[D3]
A CURVE FITTING METHOD AND A REGRESSION METHOD
Contributed by S Benjamin and L M Eagles

Introduction

This method models the run-off triangle row-by-row and then ties the rows together. Each row, or year of account, is modelled by a Weibull distribution function. This model was suggested by D H Craighead, and so the Weibull distribution function has become to be known as the Craighead curve when it is used in this context. It is not a linear model and the three parameters have to be estimated using an iterative search method. Once this has been done, the ultimate loss ratio for each year of account can be estimated.

The second part of the method relates the known loss ratios (paid or incurred) to the predicted ultimate loss ratio. Taking a particular development year, each row of the triangle has a known loss ratio for that development year and a predicted ultimate loss ratio from the first part of the method. There is thus a set of pairs of known and dependent variables: one pair for each year of account. A line of best fit is found, using standard regression methods. From this regression line, another estimate of the ultimate loss ratio for each year of account can be read off. This new estimate has the advantage that it takes into account the information from all years of account, rather than just one particular year of account. The regression line can also be used to produce a confidence interval for the estimated ultimate loss ratio, and to estimate loss ratios for future development years (ie the lower triangle).

<>

[D3.a]
THE CURVE FITTING METHOD
Contributed by S Benjamin and L M Eagles

1. Introduction

1.1 In the London Market details of numbers of claims are generally not available or not relevant. Data is usually available for each "year of account" or year of origin, i.e. for all risks written in a particular accounting year which is usually a calendar year. The items normally available are:

(i) Premiums paid to date
(ii) Claims paid to date
(iii) Claims outstanding, typically the case estimates as notified by the brokers to the companies for outstanding claims.

For each year of account separately, a past history of premium payments, claim payments and claims outstanding will be available. This information, split by year of development, may not be complete (e.g. no information may be available on the claims paid by the end of the first or second year of development). Sometimes quarterly development data is available. Data is normally available subdivided by currency and possibly by line of business.

1.2 The curve fitting method works by estimating the Ultimate Loss Ratio ("ULR") for each year of account, from which the necessary reserve is easily derived. Years of account do not need to be homogeneous.

1.3 The method will provide a reserve for each year of account for which sufficient past development data is available. As will be seen, even where little historical development for any one year is available, the information from adjacent years can be used to help.

1.4 The method lends itself well to interactive graphical illustration and is therefore easy to follow by actuaries and non-actuaries.

2. Method

2.1 Run-off triangles are drawn up for as many years of account as possible showing the development year by year (or quarter by quarter) of premiums and claims.

2.2 An estimate of the ultimate premiums receivable is made for each year of account. If it is necessary to calculate the estimate then development factors are normally applied which are calculated from the data without smoothing. Other methods, such as the Regression Method described below, could be used in appropriate circumstances. Often the underwriters' estimates are used since the underwriters have a better feel for the way in which, in practice, policies are being signed.

2.3 The estimates of ultimate premiums are divided into the relevant claims to give a run-off triangle of loss ratios.

2.4 The loss development patterns analysed may be either

(a) paid loss ratios
(b) incurred loss ratios (where incurred claims are paid claims plus claims outstanding), depending on which is believed to be more useful, or
(c) both.

2.5 For each year of accunt for which there is sufficient development a curve is fitted to the loss ratio development. The chosen curve y(t) is described below; it tends to a finite limit (the estimated ULR) as the development time, t, becomes very large.

3. Choice of curve

3.1 The curves used to fit the loss development pattern for a particular year of account are chosen:

(a) to fit the past history of claims payments as best as possible,
and
(b) to allow for additional future claim development.

3.2 Empirical considerations suggest that if a smooth curve is sought to fit the shape of the loss ratio at development time t, plotted against t, that curve would have a negative exponential shape. The actual formula is:

$$y(t) = A(1 - e^{-(t/b)^c})$$

The parameters of such a curve have specific meanings:

A: As t becomes very large y(t) tends to a value of A, i.e. A is the estimated ULR.

b: At t=b $y(t)=A(1 - e^{-1})$ i.e. b is the time taken to reach a loss ratio of about 63% of the ultimate loss ratio. b is measured in the same units of time as t; for example, if t is in years, then so is b.

c: c defines the steepness of the curve.

This curve has been called a "Craighead Curve". It was originally suggested in a paper by D H Craighead (1979).

Two graphs showing how this curve varies if the b and c parameters are varied are set out at the end of Section 4.

3.3 Other curves can be used; for example a double Craighead curve (if it is thought that the data consists of two parts, one short-tail and one long-tail), of the form:

$$y(t) = A_1(1 - e^{-(t/b_1)^{c_1}}) + A_2(1 - e^{-(t/b_2)^{c_2}})$$

can be fitted. This is useful when the data are given quarterly and hence the number of data points is larger than the number of parameters.

3.4 It is not always necessary, or advisable, to fit the curve to each year of account separately. The important more recent years of account cannot be fitted anyway. Within the same class of business it is useful to fit the same b and c (i.e. shape) to all years of account and let A (the ULR level) vary. Looking at all the curves thus generated on the same screen (graph) is helpful. Scaling them all to the same nominal ULR of 100% shows by eye whether acceptable homogeneity exists or whether, typically, one year may be an outlier and deserves to be fitted separately by allowing b or c or both to vary also. Knowledge of a structural break in the type or mix of business can be brought into the judgement.

3.5 It can also be helpful to fit A, b_1, c_1 to the paid loss ratios and A, b_2, c_2 (i.e. a common ULR, A) to the incurred loss ratios, and to view the graphical results on the same screen.

3.6 Certain classes of business have typical values of b and c. These may be imposed on data which is otherwise unhelpful.

