A STATISTICAL APPROACH TO GRADUATION BY MATHEMATICAL FORMULA

by

L. K. Chan, Ph.D.

and

H. H. Panjer, Ph.D., F.S.A., F.C.I.A.

ABSTRACT

Graduation by mathematical formula is recast as problem of statistical estimation. The method of maximum likelihood is used to determine the estimates of the parameters. Theory is developed to allow for estimation without resorting to the usual "exposure" formulas. Both single and multiple decrement models are considered. Theoretical results are obtained for some specific mortality models. Numerical procedures to obtain the estimates are considered.

The authors' address is: Statistics and Actuarial Science Group, University of Western Ontario, London, Canada. N6A 5B9

This research was supported by the National Research Council of Canada.

A STATISTICAL APPROACH TO GRADUATION BY MATHEMATICAL FORMULA

INTRODUCTION

Miller's 1946 monograph on graduation deals with a variety of methods of smoothing mortality data. It has been updated by Greville's 1974 Part 5 study notes. Kimeldorf and Jones (1967) applied Bayesian statistical principles to graduation by introducing prior information through the distribution of the "true" rates. They also pointed out that Whittaker-Henderson graduation is essentially Bayesian.

In this paper a somewhat different approach is taken. We introduce prior information by assuming an analytic form of the forces of decrement in terms of certain parameters as in Chapter 6 of Miller's monograph. The graduation is then cast into a parametric estimation problem which is solved using the method of maximum likelihood, a widely used statistical tool. Both single and multiple decrement models are considered.

This graduation process does not take a set of observed rates and smooth them, but rather determines the graduation directly from the observed data (e.g. age at insuring, age at withdrawal, age at death, etc.) which may be in a variety of forms. No additional assumptions, as are used traditionally to determine "exposures", are required since no initial set of rates are determined. The types of data considered here are very general.

The ages at decrement may be known exactly or be known to be in a certain interval, their distributions may be left-truncated at different points for different members in a sample; or, the ages at decrement may be censored, by withdrawal of a member from the sample.

The determination of the parameters is often not a simple process and requires the use of a computer to handle certain numerical procedures. This increased sophistication however should be no problem in this era of large scale computing facilities.

The maximum likelihood method of estimating the parameters will be explored in this paper. It consists of selecting the value, called the maximum likelihood estimate, of the parameter(s) that maximizes the likelihood function (the joint probability density function treated as a function of the parameter(s) given the observed information about the sample members). Intuitively, the distribution that is "most likely" to generate the observed values is selected.

The maximum likelihood method is particularly appealing as the estimates produced are in some senses optimal. It possesses certain properties such as consistency and highest efficiency when the sample size is large (cf: Cramer, 1946, Chapter 33). Incidentally, the maximum likelihood method can be thought of as a Bayesian procedure using an uninformative (i.e. improper uniform) prior distribution for the parameters being estimated and then selecting the mode of the (improper) posterior density given the observations.

Here we apply the maximum likelihood method to several forms, including the Gompertz and Makeham, of the forces of decrement for both single and multiple decrement models. We hope that being acquainted with the techniques discussed will make the reader feel comfortable about applying the methodolgy to mortality or morbidity data of parametric forms not considered here. Some of the results obtained by Panjer (1975) are given without proof to indicate the sort of analytic results that the user might obtain for his particular model. These results, when used in this paper, will be followed by the symbol [§X.Y] to indicate that the proof is given in §X.Y (e.g. §2.3) of Panjer (1975).

We shall, without loss of generality, refer only to mortality and forces of mortality, the extension to general decrements being direct.

DISTRIBUTION THEORY

It will be assumed that all lives at a given age in a mortality study are subject to the same force of mortality. The force of mortality at age $x+t$ is denoted by μ_{x+t}. Consider a life entering the mortality study at age x. Let T denote a random variable representing the time to death of this individual. The cumulative distribution function (cdf) of T is given by

$$\begin{aligned} F(t:x) &= 0 && \text{if } t < 0 \\ &= {}_tq_x && \text{if } 0 \le t \le \infty. \end{aligned}$$

This notation is similar to that used by Hickman (1964). The corresponding probability density function (pdf) of T is given by

$$f(t:x) = {}_tp_x \;\; \mu_{x+t}, \quad 0 \le t < \infty.$$

Consider now the multiple decrement case. For simplicity we shall refer only to mortality data with m associated causes of death. The corresponding forces of mortality for the kth cause is denoted by $\mu_x^{(k)}$. Since the forces of mortality are assumed to act independently we have

$$\mu_x = \sum_{k=1}^{m} \mu_x^{(k)}$$

The joint pdf of T and K (the random variable which takes on values k = 1,2,...,m) is accordingly*

$$f(t,k:x) = {}_tp_x \; \mu_{x+t}^{(k)} \qquad \begin{array}{l} 0 \le t < \infty \\ i = 1,2,\ldots,m. \end{array}$$

It can be easily seen that the marginal pdf of the random variable T over all causes is that of the single decrement model.

The conditional pdf of T and K given that $T > s$, i.e. that (x) has survived to age $x+s$, is

$$f(t,k|t > s:x) = {}_{t-s}p_{x+s} \quad \mu_{x+t}^{(k)} \qquad 0 \le s < t < \infty.$$

The probability that death was a result of the k-th cause given that death occured at time t is $\mu_{x+t}^{(k)}/\mu_{x+t}$.

THE METHOD OF MAXIMUM LIKELIHOOD

We shall assume that the force of mortality is a function of r parameters $\theta_1, \theta_2, \theta_3, \ldots, \theta_r$ and that observed information about the n sample members of and independent sample are available. For

---

*Note that we use the letter f, not as a specific function, but to denote the pdf of the variables in parentheses.

each sample member, we observe either the numerical value of age at death or some interval (possibly of infinite length) into which the age of death falls. For multiple decrement models, the cause of death may also be observed.

Let L_i be the pdf of the ith sample member or the probability that it falls in a certain interval. The <u>likelihood function</u> L is the joint pdf of the n sample members, i.e.,

$$L = C \prod_{i=1}^{n} L_i$$

where C is independent of the parameters. (Each L_i can be considered as the likelihood function of the ith member.) The <u>maximum likelihood estimates</u> are the values $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_r$ of $\theta_1, \theta_2, \ldots, \theta_r$ which maximize L when it is considered as a function of $\theta_1, \theta_2, \ldots, \theta_r$ given the observed information about the members. We shall denote L by $L(\theta_1, \theta_2, \ldots, \theta_r)$. The maximum likelihood estimates are usually obtained by differentiating the log likelihood, $l(\theta_1, \theta_2, \ldots, \theta_r) = \log L(\theta_1, \theta_2, \ldots, \theta_r)$, with respect to each of the parameters, setting the derivatives to zero and solving the system of r equations,

$$dl(\theta_1, \theta_2, \ldots, \theta_r)/d\theta_i = 0, \qquad i = 1, \ldots, r,$$

called the likelihood equations, for $\theta_1, \theta_2, \ldots, \theta_r$. (The log is of base $e = 1.71828$).

