ISBN 0 7340 2893 8

Statistical Case Estimation

by

Greg Taylor
Taylor Fry Consulting Actuaries, Sydney
and The University of Melbourne

and

Mireille Campbell
GIO Insurance Ltd, Sydney.

RESEARCH PAPER NUMBER 104

November 2002

Centre for Actuarial Studies
Department of Economics
The University of Melbourne
Victoria 3010
Australia

STATISTICAL CASE ESTIMATION

Greg Taylor

Taylor Fry Consulting Actuaries
Level 4, 5 Elizabeth Street
Sydney NSW 2000
Australia

Professorial Associate, Centre for Actuarial Studies
Faculty of Economics and Commerce
University of Melbourne
Parkville VIC 3052
Australia

greg@taylorfry.com.au

and

Mireille Campbell

GIO Insurance Ltd
Level 10
52-58 Clarence St
Sydney NSW 2000
Australia

mcampbell@gio.com.au

November 2002

Table of Contents

1. Introduction

1.1 Scope

The term **case estimate**, sometimes referred to as a **manual estimate** or **physical estimate**, is used in this paper to mean an estimate of outstanding claims liability in respect of a **single claim**.

Typically, these estimates are made by assessors who apply their experience of the claims environment to the task. The assessment of the claim will make **subjective** allowance for all its characteristics and circumstances. Occasionally, attempts are made to objectify assessment by means of an **expert system** which mimics the decision process of an experienced assessor.

An alternative means of rendering the assessment objective, and one that will be discussed in this paper, is **statistical case estimation**. This consists of:

- Codifying as many of the claim characteristics, or **attributes**, as possible;
- Constructing a function that maps the collection of these attributes to a liability value.

Because of the use of the attributes as arguments in the case estimate function, they will also be referred to as **covariates** (of the liability value). Some will be objective, eg age at claim occurrence, others may be less so, eg claim involves an injury that is "serious". Some covariates may be static, eg pre-injury earnings (for workers compensation claims); others may be dynamic, eg amount of benefit claimed to date.

The objective formulation of statistical case estimates (SCEs) makes them comparable to some extent with loss reserves obtained **in the aggregate** by statistical means. Often these can be viewed as involving very rudimentary SCEs, with only accident year and development year as covariates. Generally, such loss reserves will be referred to as **aggregate loss reserves**.

As a slightly more complex example, it is not uncommon to find workers compensation aggregate loss reserves differentiated by gender of claimant. Evidently, the more the differentiation of aggregate loss reserves, the greater their similarity to SCEs.

Indeed, aggregate loss reserves may be viewed as extreme cases of SCEs, and vice versa. In practical terms, however, techniques involved in the two are rather different, as will be discussed in the sections following.

The present paper will attempt to keep the development of SCEs as general as possible. However, the backdrop reflected by specific examples will reflect the workers compensation line of business (LOB). This is due to the authors' biases in experience, and the fact that SCEs perhaps have their greatest value in this LOB. This, however, is not to deny their feasibility and value in other LOBs.

1.2 Motivation

Case estimates (statistical or not) become useful whenever loss reserves are required for **small groups of claims**. Correspondingly, the smaller the group, the less useful become aggregate loss reserves.

For example, it might be estimated with considerable confidence that the average loss reserve per claim unsettled from the latest accident year of an entire portfolio is $15,000. However, assignment of a value of $45,000 to a group of 3 claims may not be at all sensible if one of the claims involves serious brain injury. In this case, even poor quality case estimates will out-perform high quality loss reserves.

The matter of case estimation versus aggregate loss reserving, and the bearing of claim sample size on this question, is discussed in Section 3.1 of Taylor (2000). This discussion can be largely summarised by the following points:

- The law of large numbers usually operates effectively for whole portfolios, but not for small groups of claims;
- Consequently, the information value of covariates is low in large samples, eg the various injury types will be proportionally represented in a large sample;
- Conversely, the information value of covariates is high in small samples.

There is a variety of reasons why estimates of liability may be required in respect of small groups of claims. In the case of workers compensation:

- there may be a need simply to provide small employers (with few claims each) with information on their claim costs
- experience rating of individual employers would require realistic estimates of their separate claim costs
- even within large employers, for which aggregate loss reserving methods may be effective, there may be a need to allocate claims to smaller sub-divisions, eg cost centres.

SCEs provide **objective** estimates of these liabilities. They also provide estimates that are **consistent** one with another, where the quality of manually constructed case estimates would vary with the skill of the estimators.

1.3 Structure of the paper

This paper discusses a number of the theoretical and practical considerations in the construction and maintenance of a SCE system. The discussion is developed in as generalised a context as possible. However, in the provision of illustrations of the generalisations and numerical examples, coherence has been best served by a single case.

All illustrations and examples are therefore drawn from the scheme of workers' compensation provided by one Australian state several years ago. This scheme is underwritten and supervised by the state, though claims are managed by a number of private sector insurers. The data set used in the numerical examples is scheme-wide.

2. Payment types and their dependencies

2.1 General commentary

Typically, losses paid in a specific line of business may be categorised according to **payment types**. In Liability LOBs these might be **heads of damage**, such as economic loss, medical costs, etc. More generally, a payment type constitutes a specific body of claim payments likely to have its own distinct characteristics.

These characteristics might include:

- frequency;
- total cost;
- the distribution of the payments over the lifetime of a claim, and particularly whether the payments occur predominantly as single or periodical amounts;

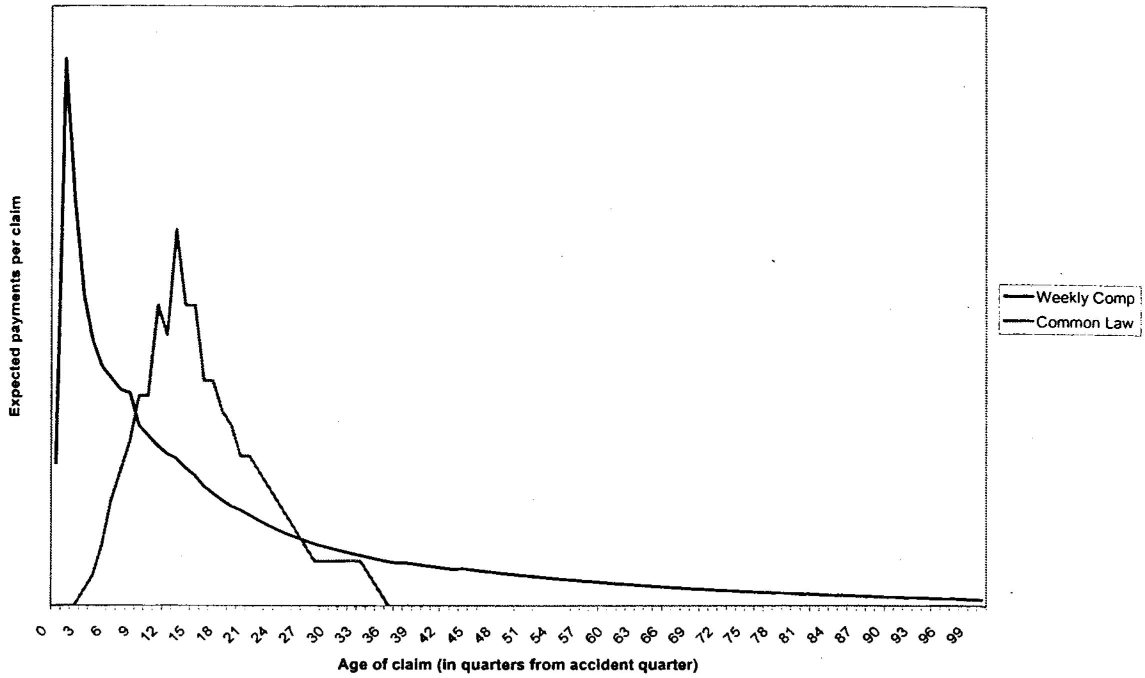
as well as perhaps their general nature and causation.

In short, payment types are categories of claim payments that deserve, for whatever reason, separate models. For the paper's workers' compensation example, the following payment types are identified:

- Weekly Compensation
- Medical, in two sub-divisions:
  - in respect of claims with some weekly compensation paid (**lost time claims**)
  - in respect of claims with no weekly compensation paid (**no lost time claims**)
- Rehabilitation
- Death
- Specific Injuries (for which tabulated amounts of benefit are paid)
- Common Law (amounts paid pursuant to issue of a writ alleging negligence of employer)
- Pain and Suffering (administratively determined, as opposed to the P&S component of a Common Law award)
- Legal, in three sub-divisions:
  - Common Law (the legal costs associated with the trial of a case and pre-trial procedures)
  - Tribunal (the costs of administrative dispute resolution bodies)
  - Medico-legal (the costs associated with the assessment of medical injuries for the purposes of legal or administrative dispute resolution)
- Investigation
- Other.

As an illustration of the differences that might arise between payment types, Figure 2.1 displays the respective distributions of claim payments by age of claim for Weekly and Common Law payment types.

Figure 2.1
Distribution of claim payments by age of claim



2.2 Dependencies

While each payment type should be modelled **separately**, it will not always be advisable that they be modelled **independently**. Where payments under payment type A are in some way provoked by payment type B, it will be preferable to recognise a dependency of the model for A on that for B.

As an example of this quite direct type of dependency, the Common Law Legal costs payable in respect of a claim will be related (on average) to the Common Law claim payments themselves (ie the size of award).

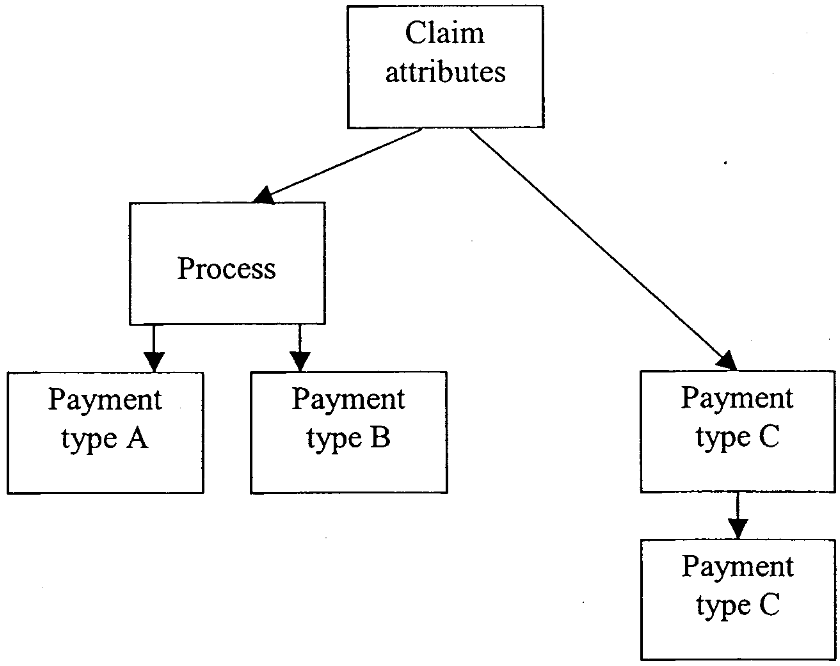
As a somewhat less direct dependency, Medical costs over the life of a time lost claim will be correlated with the continuance of the incapacity generating Weekly Compensation. In this case both Weekly Compensation and Medical costs are driven by a common process, the continuance of incapacity.

The ultimate inputs to all payment types must of course be the attributes of the claim under consideration.

Thus, the general structure of the model for an individual case estimate will be as in Figure 2.2, which illustrates the following features:

- Separate modules for separate payment types
- Some cascading of these modules (output from one module serves as input to another)
- Some modules having common inputs
- Ultimate dependency of all modules on claim attributes.

Figure 2.2
General structure of case estimate model



2.3 Example

The flowchart in Figure 2.3 illustrates the types of dependencies that may exist between processes or events and workers' compensation payment types. The types of dependencies that exist will also be affected by benefit design. A link is shown between common law payments and weekly payments, where it is common for weekly payments to cease after a common law payment has been made.

Table 2.1 also lists some examples of claim attributes that may affect liability. The table is not at all intended to be exhaustive.

Figure 2.3
Example of payment type dependencies

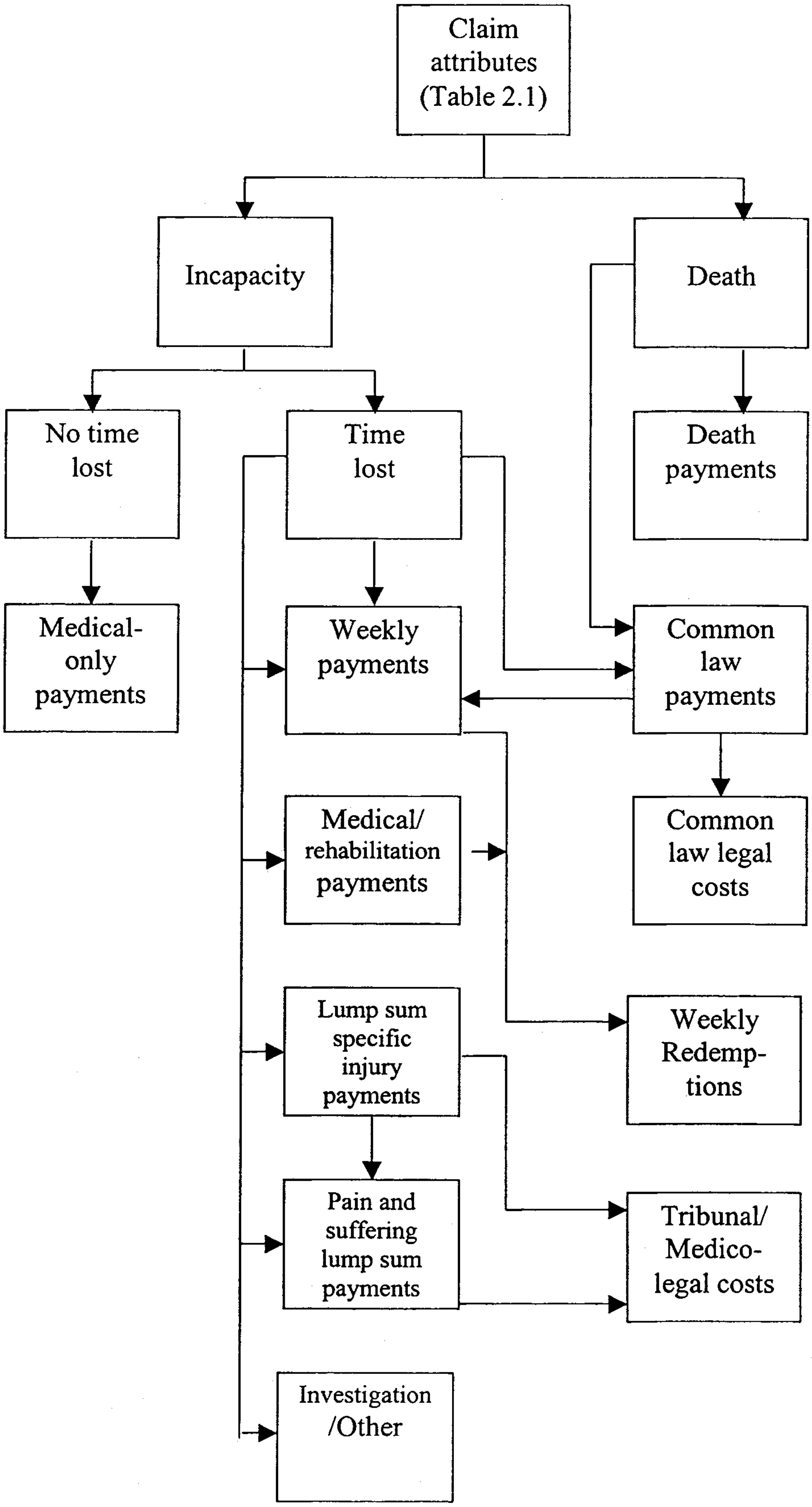


Table 2.1
Example of claim attributes affecting liability

Claimant or claim characteristics	Claim status characteristics
Injury date	Incapacitated or active (back at work)
Notification date	Number of days in current status
Claimant age	Spell number of status[(a)]
Sex	Total days of incapacity to date
Employer size	Injury severity code
Nature of injury	
Type of accident	
Cause of accident	
Bodily location of injury	
Pre-injury earnings	
Number of dependants	

Note: (a) Spells are defined and discussed in Section 5.

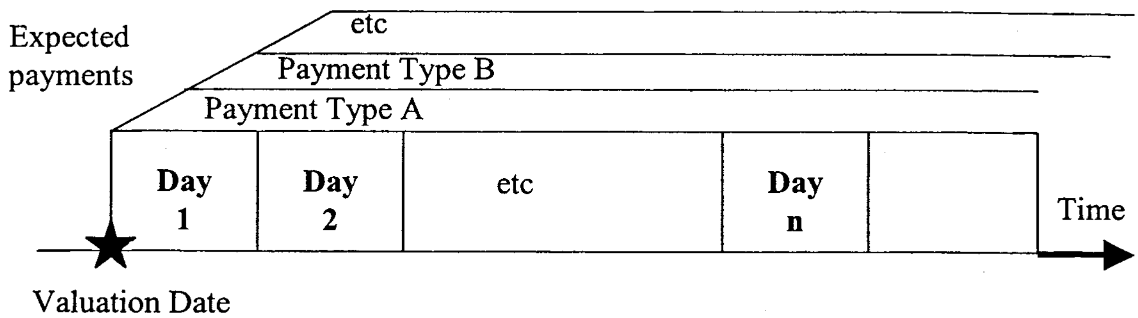
3. Statistical tools

3.1 General modelling considerations

Objective

Consider a single claim on a specific valuation date. The SCE objective is to map the claim's attributes to a forecast of claim payments day by day into the future. The forecasts are required separately by payment type. Thus, the required output is as illustrated in Figure 3.1.

Figure 3.1
Required claim payment forecast



Preferably, the payments of Figure 3.1 would be expressed in money values of the valuation date. They could then be subjected to any required manipulations, such as:

- increase to allow for future inflation of benefit levels
- discounting to present values before collation into case estimate.

Form of model

The claim payment process in respect of a particular claim under a particular payment type will be determined by:

- a claim **frequency** process
- a cost **severity** process.

It will usually be desirable to model these two processes separately, though they might be merged into a single **claim cost** process for some of the less substantial payment types.

Taken at the broadest level, payment types may be categorised as generating claim payments either:

- as long as a claim **remains in a particular status**; or
- where a claim **attains a particular status**.

As an example of the first category, the status might be "**incapacitated**", and Weekly Compensation payments would be made for the duration of this status. Such payments are continuing, or periodical, in nature.

Payments in the second category occur as isolated payments, or small groups of payments. They are not of a continuing nature. For example, a Common Law payment will occur when a claim attains a "settled" status, whether settlement be by verdict or negotiation.

The natural vehicle for modelling continuance of a defined status, as required for payment types in the first category, is **Survival Analysis** (see eg Kalbfleisch and Prentice, 1980). The probability that the claim continues in the status over the next unit of time is modelled as a function of the covariates. This is discussed in Section 3.3.

The second category of payments may be modelled in various ways. This paper uses **Generalised Linear Models** (GLMs) (McCullagh and Nelder, 1989). Again the relevant quantities, this time expected frequency and cost severity, are modelled as functions of the covariates. These models are discussed in Section 3.2.

3.2 Generalised Linear Models

As GLMs are discussed in detail elsewhere, only their most basic features will be covered here.

Consider a set of independent random variables Y_i, $i = 1, 2, ..., n$, of the form:

$$Y_i = h^{-1}\left(X_i^T \beta\right) + e_i, \qquad (3.1)$$

where

X_i^T = i-th row of an n x p **design matrix** X^T, this row containing the values of p **covariates** (or **predictors**) associated with the i-th observation

β = p-vector of **parameters** associated with the p covariates

e_i = centred **stochastic error** term (ie $E e_i = 0$)

$h: U \rightarrow \Re$ is one-one for some subset U of $\Re$.

The linear function $X_i^T\beta$ is called the model's **linear response**. The function h is called the **link function**.

Suppose that the e_i are all stochastically independent.

In more explicit form, (3.1) is:

$$Y_i = h^{-1}\left(\sum_{j=0}^{p-1} X_{ij}\beta_j\right) + e_i, \qquad (3.2)$$

where $X_i^T = \left(X_{i0},\ldots,X_{ip}\right)$.

If h were linear and e_i normal, (3.1) would be a **general linear model**. For more general choices of h and e_i, (3.1) is called a **generalised linear model** (GLM). Thus, GLMs extend the general linear model by allowing for:

- non-linear relation between the response variable and its covariates; and
- non-normal error terms.

Their best known implementations are made in the statistical packages GENSTAT (the successor to GLIM), SAS and S-Plus. These are based on specific families of h and e_i. The error terms are taken from the **exponential dispersion family** of distributions, according to which Y_i has the *pdf*

$$p(y_i) = \exp\left\{\left[y_i\theta_i - b(\theta_i)\right]/a_i(\phi) + c(y_i,\phi)\right\}, \qquad (3.3)$$

for suitable parameters θ_i,ϕ and functions $a_i(\cdot), b(\cdot)$ and $c(\cdot)$.

It may be shown that

$$E\left[Y_i\right] = b'(\theta_i) \qquad (3.4)$$

$$V\left[Y_i\right] = b''(\theta_i)a_i(\phi) \qquad (3.5)$$

where the primes denote differentiation.

The usual form of $a_i(\phi)$ is:

$$a_i(\phi) = \phi / w_i, \qquad (3.6)$$

where the w_i are known quantities. Then (3.5) and (3.6) yield:

$$V\left[Y_i\right] = \tau_i^2\phi / w_i, \qquad (3.7)$$

where

$$\tau_i^2 = b''(\theta_i). \qquad (3.8)$$

Since the variance is proportional to ϕ, this quantity is referred to as the **scale parameter**. Since the variance is inversely proportional to w_i, this quantity is referred to as the **prior weight** associated with Y_i.

Typically, one of the covariates included in X_i (denote it by X_{i0}) will be the constant 1. Then, with the other covariates denoted by X_{ij}, j = 1, 2, etc, (3.2) becomes:

$$Y_i = h^{-1}\left[\beta_0 + \sum_{j=1}^{p-1} X_{ij}\beta_j\right] + e_i \qquad (3.2)$$

with β_0 an **intercept** term.

