Applications of Mathematics

Tomáš Cipra

# Dynamic credibility with outliers and missing observations

*Applications of Mathematics*, Vol. 41 (1996), No. 2, 149--159

Persistent URL: http://dml.cz/dmlcz/134318

# DYNAMIC CREDIBILITY WITH OUTLIERS AND MISSING OBSERVATIONS

TOMÁŠ CIPRA, Praha

(Received November 9, 1994)

*Summary.* In actuarial practice the credibility models must face the problem of outliers and missing observations. If using the $M$-estimation principle from robust statistics in combination with Kalman filtering one obtains the solution of this problem that is acceptable in the numerical framework of the practical actuarial credibility. The credibility models are classified as static and dynamic in this paper and the shrinkage is used for the final ratemaking.

*Keywords*: credibility, actuarial science, outliers, missing observations, robust Kalman filter, shrinkage, time series, risk

*AMS classification*: 62P05 (62M20, 60G35, 93E11)

## 1. INTRODUCTION

From statistical point of view the credibility in actuarial science treats many (usually short) time series where each series represents insurance claims of different insurance groups. These groups (e.g. different geographical regions, different makes of cars etc.) have different historical insurance experience and the credibility models combine the individual experience of particular groups with the overall experience for sound future ratemaking. In this paper the credibility models are classified to static ones that are stable over time with deterministic trends (see Section 2 and 3) and to dynamic ones that are more flexible allowing for time-varying parameters in the claim process (see Section 4 and 5). From the statistical point of view the combination of the individual and overall experience means to imply shrinkage of group-specific estimates towards their average (see e.g. [20]).

Supported by the grant GAČR-2169 "Stochastic Models".

The objective of this paper is to modify credibility procedures for the case of data with outliers and missing observations. This is an up-to-date problem important for the actuarial practice (see e.g. [2], [10], [15], [18], [24]). The $M$-estimation principle of robust statistics seems to give results that are acceptable from the numerical point of view. In particular, if one robustifies the Kalman filter, as is done in [4], [23], one obtains a suitable instrument that, in addition, allows to treat the missing observations (see Section 5). Kremer [17] has already applied the results from [4] in this context but the problem of estimating the variance and covariance parameters stayed unsolved. Here the complete estimation procedure is suggested and demonstrated by means of a simple numerical example.

## 2. Static models

Let $y_{it}$ be the claim amount observed in the group $i$ at time $t$ ($i = 1, \ldots, k$; $t = 1, \ldots, n$) where $k$ is the number of groups under consideration and $n$ is the number of periods of observations. The ratemaking consists in forecasting the future values $y_{i,n+h}$ for $h$ steps ahead in each group, i.e. in constructing the forecast $\hat{y}_{in}(h)$.

In this static case one assumes for the observed claims $y_{it}$ linear models of the form

$$y_{it} = x_{it}'\beta_i + \varepsilon_{it} \tag{2.1}$$

($i = 1, \ldots, k$; $t = 1, \ldots, n$), where $x_{it}$ are known $p$-dimensional vectors of regressors, $\beta_i$ are $p$-dimensional regression parameters and $\varepsilon_{it}$ are independent errors fulfilling

$$\varepsilon_{it} \sim N(0, \sigma^2) \tag{2.2}$$

(the parameters $\beta_i$ and $\sigma^2$ are unknown). The classical least squares estimates of the parameters $\beta_i$ have the form

$$\hat{\beta}_i = V_i \sum_t x_{it} y_{it} \tag{2.3}$$

with

$$E\hat{\beta}_i = \beta_i, \quad \operatorname{var} \hat{\beta}_i = \sigma^2 V_i, \tag{2.4}$$

where

$$V_i = \Big(\sum_t x_{it} x_{it}'\Big)^{-1}. \tag{2.5}$$

One could consider a more general heteroscedastic case with $\operatorname{var} \varepsilon_{it} = \sigma^2/m_{it}$ where the known scalars $m_{it}$ measure the volume of the data for the global claims $y_{it}$.

Since (2.1) are standard credibility models one assumes that $\beta_i$ are independent realizations from a common distribution. Therefore one should shrink $\hat{\beta}_i$ to a common value using the shrinkage model

$$\beta_i = \beta + e_i \tag{2.6}$$

($i = 1, \ldots, k$), where $e_i$ are independent $p$-dimensional errors fulfilling

$$e_i \sim N(0, \sigma^2 B) \tag{2.7}$$

(in addition to $\beta_i$ and $\sigma^2$ the parameters $\beta$ and $B$ of appropriate dimensions are also unknown). Using the above relations one can construct the general linear model

$$\begin{pmatrix} \hat{\beta}_i \\ \beta \end{pmatrix} = \begin{pmatrix} I \\ I \end{pmatrix} \beta_i + \delta, \quad E\delta = 0, \quad \operatorname{var} \delta = \sigma^2 \begin{pmatrix} V_i & 0 \\ 0 & B \end{pmatrix} \tag{2.8}$$

which after some algebraic treatment gives the credibility estimates $b_i$ of $\beta_i$ as

$$b_i = Z_i \hat{\beta}_i + (I - Z_i)\beta \tag{2.9}$$

with

$$Z_i = B(B + V_i)^{-1}. \tag{2.10}$$

However, $b_i$ depends on unknown parameters, and one must estimate them using e.g. the moment estimates

$$\hat{\sigma} = \frac{1}{k(n-p)} \sum_i \sum_t (y_{it} - x'_{it}\hat{\beta}_i)^2, \tag{2.11}$$

$$\hat{\beta} = \Big(\sum_i Z_i\Big)^{-1} \sum_i Z_i \hat{\beta}_i, \tag{2.12}$$

$$H = \frac{1}{k-1} \sum_i Z_i(\hat{\beta}_i - \hat{\beta})(\hat{\beta}_i - \hat{\beta})', \tag{2.13}$$

$$\hat{B} = \frac{1}{2\hat{\sigma}^2}(H + H') \tag{2.14}$$

constructed e.g. in an iterative way (see [6], [7], [19]).