The system of equations may provide implicit solutions for each of the parameters. However, more typically they will have to be solved numerically, which should provide no difficulty using currently available computing equipment.

DEVELOPMENT OF THE LIKELIHOOD FUNCTION - SINGLE DECREMENT

Consider a life entering the mortality study at age x. Let Y be the random variable denoting the age at death of this individual. Here we first develop the likelihood functions L_i of individual members when the observed information is recorded in various ways.

(1) The numerical value of Y is recorded.

If this person dies at age y, i.e. if the outcome of the random variable Y is y, the corresponding likelihood function is

$$_{y-x}p_x \ \mu_y$$

and the log likelihood is

$$-\int_x^y \mu_s \, ds + \log \mu_y$$

(2) Death is only known to have occurred in a time interval.

In applications, however, often the exact age at death is not recorded but rather the time interval in which death occurs is recorded. In mortality studies such intervals are typically annual durations since issue for an insurance or annuity contract. Let s' and s'' denote the age limits of the interval in which death occurs for a life entering the study at age x.

Under the assumption that the distribution of such limits of time interval is independent of the parameters under study, the likelihood function is

$$_{s'-x|s''-s'}q_x = {}_{s'-x}p_x \cdot {}_{s''-s'}q_{s'}.$$

The log likelihood becomes

$$log\ {}_{s'-x}p_x + log\ {}_{s''-s'}q_{s'} = -\int_x^{s'} \mu_s\, ds + log\ {}_{s''-s'}q_{s'}$$

If n' sample members fall into the interval, the log likelihood is multiplied by n'.

(3) <u>Either the age at death or the age at withdrawal is recorded</u>.

In most mortality studies, however, many, if not most, of the individuals observed do not die while under observation but die after withdrawal from the sample. The time of withdrawal is normally not fixed in advance and so may be treated as a random variable. Let Z be a random variable denoting the age at withdrawal of this individual. The individual is then subject to both death and withdrawal. Let

$$T = min(Y, Z)$$

and

$$\begin{aligned} A &= 1 \qquad Y \le Z \\ &= 0 \qquad Y > Z \end{aligned}$$

Then the outcome of the two dimensional random variable (T,A) is recorded at death or withdrawal. Upon death the outcome is $(y,1)$; upon withdrawal the outcome is $(z,0)$. The joint probability density function of (T,A) given $Z = z$ is

$$\{{}_{t-x}p_x\mu_t\}^a\ \{{}_{t-x}p_x\}^{1-a}\ .$$

The unconditional probability density function of (T,A) is

$$C(z)\{{}_{t-x}p_x\ \mu_x\}^a\{{}_{t-x}p_x\}^{1-a}\ .$$

where $C(z)$ is the marginal density of Z. Under the assumption that $C(z)$ does not involve the parameters under study, $C(z)$ can be ignored in maximizing the likelihood. The log likelihood is then, ignoring terms not involving the parameters under study,

$$a\{-\int_x^t \mu_s \, ds + \log \mu_t\} + (1-a)\{-\int_x^t \mu_s \, ds\}$$

$$= -\int_x^t \mu_s \, ds + a \log \mu_t$$

Note that if $a = 1$, i.e., death occurs before withdrawal, this reduces to

$$-\int_x^y \mu_s \, dx + \log \mu_y$$

as before and if $a = 0$, i.e., death occurs after withdrawal, it reduces to

$$-\int_x^z \mu_s \, ds.$$

We now develop the likelihood function for a sample of n lives for each of the following seven types of data. Some of the types may not be of actuarial interest, becuase they are not exactly the ways traditional actuarial data have been recorded, but are of interest to statisticians and are included for the sake of completeness. It is hoped that in the future, actuaries will not feel restricted to the traditonal forms of data collection.

The likelihood function is proportional to the product of the individual likelihood functions of the sample members. The logarithm of the likelihood (log likelihood) is then the sum of the individual log likelihoods plus an additive constant which is independent of the parameters. The log likelihoods for each of the seven situations are given below. Since we are concerned with the maximization of the log likelihood with respect to the parameters, the additive constant is ignored in the expressions of the log likelihoods. The subscript i will be used to index the ith sample member.

Type I. <u>Complete Sample</u>:

The exact age at death, Y_i, is recorded for each sample member. The ages x_i may be different for different members of the sample.

$$l = -\sum_{i=1}^{n} \int_{x_i}^{y_i} \mu_s \, ds + \sum_{i=1}^{n} \log \mu_{y_i}$$

Type II. <u>Incomplete Sample</u>:

For each sample member, either the exact age at death, Y_i, or the exact age at withdrawal prior to death, Z_i, is recorded. The ages x_i may be different for different members of the sample.

$$l = -\sum_{i=1}^{n} \int_{x_i}^{t_i} \mu_s \, ds + \sum_{i=1}^{n} a_i \log \mu_{t_i}$$

Type III. <u>Grouped Sample</u>:

Suppose that all members enter the study at the same age, i.e. $x_i = x$, $i=1,2,\ldots,n$. Let $x = s_0 < s_1 < \ldots < s_J < s_{J+1} = \infty$

be the limits of $J+1$ age intervals. For a sample of n members the frequency of deaths in each interval is recorded.

$$l = -\sum_{j=1}^{J+1} n_j \int_x^{s_{j-1}} \mu_s \, ds + \sum_{j=1}^{J} n_j \log {}_{s_j - s_{j-1}}q_{s_{j-1}}$$

where n_j is the number of deaths in the j-th interval, for $j = 1, 2, \ldots, J+1$.

Type IV. Grouped Incomplete Sample:

Let the age intervals be as the grouped sample case. The ages x_i are restricted to the values s_0, s_1, ..., s_{J+1}. Withdrawals are restricted to the ages s_1, s_2, ..., s_J. The frequency of entrants and withdrawals are recorded at each age s_0, s_1, ..., s_J. The frequency of deaths in each interval is recorded.

$$l = -\sum_{j=1}^{J} \{n_{j+1} + w_j - e_j\} \int_{s_0}^{s_j} \mu_s \, ds + \sum_{j=1}^{J} n_j \log {}_{s_j - s_{j-1}}q_{s_{j-1}}$$

where w_j and e_j are the numbers of withdrawals and entrants at s_j respectively.