A modification of (3.1) sometimes used is the following:

$$Y_i = h^{-1}\left(\alpha_i + X_i^T\beta\right) + e_i, \qquad (3.1b)$$

where α_i is a **known** quantity (as opposed to an unknown intercept), called an **offset**.

Example 1

Set $a_i(\phi) = \phi, b(\theta_i) = \frac{1}{2}\theta_i^2, c(y_i,\phi) = -\frac{1}{2}\left[\log(2\pi\phi) + y_i^2/\phi\right]$ and ϕ=σ^2. Then (3.3) is the normal *pdf* with mean θ_i, variance σ^2.

Example 2

Set $a_i(\phi) = 1, b(\theta_i) = \exp\theta_i, c(y_i) = -\log y_i\,!$ for y = 0, 1, 2, etc. Then (3.3) is the Poisson *pdf* with mean $\exp\theta_i$.

Example 3

Set $a_i(\phi) = \phi, b(\theta_i) = \log(-1/\theta_i) \exp c(y_i,\phi) = y_i^{\gamma-1}\gamma^\gamma / \Gamma(\gamma)$ with $\gamma = 1/\phi$. Then (3.3) is the gamma *pdf* with mean $-1/\theta_i$, variance $= (\text{mean})^2/\gamma$.

3.3 Survival analysis

Let T be an observation on the length of time of survival in a particular status. It is supposed that T is stochastic with d.f. $F(\cdot)$. Then the complementary d.f.

$$S(t) = 1 - F(t) = \text{Prob}[T > t] \qquad (3.9)$$

is known as the **survivor function**.

For $S(\cdot)$ differentiable, the **hazard rate** associated with this survivor function is defined as

$$h(t) = -S'(t)/S(t) = -\frac{d}{dt}\log S(t). \qquad (3.10)$$

By (3.10),

$$h(t)dt = \text{Prob}[t \leq T < t + dt \mid T \geq t] \qquad (3.11)$$

$$S(t) = \exp - \int_0^t h(u)du. \qquad (3.12)$$

Now consider a sample of observations of survival times, denoted by T_i, $i = 1,2,$ etc. Suppose that T_i is sampled from a survivor function $S_i(\cdot)$ with associated hazard rate $h_i(\cdot)$, and suppose further that $h_i(\cdot)$ may be expressed in terms of a number of covariates associated with observation T_i, ie

$$h_i(t) = h(t : X_i) \qquad (3.13)$$

where X_i is a vector containing those covariates.

It is evident that there are many possible forms of the dependency of h_i on X_i. This paper will be concerned with the **proportional hazards** (PH) **regression** model in which (3.13) takes the form:

$$h_i(t) = h_0(t)\exp(X_i^T\beta) \qquad (3.14)$$

for a suitable parameter vector β. Note that X_i may depend on *t*.

The function $h_0(\cdot)$ is referred to as the **baseline hazard rate**, and the associated survivor function $S_0(\cdot)$ as the **baseline survivor function**. Note that the survivor function for the *i*-th observation is determined by the scalar quantity

$$X_i^T\beta = \sum_{j=1}^{p} X_{ij}\beta_j,$$

which therefore operates as a form of **risk score** associated with the *i*-th observation. Indeed, the quantity $X_{ij}\beta_j$ operates as a risk score for the *j*-th covariate in respect of the *i*-th observation.

The hazard rate is proportional in the sense that, if $X_i^T = (X_{i1}, \ldots, X_{ip})$, then

$$h_i(t) = h_0(t)\exp\sum_{j=1}^{p} X_{ij}\beta_j$$
$$= h_0(t)\alpha_{i1}\alpha_{i2}...\alpha_{ip} \quad (3.15)$$

with

$$\alpha_{ij} = \exp X_{ij}\beta_j \quad (3.16)$$

= multiplier in respect of j-th covariate acting on the hazard rate associated with the i-th observation.

The PH model was introduced by Cox (1972). Its baseline hazard function may take parametric or non-parametric form. In the latter case, the PH regression model is called a **Cox regression model**. The function $h_0(\cdot)$ is then to be estimated along with the parameter vector β.

Cox regression may be implemented through the SAS procedure PHREG. A SAS-oriented discussion of this, and in fact survival analysis generally, may be found in Allison (1995).

3.4 Categorical and continuous covariates

Categorical covariates

Some covariates included in model (3.1) or (3.14) are **categorical**, or discrete, in nature. The simplest example is gender of claimant, which assumes just the two values M and F.

This would appear as generating two variables in either model, which might be denoted X_{gM} and X_{gF} respectively, and defined as follows:

$$X_{gM} = 1 \text{ if gender } = M$$
$$= 0 \text{ otherwise} \quad (3.17)$$

$$X_{gF} = 1 \text{ if gender } = F$$
$$= 0 \text{ otherwise.} \quad (3.18)$$

Thus, in (3.1) and (3.14),

$$X_i^T\beta = X_{gM,i}\,\beta_{gM} + X_{gF,i}\,\beta_{gF} + ... \quad (3.19)$$

This may be expressed in the form

$$X_i^T\beta = \beta_{gM} + X_{gF,i}\left(\beta_{gF} - \beta_{gM}\right) + ... \quad (3.20)$$

If the covariates include a unit term so that (3.1) takes the "intercept form" (3.1a), then β_{gM} may be absorbed into the intercept, whence (3.20) shows that the contribution of gender to the linear response $X_i^T\beta$ is

$$0 \text{ for } g = M$$
$$\beta_{gF} - \beta_{gM} \text{ for } g = F. \qquad (3.21)$$

In GLM parlance, gender is a **factor variable** (GLIM) or a **class variable** (SAS) with 2 **levels**, *M* and *F*. When the model is expressed in the form (3.21), the level $g = M$ is called the **base level** of variable *g*. As the base level has a zero coefficient, by definition, it is said to be **aliased** from the model.

The same concepts apply to categorical variables with more than two levels:

- One level is chosen as the base and is aliased.
- The coefficient associated with each other level measures the effect on the linear response of the **difference** between that level and the base.

Grouped values of categorical variables

Some categorical variables may assume many values. For example, employers' industry codes may consist of 4 or 5 digits, generating some hundreds of values. To recognise all of these within the model would be likely to fragment the data such as to render them incapable of producing statistically significant results for other than a tiny proportion of the codes.

In this situation, greatest information will be extracted from the data by forming **groups of values** of the variable concerned. The viable number of levels of a single categorical variable is clearly dependent on the size of the data set and the number of other variables included in the model, but the authors have experienced difficulty in deriving meaningful results for more than about 10 levels, even using quite large data sets.

A **natural grouping** of variable levels may exist if the variable classifies according to a **tree structure**. An example of this would be the ANZSIC system of industry coding (Australian Bureau of Statistics, 2000) which is an alpha-numeric system forming a 5-level tree. Codes take the form Xnnnn where X is an alphabetic character, n numeric and the *i*-th character classifies the *i*-th level of the tree. For example,

A	=	agriculture, forestry and fishing
A1	=	farm production
A11	=	horticulture and fruit growing
A115	=	apple growing

An initial grouping of ANZSIC codes might consist of collapsing all classifications below the top level of the tree. The groups would be just A, B, C, etc. Suppose this lead to a model which recognised the following groupings as statistically significant in their differences: A, D, I – K, all other. Further modelling might then investigate the second level of the tree, eg subdividing A into A1, A2, etc.

An example of grouping a categorical variable is given in Appendic C.4.

Various categorical variables may have this tree structure, eg nature of injury, body site of injury, cause of accident.

Continuous covariates

Some covariates are **continuous** in the sense that:

- Their sets of admissible values are interval subsets of the real numbers; and
- The linear response is expected to be a continuous function of the covariate, ie small changes in the covariate produce small changes in the response.

If X_c is such a covariate, then the linear response takes the form:

$$X_i^T\beta = \sum_{k=1}^{r} \beta_{ck} f_{ck}\left(X_{ci}\right) + \ldots, \qquad (3.22)$$

where X_c takes the value X_{ci} for the i-th variable Y_i and the **basis functions** $f_{ck} : V \to \Re, \quad V \subset \Re, \quad k = 1,\ldots,r$ are continuous.

An example of a continuous covariate would be age of claimant at occurrence of claim. If one chose $f_{ck}\left(x\right) = x^k, \quad k = 1,2,$ then (3.22) would become:

$$X_i^T\beta = \beta_{c1}X_{ci} + \beta_{c2}X_{ci}^2 + \ldots \qquad (3.23)$$

a quadratic function of age.

Any continuous covariate may be approximated by a categorical covariate, constructed by taking intervals of its admissible values. For example, the continuous covariate age may be represented in categorical form with levels say 15-19, 20-24, etc.

This will often be useful in exploratory work aimed at determining suitable basis functions. However, a continuous representation like (3.23) will usually be more parsimonious in its use of parameters.

The modelling of a continuous covariate would usually follow a procedure consisting of the following steps:

- Group ranges of values of the covariate X to form a categorical variable, X^* say. Suppose that these ranges correspond broadly to values $X = x_1, x_2, \ldots$.
- Include X^* in the model and so obtain estimates $\hat{\beta}_1^*, \hat{\beta}_2^*$, etc of the model coefficients associated with the different ranges.
- Plot the $\hat{\beta}_j^*$ against the x_j to gain some idea of how the two are functionally related, eg linear relation, quadratic relation, etc.
- On this basis select the basis functions $f_k(X)$ $k = 1,2$, etc to be used in modelling the original covariate X. For example, if in the previous step the $\hat{\beta}_j^*$ appear to form a roughly quadratic function of the x_j, then the basis functions $f_k(x) = x^k$, $k = 1,2$ would be adopted.
- Finally, the categorical variable X^* is removed from the model, and replaced by the selected set of continuous basis functions.

An example of the modelling of a continuous covariate is given in Section 6.4.3.

4. Financial structure of SCEs and payment type components

4.1 Cash flows

The basic structure of an SCE was anticipated in Figure 3.1.

Let

t	=	t-th time period (day, week, etc) after the valuation date (t = 1, 2, etc).
X_i	=	vector of covariate values for the i-th claim for which an SCE is required
$c_{pt}(X_i)$	=	cash flow in current values (ie money values of the valuation date) predicted by model for the i-th claim in payment type p in period t.

Then the SCE in current values for the i-th claim is

$$c(X_i) = \sum_p c_{p.}(X_i) \tag{4.1}$$

with

$$c_{p.}(X_i) = \sum_{t=1}^{\infty} c_{pt}(X_i)$$
$$= \text{forecast cost of payment type } p. \tag{4.2}$$

Alternatively,

$$c(X_i) = \sum_{t=1}^{\infty} c_{.t}(X_i) \tag{4.3}$$

with

$$c_{.t}(X_i) = \sum_p c_{pt}(X_i)$$
$$= \text{forecast cash flow (from all payment types) in period } t. \tag{4.4}$$

Both forms (4.1) and (4.3) yield:

$$c(X_i) = \sum_p \sum_{t=1}^{\infty} c_{pt}(X_i). \tag{4.5}$$

4.2 Inflation and investment return

Let

$g_p(t)$ = inflation factor applying to payment type p over the t-th time period

$r(t)$ = rate of investment return at which liabilities are to be discounted over the same period.

Define

$$w_p(t)=\left[1+g_p(t)\right]/\left[1+r(t)\right] \tag{4.6}$$

$$v_p(t)=w_p(1)w_p(2)...w_p(t) \tag{4.7}$$

$$\begin{aligned}\bar{v}_p(t)&=w_p(1)w_p(2)...w_p(t-1)w_p^{1/2}(t)\\&=\left[v_p(t-1)v_p(t)\right]^{1/2}.\end{aligned} \tag{4.8}$$

Then, on the assumption that claim payments of a period occur on average at the mid-point of that period, the case estimate that corresponds to $c(X_i)$, but including future inflation and discounted for investment return, is approximately

$$\begin{aligned}c^*(X_i)&=\sum_p\sum_{t=1}^{\infty}\bar{v}_p(t)c_{pt}(X_i)\\&=\sum_{t=1}^{\infty}c_{\cdot t}^*(X_i)\end{aligned} \tag{4.9}$$

with

$$c_{\cdot t}^*(X_i)=\bar{v}_p(t)c_{\cdot t}(X_i). \tag{4.10}$$

4.3 Frequency and severity

The claim payments $c_{pt}(X_i)$ may be decomposed into frequency and severity components as follows:

$$c_{pt}(X_i)=f_{pt}(X_i)s_{pt}(X_i) \tag{4.11}$$

where

$f_{pt}(X_i)$ = probability that a payment of type p occurs in the t-th period in respect of the i-th claim

$s_{pt}(X_i)$ = expected amount (in current values) of such a payment, conditional upon its occurrence.

There are two reasons for studying frequency and severity separately.

First, they are fundamentally separate processes. In the case of a workers' compensation weekly benefit, the expected amount may change slowly over time, whereas the frequency is likely to change rapidly. Generally, a greater understanding of the claim payment process will be gained by separate modelling of the two processes.

Second, direct modelling of the quantities $c_{pt}(X_i)$ is rendered difficult by the fact that it does not have a convenient parametric distribution. Let $E_{pt}(X_i)$ denote the number of periods of exposure to payment of type p in data cell (t, X_i) (past values of t). When $E_{pt}(X_i)$ is small, $\text{Prob}\left[c_{pt}(X_i)=0\right]$ will be substantial. This means that the distribution of $c_{pt}(X_i)$ will be mixed, continuous on the strictly positive half-line, but with a probability spike at zero.

As X_i may well specify 10 to 20 covariates, the exposures $E_{pt}(X_i)$ will usually be small.

While it is true that this situation has been addressed in a GLM context by Jorgensen and Paes de Souza (1994) and Smyth and Jorgensen (1999), their approach requires the assumption that severity be subject to a gamma distribution, and also requires programming additional to the software provided by the standard packages mentioned above.

This second argument for separate modelling of frequency and severity is well known from pricing work in short tail lines of insurance (eg Motor), where it applies equally.

4.4 Benefit types

Section 3.1 classified benefits as either:

- status related; or
- event (attainment of a status) related.

Figure 2.2 indicates that payments of one type of benefit may depend upon the experience of another type. Thus, for the purpose of modelling, the various benefits may be classified as follows:

- status related benefits
- event related benefits
    - that also depend on status
    - that also depend on other events
- without any such dependencies.

Sections 5 to 8 discuss some of the issues in modelling payments under these different types of benefit.

5. Status related benefits

Status related benefits are those types of benefits that are paid as long as a claim remains in a particular status. For example, weekly compensation payments are made provided that the claimant remains in the "incapacitated" status. Such benefits are periodical in nature, adopting the characteristics of an annuity.

5.1 Frequency

In the simplest case of a status-related benefit, for example, an annuity or a pension, the benefit is paid provided that the claimant remains alive ("alive" status). The frequency of payment is related to the survival curve of that claimant. Once the status ceases (claimant dies or reaches some other defined status), benefits cease.

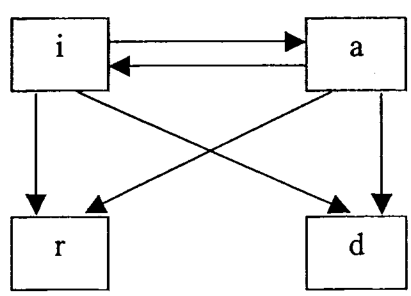
In more complex cases, claimants may make transitions between several statuses. Claimant status over time thereby becomes a stochastic process in the same manner as introduced by Taylor (1971) and used by Haberman (1983), Waters (1984) and others. This approach has been used by the Institute of Actuaries and Faculty of Actuaries (1991) in their morbidity investigations.

In the case of weekly compensation payments, there are four relevant statuses (or just **states** in the usual stochastic process terminology):

- incapacitated (i)
- active (a) (ie able to work)
- retired (r) (ie having attained retirement age after which there is no eligibility for weekly compensation)
- deceased (d).

Figure 5.1 indicates the possible transitions between states. Only changes of state are indicated. Any state may be maintained from one epoch to the next, a fact that is recognised only implicitly in the diagram.

Figure 5.1
Transitions between statuses: weekly compensation



In the case of a transition from state *m* to state *n*, *m* will be referred to as the **source state** and *n* the **destination state**.

A brief discussion of discrete time finite-dimensional stochastic processes, such as that involved here, is given in Appendix A. The graph of the process, defined there, is

$\Gamma(j) =$

	i	a	r	d
i	1	1	1	1
a	1	1	1	1
r	0	0	1	0
d	0	0	0	1

where j is used in the present section as the discrete time variable for the stochastic process (t in Appendix A).

Hence, r and d may be recognised as absorbing states.

Note that $p_{rm}(j) = 0$, m = i, a, d except at retirement age, where probabilities $p_{mn}(j)$ are defined in Appendix A. The probabilities $p_{dm}(j)$, m = i, a will usually be low, and so the distribution of a particular claim's future state occupancy will usually be dominated by transitions between i and a.

Appendix A discusses 2-state processes, and shows how the progress of a claim can be described by a sequence of **spells** in the two states.

Define a **spell of incapacity** as an unbroken period of incapacity that is maximal in the sense that it is not immediately preceded or succeeded by any other incapacity. According to this definition, the first day of incapacity experienced by a claim commences the first spell of incapacity.

A **spell of activity** is similarly defined.

Frequency of payment of weekly compensation is determined by the sequence of alternating spells of incapacity and activity set out in Table 5.1.

Table 5.1

Status (i,a)	Spell number (m)	Duration
Incapacity	1	i_1 days
Activity	1	a_1 days
Incapacity	2	i_2 days
Activity	2	a_2 days

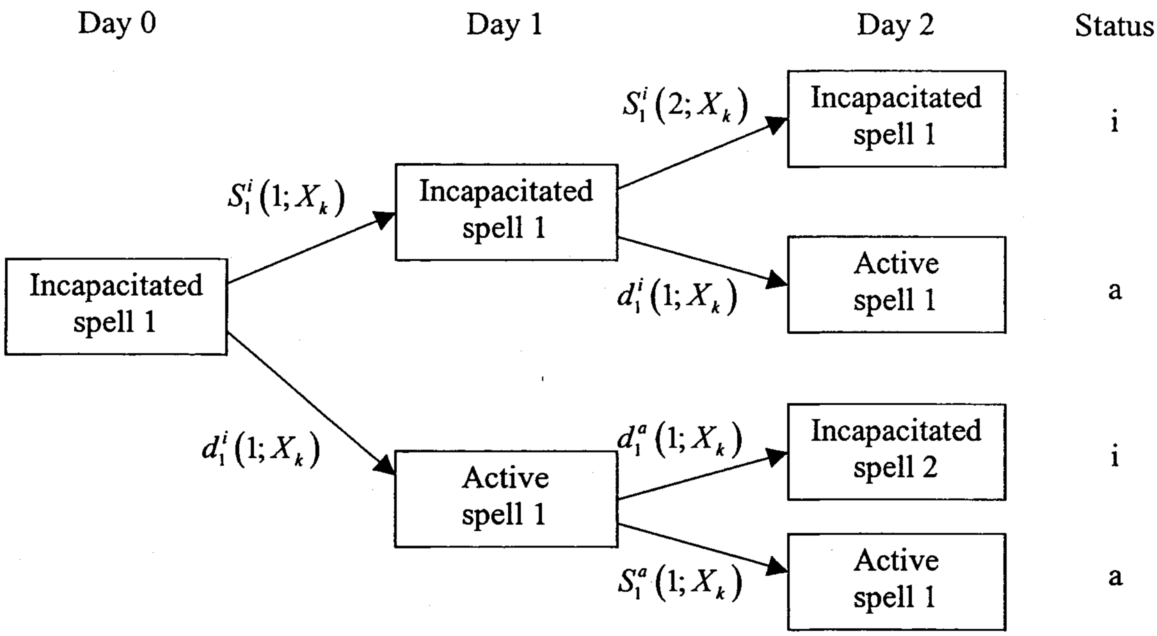
Note that the spell of activity **succeeding** Spell 1 of incapacity is designated Spell 1 of activity. Any period of activity before the commencement of incapacity may be denoted **Spell 0** of activity.

Define $S_m^s(j;X_k)$ as the probability that a claimant k with covariates X_k remains in spell number m of status s (i or a) for at least j days, and $d_m^s(j;X_k)$ as the probability that, by day $j+1$, the claimant will no longer be in status s, ie

$$d_m^s(j;X_k)=\frac{S_m^s(j;X_k)-S_m^s(j+1;X_k)}{S_m^s(j;X_k)}. \qquad (5.1)$$

Figure 5.2 considers a claim for which incapacity commences in day 0, and illustrates its possible development over days 1 and 2 as a binomial branching diagram. The binomial probabilities defined in the preceding paragraph are attached to the relevant branches.

Figure 5.2
Binomial branching of claim status



Let $f(j;X_k)$ denote the expected frequency of payment at day j for a claimant k with covariates X_k. This can be evaluated in terms of more elementary probabilities, such as displayed in Figure 5.2. Note, however, that the compounding of single-step probabilities requires a statement of the assumptions made about stochastic independence between them.

It is implicit in the above development that transition probabilities between states depend on the duration of the spell in the source state. Hence the independence assumption would **not** be that the process is fully Markovian as described in Appendix A, where transition probabilities depend on only the source state and no other history.

The simplest useful case would involve the assumption that the process has no memory beyond its current spell, ie the transition probability out of a spell depends on the spell duration as well as the source and destination states, but on nothing else. Then Figure 5.2 leads to the following evaluation:

$$f(2;X_k) = \sum \text{(Probabilities resulting in a status of } i\text{)}$$
$$= S_1^i(2;X_k) + d_1^i(1;X_k)\ d_1^a(1;X_k). \qquad (5.2)$$

Similar expressions may be written for $f(j;X_k)$. They become increasingly complex with increasing *t*, as the number of possible trajectories of the claim between days 0 and *t* increases.

In practice, the independence assumption on which (5.2) is based is rather restrictive. For example, it might be desirable that transitions from *a* to *i* depend on the total time spent in state *i* in all previous spells. In this case, (5.2) would be replaced by the following:

$$f(2;X_k) = S_1^i(2;X_k) + d_1^i(1;X_k)d_1^a(1;X_k,1) \qquad (5.2a)$$

where the final 1 in $d_1^a(1;X_k,1)$ recognises that a total of 1 day of incapacity has occurred prior to the $a \to i$ transition.

