 years at nominal rate 6%, as follows:

```
> unlist(RefExmp(100000,30,8,22,0,0,22,0.08,0.08,0.07))
 Oldpaymt  NewBal RefBal Newpaymt  Value
   733.76  91018  91018   733.76   98701
```

(The small discrepancies between the values found here and in the previous subsection are due to the rounding used there to express payment amounts to the nearest cent.)

We consider next a numerical example showing break-even point for refinancing by balancing costs versus time needed to amortize them.

Suppose that you have a 30-year mortage for \$100,000 at nominal 9% ($= i^{(12)}$), with level monthly payments, and that after 7 years of payments you refinance to obtain a new 30-year mortgage at 7% nominal interest ($= i^{(m)}$ for $m = 12$), with closing costs of \$1500 and 4 points (i.e., 4% of the total refinanced amount including closing costs added to the initial balance), also with level monthly payments. Figuring present values using the new interest rate of 7%, what is the time K (to the nearest month) such that if both loans — the old and the new — were to be paid off in exactly K years after the time (the 7-year mark for the first loan) when you would have refinanced, *then the remaining payment-streams for both loans from the time when you refinance are equivalent (i.e., have the same present value from that time) ?*

We first calculate the present value of payments under the new loan. As remarked above in the previous example, since the same interest rate is being used to value the payments as is used in figuring the refinanced loan, the valuation of the new loan *does not depend upon the time* K *to payoff.* (It is figured here as though the payoff time K were 10 years.)

```
> unlist(RefExmp(1.e5, 30,7,30, 1500,.04, 10, 0.09,0.07,0.07))
 Oldpaymt NewBal RefBal Newpaymt Value
   804.62  93640  98946   658.29 98946
```

Next we compute the value of payments under the old loan, at 7% nominal rate, also at payoff time $K = 10$. For comparison, the value under the old loan for payoff time 0 (i.e., for cash payoff at the time when refinancing would have occurred) coincides with the New Balance amount of $93640.

```
> unlist(RefExmp(1.e5, 30,7,23, 0,0, 10, 0.09,0.09,0.07))
 Oldpaymt NewBal RefBal Newpaymt  Value
   804.62  93640  93640   804.62 106042
```

The values found in the same way when the payoff time K is successively replaced by 4, 3, 3.167, 3.25 are 99979, 98946, 98593, 98951. Thus, the payoff-time K at which there is essentially no difference in present value at nominal 7% between the old loan or the refinanced loan with costs and points (which was found to have Value 98946), is 3 years and 3 months after refinancing.

2.1.5 Coupon & Zero-coupon Bonds

In finance, an investor assessing the present value of a *bond* is in the same situation as the bank receiving periodic level payments in repayment of a loan. If the payments are made every $1/m$ year, with nominal *coupon interest rate* $i^{(m)}$, for a bond with face value $1000, then the payments are precisely the interest on $1000 for $1/m$ year, or $1000 \cdot i^{(m)}/m$. For most corporate or government bonds, $m = 4$, although some bonds have $m = 2$. If the bond is *uncallable*, which is assumed throughout this discussion, then it entitles the holder to receive the stream of such payments every $1/m$ year until a fixed final *redemption date*, at which the final interest payment coincides with the repayment of the principal of \$1000. Suppose that the time remaining on a bond until redemption is R (assumed to be a whole-number multiple of $1/m$ years), and that the nominal annualized m-period-per-year interest rate, taking into account the credit-worthiness of the bond issuer together with current economic conditions, is $r^{(m)}$ which will typically not be equal to $i^{(m)}$. Then the current price P of the bond is

$$P \;=\; 1000\, i^{(m)}\, a^{(m)}_{\overline{R}|\;,\,r^{(m)}} \;+\; 1000\,(1 + \frac{r^{(m)}}{m})^{-Rm}$$

In this equation, the value P represents cash on hand. The first term on the right-hand side is the present value at nominal interest rate $r^{(m)}$ of the payments of $i^{(m)}\,1000/m$ every $1/m$ year, which amount to $1000 i^{(m)}$ every year for R years. The final repayment of principal in R years contributes the second term on the right-hand side to the present value. As an application of this formula, it is easy to check that a 10-year \$1000 bond with nominal annualized quarterly interest rate $i^{(4)} = 0.06$ would be priced at \$863.22 if the going nominal rate of interest were $r^{(4)} = 0.08$.

A slightly different valuation problem is presented by the *zero-coupon bond*, a financial instrument which pays all principal and interest, at a declared interest rate $i = i_{\text{APR}}$, at the end of a term of n years, but pays nothing before that time. When first issued, a zero-coupon bond yielding i_{APR} which will pay \$1000 at the end of n years is priced at its present value

$$\Pi_n \;=\; 1000 \,\cdot\, (1+i)^{-n} \tag{2.8}$$

(Transaction costs are generally figured into the price before calculating the yield i.) At a time, $n - R$ years later, when the zero-coupon bond has R years left to run and the appropriate interest rate for valuation has changed to $r = r_{\text{APR}}$, the correct price of the bond is the present value of a payment of 1000 R years in the future, or $1000(1+r)^{-R}$.

For tax purposes, at least in the US, an investor is required (either by the federal or state government, depending on the issuer of the bond) to declare the amount of interest income received or *deemed to have been received* from the bond in a specific calendar year. For an ordinary or coupon bond, the year's income is just the total $i^{(m)} \cdot 1000$ actually received during the year. For a zero-coupon bond assumed to have been acquired when first issued, at price Π_n, if the interest rate has remained constant at $i = i_{\text{APR}}$ since the time of acquisition, then the interest income deemed to be received in the year, also called the *Original Issue Discount* (OID), would simply be the year's interest $\Pi_n \, i^{(m)}$ on the initial effective face amount Π_n of the bond. That is because all accumulated value of the bond would be attributed to OID interest income received year by year, with the principal remaining the same at Π_n. Assume next that the actual year-by-year APR interest rate is $r(j)$ throughout the year $[j, j+1)$, for $j = 0, 1, \ldots n-1$, with $r(0)$ equal to i. Then, again because accumulated value over the initial price is deemed to have been received in the form of yearly interest, the OID should be $\Pi_n \, r(j)$ in the year $[j, j+1)$. The problematic aspect of this calculation is that, when interest rates have fluctuated a lot over the times $j = 0, 1, \ldots, n-R-1$, the zero-coupon bond investor will be deemed to have received income $\Pi_n \, r(j)$ in successive years $j = 0, 1, \ldots, n-R-1$, corresponding to a total accumulated value of

$$\Pi_n \, (1 + r(0)) \, (1 + r(1)) \, \cdots \, (1 + r(n-R-1))$$

while the price $1000 \, (1 + r(n-R))^{-R}$ for which the bond could be sold may be very different. The discrepancy between the 'deemed received' accumulated value and the final actual value when the bond is redeemed or sold must presumably be treated as a capital gain or loss. However, the present author makes no claim to have presented this topic according to the views of the Internal Revenue Service, since he has never been able to figure out authoritatively what those views are.

2.2 Force of Mortality & Analytical Models

Up to now, the function $S(x)$ called the "survivor" or "survival" function has been defined to be equal to the life-table ratio l_x/l_0 at all integer ages x, and to be piecewise continuously differentiable for all positive real values of x. Intuitively, for all positive real x and t, $S(x) - S(x+t)$ is the fraction of the initial life-table cohort which dies between ages x and $x+t$, and $(S(x) - S(x+t))/S(x)$ represents the fraction of those alive at age x who fail before $x+t$. An equivalent representation is $S(x) = \int_x^\infty f(t)\, dt$, where $f(t) \equiv -S'(t)$ is called the *failure density.* If T denotes the random variable which is the age at death for a newly born individual governed by the same causes of failure as the life-table cohort, then $P(T \geq x) = S(x)$, and according to the Fundamental Theorem of Calculus,

$$\lim_{\epsilon \to 0+} \frac{P(x \leq T \leq x + \epsilon)}{\epsilon} \;=\; \lim_{\epsilon \to 0+} \int_x^{x+\epsilon} f(u)\, du \;=\; f(x)$$

as long as the failure density is a continuous function.