Type V. Partially Grouped Sample:

This is a generalization of the grouped sample which allows for the exact ages at death to be recorded in at least one of the first J age intervals

$$l = -\sum_1 \int_x^{y_i} \mu_s \, ds + \sum_1 \log \mu_{y_i} - \sum_2 n_j \int_x^{s_{j-1}} \mu_s \, ds + \sum_2 n_j \log {}_{s_j - s_{j-1}}q_{s_{j-1}}$$

where $\sum_1$ and $\sum_2$ are the summations over outcomes in intervals in which exact ages at death are recorded and over intervals in which only frequencies are recorded, respectively.

Type VI. Partially Grouped Incomplete Sample:

This is a generalization of the grouped incomplete sample which alllows for the exact ages at death to be recorded in at least one of the first J intervals.

$$l = -\sum_{j=0}^{J} (w_j - e_j) \int_{s_0}^{s_j} \mu_s \, ds - \sum_1 \int_{s_0}^{y_i} \mu_s \, ds + \sum_1 \log \mu_{y_i}$$

$$- \sum_2 n_j \int_{s_0}^{s_{j-1}} \mu_s \, ds + \sum_2 n_j \log {}_{s_j - s_{j-1}} q_{s_{j-1}}$$

Type VII. General Sample:

For each sample member either the exact age at death, the exact age at withdrawal or an age interval in which death occurred is recorded.

$$l = - \sum_A \int_{x_i}^{y_i} \mu_s \, ds + \sum_A \log \mu_{y_i} - \sum_B \int_{x_i}^{z_i} \mu_s \, ds$$

$$- \sum_C \int_{x_i}^{s'_i} \mu_s \, ds + \sum_C \log {}_{s''_i - s'_i} q_{s'_i}$$

where $\sum_A$ is summation over sample members whose exact age of death is recorded, $\sum_B$ is the summation over those whose age at withdrawal is recorded, $\sum_C$ is the summation over those whose age at death is known to have occurred the age interval (s'_i, s''_i) $i = 1, 2, \ldots, n$.

DEVELOPMENT OF THE LIKELIHOOD FUNCTION-MULTIPLE DECREMENT

The development here parallels that of the single decrement case. Consider a life entering the mortality at age x who dies at age y as a result of the kth cause.

(1) The numerical values of Y and K are recorded.

The likelihood function is

$$ {}_{y-x}p_x \; \mu_y^{(k)} $$

The corresponding log likelihood is

$$ -\int_x^y \mu_s \, ds + \log \mu_y^{(k)} $$

(2) Death is known to have occurred in the time interval $(s', s'']$ as a result of the kth cause.

The likelihood function is

$$ {}_{s'-x}p_x \cdot {}_{s''-s'}q_{s'}^{(k)} . $$

The corresponding log likelihood is

$$ -\int_x^{s'} \mu_s ds + \log {}_{s''-s'}q_{s'}^{(k)} . $$

(3) Either the age at death and the cause or the age at withdrawal is recorded.

Let the random variables T, A and Z be the same as is the single decrement case. The likelihood function is

$$ \left\{ {}_{t-x}p_x \mu_x^{(k)} \right\}^{\alpha} \quad \left\{ {}_{t-x}p_x \right\}^{1-a} $$

We now develop the likelihood function for a sample of n lives for each of the seven types of data considered in the single decrement case. The log likelihood is given as before.

Type I. Complete Sample:

$$l = \sum_{i=1}^{n} \left\{ -\int_{x_i}^{y_i} \mu_s \, ds + \log \mu_{y_i}^{(k_i)} \right\}$$

where k_i is the cause of death of the ith sample member.

This can be rewritten as

$$l = \sum_{k=1}^{m} \left\{ -\sum_{i=1}^{n} \int_{x_i}^{y_i} \mu_s^{(k)} \, ds + \sum_k \log \mu_{y_i}^{(k)} \right\}$$

where $\sum_k$ is the sum over sample members dying of the kth cause.

Type II. Incomplete Sample:

$$l = \sum_{i=1}^{n} \left\{ -\int_{x_i}^{t_i} \mu_s \, ds + a_i \log \mu_{t_i}^{(k_i)} \right\}$$

$$= \sum_{k=1}^{m} \left\{ -\sum_{i=1}^{n} \int_{x_i}^{t_i} \mu_s^{(k)} \, ds + \sum_k \log \mu_{t_i}^{(k_i)} \right\}$$

Type III. Grouped Sample:

$$l = -\sum_{j=1}^{J+1} n_j \int_{x}^{s_{j-1}} \mu_s ds + \sum_{j=1}^{J} \sum_{k=1}^{m} n_{j,k} \log_{s_j - s_{j-1}} q_{s_{j-1}}^{(k)}$$

$$= \sum_{k=1}^{m} \left\{ -\sum_{j=1}^{J+1} n_j \int_{x}^{s_{j-1}} \mu_s^{(k)} \, ds + \sum_{j=1}^{J} n_{j,k} \log_{s_{j-1} - s_j} q_{s_{j-1}}^{(k)} \right\}$$

assuming that the cause of death of survivors to age s_J, i.e. those dying after age s_J, is unknown; where $n_{j,k}$ is the number of deaths due to the kth cause in the jth age interval.

Type IV. Grouped Incomplete Sample:

$$l = -\sum_{j=1}^{J+1} \{n_j + w_{j-1} - e_{j-1}\} \int_{s_0}^{s_{j-1}} \mu_s \, ds + \sum_{j=1}^{J} \sum_{k=1}^{m} n_{j,k} \log {}_{s_j - s_{j-1}}q_{s_{j-1}}^{(k)}$$

$$= \sum_{k=1}^{m} [- \sum_{j=1}^{J+1} \{n_j + w_{j-1} - e_{j-1}\} \int_{s_0}^{s_{j-1}} \mu_s^{(k)} ds + \sum_{j=1}^{J} n_{j,k} \log {}_{s_j - s_{j-1}}q_{s_{j-1}}^{(k)}]$$

again assuming that the cause of death for survivors to age s_J is unknown.