4. Method of curve fitting

4.1 Any appropriate method of curve fitting could be used. Typically, however, a method which chose curve parameters in such a way as to minimise the sum of weighted least squared deviations of the curve from the data has been used. For example, if an individual year is being fitted in isolation, the quantity to be minimised is D where

$$D = \sum w(t)\,(y(t) - yobs(t))^2$$

where $yobs(t)$ = observed loss ratio at time t
$w(t)$ = weights assigned to the loss ratio at time t

and $y(t)$ is the curve being fitted.

4.2 When using the Craighead curve, it is possible to determine algebraically the value of A needed to minimise the measure of deviation, D, provided b and c are known (or fixed).

4.3 When b and/or c are not known, or a more complicated curve is being fitted, it is normally necessary to determine the curve parameters using an iterative minimisation technique, such as steepest descent, or the Davidon, Fletcher and Powell algorithm. These iterative techniques normally (but not always) converge quite quickly. The parameters they converge to will define a "local" minimum of D, which may not always be the "global" minimum.

4.4 Any weights could be attached to the individual data points fitted by the above approach (thus outliers can be excluded by giving such points zero weight). Typically, however, the following could be used:

(a) equal weights of 1 to each non-zero data point,

or

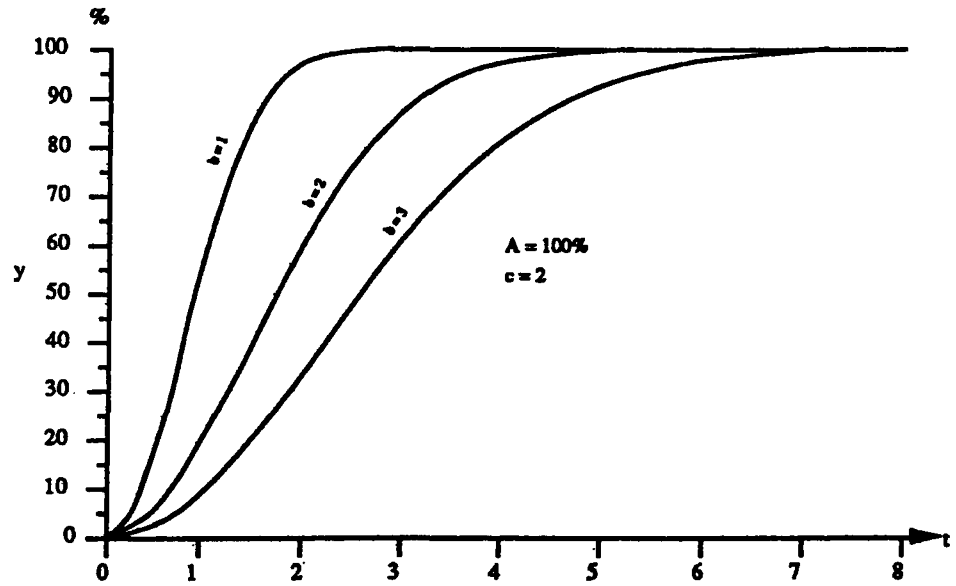
(b) $w(t)=t$ thus giving more weight to the more developed data,