Two further useful actuarial notations, often used to specify the theoretical lifetime distribution, are:

$${}_tp_x \;=\; P(T \geq x + t \,|\, T \geq x\,) \;=\; S(x+t)/S(x)$$

and

$${}_tq_x \;=\; 1 - {}_tp_x \;=\; P(T \leq x + t \,|\, T \geq x\,) \;=\; (S(x) - S(x+t))/S(x)$$

The quantity ${}_tq_x$ is referred to as the *age-specific death rate* for periods of length t. In the most usual case where $t = 1$ and x is an integer age, the notation ${}_1q_x$ is replaced by q_x, and ${}_1p_x$ is replaced by p_x. The rate q_x would be estimated from the cohort life table as the ratio d_x/l_x of those who die between ages x and $x+1$ as a fraction of those who reached age x. The way in which this quantity varies with x is one of the most important topics of study in actuarial science. For example, one important way in which numerical analysis enters actuarial science is that one wishes to interpolate the values q_x smoothly as a function of x. The topic called "Graduation Theory" among actuaries is the mathematical methodology of Interpolation and Spline-smoothing applied to the raw function $q_x = d_x/l_x$.

To give some idea what a realistic set of death-rates looks like, Figure 2.1 pictures the age-specific 1-year death-rates q_x for the simulated life-table given as Table 1.1 on page 3. Additional granularity in the death-rates can be seen in Figure 2.2, where the logarithms of death-rates are plotted. After a very high death-rate during the first year of life (26.3 deaths per thousand live births), there is a rough year-by-year decline in death-rates from 1.45 per thousand in the second year to 0.34 per thousand in the eleventh year. (But there were small increases in rate from ages 4 to 7 and from 8 to 9, which are likely due to statistical irregularity rather than real increases in risk.) Between ages 11 and 40, there is an erratic but roughly linear increase of death-rates per thousand from 0.4 to 3.0. However, at ages beyond 40 there is a rapid increase in death-rates as a function of age. As can be seen from Figure 2.2, the values q_x seem to increase roughly as a power c^x where $c \in [1.08, 1.10]$. (Compare this behavior with the Gompertz-Makeham Example (v) below.) This exponential behavior of the age-specific death-rate for large ages suggests that the death-rates pictured could reasonably be extrapolated to older ages using the formula

$$q_x \approx q_{78} \cdot (1.0885)^{x-78} \quad , \qquad x \geq 79 \tag{2.9}$$

where the number 1.0885 was found as $log(q_{78}/q_{39})/(78-39)$.

Now consider the behavior of ${}_\epsilon q_x$ as ϵ gets small. It is clear that ${}_\epsilon q_x$ must also get small, roughly proportionately to ϵ, since the probability of dying between ages x and $x+\epsilon$ is approximately $\epsilon\, f(x)$ when ϵ gets small.

<u>Definition</u>: The limiting death-rate ${}_\epsilon q_x/\epsilon$ per unit time as $\epsilon \searrow 0$ is called by actuaries the **force of mortality** $\mu(x)$. In reliability theory or biostatistics, the same function is called the *failure intensity, failure rate*, or *hazard intensity.*

The reasoning above shows that for small ϵ,

$$\frac{{}_\epsilon q_x}{\epsilon} = \frac{1}{\epsilon S(x)} \int_x^{x+\epsilon} f(u)\, du \longrightarrow \frac{f(x)}{S(x)} \, , \qquad \epsilon \searrow 0$$

Thus

$$\mu(x) = \frac{f(x)}{S(x)} = \frac{-S'(x)}{S(x)} = -\frac{d}{dx} \ln(S(x))$$

where the chain rule for differentiation was used in the last step. Replacing x by y and integrating both sides of the last equation between 0 and x, we find

$$\int_0^x \mu(y)\, dy = \Big(-\ln(S(y)) \Big)_0^x = -\ln(S(x))$$

since $S(0) = 1$. Similarly,

$$\int_x^{x+t} \mu(y)\, dy = \ln S(x) - \ln S(x+t)$$

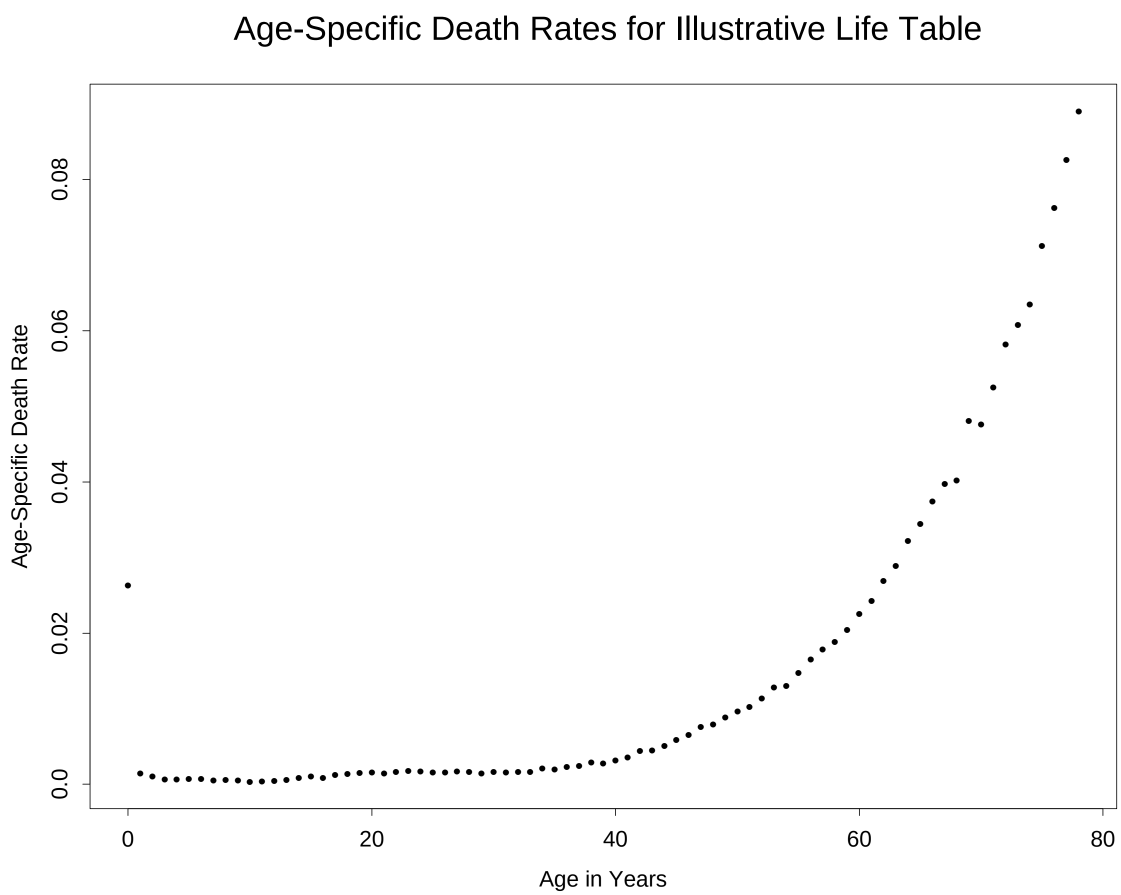

Figure 2.1: Plot of age-specific death-rates q_x versus x, for the simulated illustrative life table given in Table 1.1, page 3.