Type V. Partially Grouped Sample:

$$l = -\sum\nolimits_1 \int_x^{y_i} \mu_s \, ds + \sum\nolimits_1 \mu_{y_i}^{(k_i)}$$

$$= -\sum\nolimits_2 n_j \int_x^{s_{j-1}} \mu_s \, ds + \sum\nolimits_2 \sum_{k=1}^{m} n_{j,k} \log {}_{s_j - s_{j-1}} q_{s_{j-1}}^{(k)}$$

$$= \sum_{k=1}^{m} [-\sum\nolimits_1 \int_x^{y_i} \mu_s^{(k)} \, ds + \sum\nolimits_{1,k} \log \mu_{y_i}^{(k)}$$

$$-\sum\nolimits_2 n_j \int_x^{s_{j-1}} \mu_s^{(k)} \, ds + \sum\nolimits_2 n_{j,k} \log {}_{s_j - s_{j-1}} q_{s_{j-1}}^{(k)}]$$

where $\sum_{1,k}$ is the summation over all deaths due to the kth cause whose exact age of occurrence is recorded.

Type VI. Partially Grouped Incomplete Sample:

$$l = -\sum_{j=0}^{J} (w_j - e_j) \int_{s_0}^{s_j} \mu_s \, ds - \sum_1 \int_{s_0}^{y_i} \mu_s \, ds + \sum_1 \log \mu_{y_i}^{(k_i)}$$

$$-\sum_2 n_j \int_{s_0}^{s_{j-1}} \mu_s \, ds + \sum_2 \sum_{j=1}^{J} n_{j,k} \log {}_{s_j - s_{j-1}} q_{s_{j-1}}^{(k)}$$

$$= \sum_{k=1}^{m} [- \sum_{j=0}^{J} (w_j - e_j) \int_{s_0}^{s_j} \mu_s^{(k)} \, ds - \sum_1 \int_{s_0}^{y_i} \mu_s^{(k)} \, ds + \sum_{1,k} \log \mu_{y_i}^{(k)}$$

$$-\sum_2 n_j \int_{s_0}^{s_{j-1}} \mu_s^{(k)} \, ds + \sum_2 n_{j,k} \log {}_{s_j - s_{j-1}} q_{s_{j-1}}^{(k)}]$$

Type VII. General Sample:

$$l = -\sum_A \int_{x_i}^{y_i} \mu_s \, ds + \sum_A \log \mu_{y_i}^{(k_i)} - \sum_B \int_{x_i}^{z_i} \mu_s \, ds$$

$$-\sum_C \int_{x_i}^{s_i'} \mu_s \, ds + \sum_C \log {}_{s_i'' - s_i'} q_{s_i'}^{(k_i)}$$

$$= \sum_{k=1}^{m} [-\sum_A \int_{x_i}^{y_i} \mu_s^{(k)} \, ds + \sum_{A,k} \log \mu_{y_i}^{(k)} - \sum_B \int_{x_i}^{z_i} \mu_s^{(k)} \, ds$$

$$-\sum_C \int_{x_i}^{s_i'} \mu_s^{(k)} \, ds + \sum_{C,k} \log {}_{s_i'' - s_i'} q_{s_i'}^{(k)}]$$

where $\sum_{A,k}$ and $\sum_{C,k}$ are the corresponding summations over deaths due to the kth cause.

An interesting, although not surprising, result emerges from examination of the log likelihoods. In Types I and II the log likelihood of the sample can be written as the sum of m functions each involving only a single force of mortality. Thus it can be written as

$$l = l^{(1)} + l^{(2)} + \ldots + l^{(m)}$$

where $l^{(k)}$ involves only the kth force of mortality. Thus $dl/d\theta = dl^{(k)}/d\theta$ if θ is a paratmeter of the kth force of mortality. The estimation can then be done separately for each force of mortality.

When there is some grouping present, however, this result does not hold. Notice that each of Types III to VII involves terms of the form ${}_{s_j - s_{j-1}}q^{(k)}_{s_{j-1}}$ which can be written as

$$\int_{s_{j-1}}^{s_j} e^{-\int_{s_{j-1}}^{t} \mu_s \, ds} \, u_t^{(k)} \, dt.$$

They cannot be decomposed as the sum of terms involving only one force of mortality. The estimation of all of the parameters of all the forces of mortality must then be done simultaneously.

For Types I and II, i.e. when there is no grouping, the form of the portion of the log likelihood involving a particular force of mortality in the multiple decrement model is the same as the log likelihood for the single decrement model using the same force of mortality. Deaths due to other causes are treated as withdrawals with respect to that cause of death. Thus it suffices to study the single decrement model in the case of no grouping. Furthermore, since Type I is a special case of II, we need only study case II.

MORTALITY MODELS

In the next few sections, we shall investigate the maximum likelihood graduation procedure for some given forms of the force of mortality for some of the sampling situations described previously.

Mortality Model A.

The first model considered is the class of forces of mortality of the form $\mu_x = a(x)b(\theta)$ where $\int_y^\infty a(x)dx = \infty$ and $b(\theta)$ is a positive differentiable and strictly monotonic function of θ. Some examples follow:

AI. $\mu_x = \theta$

This is the constant force of mortality, sometimes associated with accidental deaths. The distribution of the age at death is exponential, a widely used model in life-time studies of things such as light bulbs, vacuum tubes and electronic components.

AII. $\mu_x = \theta x$

Here the force of mortality is linear. The corresponding distribution is the Rayleigh.

AIII. $\mu_x = \theta/x$

The force of mortality decreases with age in this case. The corresponding distribution is the Pareto.

In our application we need only assume that the force of mortalit is of the specified form at ages above the lowest under consideration. Alternately, if the force of mortality is to be of a specified form over a certain age, each member in the sample observed at that age can be treated as an entrant at that age. In this way data at ages below the specified age are ignored. The corresponding distributions are then truncated at that specified lower age.

Mortality Model B.

The second model is the force of mortality of the form

$$\mu_x = \frac{1}{\sigma}\, exp\{\frac{x-\mu}{\sigma}\} \ .$$

The force of mortality increases exponentially with age. Note that if we let $B = exp\{-\mu/\sigma\}/\sigma$ and $c = exp\{1/\sigma\}$ the force of mortality can be rewritten as $\mu_x = Bc^x$, the Gompertz model. We prefer to use the former notation rather than the traditional form since μ and σ are location and scale parameters, making the computations some what less complicated. Also, the mode of the distribution, i.e. the age with highest frequency of deaths is $x = \mu$. Thus if we are given μ and σ we have somewhat better intuitive feeling for the shape of the mortality frequency curve than if we are given B and C.

Model A will be investigated in both sample types where there is no grouping present, i.e. Types I and II, and, the types where there is grouping present, i.e. Types III - VII. A corresponding investigation will be done for the multiple decrement model where each of the forces of mortality is of the form $\mu_x^{(k)} = a_k(x)\ b_k(\theta)$ in the sample types involving grouping only since in the ungrouped cases the likelihood equations will each be of the same form as in the single decrement case. Corresponding studies investigation will be made of Model B. Finally, a combination of the two models, of which the Makeham model is a special case, will be studied.