The generalities of this are considered in Appendix A where it is pointed out that a Markovian assumption can be maintained if the set of states defining the process is enlarged. In the above example, the set of states would be the Cartesian product:

$$\{i,a\} \times \begin{Bmatrix}\text{possible}\\ \text{spell}\\ \text{numbers}\end{Bmatrix} \times \begin{Bmatrix}\text{possible}\\ \text{spell}\\ \text{durations}\end{Bmatrix} \text{ for (5.2)}$$

$$\{i,a\} \times \begin{Bmatrix}\text{possible}\\ \text{spell}\\ \text{numbers}\end{Bmatrix} \times \begin{Bmatrix}\text{possible}\\ \text{spell}\\ \text{durations}\end{Bmatrix} \times \begin{Bmatrix}\text{possible total}\\ \text{previous occupancies}\\ \text{of state } i\end{Bmatrix} \text{ for (5.2a)}$$

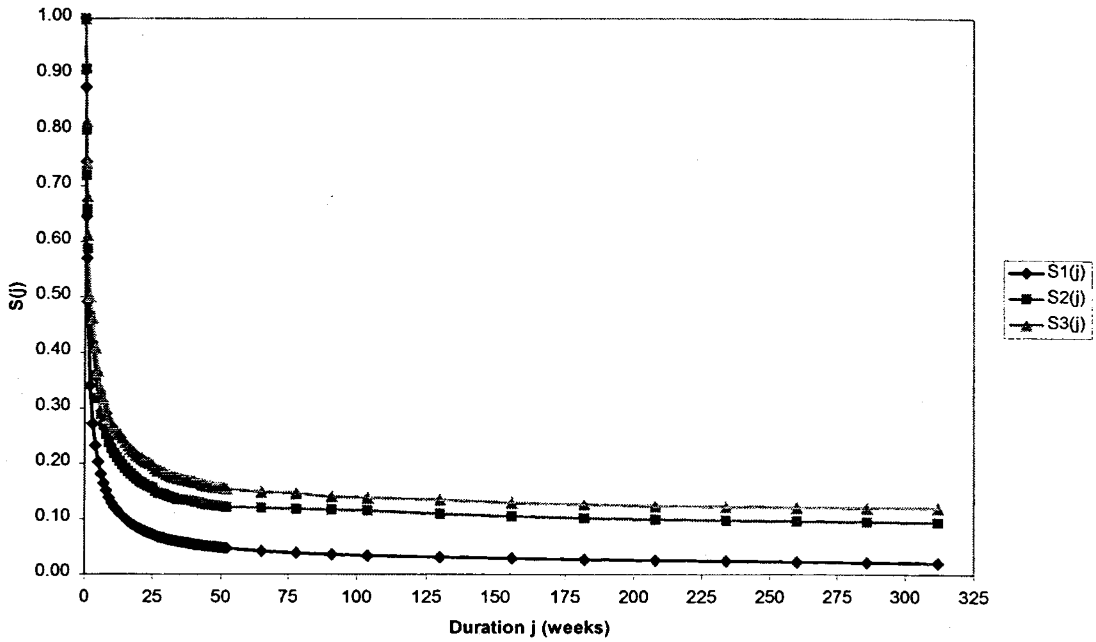
Survival of incapacity or activity status

The quantities $S_m^s(j;X_k)$ are **survival probabilities** in the sense discussed in Section 3.3. They are in fact particular cases of the function *S*(*j*) appearing there, and may be estimated from the data as discussed there. As discussed just after (3.14), these survival probabilities will be characterised by risk scores that depend on the covariates X_k. For each claim, there will be one risk score per survivor model, ie each claim will be associated with an **incapacity risk score** β_i and an **activity risk score** β_a.

Example 1

Figure 5.3 illustrates baseline survival probabilities $S_m{}^i(j;\ X_0)$ for incapacity spells $m = 1, 2,$ and 3.

Figure 5.3
Baseline incapacity survival probabilities for incapacity spells 1, 2, and 3

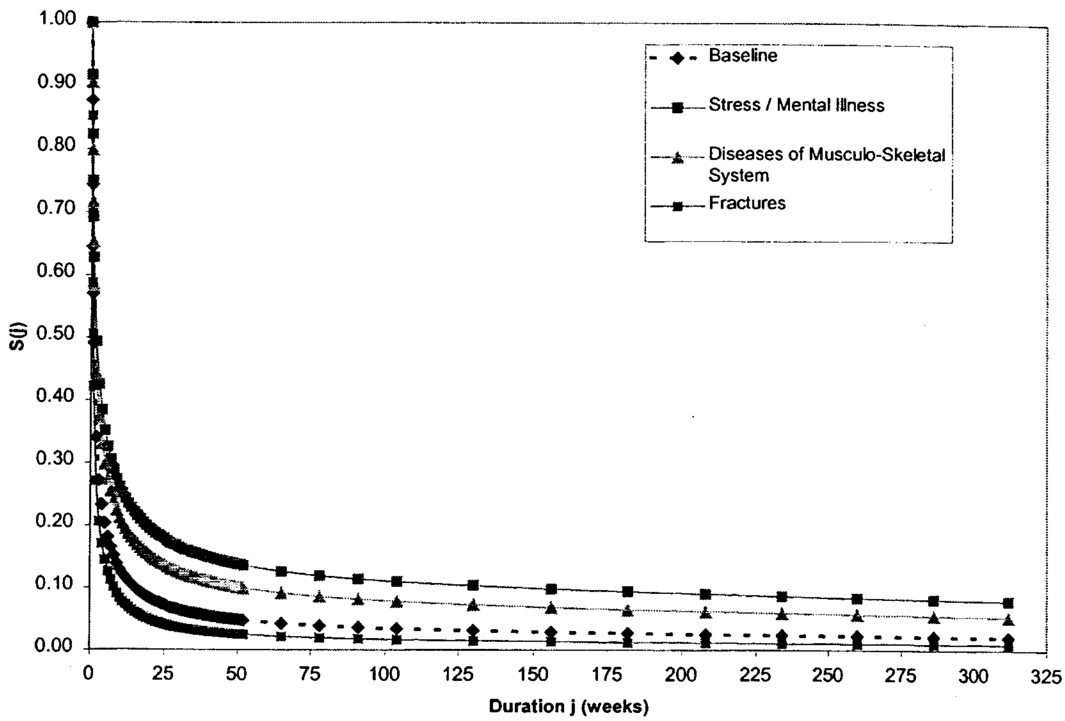


Example 2

Figure 5.4 illustrates values of $S_1{}^i(j;X_k)$ for spell 1 of incapacity, where the β_k risk scores vary by nature of injury. In this example, the scores associated with each type of injury are:

- Stress / mental illness: -0.421
- Diseases of the muscular-skeletal system and connective tissue: -0.277
- Fractures: +0.193

Figure 5.4
Survival probabilities for incapacity spell 1, by nature of injury



Covariates

Covariates affecting the shape of the survival curve may be **static** or **dynamic**.

The values of static covariates are determined permanently at the outset of the claim. Examples are age at injury, occupation, employer industry, etc. The values of dynamic variables may change as the claim progresses. An example, spell number, appears in the above example.

5.2 Severity

Typically, the periodical amount of a status related benefit will be related to the previous earnings or earning capacity of a claimant. Table 5.2 gives one example of a legislative definition of a workers' compensation weekly benefit.

Table 5.2
Example of status related benefit structure

$\leq$13 weeks' incapacity:	95% of pre-injury average weekly earnings (maximum $*X*), less notional earnings (where there is some capacity to work)
> 13 weeks' incapacity:	75% of pre-injury average weekly earnings if no current capacity for work; or if some capacity for work: 60% of [MIN ($*Y*, pre-injury earnings) less notional earnings].
> 104 weeks' incapacity:	Weekly benefits cease **Unless** worker is permanently incapacitated, with no work capacity, in which case benefits continue while this is the case until retirement.

In such a structure, it would be natural to forecast severity in the form of a multiple of pre-injury earnings. In the notation of (4.11),

$$s_{pt}(X_i) = W_i\,\mu(i_t) \qquad (5.3)$$

where

W_i = pre-injury earnings associated with the i-th claim (W_i would be a component of the vector X_i)

i_t = forecast amount of incapacity accumulated by future period t (from the model of Section 5.1)

$\mu(i_t)$ = multiple applicable when accumulated incapacity is i_t.

According to Table 5.2, one would expect $\mu(i_t)$ to be close to 95% for low values of i_t, to step down suddenly as i_t increases through 13 weeks, and again at 104 weeks.

Because weekly compensation in the example is reduced on account of notional earnings (essentially the amount the claimant is deemed capable of earning, whether actually earning or not), one would find that the multiples $\mu(i_t)$ experienced in practice would lie below their maxima of 95% and 75% specified in the table.

The best guidance as to future levels of benefit paid in respect of a particular claim may well be the present level. For example, (5.3) might be refined as follows:

$$s_{pt}(X_i) = \min\left[W_i\,\mu(i_t), B_i\, r_{t-\tau}(X_i)/r_{v-\tau}(X_i)\right] \qquad (5.4)$$

where

v = valuation period
τ = period during which claim i was reported
B_i = average weekly benefit rate paid in respect of claim i in period v
$r_k(X_i)$ = the model ratio of benefit rate payable in k-th quarter after reporting quarter to that payable in the reporting quarter (NB $r_o(X_i)=1$) for a claim with covariates X_i.

Usually, the $r_k(X_i)$ would form a decreasing sequence in k for each choice of X_i. The covariates X_i might include an indicator of injury severity. If so, it is likely that the sequence would be flatter for the more severe injuries.

In (5.4), all quantities other than the $r_k(\bullet)$ are directly available from either legislation (eg μ) or the history of individual claims (eg B). The $r_k(\bullet)$ must be modelled.

As an example, $r_k(X_i)$ might depend on X_i through some measure of injury severity included in X_i. For a particular injury severity, one might observe, as a reasonable approximation

$$r_k = \max\left[r_{\min}, 1-k(\Delta r)\right] \tag{5.5}$$

for constants $r_{\min}$ and Δr. The representation (5.5) would be obtained by some form of regression.

6. Event related benefits

6.1 General considerations

Event related benefits are those that are paid when a claim attains a particular status or when a defined event or events occur.

Types of event-related benefits

As illustrated in Section 2 and discussed in Section 4.4, the modelling of event related benefits may be classified into the following types:

(i) those that also depend on status (see Section 7)
(ii) those that also depend on other events (see Section 8)
(iii) those without any such dependencies.

This section discusses issues in modelling type (iii). The frequency of the event and the severity of these types of benefit depend on the covariates of the claim, including age of the claim. The frequency of the payment at age t of the claim is not conditional on the continuance of a particular status or some other particular event occurring.

Example 1 – Death Payments

In workers compensation, a death payment is made when an event at work (e.g. collapse of an underground mine), or the work environment (e.g. asbestos fibres), causes the claimant's death.

Example 2 – Specific Injury Lump Sums

A claimant who is injured, but is not necessarily incapacitated (ie. the claimant is active and has returned to work), may apply for a schedule lump sum which relates to the claimant's particular injury, e.g. loss of one eye. Payment is made in the event that the claimant applies for such a benefit (and is subsequently approved by the insurer). However, payment of such a benefit is not conditional on other events (eg. payment of another benefit type) or status (eg. remaining incapacitated).

Risk Scores

The natural vehicle for modelling event-related benefits is the GLM structure, as described in Section 3.2.

Re-expressing *(4.11)* in terms of age j of claim X at time t,

$$c_{pj}(X_i) = f_{pj}(X_i)s_{pj}(X_i) \qquad (6.1)$$

where $f_{pj}(X_i), s_{pj}(X_i)$ can be modelled using the following variations of (3.2):

$$f_{pj}(X_i)=h^{-1}\left(X_{fi}^T\beta_f\right)+e_{fi} \tag{6.2}$$

$$s_{pj}(X_i)=g^{-1}\left(X_{gi}^T\beta_g\right)+e_{gi} \tag{6.3}$$

where functions h, g are the link functions for $f_{pj}(X_i)$ and $s_{pj}(X_i)$ respectively, and e_{fi}, e_{gi} are the error terms for $f_{pj}(X_i)$ and $s_{pj}(X_i)$ respectively.

Define

$$\eta_{yi} = X_{yi}^T\beta_y \quad y=f,s \tag{6.4}$$

= **risk score** for claimant i for payment type p at age j for $y = f$ (frequency) or s (severity).

Then

$$f_{pj}(X_i)=h^{-1}\left(\eta_{fi}\right)+e_{fi} \tag{6.5}$$

$$s_{pj}(X_i)=g^{-1}\left(\eta_{si}\right)+e_{si} \tag{6.6}$$

Thus frequency and severity risk scores can be derived for claimant i with covariates X_i at age j, in a manner similar to that discussed in Sections 3.3 (survival analysis) and 5.1 (status related benefits).

Covariates

Covariates affecting frequency or severity may be static or dynamic, just as in Section 5.1.

An example of a dynamic covariate is the total incapacity experienced by the claimant to date which would be derived from the incapacity incidence module as specified in Sections 5.1 and 9.

6.2 Frequency

The quantity $f_{pj}(X_i)$ is an amount in the range

$$0<f_{pj}(X_i)<1. \tag{6.7}$$

In some instances, $f_{pj}(X_i)$ may be set to zero or 1 if a defined event has already occurred at the time of valuation, for example the claimant may be dead at the valuation date which then precludes payment of type p in the future.

$f_{pj}(X_i)$ is modelled from the observed statistics $N_{pj}(X_i)$ and $E_{pj}(X_i)$

where

$N_{pj}(X_i)$ = observed number of claims with covariates X_i that received payment type p at age j

$E_{pj}(X_i)$ = number of claims with covariates X_i that are exposed to receiving payment type p at age j.

GLMs have flexibility in the choice of the distribution of error term and the link function to ensure the condition (6.7) is met. A common choice for modelling such a quantity is to consider $f_{pj}(X_i)$ as a binomial proportion with N successes in E trials. The logit transformation is a common choice of link function, although the probit and complementary log-log functions may be useful.

If

$$N_{pj}(X_i) \sim Bin\left[E_{pj}(X_i), f_{pj}(X_i)\right] \quad (6.8)$$

with

$$\begin{aligned} f_{pj}(X_i) &= \mu + e_{fi} \\ &= h^{-1}(\eta_{fi}) + e_{fi} \end{aligned} \quad (6.9)$$

then

- for logit link function,
$$\eta_{fi} = \log\left(\frac{\mu}{1-\mu}\right) \quad (6.10)$$

- for the complimentary log-log link function,
$$\eta_{fi} = \log\left[-\log(1-\mu)\right] \quad (6.11)$$

- for the probit link function,
$$\eta_{fi} = \Phi^{-1}(\mu) \quad (6.12)$$

where Φ is then standard normal distribution function.

Distributions for e_{fi} other than binomial can also be chosen.

Example

$$f_{pj}(X_i) \sim \text{binomial, logit link}$$
$$= \mu + e_{fi}$$
$$\mu = \frac{e^{\eta_{fi}}}{1+e^{\eta_{fi}}} \quad (6.13)$$

$$\eta_{fi} = \sum_{k=1}^{p} X_{ik}\beta_k = \text{ frequency risk score for claimant } i \text{ at duration } j$$
$$+\beta_1 I(\text{gender} = \text{female})$$
$$+\beta_2 f(\text{age})$$
$$+\beta_3 I(\text{occupation} = \text{nurse})$$
$$+\ldots \quad (6.14)$$

where $I(.)$ is an indicator function defined as follows:

$$I(\text{condition}) = 1 \text{ if "condition" is true}$$
$$= 0 \text{ if false.} \quad (6.15)$$

Appendix C documents some of the detail of fitting a claim frequency, providing examples of some of the aspects of modelling discussed in this sub-section.

6.3 Severity

The quantity $s_{pj}(X_i)$ is typically an amount in the range

$$s_{pj}(X_i) > 0. \quad (6.16)$$

In the case of recoveries, $s_{pj}(X_i)$ may be negative. In this case, the amount $s^*_{pj}(X_i) = -s_{pj}(X_i)$ is modelled, and (6.15) still applies to $s^*_{pj}(X_i)$.

Several examples of modelling severity are given below.

Example 1

$$s_{pj}(X_i) \sim \text{Gamma, log link} \quad (6.17)$$

$$s_{pj}(X_i) = \mu_{si} + e_{si}$$
$$= g^{-1}(\eta_{si}) + e_{si} \quad (6.18)$$

ie

$$\eta_{si} = \log \mu_{si}, \quad \mu_{si} = e^{\eta_{si}} \quad (6.19)$$

Example 2

$$s_{pj}(X_i) \sim \text{Power Gamma, log link} \tag{6.20}$$

let

$$s^*_{pj}(X_i) = \left[s_{pj}(X_i)\right]^k, \quad k > 0 \tag{6.21}$$

$$\begin{aligned} s^*_{pj}(X_i) &= \mu_{si} + e_{si} \\ &= g^{-1}(\eta_{si}) + e_{si} \end{aligned} \tag{6.22}$$

with η_{si} and e_{si} as in (6.17) – (6.19).

Then

$$E\left[s_{pj}(X_i)\right] = (\mu_{si})^{\frac{1}{k}} \times b_i \tag{6.23}$$

where b_i is the bias correction factor for the inverse of transformation (6.21). It can be shown that, if $y = \mu_{si} + e_{si}$ has pdf $\left[\Gamma(\gamma)\right]^{-1} c_i^\gamma y^{\gamma-1} \exp(-c_i y)$, $\gamma > 0$ and independent of i, then

$$b_i = \Gamma(\gamma + \tfrac{1}{k}) / \gamma^{1/k} \Gamma(\gamma). \tag{6.24}$$

Example 3

$$s_{pj}(X_i) \sim \log \text{ normal} \tag{6.25}$$

$$\begin{aligned} s^*_{pj}(X_i) &= \log\left[s_{pj}(X_i)\right] \\ &= \mu_{si} + e_{si} \end{aligned} \tag{6.26}$$

$$e_{si} \sim N(0, \sigma^2). \tag{6.27}$$

Then

$$E\left[s_{pj}(X_i)\right] = \exp\left(\mu_{si} + \tfrac{1}{2}\sigma^2\right). \tag{6.28}$$

In some cases it may be more appropriate to model $s_{pj}(X_i)$ as a **proportion** of some fixed dollar maximum if legislation places dollar caps on the total benefit receivable. In such cases,

$$0 < s_{pj}(X_i) \le 1 \tag{6.16a}$$

For example, in the case of scheduled lump sums, benefits may be of the form

$r\%$ x \$M

where

$r\%$ = disability percentage
\$M = statutory maximum,

and M may be adjusted periodically in accordance with an inflation index. In the case of such adjustment, the ratio $s_{pj}(X_i)$ would be a more stable variable to model than the actual benefit paid.

When (6.16a) holds, s_{pj} can be modelled along the lines of the above examples, but the combination of data transformation, link function and error distribution must be chosen so that (6.16a) is satisfied. It may also be necessary to allow for a continuous distribution of $s_{pj}(X_i)$ for $0 < s < 1$, but a discrete mass of probability at $s = 1$. The latter can be achieved by modelling an additional frequency, conditional on $f_{pj}(X_i)$ in (6.7). This is

$$\text{Prob}\left[s_{pj}(X_i)=1 \mid s_{pj}(X_i)>0\right]. \qquad (6.29)$$

7. Event related benefits depending on status-related benefits

A relationship may exist between the status of a claim at a particular point in time and the payment of an event-related benefit. This section is explained by reference to workers compensation weekly redemption payments, although it can be applicable to any other relevant benefit types, for example common law benefits.

In workers compensation, redemptions of weekly payments are made in the event that the insurer approves of the redemption and the claimant accepts the redemption amount. The claimant must, however, be in receipt of weekly compensation payments, i.e. the redemption is made provided that the claimant has the status of incapacity.

Let

$c_{Rj}(s,m)$ = expected redemption payment ("R") at claim age j to a claimant who is

currently in spell m of status s (s equals i *or* a as defined in Section 5.1), but who may be in any status at j *(7.1)*

$c_{Rj}(s,m,k)$ = expected redemption payment at claim age j to a claimant who is

currently in spell m of status s given that the status at j is k (k equals i *or* a) *(7.2)*

Note that when $k = a$ at claim age j then

$$c_{Rj}(s,m,a) = 0 \quad (7.3)$$

$f^{s}_{mg}(d)$ = probability that a claim in day g of spell m in current status s, will be incapacitated at age of claim d days *(7.4)*

For $j \geq j_v + t$, where j_v is the age of claim at the valuation date *(7.5)*

The derivation of $f^{s}_{mg}(d)$ is the subject of Section 9, and is defined more explicitly in Section 9.2.

Assuming low mortality risk, and no benefits are payable after retirement, then

$$c_{Rj}(s,m) = \Sigma_k \, c_{Rj}(s,m,k) \, . \, Pr\{s{=}k \text{ at } j\}$$

$$= c_{Rj}(s,m,i) \, . \, f^{s}_{mg}(d) \quad (7.6)$$

with (7.3) taken into account.
The term $c_{Rj}(s,m,i)$ is derived as described in Section 6, using the subset of claimants that are incapacitated at each claim age. s and m are a subset of the covariates relating to claimant X. Therefore using similar notation to (6.1),

$c_{Rj}(s,m,i) \quad = \quad f_{Rj}(X,i)\, s_{Rj}(X,i) \qquad (7.7)$

$f_{Rj}(X,i)$ is modelled as described in Section 6.2, except that

$N_{Rj}(X,i)$ = observed number of claims in status i with covariates X that received a redemption payment at claim age j *(7.8)*

$E_{Rj}(X,i)$ = number of claims in status i with covariates X_i at claim age j *(7.9)*

$s_{Rj}(X,i)$ is modelled as described in Section 6.3. *(7.10)*

8. Event-related benefits depending on other event-related benefits

8.1 General considerations

These types of benefits are those that are paid when a claim attains a particular status or when a defined event or events occur. In the cases considered in the present section, the occurrence of the defined event in turn depends on another event-related benefit.

Example 1 – Pain and Suffering Lump Sum payments

Example 2 of Section 6.1 describes the Specific Injury lump sum benefit type. Within the statutory benefit structure of workers compensation, there may also be a lump sum Pain and Suffering (P&S) award. The P&S component is not paid unless the specific injury lump sum payment has been approved and made. Therefore, the incidence of payment of the lump sum P&S component is contingent on the specific injury payment.

In the Australian workers compensation example, the severity of the P&S payment is highly correlated with the severity of the specific injury payment. The severity of the pain and suffering component can therefore be modelled as a function of the specific injury payment.