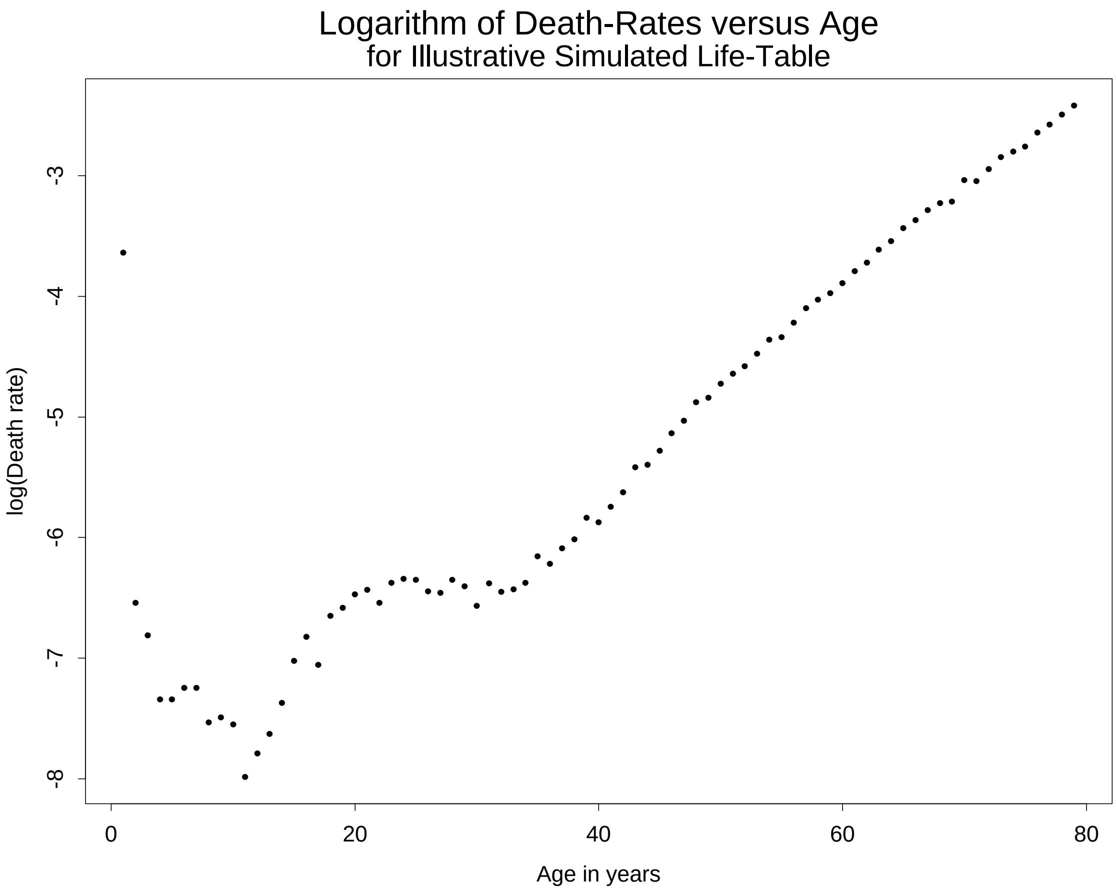

Figure 2.2: Plot of logarithm $\log(q_x)$ of age-specific death-rates as a function of age x, for the simulated illustrative life table given in Table 1.1, page 3. The rates whose logarithms are plotted here are the same ones shown in Figure 2.1.

Now exponentiate to obtain the useful formulas

$$S(x) = \exp\Big\{ - \int_0^x \mu(y)\,dy \Big\} \ , \qquad {}_tp_x = \frac{S(x+t)}{S(x)} = \exp\Big\{ - \int_x^{x+t} \mu(y)\,dy \Big\}$$

Examples:

(i) If $S(x) = (\omega - x)/\omega$ for $0 \le x \le \omega$ (the *uniform failure distribution* on $[0, \omega]$), then $\mu(x) = (\omega - x)^{-1}$. Note that this hazard function increases to ∞ as x increases to ω.

(ii) If $S(x) = e^{-\mu x}$ for $x \ge 0$ (the *exponential failure distribution* on $[0, \infty)$), then $\mu(x) = \mu$ is constant.

(iii) If $S(x) = \exp(-\lambda x^\gamma)$ for $x \ge 0$, then mortality follows the *Weibull life distribution model* with *shape parameter* $\gamma > 0$ and *scale parameter* λ. The force of mortality takes the form

$$\mu(x) \;=\; \lambda\,\gamma\,x^{\gamma-1}$$

This model is very popular in engineering reliability. It has the flexibility that by choice of the shape parameter γ one can have

(a) failure rate increasing as a function of x ($\gamma > 1$),

(b) constant failure rate ($\gamma = 1$, the exponential model again),
or

(c) decreasing failure rate ($0 < \gamma < 1$).

But what one cannot have, in the examples considered so far, is a force-of-mortality function which decreases on part of the time-axis and increases elsewhere.

(iv) Two other models for positive random variables which are popular in various statistical applications are the **Gamma**, with

$$S(x) \;=\; \int_x^\infty \beta^\alpha\, y^{\alpha-1}\, e^{-\beta y}\,dy \Big/ \int_0^\infty z^{\alpha-1}\, e^{-z}\,dz \ , \qquad \alpha,\, \beta > 0$$

and the **Lognormal**, with

$$S(x) \;=\; 1 - \Phi\Big(\frac{\ln x - m}{\sigma} \Big) \quad , \qquad m \text{ real}, \;\; \sigma > 0$$

where

$$\Phi(z) \equiv \frac{1}{2} + \int_0^z e^{-u^2/2} \frac{du}{\sqrt{2\pi}}$$

is called the *standard normal distribution function.* In the Gamma case, the expected lifetime is α/β, while in the Lognormal, the expectation is $\exp(\mu + \sigma^2/2)$. Neither of these last two examples has a convenient or interpretable force-of-mortality function.

Increasing force of mortality intuitively corresponds to aging, where the causes of death operate with greater intensity or effect at greater ages. Constant force of mortality, which is easily seen from the formula $S(x) = \exp(-\int_0^x \mu(y)\, dy)$ to be equivalent to exponential failure distribution, would occur if mortality arose only from pure accidents unrelated to age. Decreasing force of mortality, which really does occur in certain situations, reflects what engineers call "burn-in", where after a period of initial failures due to loose connections and factory defects the nondefective devices emerge and exhibit high reliability for a while. The decreasing force of mortality reflects the fact that the devices known to have functioned properly for a short while are known to be correctly assembled and are therefore highly likely to have a standard length of operating lifetime. In human life tables, infant mortality corresponds to burn-in: risks of death for babies decrease markedly after the one-year period within which the most severe congenital defects and diseases of infancy manifest themselves. Of course, human life tables also exhibit an aging effect at high ages, since the high-mortality diseases like heart disease and cancer strike with greatest effect at higher ages. Between infancy and late middle age, at least in western countries, hazard rates are relatively flat. This pattern of initial decrease, flat middle, and final increase of the force-of-mortality, seen clearly in Figure 2.1, is called a *bathtub shape* and requires new survival models.

As shown above, the failure models in common statistical and reliability usage *either* have increasing force of mortality functions *or* decreasing force of mortality, but not both. Actuaries have developed an analytical model which is somewhat more realistic than the preceding examples for human mortalty at ages beyond childhood. While the standard form of this model does not accommodate a bathtub shape for death-rates, a simple modification of it does.