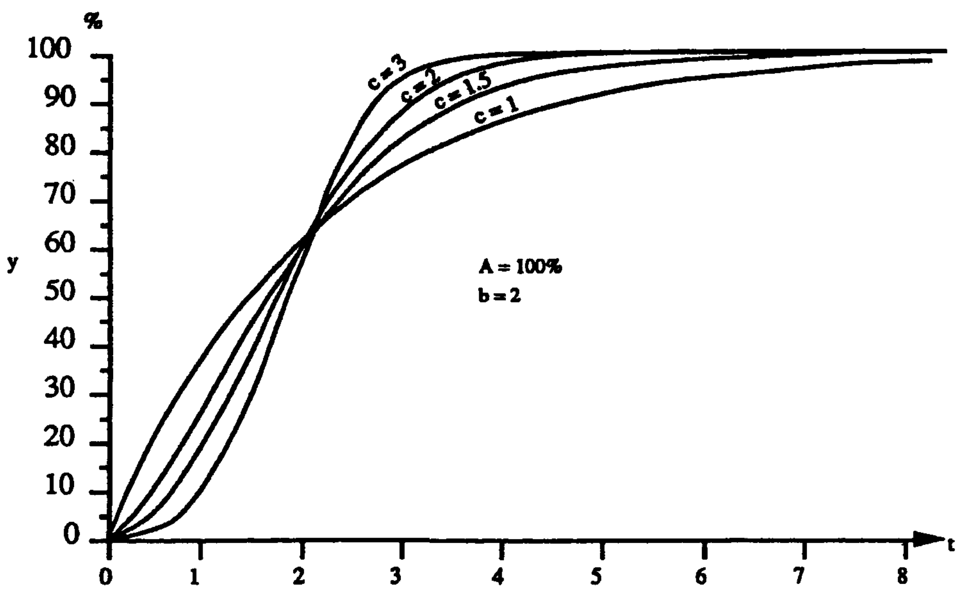
or

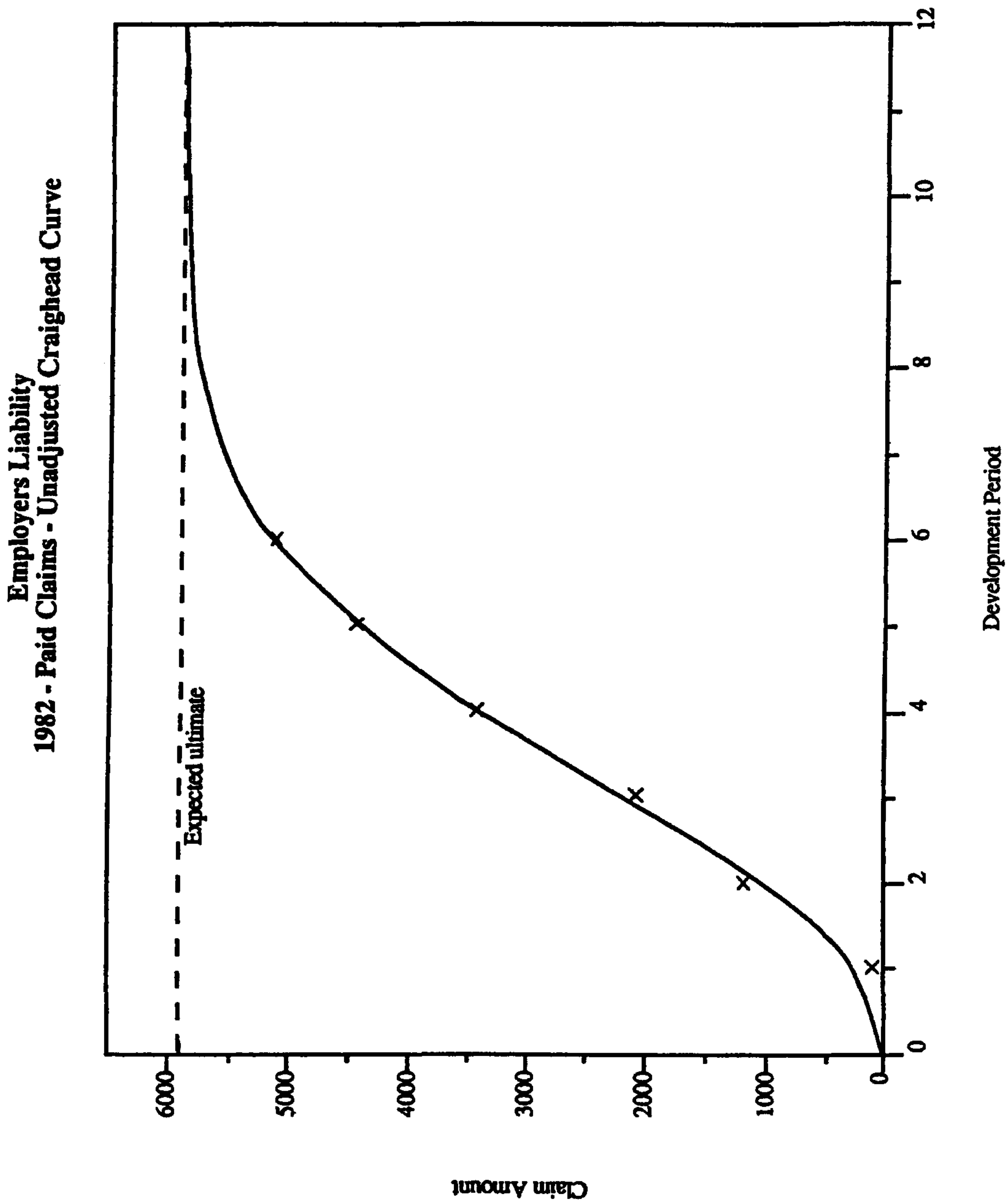
(c) A weight of 1 for the most recent data point for the given year, 0.9 for the next most recent point, 0.9^2 for the next most recent, etc.

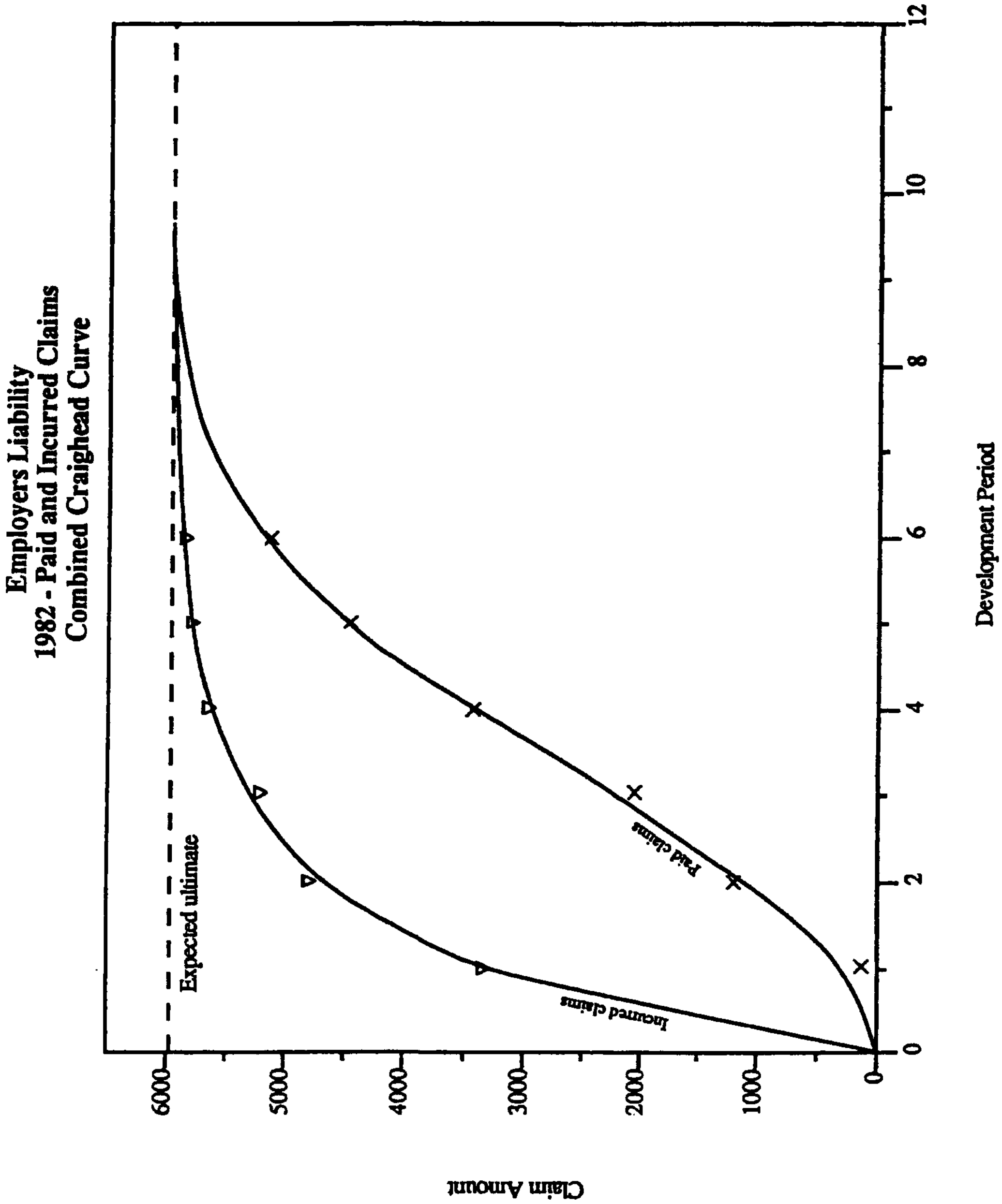
4.5 A special case, which can be considered as an "adjusted Craighead curve", is where the curve is forced through the most recent data point (e.g. by giving this point a very large weight, or to fit A, b, c, and to adjust A, keeping b and c fixed to make the curve pass through latest point). Curves fitted in this manner ensure that further development is non-negative.