Example 2 – Claimant's Common Law Legal Costs

The claimant's common law legal expenses may be payable by the insurer if the claimant is successful at common law. Payment of these expenses occurs only if a common law payment is made. Often, there is a high correlation between the amounts of legal expenses and the common law settlement (with legal fees payable as a percentage of the settlement), and the severity can likewise be modelled as a function of the size of common law settlement.

8.2 Frequency and severity

An example of how this type of benefit can be modelled is provided.

Let

$c_{pj}(X_i)$	=	expected cash flow for payment type p at age of claim j days to claimant X_i	*(6.1)*
p^*		denote the event-related payment type on which payment type p depends	*(8.1)*
k	=	expected lag between payment of p^* and p, $k \geq 0$	*(8.2)*
j'	=	$j+k$	*(8.3)*

$c_{pj'}\left(X_i \mid P_i^*\right)$ = expected cash flow for payment type p at age of claim j' days to claimant X_i given that a payment P_i^* *of* type $p*$ has been made at age of claim j *(8.4)*

$f_{pj'}\left(X_i \mid p*\right)$ = expected frequency of payment type p at age of claim j' given that $p*$ has been *(8.5)*

$N_{pj'}\left(X_i \mid p*\right)$ = observed number of claimants with covariates X_i that received a payment of type $p*$ that also received a payment of type p *(8.6)*

$E_{pj'}\left(X_i \mid p*\right)$ = observed number of claimants with covariates X_i that received a payment of type $p*$ that are exposed to receiving a payment of type p *(8.7)*

$s_{pj'}\left(X_i \mid P_i^*\right)$ = expected severity of payment type p given the severity P_i^* of $p*$ *(8.8)*

The frequency $f_{pj'}\left(X_i \mid p*\right)$ is modelled using $N_{pj'}\left(X_i \mid p*\right)$ and $E_{pj'}\left(X_i \mid p*\right)$ as described in Section 6.2.

To recognise the dependency on the severity $P*$ of payment type $p*$, $s_{pj}\left(X_i \mid P_i^*\right)$ may be modelled as a function of $P*$:

$$E\left[s_{pj}\left(X_i \mid P_i^*\right)\right] = s\left(P_i^*, X_i\right) \tag{8.9}$$

for some function $s(.)$, eg

$$s\left(P_i^*, X_i\right) = P_i^* \text{ x function } \left(X_i\right) \tag{8.10}$$

In the case of (8.10), (8.9) becomes

$$E\left[s_{pj}\left(X_i \mid P_i^*\right) / P_i^*\right] = \text{function}\left(X_i\right) \tag{8.11}$$

which means that the payment of type p is modelled as a covariate-dependent proportion of the payment of type $p*$.

Once a dependent variable, such as the left side of (8.11), has been chosen, the model is specified by choice of error and link functions, as in Section 3.

$s_{pj'}(X_i \mid p^*)$ is modelled as described in Section 6.3. It may be modelled as a proportion of p*, or as some other function of p*.

In this particular case, where payment of p is made only if p^* is made, then

$$c_{pj'}(X_i) = f_{p^*j}(X_i)\, c_{pj'}(X_i \mid P_i^*) \quad (8.12)$$

$$c_{pj'}(X_i \mid p^*) = f_{pj'}(X_i \mid p^*)\, s_{pj'}(X_i \mid P_i^*) \quad (8.13)$$

Variations of the above model may include:

- A modification of (8.12) (and (8.4) – (8.11)) to allow for payments of type p given p^* is **not** made

  That is:

  $$c_{pj'}(X_i) = f_{p^*j}(X_i)\, c_{pj'}(X_i \mid P_i^*) + \left[1 - f_{p^*j}(X_i)\right] c_{pj'}(X_i \mid \bar{p}^*)$$

  where $\bar{p}^*$ indicates no payment of type p^*.

- $s_{pj'}(X_i \mid P_i^*)$ may not necessarily be a direct function of P_i^* if there is no statistically significant relationship.

9. Incidence functions

9.1 Motivation

In the consideration of frequency of incapacity payments, Figure 5.2 illustrated the binomial branching of claim status with increasing time. It is evident that, t days after the commencement of incapacity, there are 2^t outcomes of the incapacity/activity process to be considered, each with a different probability.

It would usually be necessary to model the process day by day in this manner for some time after commencement of incapacity (eg for 5 to 10 weeks), so that even this initial stage of the claim would involve between 2^{25} and 2^{50} possible outcomes (5 working days per week). It is not computationally feasible to forecast frequency of incapacity payments for individual claims in this way.

One means of achieving computational feasibility is to process these many possibilities for a judiciously selected set of hypothetical forecasts, and to interpolate this set for actual cases.

Central to this process are the so-called **incapacity incidence functions**

$i_m(j,\beta_i,\beta_a)$ = probability that a claim, commencing spell m of incapacity with incapacity risk score β_i and activity risk score β_a, experiences incapacity on the j-th subsequent day. *(9.1)*

The incapacity and activity risk scores are those defined in Section 5.1. It is **not** required that incapacity be continuous up to the j-th day, nor that the spell of incapacity in progress then still be the m-th. In practice, i_m will usually be monotone decreasing in j, though this is not a logical necessity.

Parallel **activity incidence functions** are also defined:

$a_m(j,\beta_i,\beta_a)$ = probability that a claim, commencing spell m of activity with incapacity risk score β_i and activity risk score β_a, experiences incapacity on the j-th subsequent day. *(9.2)*

Note that both i_m and a_m measure probability of incapacity.

9.2 Application and derivation

Application

The probability of incapacity on each future day for any claim, not necessarily at the start of a spell of incapacity or activity (as in i_m and a_m), may be calculated relatively simply using those incidence functions.

For $m \geq 1$, define

$I_{mj}^{s}(d,\beta_i,\beta_a)$ = probability that a claim in day j of spell m in current status s, with risk scores β_i and β_a, will be incapacitated at age of claim d days, as measured from date of occurrence of claim.

Thus, for $m \geq 1$,

$$I_{mj}^{a}(d_m + r,\beta_i,\beta_a) = \sum_{k=0}^{r-j-2} p_m^{ai}(j,k)\, i_{m+1}\left[r - j - (k+1)\right] \text{ for } r \geq j+1 \qquad (9.3)$$

where

d_m = age of claim at commencement of spell m in status a

$p_m^{ai}(j,k)$ = probability that a claim in day j of spell m of activity makes a transition to spell m + 1 of incapacity at the end of day $j + k$ of spell m of activity. *(9.4)*

Both i_{m+1} (see (9.1)) and p_m^{ai} depend on risk scores β_i and β_a, but these arguments are suppressed in (9.3) for the sake of brevity.

Note that $p_m^{ai}(j,k)$ is a multi-step probability corresponding to the single-step probability $d_m^{s}(t)$ defined in (5.1). It may be written in the form:

$$p_m^{ai}(j,k) = \frac{S_m^{a}(j) - S_m^{a}(j+k)}{S_m^{a}(j)}$$

where, as in Section 5, $S_m^{s}(t)$ depends on risk scores β_i and β_a derived from covariates X.

Similar to (9.3):

$$I_{mj}^{i}(d_m + r) = S_m^{i}(r)/S_m^{i}(j) + \sum_{k=0}^{r-j-2} p_m^{ia}(j,k)a_m\left[r - j - (k+1)\right] \text{ for } r \geq j+1 \qquad (9.5)$$

with $p_m^{ia}(j,k)$ defined parallel to (9.4), $S_m^{i}(\bullet)$ as defined in Section 5, with d_m now defined as:

d_m = age of claim at commencement of spell m in status i

and with the convention that

$$\sum_{k=0}^{-1} = 0. \qquad (9.6)$$

Corresponding to (9.3), but for $m = 0$:

$I_{0g}^{a}(h,\beta_i,\beta_a)$ = probability that a claim, reported g days ago, currently in spell 0 of activity, with risk scores β_i and β_a, will be incapacitated h days hence

$$= \sum_{k=0}^{h} p_0^{ai}(g,k)\, i_1(h-k) \qquad (9.7)$$

where

$p_0^{ai}(g,k)$ = probability that a claim reported g days ago, currently in spell 0 of activity, makes a transition to spell 1 of incapacity at the end of k further days.

Derivation

To evaluate the incidence function i_m, note that incapacity j days after commencement of spell m can occur in either of two ways:

- the claim continues in status i for all j days; or
- it makes a transition to status a (spell m) on some day before the j-th, but is in status i again (at whatever spell > m) on the j-th day.

By this reasoning,

$$i_m(j) = S_m^i(j) + \sum_{k=0}^{j-2} S_m^i(k)\; p_m^{ia}(k)\, a_m(j-k-1) \qquad (9.8)$$

where

$$p_m^{st}(k) = p_m^{st}(k,1) \quad \text{for } s,t = i,a \qquad (9.9)$$

and the convention (9.6) applies.

It may be seen that $p_m^{st}(k)$ for $s \gamma t$ is the same as $d_m^s(k)$ defined in (5.1). The fact that these quantities are defined in terms of the $S_m^s(k)$, together with (9.8) and (9.10), indicates that the survivor functions $S_m^i(\cdot)$ and $S_m^a(\cdot)$ are the atoms of the incidence functions.

It then follows from (9.3), (9.5) and (9.7) that all incidence quantities $I_{mj}^s(\cdot)$ are derivable from these atoms.

Similarly,

$$a_m(j) = \sum_{k=0}^{j-1} S_m^a(k)\; p_m^{ai}(k)\; i_{m+1}(j-k-1). \qquad (9.10)$$

The values of the incidence functions i_m and a_m are thus derived by a recursive process that uses (9.8) and (9.10). It is initialised at

$$i_m(0) = 1. \tag{9.11}$$

Then

$$i_m(1) = S_m^i(1) \tag{9.12}$$

$$\begin{aligned} a_m(1) &= S_m^a(0)\, p_m^{ai}(0)\, i_{m+1}(0) \\ &= p_m^{ai}(0) \end{aligned} \tag{9.13}$$

$$\begin{aligned} i_m(2) &= S_m^i(2) + S_m^i(0)\, p_m^{ia}(0)\, a_m(1) \\ &= S_m^i(2) + p_m^{ia}(0)\, a_m(1) \end{aligned} \tag{9.14}$$

etc.

Note that, for m = 1, (9.12) and (9.13) illustrate the first two branches in Figure 5.2. Then (9.14), which may be put in the form $i_m(2) = S_m^i(2) + p_m^{ia}(0)\, p_m^{ai}(0)$, illustrates the two branches ending with incapacity in Day 2.

Generally,

- calculation of $i_m(j)$ requires prior calculation of $a_m(k),\ k = 0,1,\ldots,j-1$
- calculation of $a_m(j)$ requires prior calculation of $i_{m+1}(k), k = 0,1,\ldots,j-1$.

Therefore, the recursion may be carried out in the following order:

$$\begin{array}{llllllll} i_1(1) & a_1(1) & i_2(1) & a_2(1) & \cdots & a_M(1) & & \\ & i_1(2) & a_1(2) & i_2(2) & \cdots & i_M(2) & a_M(2) & \\ & & i_1(3) & a_1(3) & \cdots & a_{M-1}(3) & i_M(3) & a_M(3) \end{array}$$

where M is the maximum value of m to which the calculations are taken.

Calculations are performed row by row, and it is apparent from (9.8) and (9.10) that each entry may be obtained from entires in previous rows either directly above or to the left.

It may be that the incapacity process becomes independent of spell number at sufficiently high spells, ie $S_m^i(\bullet), S_m^a(\bullet)$ independent of m *f* some M. It then follows from (9.8) and (9.10) that $i_m(\bullet)$ and $a_m(\bullet)$ are independent of m *f* M, enabling the calculation of these quantities to be terminated at m = M, as illustrated above.

Interpolation

Recall from Section 9.1 that $i_m(j)$ is explicitly $i_m(j,\beta_i,\beta_a)$. Similarly for $a_m(j)$. The calculation scheme set out above applies to specific values of β_i,β_a.

As mentioned in Section 9.1, the incidence function may be pre-processed for a sample of values of β_i,β_a, and then interpolated for other values of these parameters. Care is needed in the interpolation as the functions are far from linear in β_i,β_a. This is illustrated in Figures 9.1 to 9.3.

Figure 9.1
Plot of $i_1(10)$ for various risk scores

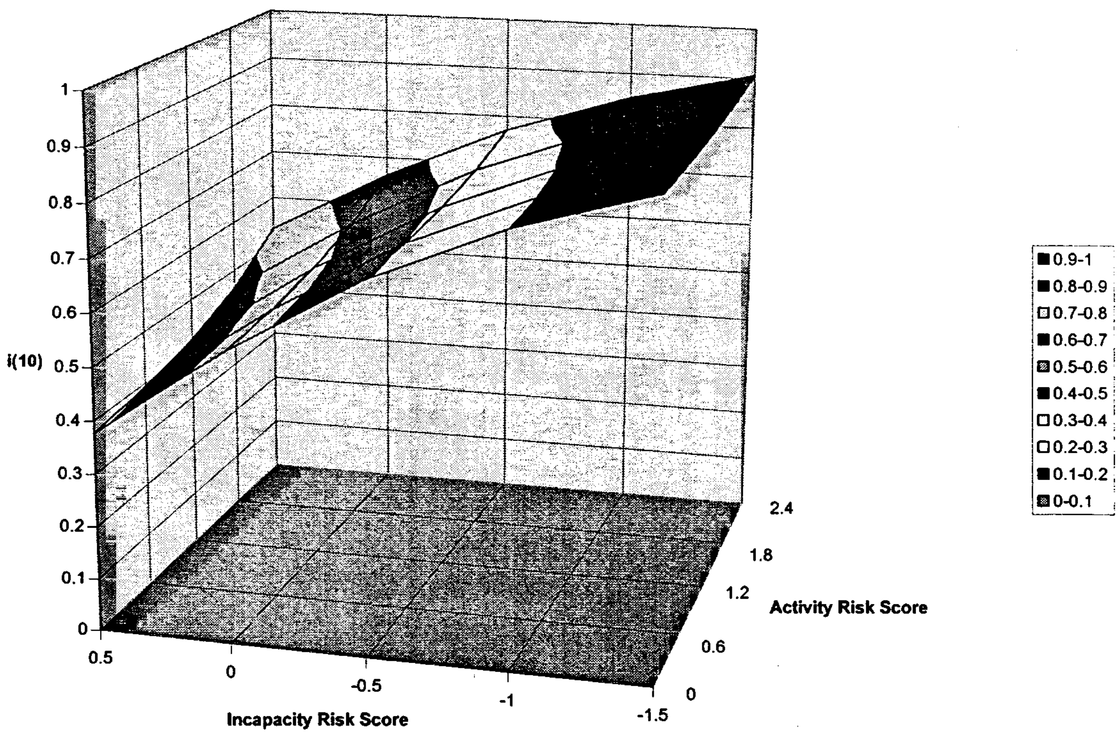


Figure 9.2
Plot of $i_1(100)$ for various risk scores

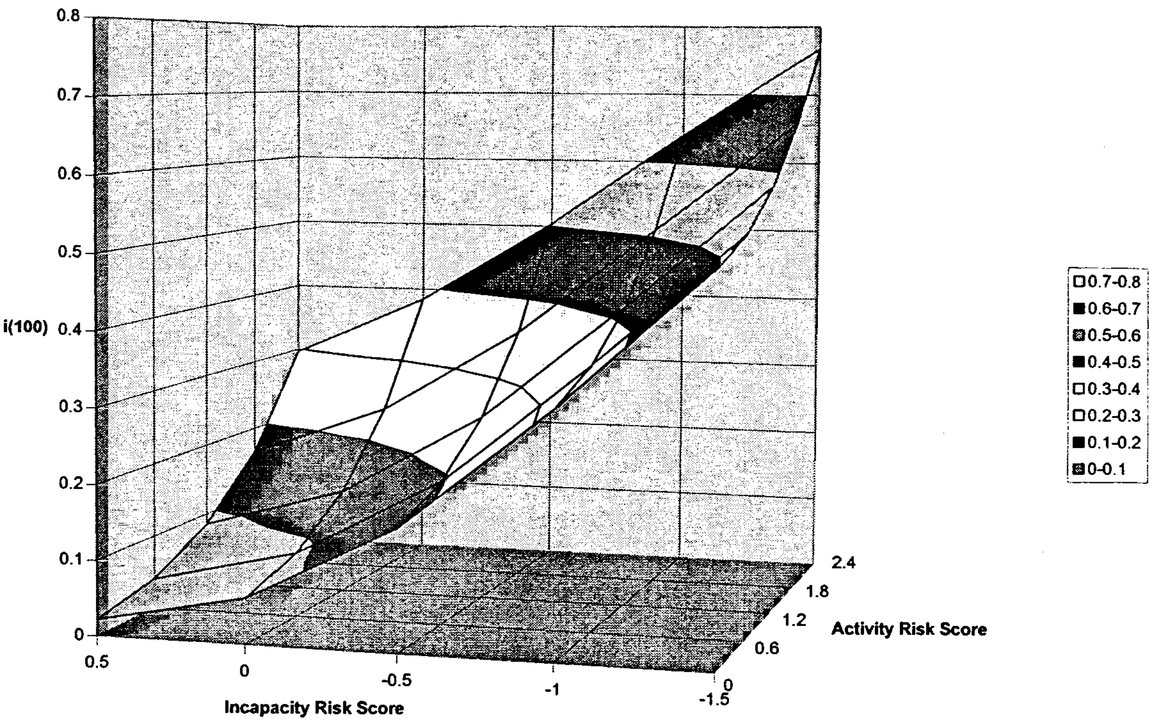
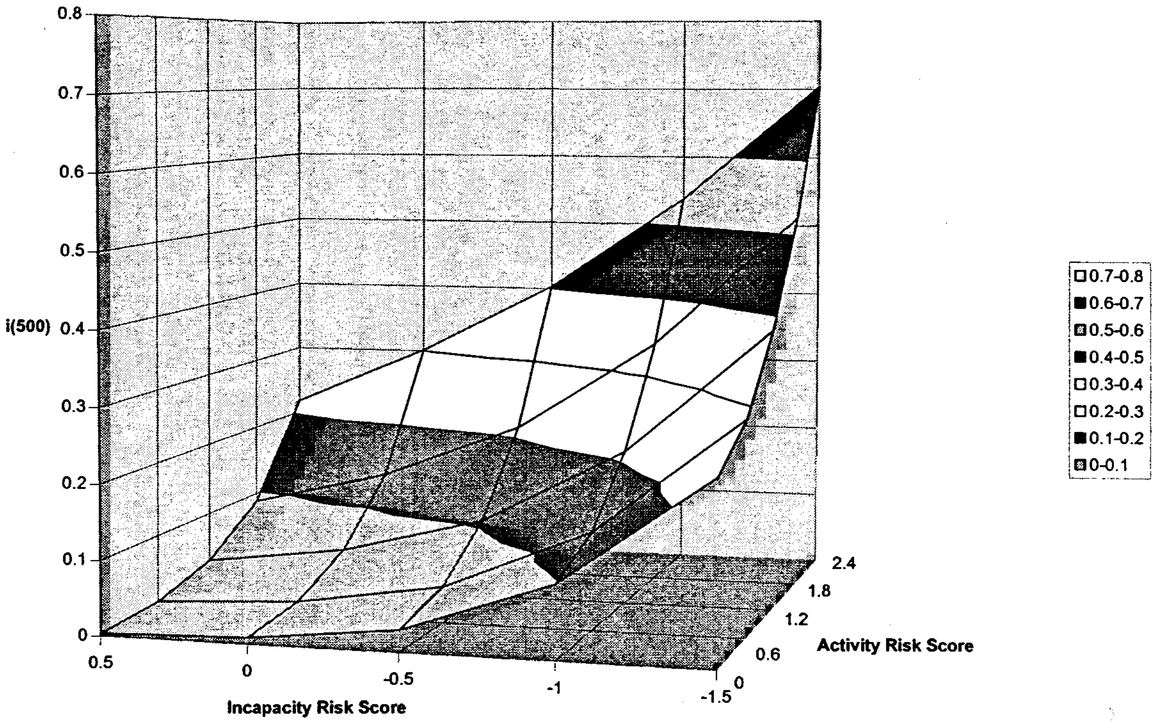


Figure 9.3
Plot of $i_1(500)$ for various risk scores



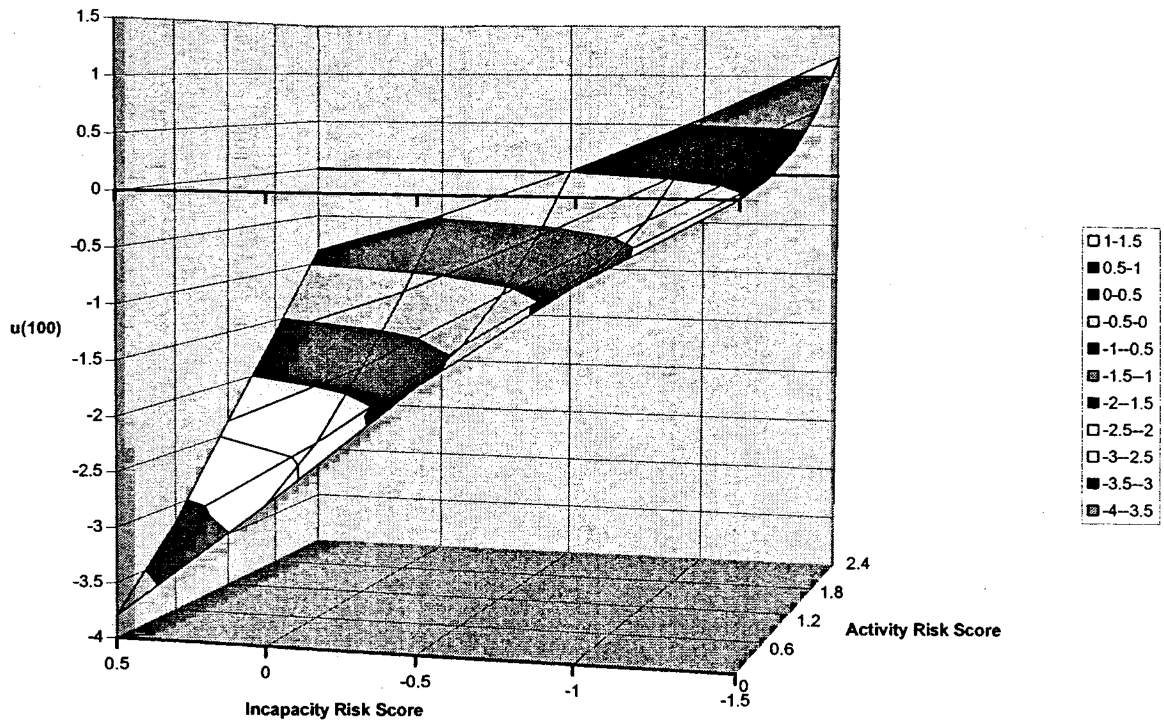
Much of the non-linearity of the incidence functions is forced by their confinement within the interval [0,1]. This suggests that a logit transform might be useful.