Example (v). (*Gompertz-Makeham* forms of the force of mortality). Suppose that $\mu(x)$ is defined directly to have the form $A + B\,c^x$. (The Bc^x term was proposed by Gompertz, the additive constant A by Makeham. Thus the *Gompertz* force-of-mortality model is the special case with $A = 0$, or $\mu(x) = Bc^x$.) By choice of the parameter c as being respectively greater than or less than 1, one can arrange that the force-of-mortality curve either be increasing or decreasing. Roughly realistic values of c for human mortality will be only slightly greater than 1: if the Gompertz (non-constant) term in force-of-mortality were for example to quintuple in 20 years, then $c \approx 5^{1/20} = 1.084$, which may be a reasonable value except for very advanced ages. (Compare the comments made in connection with Figures 2.1 and 2.2: for middle and higher ages in the simulated illustrative life table of Table 1.1, which corresponds roughly to US male mortality of around 1960, the figure of c was found to be roughly 1.09.) Note that in any case the Gompertz-Makeham force of mortality is strictly convex (i.e., has strictly positive second derivative) when $B > 0$ and $c \neq 1$. The Gompertz-Makeham family could be enriched still further, with further benefits of realism, by adding a linear term Dx. If $D < -B\ln(c)$, with $0 < A < B$, $c > 1$, then it is easy to check that

$$\mu(x) \;=\; A + B\,c^x + Dx$$

has a bathtub shape, initially decreasing and later increasing.

Figures 2.3 and 2.4 display the shapes of force-of-mortality functions (iii)-(v) for various parameter combinations chosen in such a way that the expected lifetime is 75 years. This restriction has the effect of reducing the number of free parameters in each family of examples by 1. One can see from these pictures that the Gamma and Weibull families contain many very similar shapes for force-of-mortality curves, but that the lognormal and Makeham families are quite different.

Figure 2.5 shows survival curves from several analytical models plotted on the same axes as the 1959 US male life-table data from which Table 1.1 was simulated. The previous discussion about bathtub-shaped force of mortality functions should have made it clear that none of the analytical models presented could give a good fit at all ages, but the Figure indicates the rather good fit which can be achieved to realistic life-table data at ages 40 and above. The models fitted all assumed that $S(40) = 0.925$ and that for lives aged 40, $T - 40$ followed the indicated analytical form. Parameters for all models were determined from the requirements of median age 72 at death (equal by definition to the value t_m for which $S(t_m) = 0.5$) and probability 0.04 of surviving to age 90. Thus, all four plotted survival curves have been designed to pass through the three points $(40, 0.925)$, $(72, 0.5)$, $(90, 0.04)$. Of the four fitted curves, clearly the Gompertz agrees most closely with the plotted points for 1959 US male mortality. The Gompertz curve has parameters $B = 0.00346$, $c = 1.0918$, the latter of which is close to the value 1.0885 used in formula (2.9) to extrapolate the 1959 life-table death-rates to older ages.

2.2.1 Comparison of Forces of Mortality

What does it mean to say that one lifetime, with associated survival function $S_1(t)$, has hazard (i.e. force of mortality) $\mu_1(t)$ which is a constant multiple κ at all ages of the force of mortality $\mu_2(t)$ for a second lifetime with survival function $S_2(t)$? It means that the cumulative hazard functions are *proportional*, i.e.,

$$-\ln S_1(t) \;=\; \int_0^t \mu_1(x)dx \;=\; \int_0^t \kappa\,\mu_2(x)dx \;=\; \kappa(-\ln S_2(t))$$

and therefore that

$$S_1(t) \;=\; (S_2(t))^\kappa \quad , \qquad \text{all} \quad t \ge 0$$

This remark is of especial interest in biostatistics and epidemiology when the factor κ is allowed to depend (e.g., by a regression model $\ln(\kappa) = \beta \cdot Z$) on other measured variables (*covariates*) Z. This model is called the *(Cox) Proportional-Hazards model* and is treated at length in books on survival data analysis (Cox and Oakes 1984, Kalbfleisch and Prentice 1980) or biostatistics (Lee 1980).

Example. Consider a setting in which there are four subpopulations of the general population, categorized by the four combinations of values of two binary covariates Z_1, $Z_2 = 0, 1$. Suppose that these four combinations have

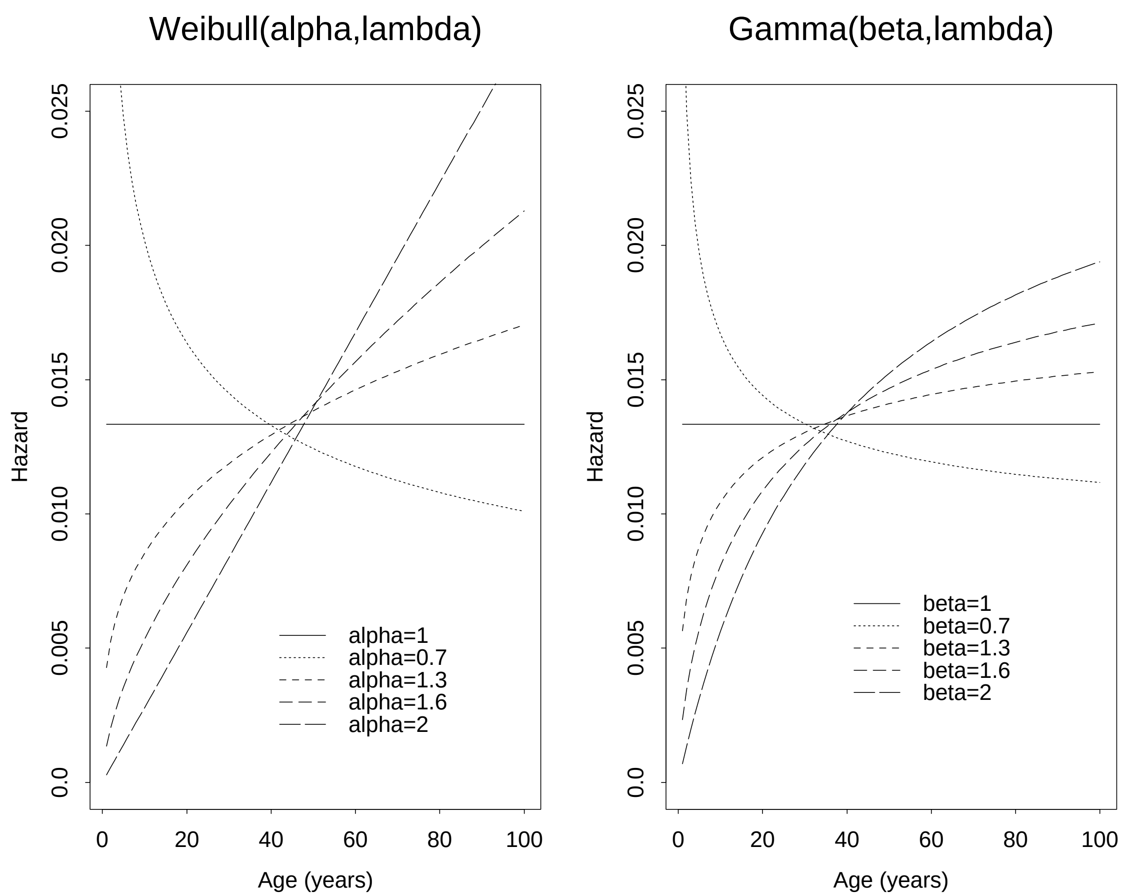

Figure 2.3: Force of Mortality Functions for Weibull and Gamma Probability Densities. In each case, the parameters are fixed in such a way that the expected survival time is 75 years.