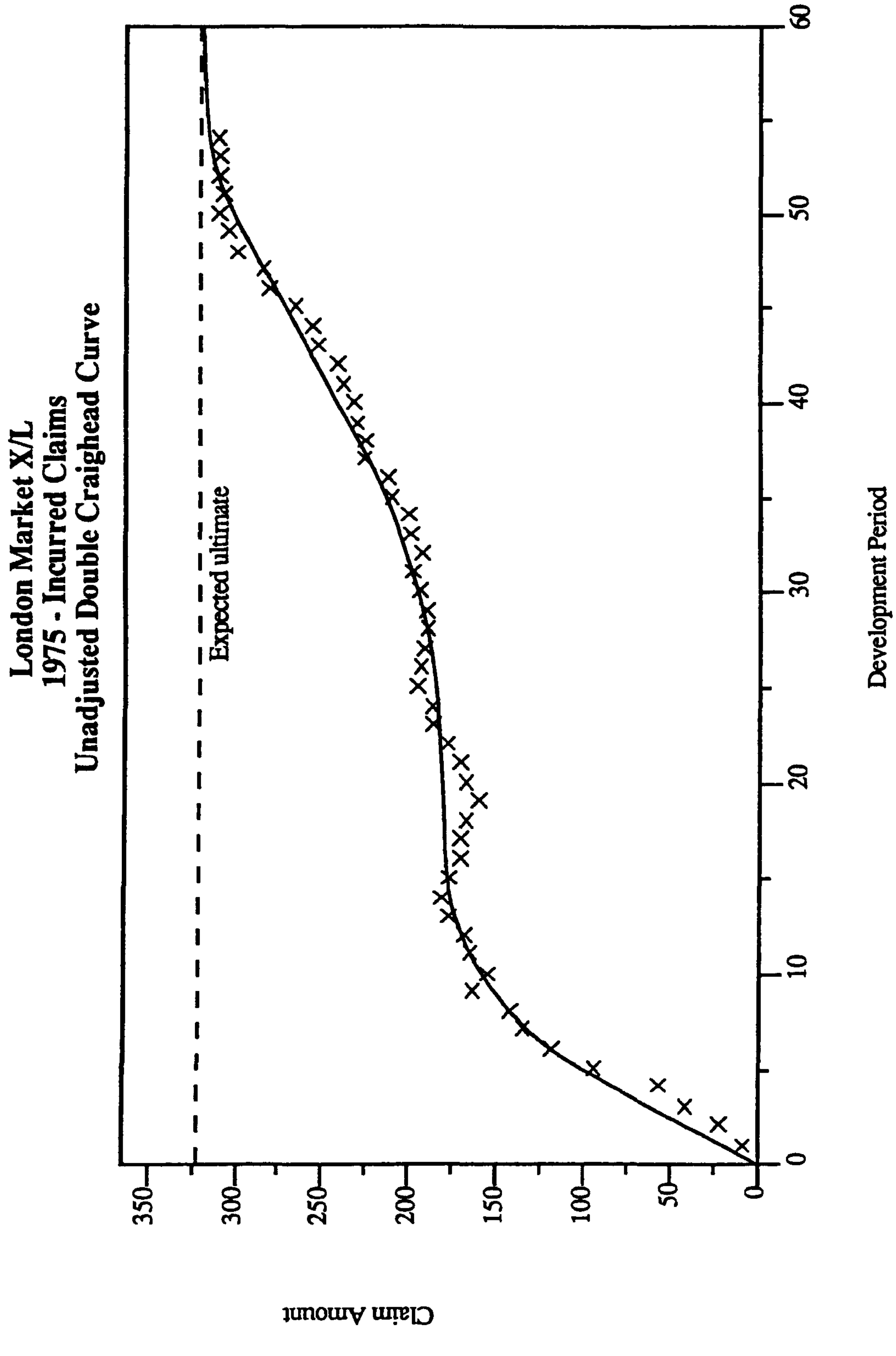
4.6 Typically, the loss ratios and the curves fitted to these loss ratios would be plotted graphically. A qualitative goodness of fit can then be determined, by visual examination. Signs of heterogeneity or structural breaks between different years of account can be seen visually. Curve fits can then be refined, if necessary, by the actuary if he or she believes that different curves would fit the data better.