MAXIMUM LIKELIHOOD SOLUTION PROCEDURE FOR MORTALITY MODEL A.

We now consider models using the force of mortality of the form $\mu_x = a(x)\ b(\theta)$. Solution procedures for obtaining the maximum likelihood estimates through solving the likelihood equations for both the single and multiple decrement cases will be proposed.

(1) <u>Types I and II samples for single and multiple decrement cases</u>

It is sufficient to consider a Type II sample since Type I is a special case of Type II. For the single decrement case, the log likelihood is

$$\ell = -\sum_{i=1}^{n} \int_{x_i}^{t_i} a(s)ds\ b(\theta) + n_A \ log\ b(\theta) + constant$$

where n_A is the number of observed deaths and the constant does not involve θ. Setting the derivative with respect to θ of the log likelihood to zero yields the likelihood equation. Solving the equation we have

$$\hat{\theta} = b^{-1}\{n_A \Big/ \sum_{i=1}^{n} \int_{x_i}^{t_i} a(s)\ ds\}$$

where b^{-1} is the inverse function of b. For the multiple decrement case the maximum likelihood estimate, $\hat{\theta}_k$, corresponding to the kth cause is of the above form with n_A substituted by n_k, the number of deaths due to the kth cause.

(2) Types III to VII samples for single and multiple decrement cases

It is sufficient to study a Type VII sample since the other types are special cases. For the single decrement case, the log likelihood is

$$\ell = -\sum_A \int_{x_i}^{y_i} a(s)\ ds\ b(\theta) + n_A \log b(\theta) - \sum_B \int_{x_i}^{z_i} a(s)\ ds\ b(\theta)$$

$$-\sum_C \int_{x_i}^{s_i'} a(s)\ ds\ b(\theta) + \sum_C \log_{s_i'' - s_i'} q_{s_i'} + \text{constant}$$

where n_A is the number of members whose exact ages at death are recorded. For notational convenience, let $c(x,y) = \int_x^y a(s)\ ds$. The term $_{s_i'' - s_i'}q_{s_i'}$ can then be written as

$$\int_{s_i'}^{s_i''} \exp\{-c(s_i',s)\ b(\theta)\}c(s_i',s)\ b(\theta)\ ds.$$

The log likelihood is then

$$\ell = -\sum_A c(x_i,y_i)\ b(\theta) + n_A \log b(\theta) - \sum_B c(x_i,z_i)\ b(\theta)$$

$$-\sum_C c(x_i s_i')b(\theta) + n_C \log b(\theta) + \sum_C \log \int_{s_i'}^{s_i''} \exp\{-c(s_i',s)b(\theta)\}c(s_i',s)ds$$

Let $$g_i = [c(x_i,s_i'') + c(x_i,s_i')]/2$$

and $$h_i = [c(x_i,s_i'') - c(x_i,s_i')]/2\ .$$

Substituting these into the log likelihood, and then differentiating log likelihood with respect to θ, the likelihood equation becomes ($db(\theta)/d\theta \neq 0$ since $b(\theta)$ is differentiable and strictly monotonic)

$$n_A/b(\theta) + \sum_C h_i \coth(b(\theta)h_i/2)/2 = \sum_A o(x_i,y_i) + \sum_B c(x_i,z_i) + \sum_C g_i$$

where $\coth(x)$ is the hyperbolic contangent function, i.e. $\coth(x) = (e^x + e^{-x})/(e^x - e^{-x})$. It can be shown [§4.2] that there exists a unique solution to this equation if at least one death occurs and that the solution corresponds to a local maximum of the likelihood function.

We now proceed to determine the solution numerically

Let $$g(y) = n_A y + \sum_C h_i \coth(h_i/y), \quad y > 0,$$

and let $y_0 = [\sum_A o(x_i,y_i) + \sum_B c(x_i,z_i) + \sum_C g_i]/(n_A + n_C)$

The likelihood equation can then be expressed as

$$g(y) = (n_A + n_C)\ y_0 \quad \text{where } y = 1/b(\theta)$$

and n_C is the total number of deaths recorded in intervals.

Suppose that its root is $y^* = 1/b(\theta^*)$. It can be shown [§4.2], that $g(y)$ is strictly increasing and convex with slope not exceeding $n_A + n_C$ for $y > 0$ and $n_A + n_C > 0$. Also it can be shown [§ 4.2] that $y_0 > y^*$.

Let $\{y_1, y_2, \ldots\}$ be a sequence defined by

$$y_{j+1} = y_j - \{g(y_j) - (n_A + n_C)/y_0\}/(dg(y_j)/dy_j)\ .$$

Then it can be shown [§4.2], that $y_0 > y_1 > \cdots > y^*$ and that the sequence $\{y_0, y_1, y_2, \ldots\}$ converges to y^*. This sequence can then be used to determine the root iteratively. The iterative procedure is based on the Newton-Raphson technique given in most texts on numerical methods. A simpler technique can be obtained by substituting $n_A + n_C$ for $g'(y_j)$.

Explicit bounds on the root of the likelihood equation can be found by using the series expansion of $coth(x)$ which is

$$coth(x) = 1/x + x/3 - x^3/45 + \ldots$$

If the first term only is used the likelihood equation becomes

$$(n_A + n_C)/b(\theta) = \textstyle\sum_A c(x_i, y_i) + \sum_B c(x_i, z_i) + \sum_C g_i$$

which has solution $\theta = \theta_0$ where $y_0 = 1/b(\theta_0)$. Using two terms of the approximation the likelihood equation is then approximately

$$(n_A + n_C)/b(\theta) + b(\theta) \textstyle\sum_C h_i^2/3 = (n_A + n_C)/b(\theta_0)$$

which has solution θ_u where

$$\frac{1}{b(\theta_u)} = \frac{1}{2b(\theta_0)} \left\{ 1 + \sqrt{1 - \frac{\sum_C 4h_i^2 b^2(\theta_0)}{3(n_A + n_C)}} \right\} .$$

Finite sample properties of $b(\theta_0)$ and $b(\theta_u)$ were studied by Kendell and Anderson (1971) in the case of a grouped sample from an exponential distribution. It can be shown [§4.2] that $b(\theta_u) \geq b(\theta^*)$. Then explicit bounds of θ^* can be determined since

$$b(\theta_0) < b(\theta_1) < \cdots < b(\theta^*) \leq b(\theta_u) .$$

Evaluation of $b(\theta_0)$ and $b(\theta_u)$ provides bounds on the true values of $b(\theta^*)$.