Define

$$\begin{aligned} u_m(j,\beta_i,\beta_a) &= \operatorname{logit} i_m(j,\beta_i,\beta_a) \\ &= \log\left[\frac{i_m(j,\beta_i,\beta_a)}{1-i_m(j,\beta_i,\beta_a)}\right]. \end{aligned} \tag{9.15}$$

Similarly for a_m. These functions are much more suitable for interpolation, as Figure 9.4 illustrates. A more linear surface is obtained for values of $i_1(100)$ close to its extremes, in this case $i_1(100)$ close to 0.

Figure 9.4
Plot of $u_1(100)$ for various risk scores



Reasonable results are obtained in the above example by means of 2-dimensional second difference polynomial interpolation. Each such interpolation requires 6 neighbouring (β_i, β_a) points for which the relevant incidence function is tabulated.

Extrapolation

Tabulated values of $i_m(j)$ and $a_m(j)$ will be subject to some upper limit *J* or *j*. Values of the incidence functions for *j* > *J* will need to be extrapolated.

This will require the identification of a functional form that fits the incidence functions reasonably accurately for values of *j* μ *J*. Generally, the continuance of incapacity seems often approximated by a Weibull survivor curve (see eg Worrall and Butler (1985)):

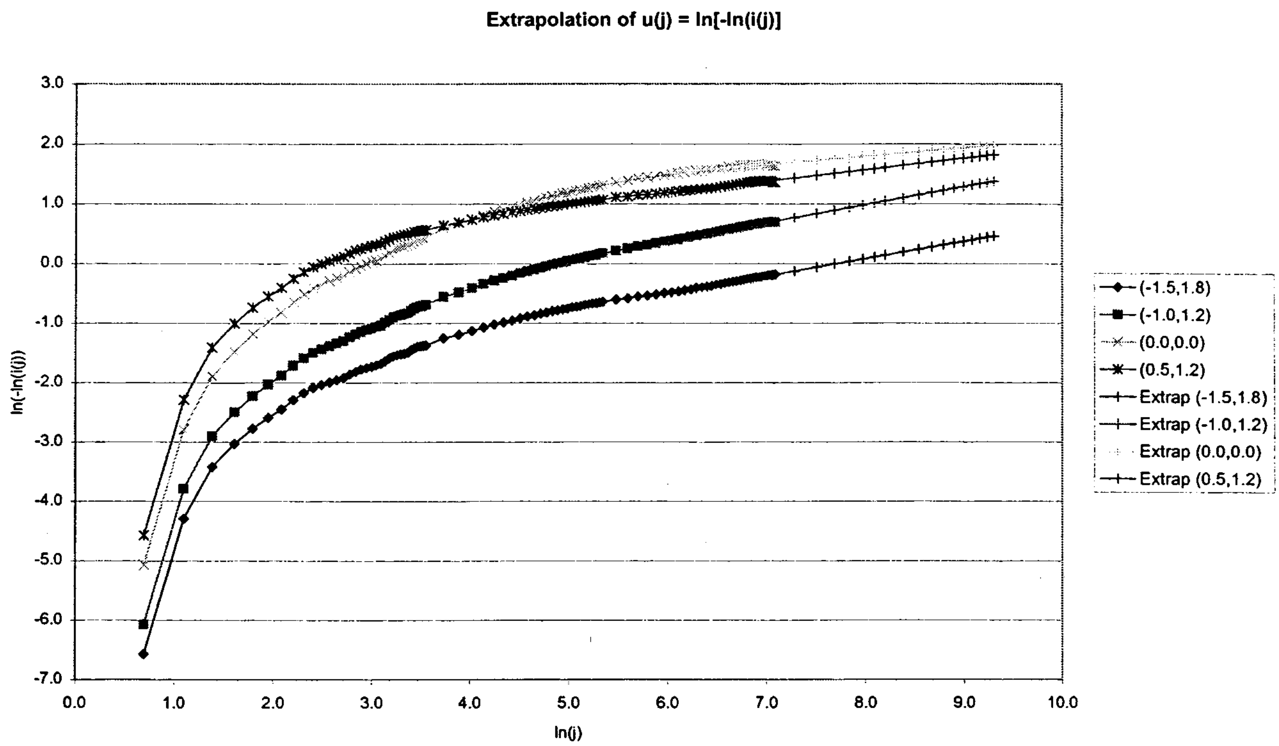
$$i_m(j) = \exp\left[-A_m j^{b_m}\right]. \qquad (9.16)$$

This yields

$$\log\left[-\log i_m(j)\right] = \log A_m + b_m \log j \qquad (9.17)$$

indicating linearity of a plot of $\log\left[-\log i_m(j)\right]$ against $\log j$. Figure 9.5 presents this diagnostic plot for a range of values of β_i, β_a, where $J = ?$ Separate plots for different m are unnecessary since the effects of m are included here in β_i and β_a.

Figure 9.5
Weibull diagnostic plots for incidence functions



The Weibull fit is seen to be satisfactory for j µ J and might be used to extrapolate for $j > J$.

Computation reduction

All discussion to this point has been in terms of computing $i_m(j)$ and $a_m(j)$ at unit intervals up to $j = J$, with j measured in **days**.

It is evident from (9.8) and (9.10) that the number of computations involved in $i_m(j)$ and $a_m(j)$ for m □ M and for a single combination β_i, β_a is of the order Mj^3. This can be time-consuming for large j.

Moreover, the incidence functions typically change only slowly at large j, and so their tabulation at daily intervals is unnecessary. Accurate estimates can be obtained from more widely spaced intervals (eg weekly).

Computation can therefore be reduced, but sufficient accuracy maintained, by means of a tabulation of incidence functions as follows (for example):

j in days up to 35 days
j in weeks up to 210 days (30 weeks)
j in months thereafter.

While this reduces computation,

- it requires that (9.8) and (9.10) be adapted to the different units of j
- it requires that the incidence functions be interpolated for untabulated values of j.

Death and retirement

Define

$S^{\{i,a\}}(x,k,\beta_i,\beta_a) =$ Probability that a claim with risk scores β_i,β_a and with status in the set $\{i,a\}$ at age (of **claimant**) x remains within that set (ie does not transfer to d or r) up to age $x+k$.

It has been implicitly assumed here that such probabilities are functions of the claimant's age (not the age of claim), and hence independent of incapacity experience.

Then allowance for death and retirement is made by the adjustment:

$$I^s_{mj}(d_m+r,\beta_i,\beta_a) \to S^{\{i,a\}}(x,r-j,\beta_i,\beta_a) I^s_{mj}(d_m+r,\beta_i,\beta_a) \qquad (9.18)$$

where age of claimant x corresponds with age of claim d_m+j.

Various simplifications of $S^{\{i,a\}}$ are possible. If death and retirement are stochastically independent over infinitesimal intervals, then

$$S^{\{i,a\}} = S^{\{i,a\}d} S^{\{i,a\}r} \qquad (9.19)$$

where the arguments x,k,β_i,β_a have been suppressed and where $S^{\{i,a\}s}$ denotes survival in status set $\{a,i\}$ when status s is the only alternative.

If retirement age is fixed at R, then

$$\begin{aligned} S^{\{i,a\}r}(x,k,\beta_i,\beta_a) &= 1 \text{ for } k \le R-x \\ &= 0 \text{ for } k > R-x. \end{aligned} \qquad (9.20)$$

The mortality factor may be expressed in the form:

$$S^{\{i,a\}d}(x,k,\beta_i,\beta_a) = \exp\left[-\int_0^k \mu(x+t,\beta_i,\beta_a)\right] dt \qquad (9.21)$$

where $\mu(y,\beta_i,\beta_a)$ is the force of mortality at age y when the risk scores are β_i,β_a.

In practice, one might set $\mu(y,\beta_i,\beta_a)$ to be some simple adjustment of the population u_y, even independent of β_i,β_a. This would ignore the fact that claimant mortality is likely to be heavier in cases of greater incapacity. On the other hand, mortality will usually affect most SCEs relatively little.

It should be noted that the adjustment (9.18) for death and retirement assumes that these effects are excluded from I^s_{mj} as calculated earlier in the present sub-section. This in turn means that the estimation of the basic survival models S^i_m and S^a_m on which I^s_{mj} is based should treat death and retirement as right-censoring events.

The alternative is not to censor such observations. In this case death and retirement would appear simply as terminations of incapacity, equivalent (from a financial standpoint) to a return to activity. If this treatment is adopted, then the adjustment (9.18) becomes unnecessary, indeed incorrect.

What is essential is **consistency** between the right censoring recognised in estimation of survivor functions and the statuses for which adjustment is made in (9.18) and (9.19).

This alternative treatment may be effective as an indirect means of incorporating the **experienced** mortality $\mu(y,\beta_i,\beta_a)$ in the survival models. It will, however, deal poorly with retirement, smearing the discrete effect (9.20) over a wide range of ages.

A reasonable compromise might be:

- to treat retirement, but not death, as a right censoring event in the estimation of survivor functions; and
- retain adjustment (9.18), but eliminate the mortality factor from the right side of (9.19).

10. Forecasting claim costs

Section 4 describes the fundamental composition of an SCE. Sections 5 to 9 discuss its parameterisation. The present section is concerned with the application of the parameterised model to the forecast of claim costs for a single claim.

Order of evaluation of forecast quantities

As indicated by Figures 2.2 and 2.3, the model is likely to comprise a cascade of payment type sub-models. It is necessary to work through this cascade **from the bottom up**, as any one component may require the output of a subsidiary component as input.

The output required of each component includes:

- a schedule of forecast cash flows by future payment period (see Section 4.1), which usually will require in turn (Section 4.3) separate schedules of (Section 4.3):
    - forecast frequencies of payment of the subject payment type; and
    - forecast severities of those payments;
- any other outputs required as subsequent inputs.

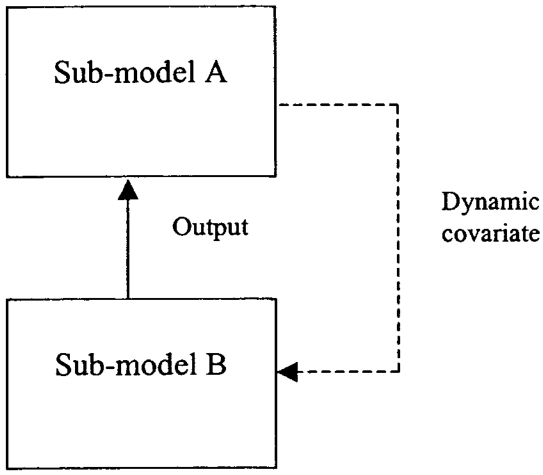
Section 1.1 mentioned that covariates may be static or dynamic. Calculation of the latter (e.g. the total incapacity experienced by the claimant to date, mentioned in Section 6.1) for any period requires forecasts (perhaps of its own value) from the preceding period. This carries implications for the architecture of the system.

Feedback loops

If the model contained only static covariates, it would be possible to process each component sub-model in turn, generating its required forecasts for all future periods on the basis of a fixed set of covariates. Likewise, if the only dynamic variables contained in the model were dependent on only forecasts from the sub-models for which they serve as covariates, of from subsidiary sub-models.

Consider, on the other hand, the situation depicted in Figure 10.1, in which an output quantity of sub-model A is required as a dynamic covariate input to subsidiary sub-model B, creating a feedback loop.

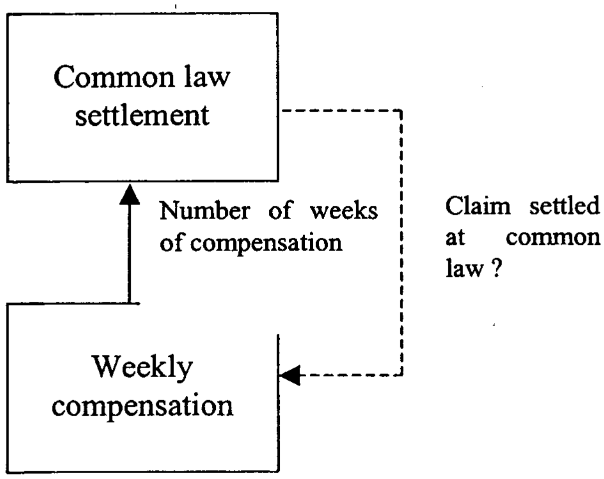
Figure 10.1
Dynamic variable feedback loop



A specific example of this might be as in Figure 10.2, which reflects two features of a hypothetical scheme of workers compensation, viz.:

- The likelihood of a settlement at common law increases with the total number of weeks of compensation paid; and
- If a claim is settled at common law, then weekly compensation will be discontinued.

Figure 10.2
Example of a feedback loop



In the case of a feedback loop it will not be possible to process the subsidiary sub-model(s) involved in the loop over all future periods prior to processing the sub-model(s) that depend on them. It will be necessary to evaluate the entire loop for time period t, then for t+1, and so on.

The architecture of the forecast system will depend on the prevalence of feedback loops. Each will need to be evaluated fully at each future time t before moving on to t+1. If there are a number of them, the simplest approach might be to evaluate all forecast quantities at time t=1, then all at t=2, then t=3, etc.

Incidence functions

The incidence functions described in Section 9 would not be evaluated in real time, but would be pre-processed and stored as a set of hard coded tables in the forecast system.

A set of tables of incapacity incidence functions $i_m(j,\beta_i,\beta_a)$ (see (9.1)) would be evaluated, each table for a fixed triple (m,β_i,β_a) and for j=1,2,etc. The pairs (β_i,β_a) would need to be chosen to produce a lattice covering the majority of the range of values of these parameters likely to arise in practice, with values of $i_m(j,\beta_i,\beta_a)$ for intermediate β_i, β_a obtained by real-time interpolation or extrapolation, as described in Section 9.

For each chosen pair β_i, β_a, a table of $i_m(j,\beta_i,\beta_a)$ would be required for each triple (m,β_i,β_a), m=1,2,...,M, with M as defined in Section 9.

Output

As Section 4 .1 indicates, the minimum requirement of the forecast system will be the production of a schedule of forecast claim payments $c_{pt}(X_i)$, p=1,2,...; t=1,2,... for each claim i. This schedule can be summarised in any way desired, e.g. the quantities $c(X_i)$, $c^*(X_i)$, etc. defined in Section 4.1.

11. Model validation

11.1 Background

It is evident from Figure 2.3 that an SCE derives from the cascading of a number of models. The combination of these models will typically involve a number of assumptions, eg

- stochastic independence of frequency and severity in (4.11)
- the precise Markovian assumption made in respect of status related benefits, as discussed in Section 5.1.

These assumptions create scope for the final model to perform poorly even when each component fits the data well. It is therefore essential to carry out a strict validation of the model, perhaps leading to adjustment (Section 12), before implementation.

Validation can take the form of **back testing**, ie comparison of past data with "predictions" (sometimes called "retrodictions") of the model. It is probably advisable to validate individual payment types. These are additive, and so validation of all of them implies validation of the entire model.

Preferably, data will have been divided into two subsets:

- a **training sample**, used for estimation of model parameters; and
- a **validation (or hold-back) sample**, used for comparison with predictions.

This enables **cross-validation**, the testing of the model against a sample of data other than that used to develop it.

The following few subsections discuss in a little detail the form validation might take. Appendix B provides a formal description.

11.2 Model components to be validated

As in all forms of stress testing, one needs to identify each potential point of weakness in the model, and test it. Thus, for each payment type, there is a need to validate frequency and severity sub-models.

Moreover, there will be a need in some cases to validate sub-sub-models, eg for status related benefits, the models of transition between statuses. In general, wherever a component model has been created it should be validated.

Thus, the specific validations to be made might include:

Weekly compensation:
- Recovering from incapacity
- Relapse from activity
- Incapacity incidence functions
- .
- .
- .

Medical costs:
- Frequency
- Severity

etc.

The weekly compensation validations might be carried out separately for spell 1, spell 2, etc.

11.3 Validation observations

Each validation will consist of a comparison of observations on claims experience with model predictions. It is therefore necessary to define these "observations".

Natural choices for the validations set out in Sections 11.2 would be as in Table 11.1.

Table 11.1
Validation observations

Model	Validation observation
Recovery from incapacity[a]	Number of days of incapacity[a]
Relapse from activity[a]	Number of days of activity[a]
Incapacity incidence functions[a]	Number of days of incapacity[a]
.	
.	
.	
Medical frequency	Number of claims generating medical costs
Medical severity	Average payment under payment type Medical per claim generating medical costs
etc.	

Note: (a) As noted in Section 11.2, these validations might be separated by spell number.

Typically, the validation sample would relate to a defined experience period (eg calendar year 2000), and so all validation observations, such as appear in Table 11.1, would relate to this period.

11.4 Data partitions

While, in the first model of Table 11.1, one might compare just the **total** number of days of incapacity (perhaps separately by spell number) with the corresponding model prediction, there will usually be value in dissecting these totals into relevant components.

For example, each model will have formulated predictions on the basis of claim attributes, and so these effects may be tested by dissection of the above totals in terms of the attributes. Thus, the observed and predicted number of days of incapacity might be tabulated by:

Gender of claimant
Age of claimant
Nature of injury
etc.

Such tabulations, according to single claim attributes, are usually referred to as **1-way comparisons**. It is also possible to produce 2-way comparisons, eg by gender x age, and in general *m*-way comparisons.

The general concept here is one of partitioning of validations in some defined manner. These are therefore referred to generically as **data partitions**. While, for expository purposes, they have been described above in terms of claim attributes, they are defined more generally in the theoretical treatment of validation given in Appendix B.

11.5 Validation output

According to Sections 11.2 to 11.4, validation will involve a collection of tables comparing claims experience with model predictions, with each table characterised by:

- the component sub-model under test (Section 11.2)
- the experience period (Section 11.3)
- the data partition (Section 11.4)
- the validation observations (Section 11.3).

Each such table might take the form set out in Table 11.2.

Table 11.2
Validation table format

Description of model: Partition: Experience period:					
Partition element	Observation				
	Actual	Model prediction	Deviation	Relative deviation	Significance
TOTAL					

Here

- "Deviation" records the difference between actual and prediction
- "Relative deviation" expresses the deviation as a percentage of the prediction
- "Significance" is the statistical significance of the deviation (against a null hypothesis of zero), and will be discussed further in Section 13.1.

A hypothetical example of Table 11.2 appears as Table 11.3. It indicates a model which over-estimates the probability of recovery from incapacity in the low order spells and under-estimates it in the high order spells.

Table 11.3
Hypothetical example of validation output

Description of model: Partition: Experience period:	Recovery from incapacity Spell number 2000 Q4				
Spell number	Number of days of incapacity				
	Actual	Model prediction	Deviation	Relative deviation	Significance
				%	%
1	3,018	2,745	+273	+10	<0.1
2	1,873	1,818	+55	+3	9
3	1,153	1,098	+55	+5	4
4	550	573	-23	-4	15
5	332	390	-58	-15	1
≥ 6	295	321	-26	-8	8
TOTAL	7,221	6,945	+276	+4	0.2

11.6 Validation output medium

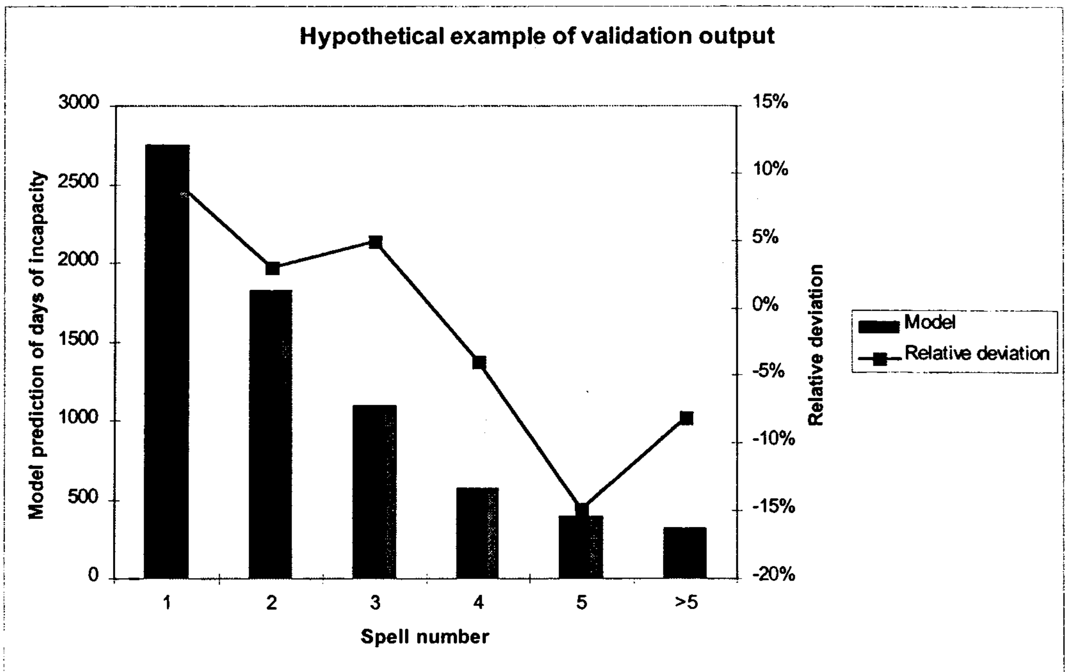
It is evident from Sections 11.2 and 11.4 that many comparisons of experience with model will be required to constitute a comprehensive validation. The comparison process will be assisted if tables of the form of 11.3 are available on demand.

The ideal medium for this is a data base for each model organised in what is sometimes referred to as **data mart** form. The data base is considered an n-cube where each of the n dimensions represents one value range for one variable, or equivalently one row of one validation output table.

The different output tables then represent slicings of the cube along different dimensions. The slicings can conceivably be m-dimensional for any $m < n$. Table 11.3 illustrates the case $m = 1$. The case of general m corresponds to the m-way comparisons mentioned in Section 11.4. A data base which pre-packages the various m-way comparisons is a data mart.