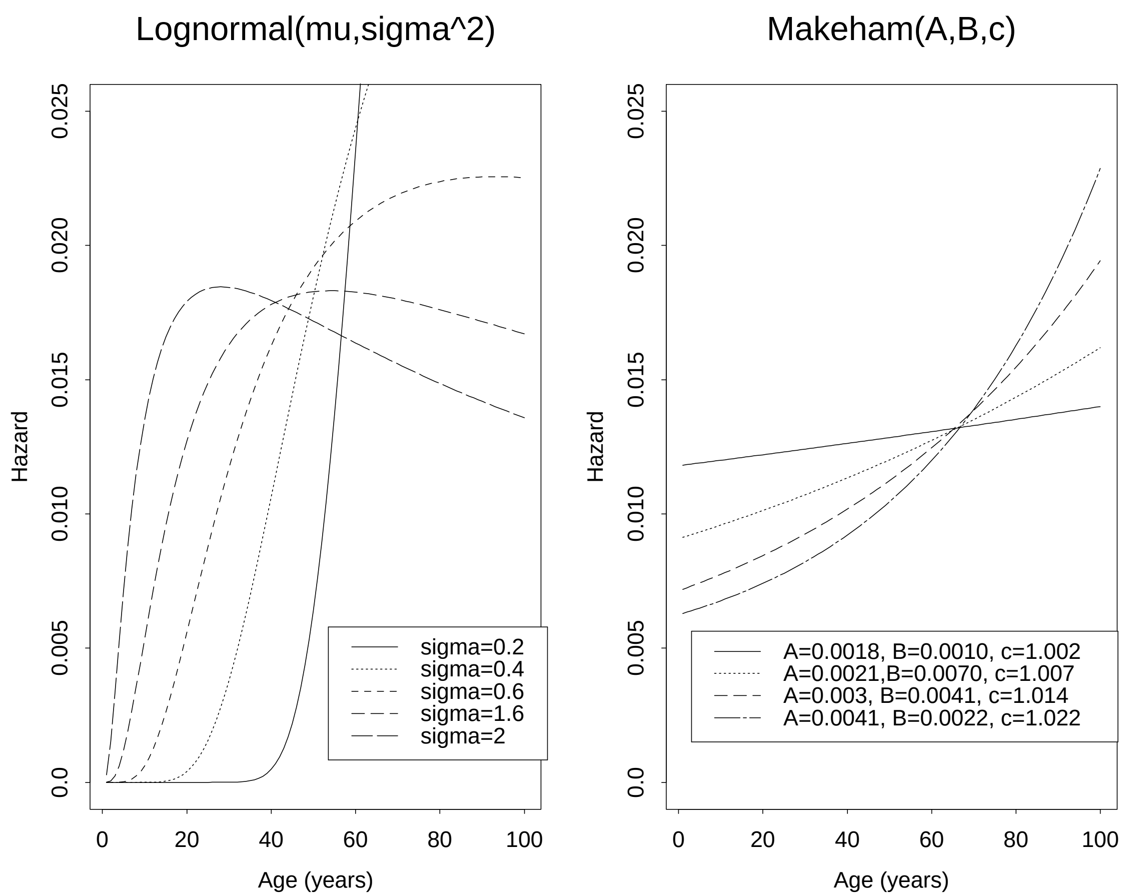

Figure 2.4: Force of Mortality Functions for Lognormal and Makeham Densities. In each case, the parameters are fixed in such a way that the expected survival time is 75 years.

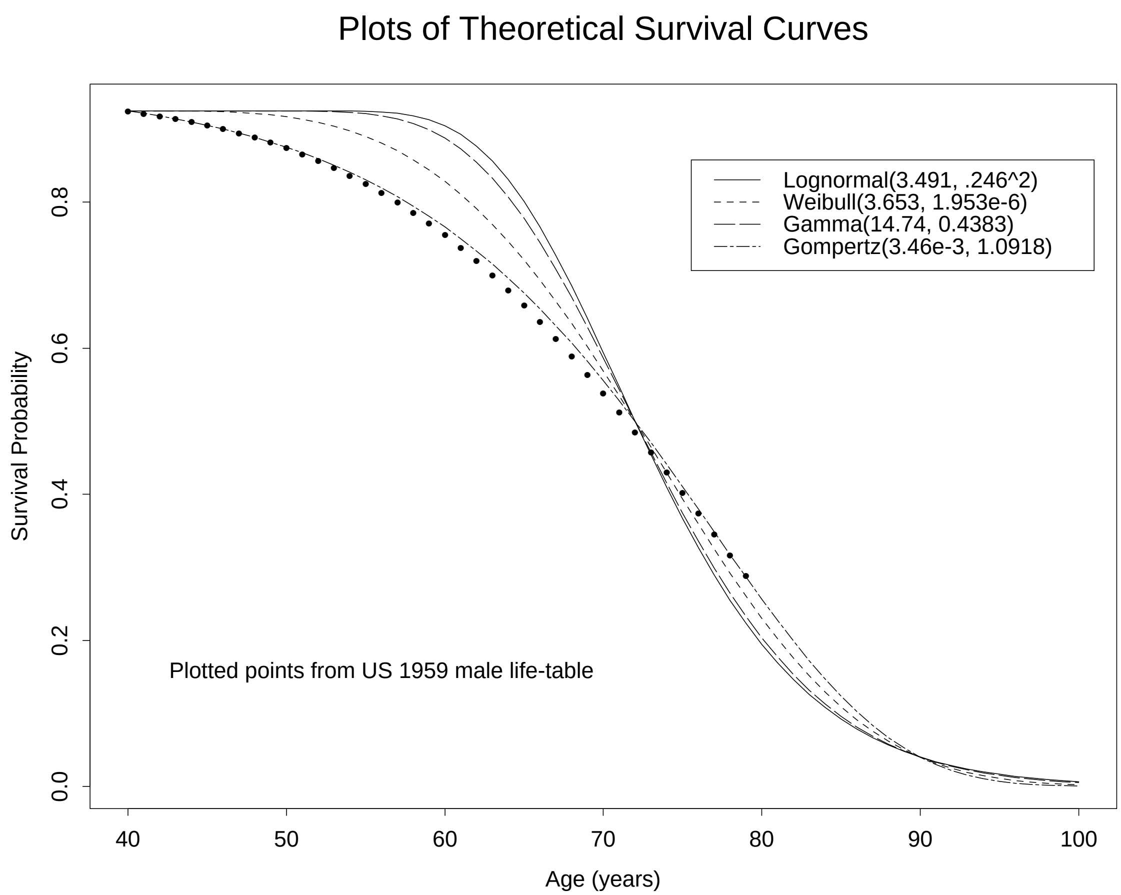

Figure 2.5: Theoretical survival curves, for ages 40 and above, plotted as lines for comparison with 1959 US male life-table survival probabilities plotted as points. The four analytical survival curves — Lognormal, Weibull, Gamma, and Gompertz — are taken as models for age-at-death minus 40, so if $S_{\rm theor}(t)$ denotes the theoretical survival curve with indicated parameters, the plotted curve is $(t,\ 0.925 \cdot S_{\rm theor}(t-40))$. The parameters of each analytical model were determined so that the plotted probabilities would be $0.925,\ 0.5,\ 0.04$ respectively at $t = 40,\ 72,\ 90$.

respective conditional probabilities for lives aged x (or relative frequencies in the general population aged x)

$$P_x(Z_1 = Z_2 = 0) = 0.15 \quad , \quad P_x(Z_1 = 0,\ Z_2 = 1) = 0.2$$
$$P_x(Z_1 = 1,\ Z_2 = 0) = 0.3 \quad , \quad P_x(Z_1 = Z_2 = 1) = 0.35$$

and that for a life aged x and all $t > 0$,

$$P(T \geq x + t \,|\, T \geq x,\ Z_1 = z_1,\ Z_2 = z_2) \;=\; \exp(-2.5\, e^{0.7z_1 - .8z_2}\, t^2/20000)$$