We now consider the Type VII sample for the multiple decrement case when the force of mortality for each cause of death has the form

$$\mu_x^{(k)} = a_k(x)\ b_k(\theta_k)\ .$$

Let n_{Ak} and n_{Ck} denote the numbers of deaths of the kth cause where the exact ages of death are recorded and where intervals of death are recorded respectively. If $a_k(x) = a(x)$, $k = 1, 2, 3, \ldots, m$, then $\sum_{k=1}^{m} \mu_x^{(k)} = a(x) \sum_{k=1}^{m} b_k(\theta_k)$. Let $b = \sum_{k=1}^{m} b_k(\theta_k)$. Then the likelihood equations are

$$(n_{Ak} + n_{Ck})/b_k(\theta_k) - n_C/b + \textstyle\sum_C h_i \coth(b\ h_i)$$

$$= \textstyle\sum_A c(x_i, y_i) + \sum_B c(x_i, z_i) + \sum_C g_i, \qquad k = 1, \ldots, m.$$

As a result we also have

$$(n_{Ak} + n_{Ck})/b_k(\theta_k) = (n_A + n_C)/b, \qquad k = 1, \ldots, m.$$

It can be shown [§4.3], that the solution of equations is unique and provides a local maximum of the likelihood function.

If $a_1(x)$, $a_2(x), \ldots, a_k(x)$ are not the same, the system of likelihood equations is

$$(n_{Ak} + n_{Ck})/b_k(\theta_k) = \sum_A c_k(x_i,y_i) + \sum_B c_k(x_i,z_i)$$

$$+ \sum_{\alpha=1}^{m} \sum_{C\alpha} \frac{\int_{s_i'}^{s_i''} a_\alpha(s)\; c_k(x_i,s)\; exp\{-\sum_{\beta=1}^{m} c_\beta(x_i,s)\; b_\beta(\theta_\beta)\}\; ds}{\int_{s_i'}^{s_i''} a_\alpha(s)\; exp\{-\sum_{\beta=1}^{m} c_\beta(x_i,s)\; b_\beta(\theta_\beta)\}\; ds}, \quad k = 1,\ldots, m$$

where $c_k(x,y) = \int_x^y a_k(s)\; ds.$

One way of solving this system of equations approximately is to substitute for the ratio either the average of $c_k(x_i,s_i')$ and $c_k(x_i,s_i'')$ or the value of $c_k(x_i,s)$ at $s = (s_i' + s_i'')/2$. In either case, the approximate likelihood equations have explicit solutions.

Boardman (1973) studies finite sample properties of the estimates when $\mu_x^{(k)} = \lambda_k$, $k = 1,\ldots, m$ for grouped samples.

MAXIMUM LIKELIHOOD SOLUTION PROCEDURE FOR MORTALITY MODEL B

We now consider models using the force of mortality of the form $\mu_x = \frac{1}{\sigma} exp(\frac{x - \mu}{\sigma})$, the Gompertz model.

(1) <u>Types I and II samples for the single decrement case</u>

The log likelihood for the Type I sample is

$$-\sum_{i=1}^{n} \{exp(\frac{y_i - \mu}{\sigma}) - exp(\frac{x_i - \mu}{\sigma})\} + \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma} - n\; log\; \sigma$$

and for the Type II sample is

$$-\sum_{i=1}^{n} \{exp(\frac{t_i - \mu}{\sigma}) - exp(\frac{x_i - \mu}{\sigma})\} + \sum_A \frac{x_i - \mu}{\sigma} - n_A\; log\; \sigma$$

The corresponding likelihood equations for Type I samples are

$$\frac{\sum\{y_i\ exp(y_i/\sigma)\ -\ x_i\ exp(x_i/\sigma)\}}{\sum\{exp(y_i/\sigma)\ -\ exp(x_i/\sigma)\}} = \bar{y} + \sigma,\ where\ \bar{y} = \sum y_i/n$$

$$\sum\{exp(y_i/\sigma)\ -\ exp(x_i/\sigma)\} = n\ exp(\mu/\sigma)$$

and for Type II samples are

$$\frac{\sum\{t_i\ exp(t_i/\sigma)\ -\ x_i\ exp(x_i/\sigma)\}}{\sum\{exp(t_i/\sigma)\ -\ exp(x_i/\sigma)\}} = \bar{y} + \sigma,\ where\ \bar{y}_A = \sum_A t_i/n_A$$

$$\sum\{exp(t_i/\sigma)\ -\ exp(x_i/\sigma)\} = n_A\ exp(\mu/\sigma)$$

It can be shown [§2.2] that:

1. Any solution provides a local maximum of the likelihood function;

2. There exists a solution to the likelihood equations if and only if

$$\sum_{i=1}^{n} (x_i - \bar{y})^2 > \sum_{i=1}^{n} (y_i - \bar{y})^2 \quad for\ a\ Type\ I\ sample$$

and

$$\sum_{i=1}^{n} (x_i - \bar{y})^2 > \sum_{i=1}^{n} (t_i - \bar{y})^2\ for\ a\ Type\ II\ sample;$$

3. If a solution exists, it is unique.

So, if the condition in (2) is satisfied, the solution of the likelihood equations provides the maximum likelihood estimates of the parameters.

The conditions in (2) involve Euclidean distance functions and as such can be interpreted as follows. The distances of the ages at entry from the average at death must be greater than the distance of the ages at withdrawal or death from the average age at death. Only in extreme cases is this condition not satisfied.

To numerically obtain the solutions of the likelihood equations standard methods can be used. Note that the first equation involves only σ. It can be rewritten as

$$g(\sigma) = \sigma$$

by transposition of $\bar{y}$ to the left side of the equation. Two iterative procedures are considered. The first is

$$\sigma_{j+1} = \sigma_j - \{g(\sigma_j) - \sigma_j\}/\{g'(\sigma_j) - 1\} \quad where\ g' = dg/d\sigma$$

which is based on the Newton-Raphson method. The second is

$$\sigma_{j+1} = g(\sigma_j)$$

which is simpler in the form but converges at a lower rate since convergence is linear while for the Newton-Raphson method, convergence is quadratic. However the second procedure converges in some cases for which the first diverges. For example $f(\sigma) - \sigma$ for a Type I sample is a decreasing function that is asymptotic to

$$\sum_{i=1}^{n} (y_i^2 - x_i^2)/\sum_{i=1}^{n} (y_i - x_i) - \bar{y}$$

for large values of σ. As a result its slope approaches zero from below. If any of the successive values of σ_j using the Newton-Raphson method is large, the tangent at that point may intersect the σ-axis at large negative value of σ, resulting in divergence. When the initial value σ_0 is chosen sufficiently close to the root, the Newton-Raphson method is known to converge.