%
100
90
80
70
60
50
40
30
20
10
0
y
b=1
b=2
b=3
A = 100%
c = 2
0
1
2
3
4
5
6
7
8
t


%
100
90
80
70
60
50
40
30
20
10
0
y
c=3
c=2
c=1.5
c=1
A = 100%
b = 2
0
1
2
3
4
5
6
7
8
t


Employers Liability
1982 - Paid Claims - Unadjusted Craighead Curve
Expected ultimate
Claim Amount
0
1000
2000
3000
4000
5000
6000
Development Period
0
2
4
6
8
10
12


Employers Liability
1982 - Paid and Incurred Claims
Combined Craighead Curve
Expected ultimate
Incurred claims
Paid claims
Claim Amount
Development Period
0
1000
2000
3000
4000
5000
6000
0
2
4
6
8
10
12


London Market X/L
1975 - Incurred Claims
Unadjusted Double Craighead Curve
Expected ultimate
Claim Amount
Development Period
0
50
100
150
200
250
300
350
10
20
30
40
50
60

5. Comparison of the curve fitting method with other methods

5.1 The curve fitting method can model claims development (or even premium development) directly, rather than modelling loss ratios. Graphically, however, it is easier to visualise and present loss ratios.

5.2 The link ratio method may technically be reformulated in a manner closely analogous to the curve fitting method. If a single curve is chosen for every year of account so that each successive link ratio matches those derived from a chain ladder or other link ratio method and if, when fitting this curve to the data no weight is given to any data point for a given year of account other than the most recent then the curve fitting approach will produce the same estimated reserves as the link ratio method.

5.3 Three potential weaknesses inherent in the link ratio method are:

(i) over parameterisation of the data (analogous to the curve chosen being fitted with too many parameters).

(ii) undue weight given to the most recent data point in each account year, and

(iii) inability to reflect structural breaks in the underlying business run-off patterns.

It can be seen from 5.2 that the first two weaknesses can be partially overcome using a curve fitting method which chooses curves according to fewer parameters than those implied by a link ratio method, and by giving some weight in the curve fitting to data points other than the most recent. The curve fitting method can also accommodate structural breaks (as mentioned in 3.6).

5.4 The curve fitting method as we have applied it uses premiums as a measure of exposure (by analysing loss ratios). In principle the curve fitting method could instead use other exposure measures, such as number of policies, if they were appropriate to the type of business being analysed.

[D3.b]
THE REGRESSION METHOD
Contributed by S Benjamin and L M Eagles

1. Introduction

1.1 The regression method is used to refine estimated ULR's (Ultimate Loss Ratios), particularly those derived from the curve fitting method.

1.2 The method can provide explicit confidence limits for the ULR thus estimated. These confidence limits are not those in the strict statistical sense but do give a practical range in which the ULR can be expected to lie.

1.3 This method also lends itself well to graphical illustration and is therefore easy to follow by actuaries and non-actuaries.

1.4 Further details of the method, and of its potential application to a minimum reserving basis used by Lloyd's were set out in 1986 in a paper by S Benjamin and L M Eagles.

2. Method

2.1 ULRs for each year of account are estimated using an alternative reserving method. It is not necessary to have estimates for every single year of account, although it is usually helpful to have them for as many years as possible.

2.2 IBNR loss ratios are calculated as the difference between the ULRs and the incurred loss ratios.

2.3 For each development year in turn, a regression line is fitted to the set of points obtained from plotting the IBNR loss ratios against the incurred loss ratios at the development year under consideration; each year of account being represented by a separate point.

2.4 At one extreme, the points may all lie on a straight line so that the relationship may be deduced with little uncertainty. At the other extreme, the points may appear to be scattered at random which suggests that there is little relationship between losses at the development year under consideration and the eventual level of future losses. In fitting this regression line the fitting can be found as a "confidence interval"; this is a region about the line where future results are likely to fall.

2.5 The regression line can be expressed in terms of its "Slope" and "Constant" and leads to an expression of the form:

Estimated IBNR loss ratio = Slope × (incurred Loss Ratio at development year under consideration) + Constant %

2.6 This expression can thus be used to estimate the IBNR loss ratio and hence the ULR for the year of account whose latest development year corresponds to the development year under consideration. The regression lines arising from other years of development can similarly be applied to estimate the ULRs for the other years of account.

2.7 Where very few years of account have advanced to a particular development year, a curve fitted to the underlying loss ratios can be used to provide estimates not only of IBNR loss ratios but also of the expected incurred loss ratios to each development year. These expected loss ratios for any development year can then be plotted against estimated IBNR loss ratios in exactly the same way as actual loss ratios. A line of best fit and a confidence interval can then be derived. For some recent years of account this approach may also be adopted where there is a large amount of fluctuation in the data for the early development years generally.

2.8 It is possible to estimate ULRs (or to estimate IBNR and/or outstandings) by regressing one of these items against either paid, incurred or outstanding loss ratios. It is also possible to use more complicated regression, e.g. log-linear regressions rather than linear regressions. The mechanics of the method remain unchanged.

2.9 The decision as to which regression to make is made after looking at the plots of residuals in each particular case.

3. Comments

3.1 A qualitative degree of confidence in the reserves established by this reserving method can be gained by visual examination of the plots mentioned in 2.3.

3.2 Plots consisting of points scattered apparently at random (such as described in 2.4) do occur, particularly for more recent years of account. In such cases it may be appropriate to take the regression line as the Estimated IBNR loss ratio = Constant (where the constant is chosen as the average IBNR loss ratio, or from other considerations).

In many other cases confidence intervals, again particularly for more recent years of account, can be quite large. It is believed that this usually reflects inherent difficulties facing any loss reserving method, rather than weaknesses of this particular method.