It can be seen from the conditional survival function just displayed that the forces of mortality at ages greater than x are

$$\mu(x + t) \;=\; (2.5\, e^{0.7z_1 - .8z_2})\, t/10000$$

so that the force of mortality at all ages is multiplied by $e^{0.7} = 2.0138$ for individuals with $Z_1 = 1$ versus those with $Z_1 = 0$, and is multiplied by $e^{-0.8} = 0.4493$ for those with $Z_2 = 1$ versus those with $Z_2 = 0$. The effect on age-specific death-rates is approximately the same. Direct calculation shows for example that the ratio of age-specific death rate at age x+20 for individuals in the group with $(Z_1 = 1, Z_2 = 0)$ versus those in the group with $(Z_1 = 0, Z_2 = 0)$ is not precisely $e^{0.7} = 2.014$, but rather

$$\frac{1 - \exp(-2.5e^{0.7}((21^2 - 20^2)/20000)}{1 - \exp(-2.5((21^2 - 20^2)/20000)} \;=\; 2.0085$$

Various calculations, related to the fractions of the surviving population at various ages in each of the four population subgroups, can be performed easily . For example, to find

$$P(Z_1 = 0,\ Z_2 = 1 \,|\, T \geq x + 30)$$

we proceed in several steps (which correspond to an application of Bayes' rule, *viz.* Hogg and Tanis 1997, sec. 2.5, or Larson 1982, Sec. 2.6):

$$P(T \geq x{+}30,\ Z_1 = 0\, Z_2 = 1 |\, T \geq x) \;=\; 0.2\, \exp(-2.5e^{-0.8}\frac{30^2}{20000}) \;=\; 0.1901$$

and similarly

$$P(T \geq x + 30 \,|\, T \geq x) \;=\; 0.15\, \exp(-2.5(30^2/20000)) \;+\; 0.1901 \;+$$

$$+ \, 0.3 \, \exp(-2.5 * e^{0.7} \frac{30^2}{20000}) \; + \; 0.35 \, \exp(-2.5 e^{0.7-0.8} \frac{30^2}{20000}) \; = \; 0.8795$$

Thus, by definition of conditional probabilities (restricted to the cohort of lives aged x), taking ratios of the last two displayed quantities yields

$$P(Z_1 = 0, \, Z_2 = 1 \, | \, T \geq x + 30) \; = \; \frac{0.1901}{0.8795} \; = \; 0.2162$$

$\square$.

In biostatistics and epidemiology, the measured variables $\underline{Z} = (Z_1, \ldots, Z_p)$ recorded for each individual in a survival study might be: indicator of a specific disease or diagnostic condition (e.g., diabetes, high blood pressure, specific electrocardiogram anomaly), quantitative measurement of a risk-factor (dietary cholesterol, percent caloric intake from fat, relative weight-to-height index, or exposure to a toxic chemical), or indicator of type of treatment or intervention. In these fields, the objective of such detailed models of covariate effects on survival can be: to correct for incidental individual differences in assessing the effectiveness of a treatment; to create a prognostic index for use in diagnosis and choice of treatment; or to ascertain the possible risks and benefits for health and survival from various sorts of life-style interventions. The multiplicative effects of various risk-factors on age-specific death rates are often highlighted in the news media.

In an insurance setting, categorical variables for risky life-styles, occupations, or exposures might be used in *risk-rating*, i.e., in individualizing insurance premiums. While risk-rating is used routinely in casualty and property insurance underwriting, for example by increasing premiums in response to recent claims or by taking location into account, it can be politically sensitive in a life-insurance and pension context. In particular, while gender differences in mortality can be used in calculating insurance and annuity premiums, as can certain life-style factors like smoking, it is currently illegal to use racial and genetic differences in this way.

All life insurers must be conscious of the extent to which their policyholders as a group differ from the general population with respect to mortality. Insurers can collect special mortality tables on special groups, such as employee groups or voluntary organizations, and regression-type models like the Cox proportional-hazards model may be useful in quantifying group mortality differences when the special-group mortality tables are not based upon large enough cohorts for long enough times to be fully reliable. See Chapter 6, Section 6.3, for discussion about the modification of insurance premiums for select groups.

2.3 *Exercise Set 2*

(1). The sum of the present value of \$1 paid at the end of n years and \$1 paid at the end of $2n$ years is \$1. Find $(1+r)^{2n}$, where $r =$ annual interest rate, compounded annually.

(2). Suppose that an individual aged 20 has random lifetime Z with continuous density function

$$f_Z(t) \;=\; \frac{1}{360}\left(1+\frac{t}{10}\right) \;, \qquad \text{for} \qquad 20 \le t \le 80$$

and 0 for other values of t.

(a) If this individual has a contract with your company that you must pay his heirs $\$10^6 \cdot (1.4 - Z/50)$ at the exact date of his death if this occurs between ages 20 and 70, then what is the expected payment ?

(b) If the value of the death-payment described in (a) should properly be discounted by the factor $\exp(-0.08 \cdot (Z-20))$, i.e. by the nominal interest rate of $e^{0.08} - 1$ per year) to calculate the present value of the payment, then what is the expected present value of the payment under the insurance contract ?

(3). Suppose that a continuous random variable T has hazard rate function (= force of mortality)

$$h(t) \;=\; 10^{-3} \cdot \Big[7.0 - 0.5t + 2e^{t/20}\Big] \;, \qquad t > 0$$

This **is** a legitimate hazard rate of Gompertz-Makeham type since its minimum, which occurs at $t = 20\ln(5)$, is $(17 - 10\ln(5)) \cdot 10^{-4} = 9.1 \cdot 10^{-5} > 0$.

(a) Construct a cohort life-table with $h(t)$ as "force of mortality", based on integer ages up to 70 and cohort-size (= "radix") $l_0 = 10^5$. (Give the numerical entries, preferably by means of a little computer program. If you do the arithmetic using hand-calculators and/or tables, stop at age 20.)

(b) What is the probability that the random variable T exceeds 30, given that it exceeds 3 ? **Hint:** find a closed-form formula for $S(t) = P(T \geq t)$.

(4). Do the Mortgage-Refinancing exercise given in the Illustrative on mortgage refinancing at the end of Section 2.1.

(5). (a) The mortality pattern of a certain population may be described as follows: out of every 98 lives born together, one dies annually until there are no survivors. Find a simple function that can be used as $S(x)$ for this population, and find the probability that a life aged 30 will survive to attain age 35.

(b) Suppose that for x between ages 12 and 40 in a certain population, 10% of the lives aged x die before reaching age $x+1$. Find a simple function that can be used as $S(x)$ for this population, and find the probability that a life aged 30 will survive to attain age 35.

(6). Suppose that a survival distribution (i.e., survival function based on a cohort life table) has the property that ${}_1p_x = \gamma \cdot (\gamma^2)^x$ for some fixed γ between 0 and 1, for every real ≥ 0. What does this imply about S(x) ? (*Give as much information about S as you can.*)

(7). If the instantaneous interest rate is $r(t) = 0.01\,t$ for $0 \leq t \leq 3$, then find the equivalent single *effective rate of interest* or APR for money invested at interest over the interval $0 \leq t \leq 3$.

(8). Find the accumulated value of \$100 at the end of 15 years if the nominal interest rate compounded quarterly (i.e., $i^{(4)}$) is 8% for the first 5 years, if the effective rate of discount is 7% for the second 5 years, and if the nominal rate of discount compounded semiannually ($m = 2$) is 6% for the third 5 years.