(2) Types III and IV samples from the single decrement case

We shall consider the grouped incomplete sample, i.e. the Type IV sample. To simplify notation, for $j = 1, \ldots, J$, let

$$g_j = [exp(s_j/\sigma) + exp(s_{j-1}/\sigma)]/2,$$

$$g_0 = exp(s_0/\sigma),$$

$$g_j = exp(s_j/\sigma),$$

$$h_j = [exp(s_j/\sigma) - exp(s_{j-1}/\sigma)]/2,$$

$$l_j = [s_j \; exp(s_j/\sigma) + s_{j-1} \; exp(s_{j-1}/\sigma)]/2,$$

$$l_0 = s_0 \; exp(s_0/\sigma),$$

$$l_j = s_j \; exp(s_j/\sigma),$$

$$m_j = [s_j \; exp(s_j/\sigma) - s_{j-1} \; exp(s_{j-1}/\sigma)]/2 \; .$$

Also let $g_{J+1} = g_J$ and $l_{J+1} = l_J$.

Using these substitutions the likelihood equations become

$$\sum_{j=1}^{J+1} (n_j + w_j - e_j) \; g_i + \sum_{j=1}^{J} (w_j - e_j) \; h_j - \sum_{j=1}^{J} n_j h_j \; coth(c \; h_j) = 0$$

$$\sum_{j=1}^{J+1} (n_j + w_j - e_j) \quad {}_j + \sum_{j=1}^{J} (w_j - e_j) \; m_j - \sum_{j=1}^{J} n_j m_j \; coth(c \; h_j) = 0$$

where $c = exp(\mu/\sigma)$.

It can be shown [§2.5] that any solution to the above equations provides either a local maximum or a saddle point of the likelihood function. By examining the likelihood function in the neighbourhood of the solution of the likelihood equations one can determine if in fact a maximum is obtained.

To obtain the solution of the likelihood equations numerically, the Newton-Raphson method in two variables can be used. If the first approximation to the root is not sufficiently close to the true root, the procedure may diverge. A first approximation may be made by treating all deaths in intervals as having occurred at the corresponding midpoints and using the methods for the ungrouped cases.

Another method of obtaining the maximum of the log likelihood function is to directly compute its value in the neighbourhood of the initial guess and plotting its contours. This method will clearly show where the maximum occurs.

For the sake of brevity, other types of samples and the multiple decrement case shall not be considered here; however, similar results were given in Chapters 2 and 3 of Panjer (1975) can be assumed to hold.

THE TWO CAUSE MAKEHAM MODEL AND THE MAXIMUM LIKELIHOOD SOLUTION PROCEDURE

In this section we consider, as a further example, the Makeham model with force of mortality

$$\mu_x = a + Bc^x = a + \frac{1}{\sigma} exp(\frac{x - \mu}{\sigma}) \ .$$

We shall assume that deaths are separated by cause into two groups, the first with force of mortality

$$\mu_x^{(1)} = a$$

and the second with force of mortality

$$\mu_x^{(2)} = \frac{1}{\sigma} \exp\left(\frac{x-\mu}{\sigma}\right) .$$

For the general sample, i.e., Type VII, when there are no outcomes in intervals, the likelihood equations are

$$n_{A1}/a = \textstyle\sum_A (y_i - x_i) + \sum_B (z_i - x_i)$$

$$n_{A2} \exp(\mu/\sigma) = \textstyle\sum_A \exp(y_i/\sigma) + \sum_B \exp(z_i/\sigma) - \sum_{i=1}^{n} \exp(x_i/\sigma)$$

$$\frac{\sum_A y_i \exp(y_i/\sigma) + \sum_B z_i \exp(z_i/\sigma) - \sum_{i=1}^{n} x_i \exp(x_i/\sigma)}{\sum_A \exp(y_i/\sigma) + \sum_B \exp(z_i/\sigma) - \sum_{i=1}^{n} \exp(x_i/\sigma)} = \frac{\sum_{A2} y_i}{n_{A2}} + \sigma$$

It can be shown [§4.4] that the likelihood function possesses at most one local maximum which exists if and only if $n_{A1} > 0$, $n_{A2} > 1$ and

$$\sum_{i=1}^{n} (x_i - \bar{y}_{A2})^2 > \textstyle\sum_A (y_i - \bar{y}_{A2})^2 + \sum_B (z_i - \bar{y}_{A2})^2 \text{ where } \bar{y}_{A2} = \sum_{A2} y_i/n_{A2} .$$

When there are outcomes in intervals, the likelihood equations are somewhat more complicated since the two forces of mortality cannot be isolated. Let n_{C1}, n_{C2}, $\sum_{C1}$ and $\sum_{C2}$ correspond to the previous symbols but for outcomes in intervals. The likelihood equations are

$$(n_{A1} + n_{C1})/a = \sum_A (y_i - x_i) + \sum_B (z_i - x_i)$$

$$+ \sum_{C1} I_{3i}/I_{1i} + \sum_{C2} I_{4i}/I_{2i} ,$$

$$\{\sum_{A2} y_i + (n_{A2} + n_{C2})\sigma\} \exp(\mu/\sigma)$$

$$= \sum_A \exp(y_i/\sigma) + \sum_B \exp(z_i/\sigma) - \sum_{i=1}^{n} \exp(x_i/\sigma)$$

$$+ \sum_{C1} I_{2i}/I_{1i} + \sum_{C2} I_{5i}/I_{2i} ,$$

$$(n_{A1} + n_{C1})\sigma \exp(\mu/\sigma) = \sum_A y_i \exp(y_i/\sigma) + \sum_B z_i \exp(z_i/\sigma)$$

$$- \sum_{i=1}^{n} x_i \exp(x_i/\sigma) + \sum_{C1} I_{6i}/I_{1i} + \sum_{C2} I_{7i}/I_{2i}$$

where

$$I_{ji} = \int_{c_i}^{d_i} J_{ji} \exp\{-a(x - x_i) - \exp(\frac{x - \mu}{\sigma})\} \, dx$$

and

$$J_{1i} = 1, \qquad J_{2i} = \exp(x/\sigma), \qquad J_{3i} = x - x_i$$

$$J_{4i} = (x - x_i)\exp(x/\sigma), \qquad J_{5i} = \exp(2x/\sigma),$$

$J_{6i} = x \exp(x/\sigma)$ and $J_{7i} = \{\exp(x/\sigma) - \exp(\mu/\sigma)\} x \exp(x/\sigma)$

To obtain an initial estimate one could evaluate each of the ratios of integrals at the midpoints of the corresponding intervals. The likelihood equations then reduce to those of the case without grouping.