3.3 The regression line stated in 2.5 may be algebraically reformulated in terms of credibility theory, as:

Future Claims = $z \times k \times$ claims to date + $(1 - z) \times 1 \times$ premiums
(z is the degree of credibility given to the claims data).

As the development year (and the credibility given to claims data) increases we would therefore expect the constant and the slope of the regression line both to tend to zero.

The actual regression lines adopted may be adjusted if desired to exhibit this behaviour. For years of account when the regression line constant has become zero the method becomes analogous to the link ratio method.

3.4 In practice, for discussion with underwriters, it is more effective to plot separately for each year of development the ULR as ordinate and the paid (or incurred) loss ratio as abscissa, and to fit the line of regression.

3.5 A simple measure of the inherent variability can then be demonstrated visually by:

(a) drawing a line parallel to the regression line and passing through the data point furthest from it, followed by

(b) drawing a further line parallel to it at the same distance on the other side.

Thus, a path symmetrical about the line of regression is formed which just encloses all the data points. The width of the path is an intuitively appealing display and measure of the inherent variability. The width of the path obtained by regressing ULRs against paid loss ratios can be compared with the width obtained by using incurred loss ratios.

3.6 Outliers can be identified visually and lead to useful discussion.

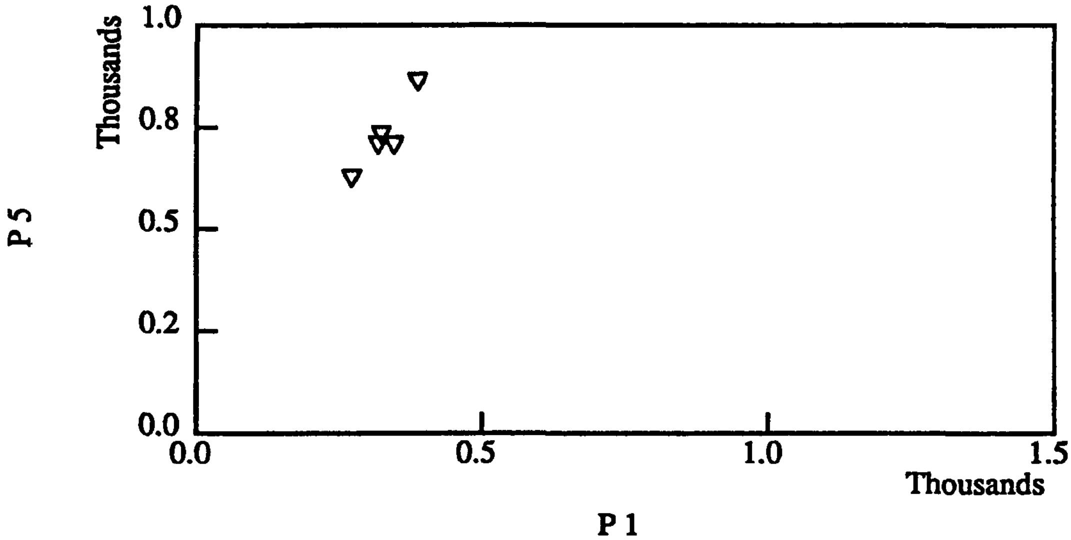
3.7 At the first level of presentation, if curve fitting is to be avoided, a sufficient number of ULRs must be available. Hence either sufficient history must be available or the underwriter's own estimates of ULRs on partly developed years of account must be used.