(9). Suppose that you borrow \$1000 for 3 years at 6% APR, to be repaid in level payments every six months (twice yearly).

(a) Find the level payment amount P.

(b) What is the present value of the payments you will make if you skip the 2nd and 4th payments ? (You may express your answer in terms of P.)

(10). A survival function has the form $S(x) = \frac{c-x}{c+x}$. If a mortality table is derived from this survival function with a radix l_0 of 100,000 at age 0, and if $l_{35} = 44,000$:

(i) What is the terminal age of the table ?

(ii) What is the probability of surviving from birth to age 60 ?

(iii) What is the probability of a person at exact age 10 dying between exact ages 30 and 45 ?

(11). A separate life table has been constructed for each calendar year of birth, Y, beginning with $Y = 1950$. The mortality functions for the various tables are denoted by the appropriate superscript ^{Y}. For each Y and for all ages x

$$\mu_x^Y \;=\; A \cdot k(Y) \,+\, B\, c^x \quad , \qquad p_x^{Y+1} \;=\; (1+r)\, p_x^Y$$

where k is a function of Y alone and A, B, r are constants (with $r > 0$). If $k(1950) = 1$, then derive a general expression for $k(Y)$.

(12). A standard mortality table follows Makeham's Law with force of mortality

$$\mu_x \;=\; A \,+\, B\, c^x \qquad \text{at all ages } x$$

A separate, higher-risk mortality table also follows Makeham's Law with force of mortality

$$\mu_x^* \;=\; A^* \,+\, B^*\, c^x \qquad \text{at all ages } x$$

with the same constant c. If for all starting ages the probability of surviving 6 years according to the higher-risk table is equal to the probability of surviving 9 years according to the standard table, then express each of A^* and B^* in terms of A, B, c.

(13). A homeowner borrows \$100,000 at 7% APR from a bank, agreeing to repay by 30 equal yearly payments beginning one year from the time of the loan.

(a) How much is each payment ?

(b) Suppose that after paying the first 3 yearly payments, the homeowner misses the next two (i.e. pays nothing on the 4^{th} and 5^{th} anniversaries of the loan). Find the outstanding balance at the 6^{th} anniversary of the loan, figured at 7%). This is the amount which, if paid as a lump sum at time

6, has present value together with the amounts already paid of \$100,000 at time 0.

(14). A deposit of 300 is made into a fund at time $t = 0$. The fund pays interest for the first three years at a nominal monthly rate $d^{(12)}$ of discount. From $t = 3$ to $t = 7$, interest is credited according to the force of interest $\delta_t = 1/(3t + 3)$. As of time $t = 7$, the accumulated value of the fund is 574. Calculate $d^{(12)}$.

(15). Calculate the price at which you would sell a \$10,000 30-year coupon bond with nominal 6% semi-annual coupon ($n = 30$, $m - 2$, $i^{(m)} = 0.06$), 15 years after issue, if for the next 15 years, the effective interest rate for valuation is $i_{\text{APR}} = 0.07$.

(16). Calculate the price at which you would sell a 30-year zero-coupon bond with face amount \$10,000 initially issued 15 years ago with $i = i_{\text{APR}} = 0.06$, if for the next 15 years, the effective interest rate for valuation is $i_{\text{APR}} = 0.07$.

2.4 Worked Examples

Example 1. How large must a half-yearly payment be in order that the stream of payments starting immediately be equivalent (in present value terms) at 6% interest to a lump-sum payment of \$5000, if the payment-stream is to last (a) 10 years, (b) 20 years, or (c) forever ?

If the payment size is P, then the balance equation is

$$5000 = 2P \cdot \ddot{a}^{(2)}_{\overline{n}|} = 2P \, \frac{1 - 1.06^{-n}}{d^{(2)}}$$

Since $d^{(2)} = 2(1 - 1/\sqrt{1.06}) = 2 \cdot 0.02871$, the result is

$$P = (5000 \cdot 0.02871)/(1 - 1.06^{-n}) = 143.57/(1 - 1.06^{-n})$$

So the answer to part (c), in which $n = \infty$, is \$143.57. For parts (a) and (b), respectively with $n = 10$ and 20, the answers are \$325.11, \$208.62.

Example 2. Assume m is divisible by 2. Express in two different ways the present value of the perpetuity of payments $1/m$ at times $1/m$, $3/m$, $5/m$, $\ldots$, and use either one to give a simple formula.

This example illustrates the general methods enunciated at the beginning of Section 2.1. Observe first of all that the specified payment-stream is exactly the same as a stream of payments of $1/m$ at times $0, 2/m, 4/m, \ldots$ forever, deferred by a time $1/m$. Since this payment-stream starting at 0 is exactly one-half that of the stream whose present value is $\ddot{a}^{(m/2)}_{\overline{\infty}|}$, a first present value expression is

$$v^{1/m}\, \frac{1}{2}\, \ddot{a}^{(m/2)}_{\overline{\infty}|}$$

A second way of looking at the payment-stream at odd multiples of $1/m$ is as the perpetuity-due payment stream ($1/m$ at times k/m for all $k \ge 0$) **minus** the payment-stream discussed above of amounts $1/m$ at times $2k/m$ for all nonnegative integers k. Thus the present value has the second expression

$$\ddot{a}^{(m)}_{\overline{\infty}|} \;-\; \frac{1}{2}\, \ddot{a}^{(m/2)}_{\overline{\infty}|}$$

Equating the two expressions allows us to conclude that

$$\frac{1}{2}\, \ddot{a}^{(m/2)}_{\overline{\infty}|} \;=\; \ddot{a}^{(m)}_{\overline{\infty}|} \Big/ (1 \,+\, v^{1/m})$$

Substituting this into the first of the displayed present-value expressions, and using the simple expression $1/d^{(m)}$ for the present value of the perpetuity-due, shows that that the present value requested in the Example is

$$\frac{1}{d^{(m)}} \cdot \frac{v^{1/m}}{1 \,+\, v^{1/m}} \;=\; \frac{1}{d^{(m)}\,(v^{-1/m} \,+\, 1)} \;=\; \frac{1}{d^{(m)}\,(2 + i^{(m)}/m)}$$

and this answer is valid whether or not m is even.

Example 3. Suppose that you are negotiating a car-loan of $\$10,000$. *Would you rather have an interest rate of* 4% *for 4 years,* 3% *for 3 years,* 2% *for 2 years, or a cash discount of* $\$500$ *? Show how the answer depends upon the interest rate with respect to which you calculate present values, and give numerical answers for present values calculated at 6% and 8%. Assume that all loans have monthly payments paid at the beginning of the month (e.g., the 4 year loan has 48 monthly payments paid at time 0 and at the ends of 47 succeeding months).*

The monthly payments for an n-year loan at interest-rate i is $10000/(12\,\ddot{a}^{(12)}_{\overline{n}|}) \;=\; (10000/12)\,d^{(12)}/(1-(1+i)^{-n})$. Therefore, the present value at interest-rate r of the n-year monthly payment-stream is

$$10000 \cdot \frac{1-(1+i)^{-1/12}}{1-(1+r)^{-1/12}} \cdot \frac{1-(1+r)^{-n}}{1-(1+i)^{-n}}$$

Using interest-rate $r = 0.06$, the present values are calculated as follows:

For 4-year 4% loan: \$9645.77

For 3-year 3% loan: \$9599.02

For 2-year 2% loan: \$9642.89

so that the most attractive option is the cash discount (which would make the present value of the debt owed to be \$9500). Next, using interest-rate $r = 0.08$, the present values of the various options are:

For 4-year 4% loan: \$9314.72

For 3-year 3% loan: \$9349.73

For 2-year 2% loan: \$9475.68

so that the most attractive option in this case is the 4-year loan. (The cash discount is now the least attractive option.)