ASYMPTOTIC PROPERTIES OF THE MAXIMUM LIKELIHOOD ESTIMATES

Uvell (1973) showed that if certain regularity conditions analogous to those imposed by Cramer (1946), are satisfied, there exists with probability tending to unity as $n \to \infty$, one and only one solution of the likelihood equation which is a consistent estimate. Moreover, he shows that this solution provides a local maximum of the likelihood function. He also proves that the solution is asymptotically distributed as the (multivariate) normal distribution.

It can be shown [§1.5, 2.10, 3.7, 4.5] that each of the models considered in this paper satisfy the regularity conditions. In general, models with continuous smooth forces of decrement will satisfy the conditions of Uvell's theory. As a result the user can assume that the solution corresponding to the maximum of the likelihood function has the properties described above. The elements in the inverse of the asymptotic covariance matrix of the maximum likelihood estimates can be estimated by the negative of the second order partial derivatives of the log likelihood function divided by the sample size.

CONCLUSIONS

The main purpose of this paper was to point out and demonstrate the use of the maximum likelihood approach to graduation when an analytic form of the forces of decrement is used. The approach is a unified one that can be applied to any form of data. It can be used for mortality models as well as models involving more general decrements. The general technique is simple; write down the likelihood function and find where it is maximized.

The techniques for specific models can be somewhat complicated depending on the form of the forces of decrement involved and the form of the data.

The maximum likelihood method can be applied to situations in which the force of mortality is not of a single analytic form over the whole age range considered, but of separate forces over the various subranges. Data with a select period can be handled by treating forces of decrement during the select period for each age at entry as being of a specific analytic form different from that of the ultimate period; or the select period can be ignored altogether by treating each survivor to the ultimate period as an entrant at that time.

No criteria have been given for the selection of the appropriate mortality or morbidity model. This is outside the scope of this paper but, of course, a problem that must be dealt with before using the methods given here.

Department of Mathematics
The University of Western Ontario
London, Ontario, Canada N6A 5B9

REFERENCES

1. Boardman, Thomas J. (1973). "Estimation in Compound Exponential Failure Models--When the Data are Grouped", *Technometrics*, 15, No. 2, pp. 271-277.

2. Cramer, H. (1946). *Mathematical Methods of Statistics*, Princeton University Press.

3. Greville, T.N.E. (1974). *Part 5 Study Notes-Graduation*, Society of Actuaries.

4. Hickman, James C. (1964). "A Statistical Approach to Premiums and Reserves in Multiple Decrement Theory", *Transactions of the Society of Actuaries*, 16, pp. 1-16.

5. Kendall, P.J. and R.L. Anderson (1971). "An Estimation Problem in Life-Testing", *Technometrics*, 13, No. 2, pp. 289-301.

6. Kimeldorf, George S. and Donald A. Jones (1967). "Bayesian Graduation", *Transactions of the Society of Actuaries*, 19, pp. 66-112.

7. Miller, Morton D. (1946). *Elements of Graduation*. (Actuarial Monographs", No. 1) Actuarial Society of America and American Institute of Actuaries.

8. Panjer, H.H. (1975). *Maximum Likelihood Estimation from Distributions Useful in Actuarial Applications*, Ph.D. Thesis. University of Western Ontario.

9. Uvell, S. (1973). *Maximum Likelihood Estimates from Partially Grouped Samples from Independent not Identically Distributed Random Variables I.* Technical Report No. 1973-2, University of Umeå, Sweden.

SOME INSIGHTS INTO ALLOCATING

THE FEDERAL INCOME TAXES

by
Charles E. Johnson

[Author's Note: The author does not intend to be original with the material presented in this paper but only intends to express the views contained herein.]

EXTRACT AND INTRODUCTION

The purpose of this paper is to look at some approaches to allocating the Federal Income Taxes among various lines of business. For this purpose, the Federal Income Taxes for a company may be broken down into a Separate Company Tax for each line of business and a Marriage Tax for the company as a whole.

ALLOCATING THE FEDERAL INCOME TAXES

For allocation purposes, the Federal Income Taxes may be divided into two parts:

(1) a Separate Company Tax and

(2) a Marriage Tax.

The Separate Company Tax is the tax calculated for each line of business treating the line of business as if it was a separate company in the same tax position* as the company as a whole. The Marriage Tax (positive or negative) is the difference between the tax for the company as a whole and the sum of the Separate Company Taxes. Therefore, the main problem in allocating the Federal Income Taxes among various lines of business is how to allocate the Marriage Tax.

How does each line of business contribute to the overall tax for the company as a whole? The answer to this question may lie in the All But Method. This method determines a tax for each line of business by taking the difference between the tax for the company as a whole and the tax for the same company but which excludes the contribution by that line of business to each element in calculating the tax assuming the same tax position as that for the company as a whole. This All But Tax may be broken down into a Separate Company Tax and a Marriage Tax from the All But Method. This Marriage Tax for each line of business is the difference between the All But Tax for that line of business and its Separate Company Tax.

Can the Marriage Tax for the company as a whole be allocated among various lines of business according to the All But Marriage Taxes? Fact: An allocation method should have at least two characteristics in common with the thing it is allocating:

(1) a directional characteristic, for example, it should allocate the Marriage Tax to each line of business according to the direction the Marriage Tax is actually taking for that line of business, and

(2) a magnitudinal characteristic, for example, it should also allocate the Marriage Tax to each line of business according to how far in either direction the Marriage Tax is actually going for that line of business.

Therefore, allocating the Marriage Tax for the company as a whole among various lines of business according to the All But Marriage Taxes would appear to be theoretically plausible. But, this may not be practical because the Marriage Tax for the company as a whole may be large compared to the actual tax (unknown, but supposedly approximated by the Separate Company Tax) for any one line of business, causing an allocated tax which is a wide deviation from the Separate Company Tax.

Another approach is to use the Separate Company Tax as a smoothing factor by allocating the Marriage Tax for the company as a whole according to the All But Taxes. This approach tends to give a tax for each line of business which is a smooth ascent (or descent) from the Separate Company Tax--as well as being theoretically plausible.

And, of course, there are other techniques...such as allocating the Federal Income Taxes according to the Separate Company Taxes.

CONCLUSION

For allocation purposes, the Federal Income Taxes for a company may be broken down into a Separate Company Tax for each line of business and a Marriage Tax for the company as a whole. The main problem in allocating the Federal Income Taxes among various lines of business is how to allocate the Marriage Tax.

*A tax position defines the elements through which taxes are paid.