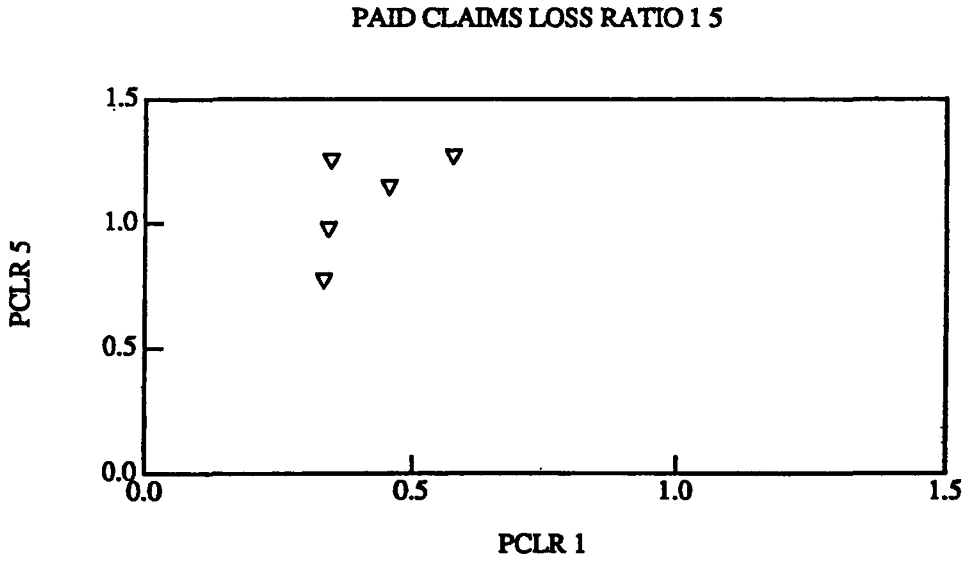
References

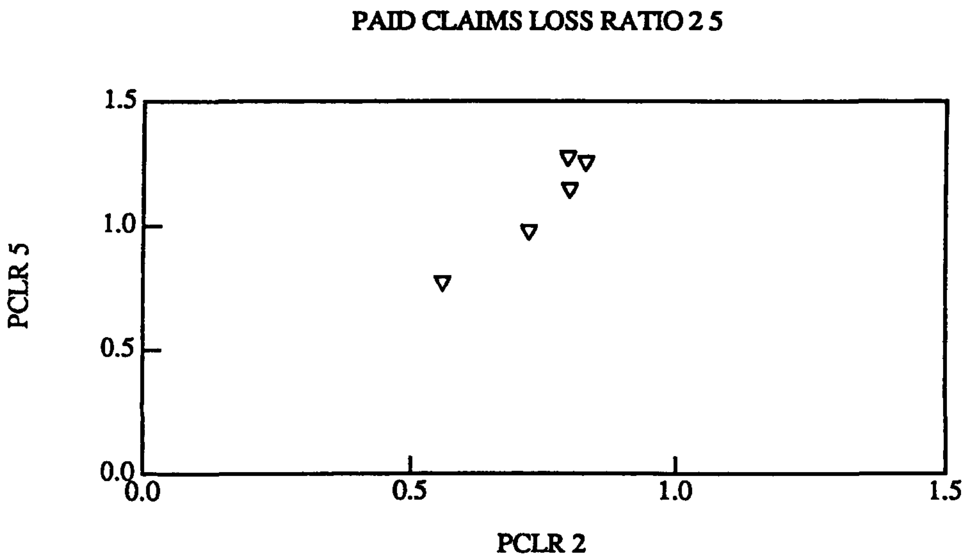
(1) 1979: Craighead, D.H. "Some aspects of the London reinsurance market in worldwide short-term business", *Journal of the Institute of Actuaries*, Vol. 106 Part III p 286.

(2) 1986: Benjamin, S. and Eagles, L.M. "Reserves in Lloyd's and the London Market", *Journal of the Institute of Actuaries*, Vol. 113 Part II p 197.

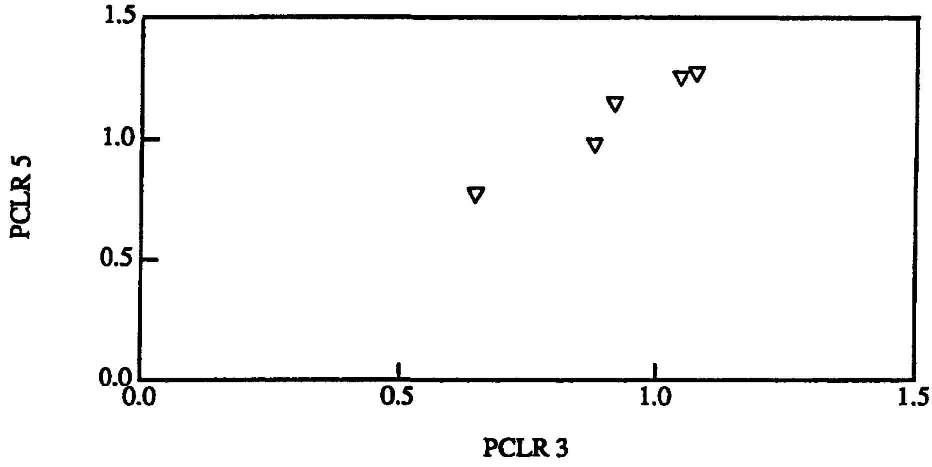
PREMIUMS 1 5



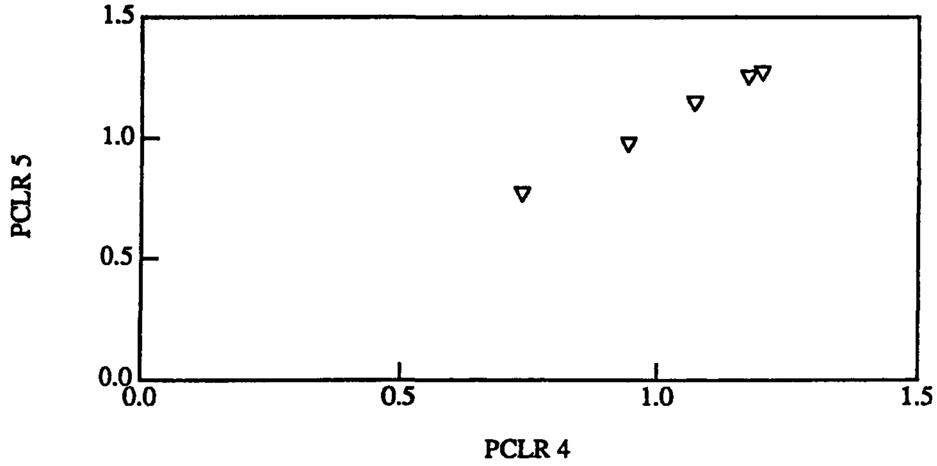

PAID CLAIMS LOSS RATIO 1 5
PCLR 5
PCLR 1
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5


PAID CLAIMS LOSS RATIO 2 5
PCLR 5
PCLR 2
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5