Example 4. Suppose that the force of mortality $\mu(y)$ is specified for exact ages y ranging from 5 to 55 as

$$\mu(y) \;=\; 10^{-4} \cdot (20 - 0.5|30-y|)$$

Then find analytical expressions for the survival probabilities $S(y)$ for exact ages y in the same range, and for the (one-year) death-rates q_x for integer ages $x = 5, \ldots, 54$, assuming that $S(5) = 0.97$.

The key formulas connecting force of mortality and survival function are here applied separately on the age-intervals $[5, 30]$ and $[30, 55]$, as follows. First for $5 \le y \le 30$,

$$S(y) \;=\; S(5)\,\exp(-\int_5^y \mu(z)\,dz) \;=\; 0.97\,\exp\Big(-10^{-4}(5(y-5)+0.25(y^2-25))\Big)$$

so that $S(30) \;=\; 0.97\,e^{-0.034375} \;=\; 0.93722$, and for $30 \le y \le 55$

$$S(y) \;=\; S(30)\,\exp\Big(-10^{-4}\int_{30}^y (20+0.5(30-z))\,dz\Big)$$

$$= \; 0.9372 \, \exp\Big(-\, .002(y-30) + 2.5 \cdot 10^{-5}(y-30)^2 \Big)$$

The death-rates q_x therefore have two different analytical forms: first, in the case $x = 5, \ldots, 29$,

$$q_x \; = \; S(x+1)/S(x) \; = \; \exp\Big(-\, 5 \cdot 10^{-5}\,(10.5 + x) \Big)$$

and second, in the case $x = 30, \ldots, 54$,

$$q_x \; = \; \exp\Big(-\, .002 + 2.5 \cdot 10^{-5}(2(x-30)+1) \Big)$$

2.5 Useful Formulas from Chapter 2

$$v \; = \; 1/(1+i)$$

p. 26

$$a^{(m)}_{\overline{n}|} \; = \; \frac{1-v^n}{i^{(m)}} \quad , \qquad \ddot{a}^{(m)}_{\overline{n}|} \; = \; \frac{1-v^n}{d^{(m)}}$$

pp. 27–27

$$a_{\overline{n}|} m \; = \; v^{1/m} \, \ddot{a}_{\overline{n}|} m$$

p. 27

$$\ddot{a}^{(\infty)}_{\overline{n}|} \; = \; a_{\overline{n}|} \infty \; = \; \overline{a}_n \; = \; \frac{1-v^n}{\delta}$$

p. 27

$$a^{(m)}_{\overline{\infty}|} \; = \; \frac{1}{i^{(m)}} \quad , \qquad \ddot{a}_{\overline{\infty}|} m \; = \; \frac{1}{d^{(m)}}$$

p. 28

$$(I^{(m)}\ddot{a})^{(m)}_{\overline{n}|} \; = \; \ddot{a}^{(m)}_{\overline{\infty}|} \left(\ddot{a}^{(m)}_{\overline{n}|} \; - \; n \, v^n \right)$$

p. 30

$$(D^{(m)}\ddot{a})^{(m)}_{\overline{n}|} \; = \; (n + \frac{1}{m}) \; \ddot{a}^{(m)}_{\overline{n}|} \; - \; (I^{(m)}\ddot{a})^{(m)}_{\overline{n}|}$$

p. 30

n-yr m'thly Mortgage Paymt : $\dfrac{Loan\ Amt}{m \, \ddot{a}^{(m)}_{\overline{n}|}}$

p. 31

n-yr Mortgage Bal. at $\frac{k}{m}+$: $\quad B_{n,k/m} \;=\; \frac{1-v^{n-k/m}}{1-v^n}$

p. 32

$$ {}_tp_x \;=\; \frac{S(x+t)}{S(x)} \;=\; \exp\left(-\int_0^t \mu(x+s)\,ds\right) $$

p. 40

$$ {}_tp_x \;=\; 1 \;-\; {}_tp_x $$

p. 40

$$ q_x \;=\; {}_1q_x \;=\; \frac{d_x}{l_x} \quad , \qquad p_x \;=\; {}_1p_x \;=\; 1 \;-\; q_x $$

p. 40

$$ \mu(x+t) \;=\; \frac{f(x+t)}{S(x+t)} \;=\; -\frac{\partial}{\partial t}\ln S(x+t) \;=\; -\frac{\partial}{\partial t}\ln l_{x+t} $$

p. 41

$$ S(x) \;=\; \exp(-\int_0^x \mu(y)\,dy) $$

p. 44

Unif. Failure Dist.: $\quad S(x) \;=\; \frac{\omega-x}{\omega} \;,\quad f(x) \;=\; \frac{1}{\omega} \;,\quad 0 \le x \le \omega$

p. 44

Expon. Dist.: $\quad S(x) \;=\; e^{-\mu x}\;,\quad f(x) \;=\; \mu e^{-\mu x}\;,\quad \mu(x) \;=\; \mu\;,\quad x > 0$

p. 44

$$\text{Weibull. Dist.:} \quad S(x) = e^{-\lambda x^\gamma} \quad , \quad \mu(x) = \lambda\gamma x^{\gamma-1} \quad , \quad x > 0$$

p. 44

$$\text{Makeham:} \qquad \mu(x) = A + Bc^x \qquad , \quad x \geq 0$$

$$\text{Gompertz:} \qquad \mu(x) = Bc^x \qquad , \quad x \geq 0$$

$$S(x) = \exp\left(-Ax - \frac{B}{\ln c}\,(c^x - 1)\right)$$

p. 46

Bibliography

[1] Bowers, N., Gerber, H., Hickman, J., Jones, D. and Nesbitt, C. **Actuarial Mathematics** Society of Actuaries, Itasca, Ill. 1986

[2] Cox, D. R. and Oakes, D. **Analysis of Survival Data**, Chapman and Hall, London 1984

[3] Feller, W. **An Introduction to Probability Theory and its Applications**, vol. I, 2nd ed. Wiley, New York, 1957

[4] Feller, W. **An Introduction to Probability Theory and its Applications**, vol. II, 2nd ed. Wiley, New York, 1971

[5] Gerber, H. **Life Insurance Mathematics**, 3rd ed. Springer-Verlag, New York, 1997

[6] Hogg, R. V. and Tanis, E. **Probability and Statistical Inference**, *5th ed.* Prentice-Hall Simon & Schuster, New York, 1997

[7] Jordan, C. W. **Life Contingencies**, 2nd ed. Society of Actuaries, Chicago, 1967

[8] Kalbfleisch, J. and Prentice, R. **The Statistical Analysis of Failure Time Data**, Wiley, New York 1980

[9] Kellison, S. **The Theory of Interest**. Irwin, Homewood, Ill. 1970

[10] Larsen, R. and Marx, M. **An Introduction to Probability and its Applications**. Prentice-Hall, Englewood Cliffs, NJ 1985

[11] Larson, H. **Introduction to Probability Theory and Statistical Inference**, 3rd ed. Wiley, New York, 1982

[12] Lee, E. T. **Statistical Models for Survival Data Analysis**, Lifetime Learning, Belmont Calif. 1980

[13] The R Development Core Team, **R: a Language and Environment**.

[14] **Splus**, version 7.0. MathSoft Inc., 1988, 1996, 2005

[15] Spiegelman, M. *Introduction to Demography*, Revised ed. Univ. of Chicago, Chicago, 1968

[16] Venables, W. and Ripley, B. **Modern Applied Statistics with S**, 4th ed. Springer-Verlag, New York, 2002