There is clear scope for graphical presentation of output as an aid to decision making. Thus, Table 11.3 might be presented in the form of Figure 11.1.

Figure 11.1



This graph clearly exhibits the trend in relative deviation with increasing spell number. The model predictions have been included as an indicator of volume experience.

12. Model calibration

12.1 Background

Section 11.1 points out that an SCE model might prove inaccurate in its totality even if all components were well fitted to data. For example, perfectly reasonable models of transitions between incapacity and activity may have been obtained. But, if the Markovian property assumed (see Section 5.1) breaks down, the associated incidence functions may fail to be validated.

In such circumstances a simple device for adjusting one or more of the fundamental models will be useful.

12.2 Control parameters

The atoms of the model are defined by (3.1) (GLM) and (3.14) (survival model). In each case, it is possible to insert a small number of additional parameters which will cause a fundamental shift in the model.

For example, (3.1) might be extended to:

$$\hat{Y}_i = h^{-1}\left(X_i^T\hat{\beta}+\theta\right) \tag{12.1}$$

where

$\hat{\beta}$ is an estimate of β
$\hat{Y}_i$ is the model prediction of Y_i
θ is the additional parameter.

Note that, if the model has an intercept term, as in (3.2a), changing θ is equivalent to shifting the intercept β_0. This may be used to shift the **level** of all predictions $\hat{Y}_i$ up or down by a constant factor.

The introduced quantity θ will be called a **control parameter**.

Consider the other example, (3.14). Here simple adjustments might be applied to modify the shape of the survivor functions. For example, (3.14) might be extended to:

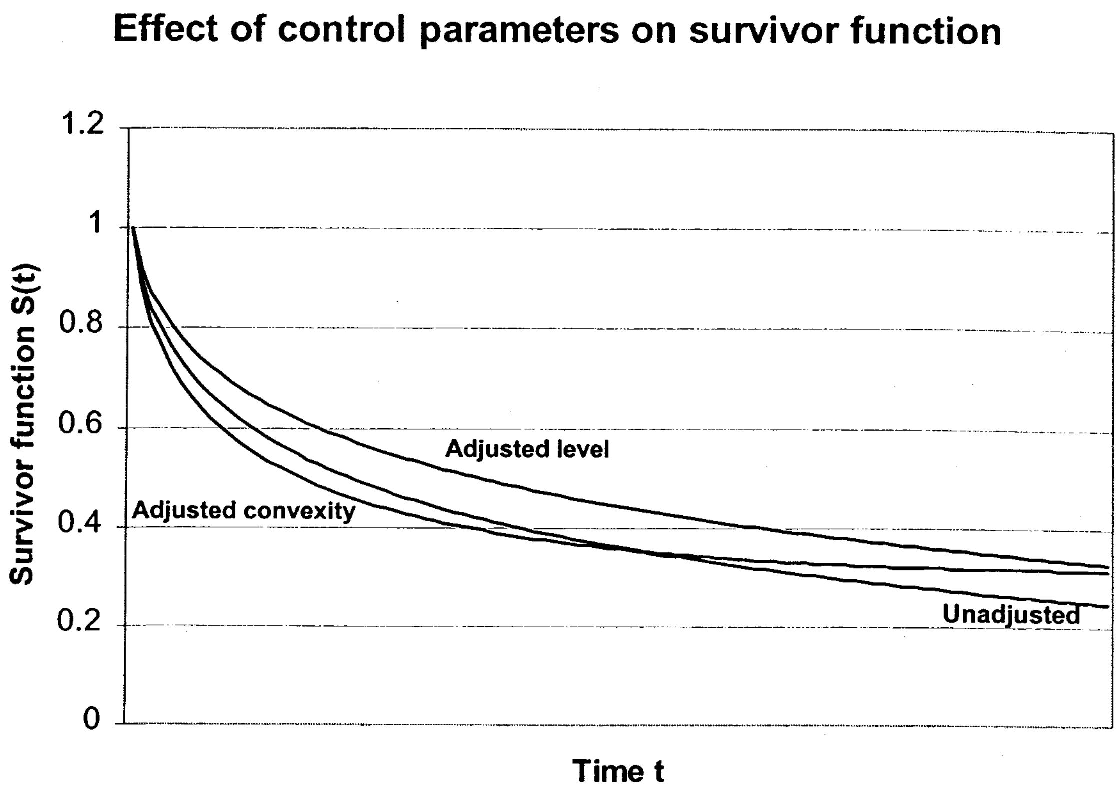
$$\hat{h}_i(t) = \theta\hat{h}_0(t)\exp\left(X_i^T\hat{\beta}\right) \tag{12.2}$$

or even

$$\hat{h}_i(t) = [\theta_0+\theta_1 t]\hat{h}_0(t)\exp\left(X_i^T\beta\right). \tag{12.3}$$

Figure 12.1 illustrates how the control parameters θ_0,θ_1 change a survivor function.

Figure 12.1



The whole curve can be shifted up or down by means of θ_0. It may be made more or less convex by means of θ_1.

12.3 Calibration

The term calibration is used here to refer to the adjustment of a small number of control parameters to correct any failures observed at the model validation stage. It should typically involve simple adjustments of modest size. If these requirements cannot be met, the validity of the whole model framework might be questionable.

13. Stochastic properties

13.1 Theoretical

The basic structure of an SCE was given by (4.5), repeated for convenience here:

$$c(X_i) = \sum_p \sum_{t=1}^{\infty} c_{pt}(X_i). \tag{13.1}$$

The $c_{pt}(X_i)$ relate to a collection of models of form (3.1) or (3.9), (3.12) and (3.14), each of which is stochastic. The SCEs are in fact predictions of stochastic quantities.

Write $C_{pt}(X_i)$ and $C(X_i)$ for the stochastic quantities relating to $c_{pt}(X_i)$ and $c(X_i)$, and now let $c_{pt}(X_i)$ and $c(X_i)$ be their expected values:

$$C_{pt}(X_i) = c_{pt}(X_i) + \xi_{pti} \tag{13.2}$$

with

$$E[\xi_{pti}] = 0. \tag{13.3}$$

Define

$$\begin{aligned} C(X_i) &= \sum_p \sum_{t=1}^{\infty} C_{pt}(X_i) \\ &= c(X_i) + \xi_i \end{aligned} \tag{13.4}$$

with $c(X_i)$ given by (13.1) and

$$\xi_i = \sum_p \sum_{t=}^{\infty} \xi_{pti} \tag{13.5}$$

$$E[\xi_i] = 0. \tag{13.6}$$

The SCEs realised by the system are **predictions** of the $C(X_i)$, and are also random variables. Write $\hat{C}_{pt}(X_i)$ to denote the prediction of $C_{pt}(X_i)$, and assume that all predictions are unbiased, ie

$$\hat{C}_{pt}(X_i) = c_{pt}(X_i) + \zeta_{pti} \tag{13.7}$$

with

$$E\left[\zeta_{pti}\right]=0. \tag{13.8}$$

Define

$$\begin{aligned}\hat{C}(X_i) &= \sum_p \sum_{t=1}^{\infty} \hat{C}_{pt}(X_i) \\ &= c(X_i)+\zeta_i\end{aligned} \tag{13.9}$$

where (13.1) has been used and

$$\zeta_i = \sum_p \sum_{t=1}^{\infty} \zeta_{pti} \tag{13.10}$$

$$E[\zeta_i]=0. \tag{13.11}$$

Consider any collection of SCEs:

$$\hat{C} = \sum_i \hat{C}(X_i) \tag{13.12}$$

with i varying over some specified subset of all claims. The **prediction error** associated with $\hat{C}$ is

$$\phi = C - \hat{C} \tag{13.13}$$

where

$$C = \sum_i C(X_i). \tag{13.14}$$

By (13.12) and (13.14)

$$\phi = \sum_i \phi_i \tag{13.15}$$

$$\begin{aligned} &= \sum_i \left[C(X_i) - \hat{C}(X_i)\right] \\ &= \xi - \zeta \end{aligned} \tag{13.16}$$

by (13.4) and (13.9) and where

$$\xi = \sum_i \xi_i, \qquad \zeta = \sum_i \zeta_i. \tag{13.17}$$

This is the standard time series formulation in which prediction error is equal to the difference between:

- the noise in the system (ξ); and
- the error in the model (ζ) due to parameter estimation error.

Note that ξ relates to future observations, whereas ζ relates to predictions, which will have been generated by a model, which will depend in turn on past observations. It will usually be valid, therefore, to assume that ξ and ζ are **stochastically independent**.

In this case

$$V[\phi]=V[\zeta]+V[\xi]. \tag{13.18}$$

This is the time series result that is commonly expressed as:

Mean square error of prediction	=	parameter error	+	process error	(13.19)

Note also, however, that the correlation structures **within** ξ and ζ will usually be complex. By (13.10) and (13.17),

$$\zeta=\sum_i\sum_p\sum_t\zeta_{pti}. \tag{13.20}$$

So

$$\begin{aligned}V[\zeta] &= \sum_{i,j}\sum_p\sum_t \mathrm{Cov}\left[\zeta_{pti},\zeta_{ptj}\right]\\ &+\sum_i\sum_{\substack{p,q\\p\neq q}}\sum_t \mathrm{Cov}\left[\zeta_{pti},\zeta_{qti}\right]\\ &+ \text{various other terms.}\end{aligned} \tag{13.21}$$

Although i and j relate to different claims, they will be subject to the same parameter error, and so $\mathrm{Cov}\left[\zeta_{pti},\zeta_{ptj}\right]\not\equiv 0$.

Because of payment type dependencies of the sort illustrated in Figure 2.3, $\mathrm{Cov}\left[\zeta_{pti},\zeta_{qti}\right]$ may be non-zero for some distinct p, q.

Similarly for other terms in (13.21), and consequently general statements about $V[\zeta]$ are difficult. One generality can be deduced, however, by considering the special case of (13.20) in which all claimants are identical. The claims themselves are separate and in fact stochastically independent but, as pointed out above, their parameter errors are identical.

Then (13.20) becomes

$$\zeta=n\sum_p\sum_t\zeta_{pti} \tag{13.22}$$

where n is the number of values of i, and hence

$$V[\zeta] = n^2 \sigma_{pa}^2 \tag{13.23}$$

where

$$\begin{aligned}\sigma_{pa}^2 &= V\left[\sum_p \sum_t \zeta_{pti}\right] \\ &= \text{parameter error for a single claim.}\end{aligned} \tag{13.24}$$

Similar arguments can be applied to $V[\xi]$. Again one can reason, for example, that some $\text{Cov}[\xi_{pti}, \xi_{qti}]$ may be non-zero for distinct p, q.

Note, however, that if distinct claims are stochastically independent, then

$$\text{Cov}\left[\sum_p \sum_t \xi_{pti}, \sum_p \sum_t \xi_{ptj}\right] = 0 \tag{13.25}$$

and so

$$V[\xi] = V\left[\sum_i \sum_p \sum_t \xi_{pti}\right] = \sum_i V\left[\sum_p \sum_t \xi_{pti}\right]. \tag{13.26}$$

In the case of identical claims, this becomes

$$V[\xi] = n\sigma_{pr}^2 \tag{13.27}$$

with

$$\begin{aligned}\sigma_{pr}^2 &= V\left[\sum_p \sum_t \xi_{pti}\right] \\ &= \text{process error for a single claim.}\end{aligned} \tag{13.28}$$

Combining (13.18), (13.23) and (13.27) gives

$$\begin{aligned}V[\phi] &= n^2 \sigma_{pa}^2 + n\sigma_{pr}^2 \\ &= n\sigma_{pa}^2 (n + \tau)\end{aligned} \tag{13.29}$$

with

$$\tau = \sigma_{pr}^2 / \sigma_{pa}^2. \tag{13.30}$$

If the training sample is large, parameter error will be small, and τ will be large. Hence (13.29) demonstrates that:

- for individual claims, or small subsets of claims, prediction error is dominated by process error (noise);

- for large sets of claims, prediction error is dominated by parameter error;
- prediction error changes roughly linearly with sample size for small samples, but roughly quadratically for large samples.

Reasoning like this in terms of groups of identical claims is, of course, restrictive. However, if one views large sets of claims as the aggregation of subsets that are approximately replicates of one another, then the above reasoning goes through in an approximate form with σ^2_{pa} and σ^2_{pr} now relating to a single "typical" claim, ie a claim with claim attributes that are in some sense average.

13.2 Application

As noted in Section 13.1, each sub-model of the SCE model discussed here is stochastic, and so the distributions of predictions generated by it are in principle derivable.

This is unlikely to be a practical course, however. For example, the complexities of incidence functions, such as seen in (9.8) and (9.10), would render any attempt to derive these distributions analytically from the stochastic properties of the underlying survivor models highly laborious.

Add to this the additional complexity noted in Section 13.1 as arising from the various stochastic dependencies within the model, and a project of parametric calculation of distribution of predictions does not appear feasible.

Resort needs to be had to other means, of which the **bootstrap** (Efron and Tibshirani, 1995) is a prime candidate. A basic knowledge of this technique will be assumed here. There are a few points to be made in connection with it.

First, the re-sampling of data required by the bootstrap should be relatively routine, but each observation in the re-sampling needs to comprise **all data** (ie in respect of all payment types) for the claim concerned. Thus, the pseudo-observations generated would contain, for example, a pseudo-history of weekly compensation as well as medical payments, etc. Proceeding in this way ensures that the correlations between the various SCE sub-models are properly bootstrapped.

Second, it would be useful to retain the whole ensemble of values $\hat{C}_{pt}(X_i)$ for each bootstrap replication of the data set. This would enable one to calculate the distribution of each of these component predictions rather than just $\hat{C}$. Hence, one would be able to examine the distributions of various relevant dissections of $\hat{C}$, eg

- by payment type
- by accident period

etc.

Third, (13.29) may be used to estimate the effect of sample size on prediction error. Indeed, it may also be used to enable bootstrapping to be performed on a sample of the data set, so reducing computation. If m is the size of this sample, then the prediction error for a data set of size n is:

$$\left(\frac{n}{m}\right)^2 \text{ x parameter error} + \left(\frac{n}{m}\right) \text{ x process error} \qquad (13.31)$$

where the parameter and process errors relate to the m-sample. It is necessary here that the m- and n-samples have the same structure in terms of covariates, ie can be regarded as samplings of the same population.

Fourth, it must be realised that the sort of bootstrapping discussed here will be computationally intensive **in the extreme**. The choice of a modest value of m can greatly reduce computation, but the derivation of incidence functions (Section 9) is typically the most intensive part of the entire SCE procedure and it is independent of sample size. This creates a formidable obstacle to bootstrapping, and it is proper for the authors to confess that they have not implemented the techniques discussed in this section.

14. Miscellaneous matters

14.1 Performance tracking

Section 11 discussed model validation. This was carried out by means of comparison of:

- claims experience from a validation data set; with
- the corresponding predictions of that experience on the basis of a training data set.

Table 11.3 gave a numerical example of validation.

This testing is carried out prior to the implementation of the model, but exactly the same concepts extend to post-implementation periods. In this case, the model predictions of a period's claims experience included in the SCEs made at the start of that period are compared with the actual experience as it emerges.

For example, SCEs made at 31 December 2000 comprise predicted claim payments for 2001, 2002, etc. At the end of the first quarter of 2001 it will be possible to compare claim payments made in that quarter with those predicted at 31 December.

These comparisons can be carried out in exactly the same detail as in Section 11. They then provide a test of whether, after its initial validation, the model continues to track well against experience.

As at the validation stage, decisions on this question will be aided by the significance testing contemplated in Table 11.3. This requires an evaluation of the model's stochastic properties, discussed in Section 13.

14.2 System maintenance

The possible outcomes of performance tracking are:

- no tracking failure;
- isolated failures (eg a single age group within a single sub-model);
- systematic but localised failure (eg incidence of incapacity consistently higher or lower than predicted); or
- widespread and substantial failure of one or more sub-models.

The appropriate responses to these outcomes would be (in order):

- no action;
- isolated adjustment (in the example cited, a single parameter relating to the offending age group might be adjusted);
- **re-calibration**, ie adjustment of control parameters (Section 12.2) so as to restore the model to alignment with experience;
- **re-parameterisation**, or even structural **re-design**, of the failed sub-models.

In the final option, "re-parameterisation" refers to the re-estimation of model parameters on the basis of more recent experience, but with the algebraic structure of the model retained.

In the most extreme circumstances, the algebraic structure of the model will be found to have failed, in which case re-design will be required. For example, gender might have been included as a covariate in (3.14) for incapacity modelling (ie both genders are subject to the same baseline hazard rate $h_o(t)$), whereas more recent experience indicates that separate baselines are required.

14.3 Practicalities

It may be advisable to overlay the statistical part of the SCE model with a number of additional rules that capture aspects of the governing legislation not directly included in the model, or simply to recognise practicalities.

Suppose, for example, that legislation limited the lump sum specific injury payments displayed in Figure 2.3 to a single settlement in respect of any one claim. Then this payment type's contribution to a particular claim's SCE should become zero immediately such a settlement is flagged as having been made in respect of that claim.

As an example of a simple practicality, consider a claim for which there has been no payment activity for several years. Effectively, the claim is closed. However, the SCE system will typically recognise a small but non-zero probability of relapse, and assign it a small SCE.

This can lead to the system's carrying a large number of trivial SCEs. One might introduce a rule that:

- all SCEs for compensation payments of less than $50 will be ignored; or alternatively
- the total SCE will be set to zero for all claims with no payment activity in the last 2 years.

Such SCEs, while individually trivial, might contribute a material total. If so, some re-calibration of the system (in the sense of Section 14.2) might be required to correct for this.

15. Implementation issues

A project for construction of an SCE system will need to incorporate the following phases:

- Model design;
- Parameter estimation;
- System construction;
- Model validation;
- Calibration;
- Bench testing;
- Documentation;
- Implementation;
- Performance tracking;
- Maintenance.

The phase of **model design** will need to include consideration of which payment types are to be separately recognised. Consideration will also need to be given to control parameters (Section 12.2.), specifically how many are to be included in the system, and where and in precisely what form.

There may be some subjective matters to be considered. For example, if a new benefit type has recently been introduced, or an existing one re-structured in some way, there will not be adequate experience from which to parameterise a forecast model.

It may be necessary to plug some temporary and *ad hoc* structure into the slot for this component. It can be unplugged, and replaced with something more permanent at a later date. For reasons such as this, the implementation of the system in a strictly modular form is essential.

Temporary structures for changed benefits may be definable in terms of those that have become obsolete, with some broad and subjective changes to parameter, such as a reduced frequency of access to the benefit type, or a scaling upward or downward of the entire distribution of the amount payable under that type.

Parameter estimation may require application of both GLMs and Survival Analysis, as in Sections 5 to 8.

Calibration (Section 12) might include adjustment of the model so that its aggregate results are consistent with those of some other valuation of claim liabilities, based on conventional actuarial techniques that use aggregate, as opposed to individual, claim data. Any calibration of this type would need to recognise that the total of SCEs in respect of all currently open claims will include nothing for IBNR claims.

Bench testing will be concerned with establishing system characteristics, particularly performance in terms of timing. Hardware considerations may be critical to the achievement of acceptable run times, if large numbers of claims are to processed in batch mode.

The evaluation of incidence tables can be a time consuming procedure. It is useful to establish at bench testing stage the time involved in each re-calculation and replacement of these.

The **bench testing, documentation and implementation** phases may well be contemporaneous. Together they will form a lengthy procedure. Delivery expectations need to be managed to accommodate this.

Time needs to be allowed for adequate testing of such a complex system. As with all software systems, documentation must be adequate to ensure that successive users/programmers can maintain it when the model structure is changed and/or its parameters re-estimated. If these requirements are not observed, the user will be liable to repent at leisure.

A major issue in the **maintenance** of the system is the period between successive major re-calibrations. As with all data-driven systems, the risk of parametric obsolescence needs to be recognised. This would be monitored by a system of **performance tracking**, in which all the validation procedures of Section 11 would be applied to new data as it accumulated.

A default plan might be to subject the system to major re-calibration every 3 to 5 years. This intention might, however, be moderated by the outcome of performance tracking in the interim. The better the model tracks experience, the less the need for re-calibration. Conversely, the emergence of substantial deviation of experience from model predictions might precipitate early re-calibration.

Between major re-calibrations, one might either:

- Carry out re-calibration of one or more modules of the system; or
- Effect minor adjustments by means of the system control parameters.

Choice between these alternatives would be determined by reference to the extent of departure of experience from predictions. Likewise, the tracking system would identify those modules of the SCE system which were predicting poorly.

Appendix A
Discrete time finite-dimensional stochastic processes

Consider a discrete time process $\{X_t,\ t = 0,1,\text{etc}\}$ such that $X_t \in S = \{1,2,...,S\}$.

The interpretation for present purposes is that $X_t = m$ indicates the process to be in **state** *m* at **epoch** *t* (see eg [reference?]).

Let

$$p_{mn}(t) = \text{Prob}\left[X_{t+1} = n \mid X_t = m, I_t\right] \tag{A.1}$$

where I_t denotes full information about epochs prior to *t*. If $p_{mn}(t)$ is independent of I_t (the process has no memory beyond the immediately preceding state), the process is called **Markovian**.

The probabilities $p_{mn}(t)$ are called **transition probabilities**. Let $P(t)$ denote the *S* x *S* matrix with $p_{mn}(t)$ as its (*m*, *n*)-element. This is called the **transition matrix** between epochs *t* and *t* + 1.

Define the **graph** of the process at epoch *t* as the matrix $\Gamma(t)$ with (*m*, *n*)-element

$$\begin{aligned} g_{mn}(t) &= 1, \text{ if } p_{mn}(t) > 0 \\ &= 0, \text{ if } p_{mn}(t) = 0. \end{aligned} \tag{A.2}$$

The graph indicates which transitions are possible and which are not.

Since the process must go to some state at each epoch,

$$\sum_{n=1}^{S} p_{mn}(t) = 1 \text{ for each } m \tag{A.3}$$

ie row sums of $P(t)$ are each unity.

A state *m* for which

$$p_{mm}(t) = 1 \text{ for each } t, \tag{A.4}$$

and therefore

$$p_{mn}(t) = 0, \text{ for each } n \neq m \text{ and each } t, \tag{A.5}$$

is called an **absorbing state**. Once the process enters such a state, it never emerges from it. Equivalent to (A.4) is the fact that the m-th row of each $\Gamma(t)$ contains 1 on the diagonal and zeros elsewhere.