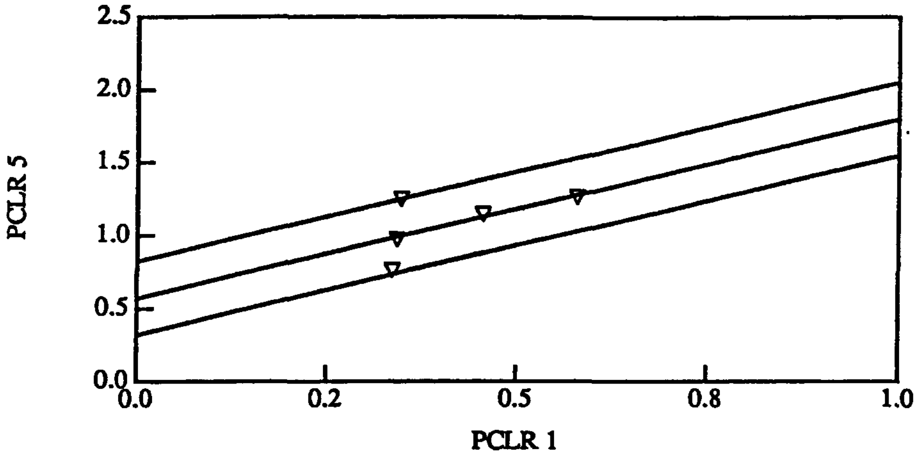
PAID CLAIMS LOSS RATIO 3 5



PAID CLAIMS LOSS RATIO 4 5

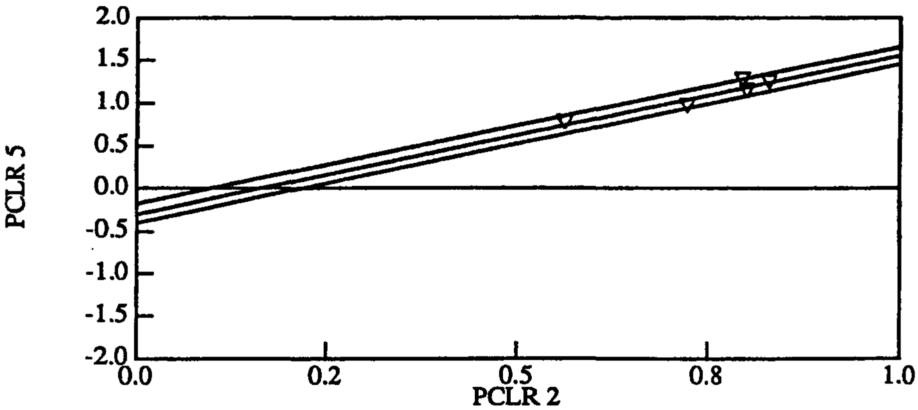


GRAPH PATH 1 5



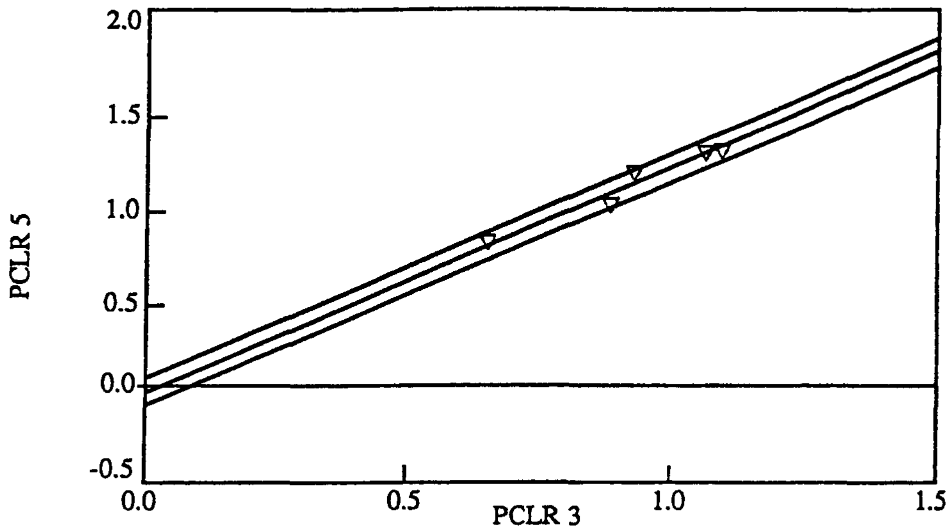
LINE 1 5: PCLR 5 = 1.224 * PCLR1 + 0.575
UPPER LINE 1 5: = LINE 1 5 + 0.250
LOWER LINE 1 5: = LINE 1 5 - 0.250

GRAPH PATH 2 5



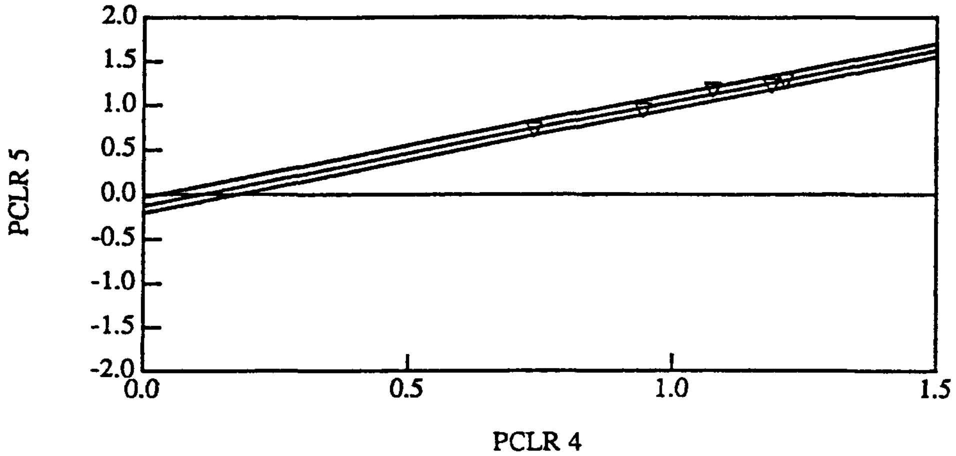
LINE 2 5: PCLR 5 = 1.827 * PCLR2 - 0.277
UPPER LINE 2 5: = LINE 2 5 + 0.087
LOWER LINE 2 5: = LINE 2 5 - 0.087

GRAPH PATH 3 5



LINE 3 5: PCLR 5 = 1.197 * PCLR3 - 0.013
UPPER LINE 3 5: = LINE 3 5 + 0.069
LOWER LINE 3 5: = LINE 3 5 - 0.069

GRAPH PATH 4 5



LINE 4 5: PCLR 5 = 1.096 * PCLR34 - 0.046
UPPER LINE 4 5: = LINE 4 5 + 0.022
LOWER LINE 4 5: = LINE 4 5 - 0.022