The quantities (A.1) are **single-step probabilities**. One can also consider **multi-step** probabilities.

$$p_{mn}(s,t) = \text{Prob}[X_t = n \mid X_s = m, I_s]. \qquad (A.6)$$

Note that the Markovian property allows this to be decomposed as follows:

$$p_{mn}(s,t) = \sum_{j=1}^{S} p_{mj}(s,u)\, p_{jn}(u,t), \qquad (A.7)$$

for any $u = s+1, s+2, \ldots, t-1$.

This is a version of the Chapman-Kolmogorov equation (Feller, 1965, p.370).

It may be expanded to express a multi-step probability in terms of just single-step.

$$p_{mn}(s,t) = \sum_{j_1,\ldots,j_r=1}^{S} p_{mj_1}(s,s+1)\, p_{j_1 j_2}(s+1,s+2) \ldots p_{j_r n}(t-1,t) \qquad (A.8)$$

where $r = t - s - 1$.

Insurance defined on the stochastic process

One may consider an insurance that pays a benefit of $B_n(t)$ at time t (= s + 1, s + 2, etc) if the process is then in state n. The expected total benefit payable under this insurance when the system commences in state m at time s is:

$$V_m^{(n)}(s) = \sum_{t=s+1}^{\infty} p_{mn}(s,t) B_n(t). \qquad (A.9)$$

2-state processes

Consider (A.8) in the case $S = 2$. Each of the states $m, j_1, \ldots, j_r, n$ takes the value either 1 or 2. Hence each of the single-step probabilities involves either remaining in the same state or making a transition to the other.

It is therefore possible to re-express (A.8) in the form:

$$\begin{aligned} p_{mn}(s,t) = \sum_{u_1,\ldots,u_q} & p_m(s,u_1)p_{mm^*}(u_1,u_1+1) \\ & \text{x}\, p_{m^*}(u_1+1,u_2)\, p_{m^*m}(u_2,u_2+1) \\ & \vdots \\ & \text{x}\, p_{n^*}(u_{q-1}+1,u_q)\, p_{n^*n}(u_q,u_q+1) \\ & \text{x}\, p_n(u_q+1,t) \end{aligned} \qquad (A.10)$$

where $p_m(s,u_1)$ denotes the probability of remaining in state *m* from epoch *s* to u_1 inclusive, *m** denotes the state other than *m*, and the summation runs over all $u_1,\ldots,u_q$ such that $u_1 \geq s$, $u_2 > u_1,\ldots u_q > u_{q-1}$, $t > u_q$.

Each summand in (A.10) represents:

- a period of occupancy of state *m* (called here a **spell**) from epoch *s* to u_1
- a period of occupancy of state *m** from u_1+1 to u_2
- another period of occupancy of state *m*

etc.

The process, as represented by (A.10), passes through alternating spells in status *m* and *m**. The spells may be assigned **spell numbers**, dating from the first occupancy of a particular state, say *m**. Thus, in (A.10),

- spell 1 in state *m** extends for epoch u_1+1 to u_2
- spell 1 in state *m* extends for epoch u_2+1 to u_3
- spell 2 in state *m** extends for epoch u_3+1 to u_4

etc.

The formulation (A.1) of transition probabilities, together with the Markovian property, requires them to be independent of spell duration. Note, however, that the formulation (A.10) allows this requirement to be weakened. Transition probabilities may be allowed to depend on spell duration.

Indeed, one may generalise further, allowing probabilities to be dependent on spell number as well. Technically this is done by expanding the available states of the system to the Cartesian product
{original state} x {spell numbers} x {spell duration}.

Within this expanded set of states, the Markovian property is still assumed, in which case (A.10) is replaced by the following:

$$p_{mn}(s,t) = \sum_{v_1,v_2,\dots,v_q} p_m(s,u_1)\, d_r^m(x+v_1)$$

$$\text{x}\, S_{r+1}^{m*}(v_2)\, d_{r+1}^{m*}(v_2)$$

$$\text{x}\, S_{r+1}^{m}(v_3)\, d_{r+1}^{m}(v_3)$$

$$\text{x}\, S_{r+2}^{m*}(v_4)\, d_{r+2}^{m*}(v_4)$$

$$\text{x}\dots \qquad (A.11)$$

where the process was at duration x of spell r in state m at epoch s, and where

$S_k^s(v) = \text{Prob}\left[\text{duration of spell } k \text{ in state } s \geq v\right]$ *(A.12)*

$d_k^s(v) = \text{Prob}$ [duration of spell k in state s terminates at duration v, given that duration f v] *(A.13)*

$v_k = u_k - u_{k-1}$ for $k = 2, 3$ etc

$= u_k - s$ for $k = 1$.

In the summation in (A.11), the indexes $v_1,\dots,v_q$ run over all strictly positive integers such that $v_1 + \dots + v_q = t - s$ and q also takes all possible values.

It is possible to generalise (A.11) further, eg by declaring the probabilities there to be dependent on the total time spent in a particular one of the original two states up to the commencement of the present spell, but taken over all previous spells. Such generalisations create further expansion of the set of available states but, within this expanded set, the Markovian property is retained.

Appendix B
General validation framework

B.1 Validating a regression model

Consider a data set $\mathcal{Y} = \{Y_i\}$ subject to model (B.1) below. If μ_i denotes $E[Y_i]$, then

$$\mu_i = h^{-1}(X_i^T \beta) \tag{B.1}$$

where X_i^T is a design matrix applicable to observation Y_i, and β a vector of constant coefficients.

Define the **deviation**

$$D_i = Y_i - \mu_i. \tag{B.2}$$

Let $\mathcal{P} = \{\mathcal{P}_1, ..., \mathcal{P}_m\}$ be a partition of $\mathcal{Y}$ and define

$$D(j) = \sum_{i \in \mathcal{P}_j} D_i \tag{B.3}$$

$$R(j) = \sum_{i \in \mathcal{P}_j} D_i / \sum_{i \in \mathcal{P}_j} \mu_i \tag{B.4}$$

for $j = 1, 2, \ldots, m$.

The $D(j)$ in (B.3) are **grouped deviations**. The $R(j)$ are **grouped relative deviations**, which express the $D(j)$ relative to the associated model values.

Validation would consist of:

- calculating $D(j)$ and/or $R(j)$ for various $\mathcal{Y}$ and various $\mathcal{P}$ within $\mathcal{Y}$
- assessing whether these are statistically significant deviations of data from model.

Example

Considering again the incapacity continuance example used in Section 11.4, let

$\mathcal{Y}$ = observed values of Y_i as defined in Section 11.4

$\mathcal{P}_j$ = j-th defined value assumed by the covariate, gender of claimant ($j = M, F$).

Then $D(M)$ and $D(F)$ measure the total deviation for male and female claimants respectively.

There are many alternative definitions of $\mathcal{P}$, eg one defined for each covariate in the same way as was done for gender in the above example.

B.2 Validating a survival analysis model

B.2.1 Background

Consider a **survival analysis model** whose survivor function in respect of the k-th claim is $S_k(t)$ for some suitable age measure t. Suppose that failure occurs only at integral values of t, and that $S_k(\cdot)$ is right continuous, ie $S_k(t)$ records Prob [survival to age $\mu \geq t$].

Define the discretised hazard function

$$h_k(t) = 1 - S_k(t)/S_k(t-1) = \text{Prob}\left[\text{failure at exact age } t\right], \qquad t = 1, 2, \text{ etc.} \qquad (B.5)$$

Suppose the model is of the proportional hazards type, so that

$$h_k(t) = h_0(t)\exp\left(X_k^T\beta\right) \qquad (B.6)$$

where

$h_0(t)$ = baseline hazard function
X_k^T = design matrix applicable to k-th claim
β = vector of risk scores for covariates.

"Failure" might be recovery from incapacity, for example. This would be so if (B.5) were the model of incapacity survival.

Estimate $h_k(t)$ by

$$\hat{h}_k(t) = \hat{h}_0(t)\exp\left(X_k^T\hat{\beta}\right) \qquad (B.7)$$

where $\hat{h}_0(t)$, $\hat{\beta}$ are estimates of $h_0(t)$, β.

There is a need for separate validation of $\hat{h}_0(\cdot)$ and $\hat{\beta}$.

Consider a specific interval of time. Within that interval, consider claims experience in the interval $(t-1, t]$, ie failures at age t. Let

$E_k(t-1)$ = number of days exposure of the k-th claim to failure at age measure t
$N_k(t)$ = number such failures.

Then the model prediction of $N_k(t)$ is

$$\begin{aligned}\hat{N}_k(t) &= E_k(t-1)\hat{h}_k(t)\\ &= E_k(t-1)\hat{h}_0(t)\exp\left(X_k^T\hat{\beta}\right).\end{aligned} \tag{B.8}$$

Corresponding to (B.2), define the deviation

$$D_k(t) = N_k(t) - \hat{N}_k(t). \tag{B.9}$$

B.2.2 Validating risk scores

The deviations constructed in (B.9) are labelled by both k and t, rather than by the single label i as in Appendix B1. However, as far as the partitioning of the data set is concerned, this does not create any difference in principle from Appendix B.1.

Let $P = \{P_j\}$ be a partition of the data set such that each P_j **includes all possible values of t.** Then, for $D(j)$ defined in (B.3),

$$D(j) = \sum_k \sum_{t=0}^{\infty} D_k(t) \tag{B.10}$$

with summation for k in some subset of all claims.

Example

The example given in Appendix B.1 will serve again. Define P_j in terms of gender, as there. Then, for example, $D(M)$ is obtained from (B.10) by summing over all claims k for which claimant is male.

B.2.3 Validating the baseline hazard function

Let $P = \{P_j\}$ be a partition of the data set such that each P_j **includes all claims.** Then,

$$D(j) = \sum_t \sum_{\text{all } k} D_k(t) \tag{B.11}$$

with summation for t in some subset of the non-negative integers.

Example

For example, $D(j)$ might be defined:

$$D(j) = \sum_{t=u}^{v-1} \sum_{\text{all } k} D_k(t). \tag{B.12}$$

Then $D(j)$ would represent the grouped deviation over the interval $(u,v]$ of age measure.

B.3 Validating a weekly incidence model

Consider a data set $\{N_k^w(t)\}$ relating to a specific time interval, where $N_k^w(t)$ is the indicator function:

$N_k^w(t)$ = 1, if the k-th claim experiences status "incapacity" in development period t;
= 0, otherwise;

where development period is measured from date of accident, and the first is labelled zero.

A model prediction of this quantity is

$$\hat{N}_k^w(t) = E_k^w(t)\hat{I}_k^w(t) \qquad \text{(B.13)}$$

where

$E_k^w(t)$ = 1, if it is logically possible for the k-th claim to experience status "incapacitated" in development period t within the designated interval;
= 0, otherwise

and $\hat{I}_k^w(t)$ is a model estimate of

$I_k^w(t)$ = Prob [k-th claim has status "incapacitated" in development period t].

Validation would be as for a baseline hazard function, ie according to intervals of age measure (development time in the present context) (see Appendix B.2.3).

Let $P = \{P_j\}$ be a partition of the data set such that each P_j **includes all claims.** Then,

$$D(j) = \sum_t \sum_{\text{all } k} D_k(t) \qquad \text{(B.14)}$$

with summation for t in some subset of the non-negative integers.

Appendix C
Example of claim frequency modelling

C.1 Preliminary data manipulation

This appendix provides an example of how the final structure of the frequency of medico-legal expense payment might be determined within a workers compensation framework.

Let

$$f_{it}(X_i) = \text{probability that a medico-legal payment occurs in week } t \text{ in respect of the } i\text{-th claim,} \qquad (C.1)$$

where X_i is based on the following covariates.

Table C.1
Covariates of medico-legal claim frequency

Attribute	Abbreviated name	Categorical or continuous	# categorical levels (initial)
Claims characteristics			
Age at injury	age	Continuous	
Sex	sex	Categorical	2
Nature of injury	noi	Categorical	9
Type of accident	toa	Categorical	10
Injury severity code	sev	Categorical	5
Industry code	ind	Categorical	12
Bodily location of injury	loi	Categorical	10
Claims status characteristics			
Development week (= # weeks from injury date to payment date)	w	Continuous	
Total (categorical)			**48**

To formulate a relationship between $f_{it}(X_i)$ and the listed continuous covariates, one might begin by grouping the latter, as shown in the following table.

Table C.2
Grouping of continuous covariates

Covariate	Intervals for grouping covariate values
Development week (w)	
0 – 40	Unit intervals (41 categories)
41-52	4-week intervals (3 categories)
53-724	12-week intervals (56 categories)
	(Total 100 categories)
Age at injury (age)	< 20
	20-24
	25-29
	30-34
	35-39
	40-44
	45-49
	50-54
	55-59
	60-64
	65+
	Unknown
	(Total 12 categories)

The total number of categories is now 48 + 100 + 12 = 160. Inclusion of all these in a model would be quite prodigal.

It is preferable to gain a rough idea of the functional dependence of $f_{lt}(X_i)$ on w before proceeding. This is done by examining **one-way tables** of observed claim frequency by w, ie claim frequencies observed for various values of w, but otherwise averaged across the entire portfolio.

C.2 One-way tabulation by development week

Define the estimator

$$\hat{f}_l(w) = l(w)/n(w) \qquad (C.2)$$

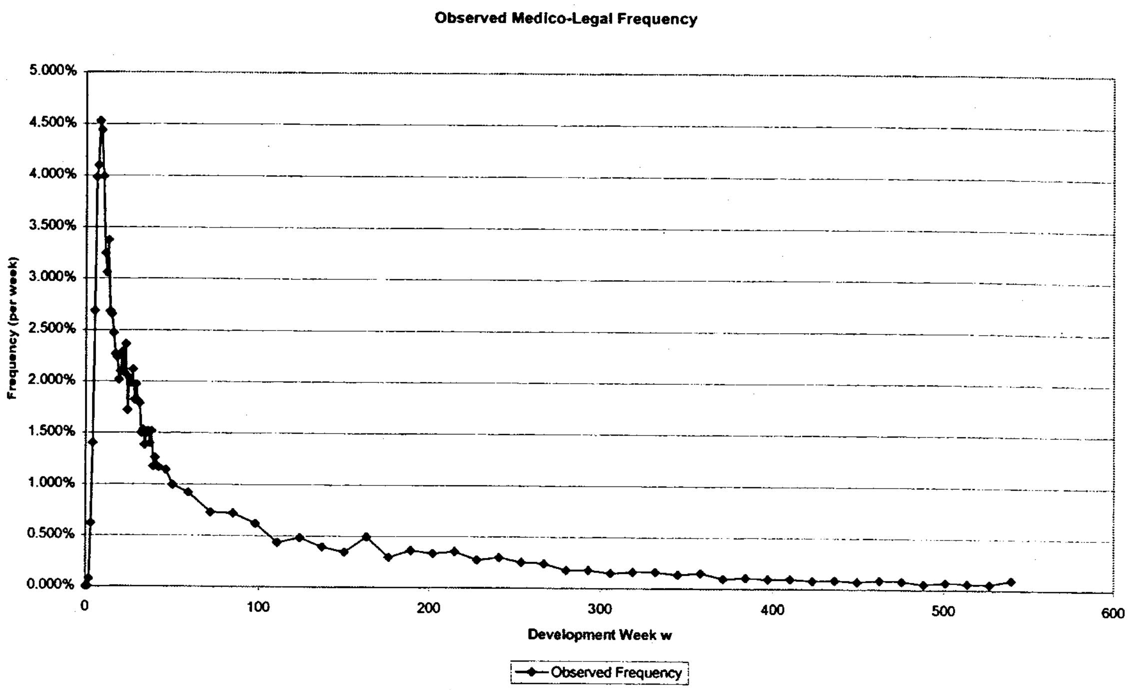
where

$l(w)$ = observed number of claims receiving a medico-legal payment in week w

$n(w)$ = total number of claims in the data set reported by end of week w.

The quantities $\hat{f}_l(w)$ are the entries in the one-way table (labelled by w) described in Appendix C.1. Figure C.1 charts the table.

Figure C.1



C.3 Modelling continuous covariates

C.3.1 Development week

Of the two continuous covariates identified in Table C.1, development week appears the more influential. It is therefore singled out as the first covariate to be converted from categorical to continuous form.

Since $f_{lt}(X_i)$ will be modelled with a binomial distribution using a logit transformation, a continuous functional form of w is developed using

$$f^*(w) = \log\left[\frac{\hat{f}(w)}{1-\hat{f}(w)}\right]. \qquad (C.3)$$

Figure C.2 charts the observed values of $f^*(w)$, against w, and Figure C.3 expands the scale for the lower values of w.

Figure C.2

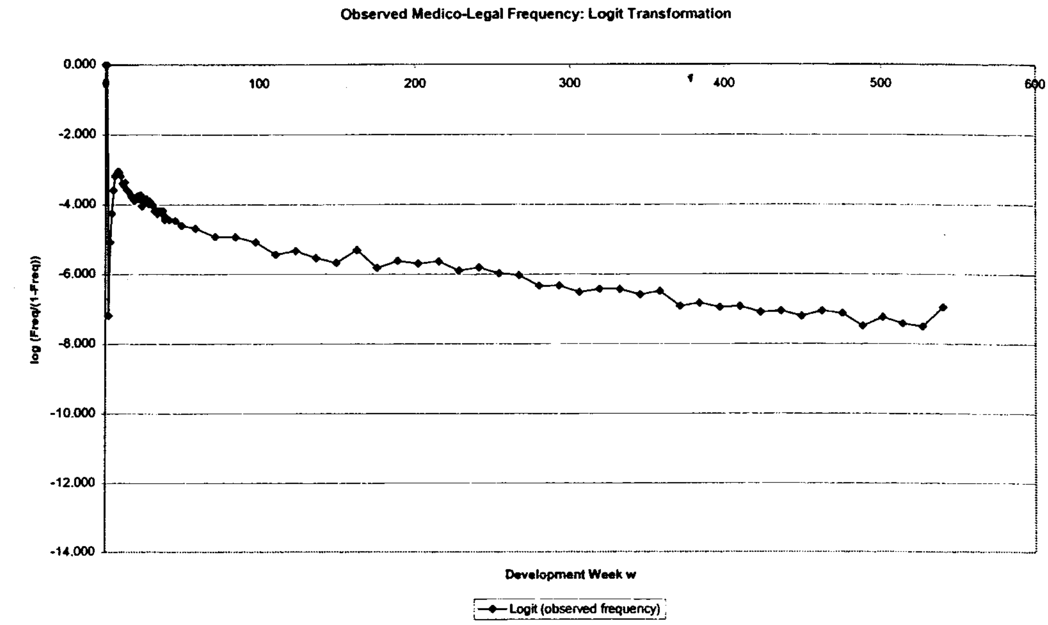
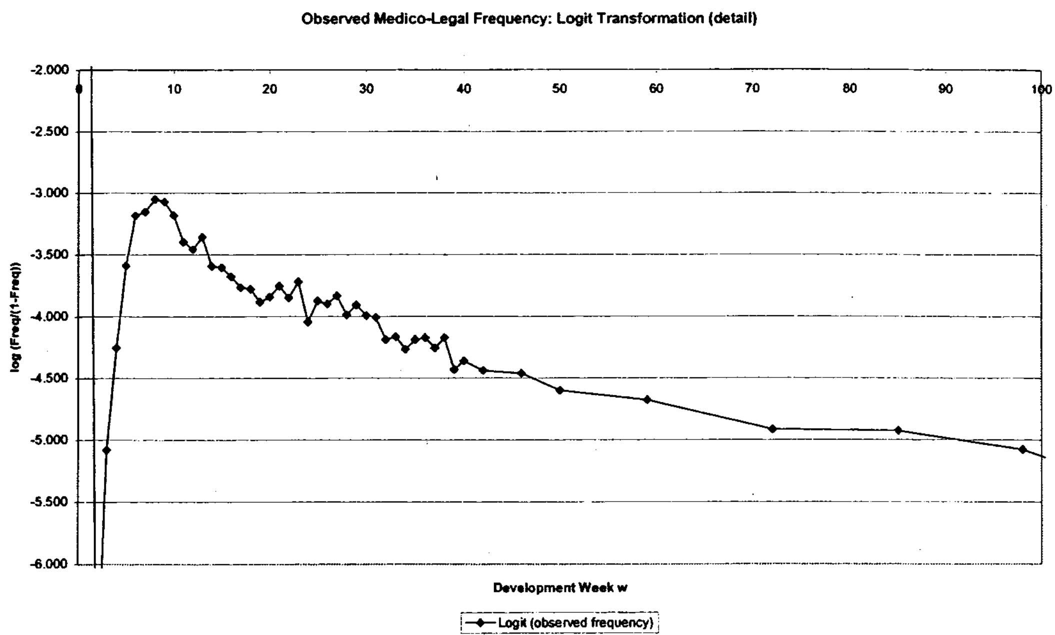


Figure C.3



Local extrema are observed at weeks 4 and 9, and hence $\hat{f}$ is estimated with the following:

$$\begin{aligned}\hat{f}(w) &= \alpha \\ &+\beta_1 \max(0,4-w)^{1.5} \\ &+\beta_2 \max\left(0,\min(w,9)-4\right)^{0.2} \\ &+\beta_3 \max(0,w-9)^{0.4}\end{aligned} \qquad (C.4)$$

with the result (α = -4.2491, β_1 = -1.01454, β_2 = 0.9813, β_3 = -0.3613) illustrated in Figures C.4 and C.5.

Figure C.4

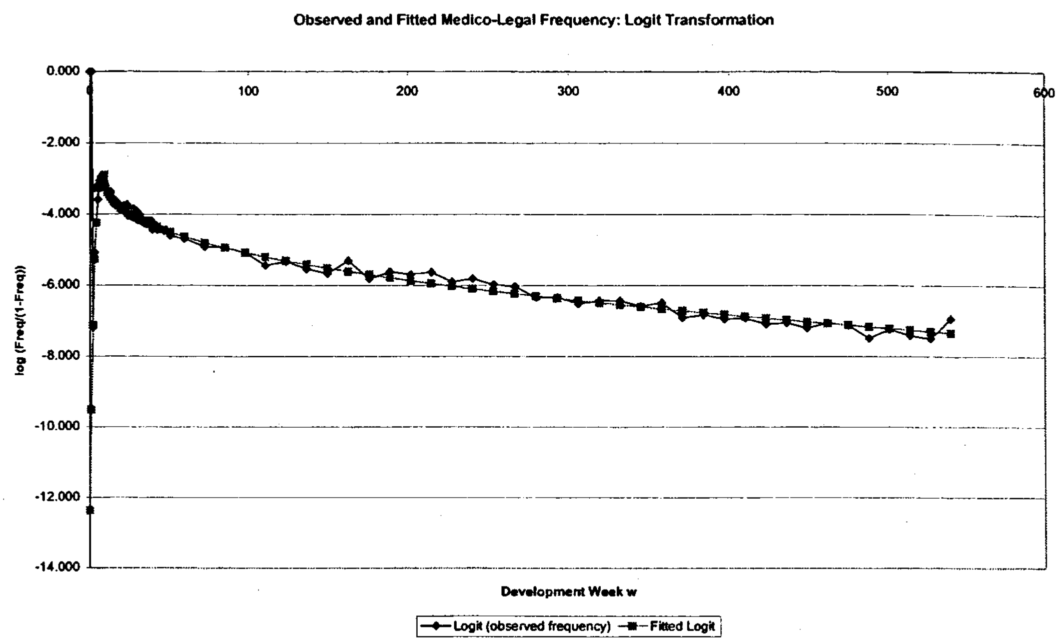
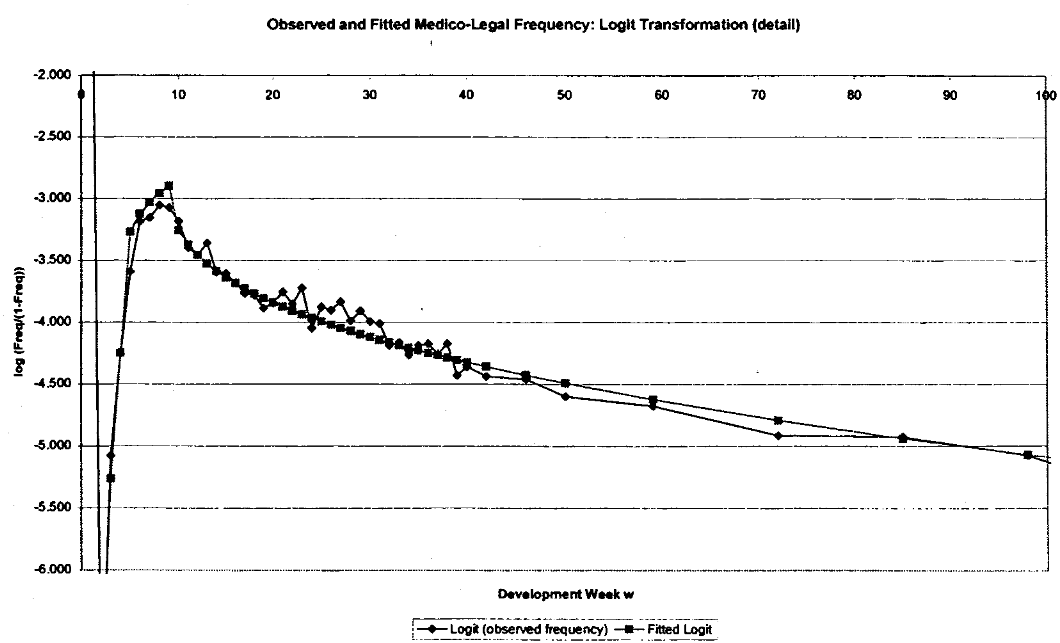


Figure C.5



C.3.2 Age of claimant

The conversion of development week to a continuous variate in Appendix C.3.1 reduces the number of parameters required in a model of Medico-Legal claim frequency to manageable proportions. An initial model based on the covariates identified in Table C.1 and (C.4) is given by the following GLM output.

Table C.2
Model with development week continuous and age categorical

Analysis Of Parameter Estimates

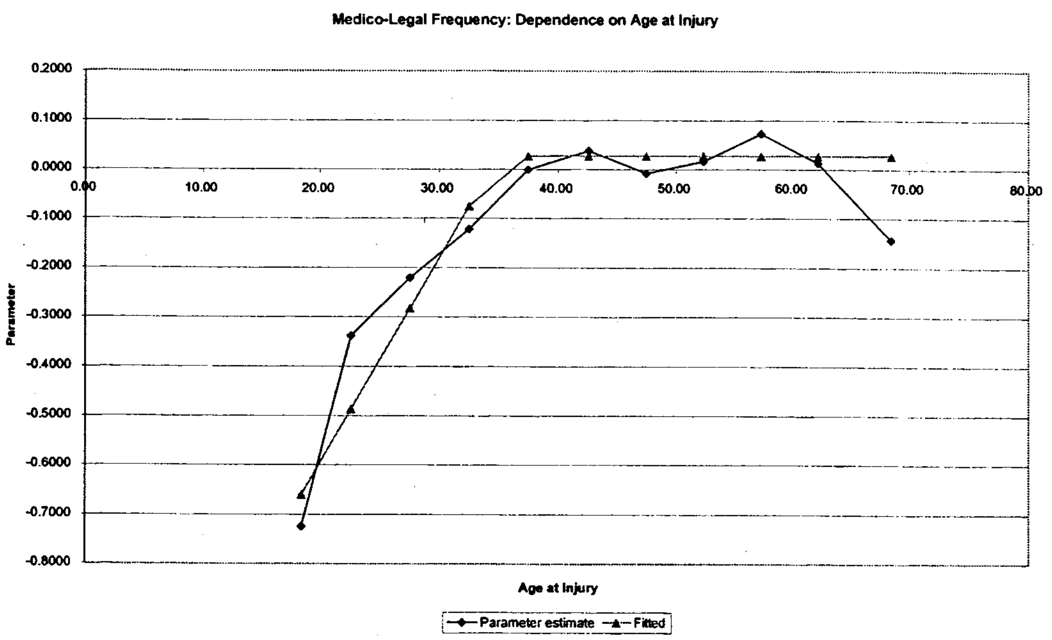
Parameter	Value	DF	Estimate	Std Err	ChiSquare	Pr>Chi	Formatted Value
INTERCEPT		1	-4.208	0.1512	774.6763	0.0001	
D1		1	-0.7449	0.2063	13.0423	0.0003	MAX(0,4-w)$^{1.5}$
D2		1	1.1536	0.1056	119.3823	0.0001	MAX(0,MIN(w,9)-4)$^{0.2}$
D3		1	-0.3746	0.0062	3667.8394	0.0001	MAX(0,w-9)$^{0.4}$
TOA	1	1	-0.0855	0.0392	4.7474	0.0293	Person Falling
TOA	2	1	-0.0374	0.0946	0.1567	0.6922	Object Falling
TOA	3	1	-0.2015	0.0482	17.4827	0.0001	Struck by moving objects
TOA	4	1	-0.3556	0.0815	19.0251	0.0001	Caught
TOA	6	1	-0.4256	0.1728	6.0633	0.0138	Temperature
TOA	7	1	-0.6111	0.297	4.2342	0.0396	Electricity
TOA	8	1	-0.1711	0.1172	2.1326	0.1442	Toxic substances
TOA	9	1	-0.4754	0.1737	7.4871	0.0062	Explosion
TOA	10	1	0.084	0.0362	5.404	0.0201	Other
TOA	15	0	0	0	.	.	Exertion
AGEGRP	1	1	-0.7238	0.0874	68.5062	0.0001	Under 20
AGEGRP	3	1	-0.338	0.0519	42.4311	0.0001	20 - 24
AGEGRP	4	1	-0.2197	0.0483	20.6537	0.0001	25 - 29
AGEGRP	5	1	-0.1213	0.0459	6.9871	0.0082	30 - 34
AGEGRP	7	1	0.0384	0.0428	0.8029	0.3702	40 - 44
AGEGRP	8	1	-0.0075	0.043	0.0305	0.8614	45 - 49
AGEGRP	9	1	0.0167	0.0453	0.1364	0.7118	50 - 54
AGEGRP	10	1	0.073	0.0481	2.3016	0.1292	55 - 59
AGEGRP	11	1	0.0133	0.0576	0.0533	0.8175	60 - 64
AGEGRP	12	1	-0.1435	0.1162	1.5241	0.217	65 +
AGEGRP	13	1	0.2287	0.2107	1.1784	0.2777	Unknown
AGEGRP	99	0	0	0	.	.	35 - 39
SEV	0	1	-0.5559	0.0389	203.7179	0.0001	Not Coded
SEV	1	1	0.8463	0.0323	687.3161	0.0001	Partial Temporary
SEV	4	1	1.7096	0.045	1441.2703	0.0001	Serious
SEV	6	1	1.5773	0.0885	317.779	0.0001	Total Permanent
SEV	15	0	0	0	.	.	Total Temporary
LOI	0	1	-1.0823	0.4596	5.5452	0.0185	OTHER
LOI	1	1	-0.5841	0.0876	44.4264	0.0001	HEAD
LOI	2	1	0.0729	0.0559	1.6981	0.1925	NECK
LOI	3	1	-0.7113	0.2779	6.5515	0.0105	UNSPECIFIED
LOI	4	1	-0.0884	0.0736	1.4418	0.2298	TRUNK
LOI	5	1	-0.1167	0.0359	10.5415	0.0012	UPPER LIMBS
LOI	6	1	-0.364	0.0421	74.9217	0.0001	LOWER LIMBS
LOI	7	1	-0.1849	0.1135	2.6537	0.1033	MULTIPLE
LOI	8	1	-0.1266	0.1312	0.9308	0.3347	GENERAL
LOI	9	0	0	0	.	.	BACK
NOI	0	1	-0.0026	0.0653	0.0016	0.9678	Other
NOI	1	1	-0.4774	0.057	70.0768	0.0001	100 Fractures
NOI	2	1	0.4171	0.0609	46.9469	0.0001	1722 - 1739 M-Skel
NOI	3	1	-0.6377	0.119	28.7066	0.0001	351 - 359 Internal
NOI	4	1	-0.6816	0.072	89.4894	0.0001	401 - 409 Open Wound
NOI	5	1	-0.1514	0.0641	5.5778	0.0182	550 Contusion
NOI	6	1	0.9004	0.0988	83.0047	0.0001	1290 Mental Illness
NOI	7	1	1.0393	0.0978	112.9087	0.0001	Deafness
NOI	8	0	0	0	.	.	250 Sprains
SEX	1	1	0.2694	0.0279	92.8888	0.0001	Female
SEX	2	0	0	0	.	.	Male
IND	A	1	-0.0017	0.0869	0.0004	0.9848	Agriculture, Forestry, Fishing and Hunting
IND	B	1	0.0036	0.1682	0.0005	0.9828	Mining
IND	C	1	0.1397	0.0518	7.2837	0.007	Manufacturing
IND	D	1	-0.2519	0.1037	5.9011	0.0151	Electricity, Gas and Water
IND	E	1	-0.0528	0.0619	0.7289	0.3932	Construction
IND	F	1	-0.1266	0.0578	4.8012	0.0284	Wholesale and Retail Trade
IND	G	1	-0.2091	0.064	10.6753	0.0011	Transport & Storage
IND	H	1	0.1244	0.5181	0.0577	0.8102	Communication
IND	I	1	0.0889	0.0664	1.7957	0.1802	Finance, Property and Business Services
IND	J	1	0.0503	0.0893	0.3171	0.5733	Public Administration
IND	K	1	-0.2108	0.0559	14.2195	0.0002	Community Services
IND	L	0	0	0	.	.	Recreation, Personal and Other Services
SCALE		0	1	0	.	.	

Wald Statistics For Type 3 Analysis

Source	DF	ChiSquare	Pr>Chi
D1	1	13.0423	0.0003
D2	1	119.3823	0.0001
D3	1	3667.8394	0.0001
TOA	9	69.0253	0.0001
AGEGRP	11	161.7338	0.0001
SEV	4	2615.8366	0.0001
LOI	9	120.9532	0.0001
NOI	8	397.0506	0.0001
SEX	1	92.8888	0.0001
IND	11	152.8959	0.0001

The age variable may now be converted to continuous form by examining the trend of its associated parameter estimates in Table C.2. These are plotted in Figure C.6, and indicate that the frequency increases until age 35, and flattens thereafter.

Figure C.6



The frequency $f_{it}(X_i)$ can therefore be modelled as a function of:

age:	min(age, 35)
development week:	max $(0, 4\text{-}w)^{1.5}$
	max $(0, \min(w,9) - 4)^{0.2}$
	max $(0, w\text{-}9)^{0.4}$
other categorical covariates:	48 levels (Table C.1).

The dimension of the covariate vector has now been reduced from 160 to 52.

C.4 Grouping categorical variables

The final model selected is illustrated by the further GLM output in Table C.3, which contains further grouping of categorical variables that are not statistically different from one another. For example, neck and back injuries are grouped together, and seriously injured and total-permanent injury severities have been grouped together.

Table C.3
Model with development week and age both continuous, and grouped categorical variables

Analysis of Parameter Estimates

Parameter		DF	Estimate	Std Err	ChiSquare	Pr>Chi	Formatted value
INTERCEPT		1	-5.5403	0.1671	1099.559	0.0001	INTERCEPT
D1		1	-0.7472	0.2064	13.1094	0.0003	$MAX(0,4-w)^{1.5}$
D2		1	1.1551	0.1055	119.8171	0.0001	$MAX(0,MIN(w,9)-4)^{0.2}$
D3		1	-0.3751	0.0062	3687.0688	0.0001	$MAX(0,W-9)^{0.4}$
AGEVAR		1	0.0318	0.0026	144.4414	0.0001	MIN(age, 35)
TOA	1	1	-0.1337	0.0326	16.8596	0.0001	Per Fall & Struck & Toxic
TOA	4	1	-0.3886	0.0646	36.2191	0.0001	Caught & Temp & Elect & Expl
TOA	10	1	0.0864	0.0339	6.5013	0.0108	Other
TOA	15	0	0	0	.	.	Obj Fall & Exert
SEV	0	1	-0.5501	0.0387	202.5036	0.0001	Not Coded
SEV	1	1	0.839	0.032	688.0601	0.0001	Partial Temporary
SEV	4	1	1.6893	0.0408	1716.6589	0.0001	Serious & Tot Perm
SEV	15	0	0	0	.	.	Total Temporary
LOI	1	1	-0.2901	0.0792	13.4274	0.0002	HEAD & OTHER
LOI	2	1	0.365	0.0403	82.2012	0.0001	NECK & BACK
LOI	3	1	-0.3679	0.2711	1.842	0.1747	UNSPECIFIED
LOI	4	1	0.2358	0.0403	34.2084	0.0001	TRUNK & UPP LIMBS & MULTIPLE & GENERAL
LOI	6	0	0	0	.	.	LOWER LIMBS
NOI	1	1	-0.4693	0.0541	75.1088	0.0001	Fractures
NOI	2	1	0.4156	0.0593	49.1498	0.0001	Diseases of musculo-skeletal and connective tissue
NOI	3	1	-0.6723	0.0563	142.5687	0.0001	Internal injuries and open wound
NOI	5	1	-0.1516	0.0593	6.5246	0.0106	Contusions
NOI	6	1	0.9482	0.0881	115.9328	0.0001	Mental illness
NOI	7	1	1.0953	0.085	166.2232	0.0001	Deafness
NOI	8	0	0	0	.	.	Sprains and other
NSEX	1	1	0.2779	0.026	114.4721	0.0001	Female
NSEX	2	0	0	0	.	.	Male
IND	1	1	-0.1521	0.031	24.1269	0.0001	Agriculture, Forestry, Fishing and Hunting & Mining & Construction & Public Admin & Recreation, Personal services
IND	4	1	-0.3553	0.029	150.3984	0.0001	Elect, Gas and Water & Trans and Storage & Comm Serv
IND	6	1	-0.2649	0.0371	51.0193	0.0001	Wholesale and Retail Trade
IND	13	1	0.1655	0.3845	0.1852	0.6669	Unknown
IND	30	0	0	0	.	.	Communications, Financial services, Property and Business Services, Manufacturing
SCALE		0	1	0	.	.	

Wald Statistics For Type 3 Analysis

Source	DF	ChiSquare	Pr>Chi
D1	1	13.1094	0.0003
D2	1	119.8171	0.0001
D3	1	3687.0688	0.0001
AGEVAR	1	144.4414	0.0001
TOA	3	69.7379	0.0001
NCAC	4	153.8036	0.0001
SEV	3	2674.2753	0.0001
LOI	4	135.2693	0.0001
NOI	6	461.9198	0.0001
NSEX	1	114.4721	0.0001
IND	4	165.7021	0.0001

The dimension of the covariate vector has now been reduced from 52 (originally 160) to 31, consisting of:

Age: 1 parameter
Development week: 3 parameters
Categorical covariates: 27 parameters

References

Allison, P.D. (1995). **Survival analysis using the SAS system: a practical guide.** SAS Institute Inc., Cary NC.

Australian Bureau of Statistics (2000). **ANZSIC codes – version 2.0, 1 July 2000.** Australian Bureau of Statistics, Canberra. Catalogue No 1290.0.

Cox, D.R. (1972). Regression models and life tables. **Journal of the Royal Statistical Society, B**, 34, 187-220.

Efron, B. and Tibshirani, R. J. (1995). **An introduction to the Bootstrap**. Chapman and Hall, New York.

Feller, W. (1965). **An introduction to probability theory and its applications, vol. I** (2nd edition). John Wiley & Sons, Inc., New York.

Haberman, S. (1983). Decrement tables and the measurement of morbidity. **Journal of the Institute of Actuaries**, 110, 361-381.

Institute of Actuaries and Faculty of Actuaries (1991). **Continuous mortality investigation reports**, 12.

Jørgensen, B. and Paes de Souza, M.C. (1994). **Fitting Tweedie's compound Poisson model to insurance claims data.** Scandinavian Actuarial Journal, 1994 (1), 69-93.

Kalbfleisch, J.D. and Prentice, R.L. (1980). **The statistical analysis of failure time data**. John Wiley & Sons Inc, New York.

McCullagh, P. and Nelder, J.A. (1989). **Generalised linear models (3rd edition).** Chapman and Hall, London.

Smyth, G.K. and Jørgensen, B. (1999). Fitting Tweedie's compound Poisson model to insurance claims data: dispersion modelling. Unpublished.

Taylor, G. (1971). Sickness: a stochastic process. **Journal of the Institute of Actuaries**, 97, 69-83.

Taylor, G. (2000). **Loss reserving: an actuarial perspective**. Kluwer Academic Publishers, Boston.

Waters, H.R. (1984). An approach to the study of multiple state models. **Journal of the Institute of Actuaries**, 111, 363-374.

Worrall, J.D. and Appel, D. (1985) (eds). **Workers' compensation benefits: adequacy, equity, and efficiency**. ILR Press, New York State School of Industrial and Labor Relations, Cornell University.

Worrall, J.D. and Butler, R.J. (1985). Benefits and claim duration. Appears in Worrall and Appel (1985), 57-70.

RECENT ADDITIONS TO THE RESEARCH PAPER SERIES

No.	Date	Subject	Author
84	FEBRUARY 2001	DISCRETE TIME RISK MODELS UNDER STOCHASTIC FORCES OF INTEREST	Jun Cai
85	FEBRUARY 2001	MODERN LANDMARKS IN ACTUARIAL SCIENCE Inaugural Professorial Address	David C M Dickson
86	JUNE 2001	LUNDBERG INEQUALITIES FOR RENEWAL EQUATIONS	Gordon E Willmot Jun Cai X Sheldon Lin
87	SEPTEMBER 2001	VOLATILITY, BETA AND RETURN WAS THERE EVER A MEANINGFUL RELATIONSHIP?	Richard Fitzherbert
88	NOVEMBER 2001	EXPLICIT, FINITE TIME RUIN PROBABILITIES FOR DISCRETE, DEPENDENT CLAIMS	Zvetan G Ignatov Vladimir K Kaishev Rossen S Krachunov
89	NOVEMBER 2001	ON THE DISTRIBUTION OF THE DEFICIT AT RUIN WHEN CLAIMS ARE PHASE-TYPE	Steve Drekic David C M Dickson David A Stanford Gordon E Willmot
90	NOVEMBER 2001	THE INTEGRATED SQUARE-ROOT PROCESS	Daniel Dufresne
91	NOVEMBER 2001	ON THE EXPECTED DISCOUNTED PENALTY FUNCTION AT RUIN OF A SURPLUS PROCESS WITH INTEREST	Jun Cai David C M Dickson
92	JANUARY 2002	CHAIN LADDER BIAS	Greg Taylor
93	JANUARY 2002	FURTHER OBSERVATIONS ON CHAIN LADDER BIAS	Greg Taylor
94	JANUARY 2002	A GENERAL CLASS OF RISK MODELS	Daniel Dufresne
95	JANUARY 2002	THE DISTRIBUTION OF THE TIME TO RUIN IN THE CLASSICAL RISK MODEL	David C M Dickson Howard R Waters
96	MAY 2002	A NOTE ON THE MAXIMUM SEVERITY OF RUIN AND RELATED PROBLEMS	David C M Dickson
97	JUNE 2002	UPPER BOUNDS FOR ULTIMATE RUIN PROBABILITIES IN THE SPARRE ANDERSEN MODEL WITH INTEREST	Jun Cai David C M Dickson
98	JUNE 2002	CONTINUOUS COMPOUNDING, VOLATILITY AND THE EQUITY PREMIUM	Richard Fitzherbert

No.	Date	Subject	Author
99	JUNE 2002	THE DEFICIT AT RUIN IN THE STATIONARY RENEWAL RISK MODEL	Gordon E Willmot David C M Dickson Steve Drekic David A Stanford
100	AUGUST 2002	ASIAN AND BASKET ASYMPTOTICS	Daniel Dufresne
101	AUGUST 2002	RUIN PROBABILITIES WITH A MARKOV CHAIN INTEREST MODEL	Jun Cai David C M Dickson
102	AUGUST 2002	THE GERBER-SHIU DISCOUNTED PENALTY FUNCTION IN THE STATIONARY RENEWAL RISK MODEL	Gordon E Willmot David C M Dickson
103	NOVEMBER 2002	INITIAL CAPITAL AND MARGINS REQUIRED TO SECURE A JAPANESE LIFE INSURANCE POLICY PORTFOLIO UNDER VARIABLE INTEREST RATES	Manabu Sato David C M Dickson Richard M Fitzherbert
104	NOVEMBER 2002	STATISTICAL CASE ESTIMATION	Greg Taylor Mireille Campbell

A complete list of papers can be found at http://www.economics.unimelb.edu.au/actwww/papers.html