Then the credibility forecasts of the future claims have the form

$$\hat{y}_{in}(h) = x'_{i,n+h} b_i. \tag{2.15}$$

As special cases of the static models described one obtains

(i) Bühlmann–Straub model for $x_{it} = 1$ (see [3]);

(ii) Hachemeister model for $x_{it} = (1, t)'$ (see [11]).

## 3. Static models with outliers

Robustification of the static models which should be insensitive to outliers can be based on the $M$-estimation approach using IWLS scheme (Iterated Weighted Least Squares, see e.g. [13]):

The corresponding normal equations have the form

$$\sum_t \psi\Big(\frac{y_{it} - x'_{it}\beta_i}{\sigma}\Big) x_{it} = 0, \tag{3.1}$$

$$\frac{1}{n-p} \sum_t \chi\Big(\frac{y_{it} - x'_{it}\beta_i}{\sigma}\Big) = d, \tag{3.2}$$

where $\varrho(z)$ is a suitable robustifying loss function, $\psi(z) = \varrho'(z)$ is the corresponding psi-function, $\chi(z) = z\psi(z) - \varrho(z)$ and the constant $d = \frac{1}{2}\{\psi^2(Z)\}$ with $Z \sim N(0,1)$. If one uses the usual Huber's psi-function

$$\psi_H(z) = \begin{cases} z & \text{for } |z| \leqslant c, \\ c\,\mathrm{sgn}(z) & \text{for } |z| > c \end{cases} \tag{3.3}$$

(the constant $c$ is chosen according to contamination by outliers, e.g. one recommends $c = u_{0.95} = 1.645$ for 5%-contamination) then (3.2) rewrites to the form

$$\frac{1}{n-p} \sum_t \psi_H^2\Big(\frac{y_{it} - x'_{it}\beta_i}{\sigma}\Big) = 0.7785 \tag{3.4}$$

(see e.g. [1]).

The IWLS scheme solves the normal equations (3.1) and (3.2) by means of an iterative procedure that converges under weak assumptions (see [8]). It seems convenient to combine this scheme with the iterative estimation of unknown parameters from Section 2 so that one obtains for the $m$-th iteration

$$\hat{\beta}_i^{(m)} = V_i^{(m-1)} \sum_t w_{it}^{(m-1)} x_{it} y_{it}, \tag{3.5}$$

$$[\hat{\sigma}^{(m)}]^2 = \frac{[\hat{\sigma}^{(m-1)}]^2}{k(n-p)d} \sum_i \sum_t \chi\Big(\frac{y_{it} - x_{it}'\hat{\beta}_i^{(m-1)}}{\hat{\sigma}^{(m-1)}}\Big), \tag{3.6}$$

$$Z_i^{(m)} = \hat{B}^{(m-1)}(\hat{B}^{(m-1)} + V_i^{(m-1)})^{-1}, \tag{3.7}$$

$$\hat{\beta}^{(m)} = \Big(\sum_i Z_i^{(m)}\Big)^{-1} \sum_i Z_i^{(m)} \hat{\beta}_i^{(m)}, \tag{3.8}$$

$$H^{(m)} = \frac{1}{k-1} \sum_i Z_i^{(m)} (\hat{\beta}_i^{(m)} - \hat{\beta}^{(m)})(\hat{\beta}_i^{(m)} - \hat{\beta}^{(m)})', \tag{3.9}$$

$$B^{(m)} = \frac{1}{2[\hat{\sigma}^{(m)}]^2}(H^{(m)} + H^{(m)\prime}), \tag{3.10}$$

where

$$V_i^{(m)} = \Big(\sum_t w_{it}^{(m)} x_{it} x_{it}'\Big)^{-1} \tag{3.11}$$

and the weights $w_{it}^{(m)}$ are defined as

$$w_{it}^{(m)} = \psi\Big(\frac{y_{it} - x_{it}'\hat{\beta}_i^{(m)}}{\hat{\sigma}^{(m)}}\Big) \Big/ \frac{y_{it} - x_{it}'\hat{\beta}_i^{(m)}}{\hat{\sigma}^{(m)}}. \tag{3.12}$$

In the special case of Huber's psi-function one uses $\psi_H$ in (3.12) and replaces the relation (3.6) by

$$[\hat{\sigma}^{(m)}]^2 = \frac{[\hat{\sigma}^{(m-1)}]^2}{0.7785k(n-p)} \sum_i \sum_t \psi_H^2\Big(\frac{y_{it} - x_{it}'\hat{\beta}_i^{(m-1)}}{\hat{\sigma}^{(m-1)}}\Big). \tag{3.13}$$

As the missing observations are concerned it is no problem to treat them (in addition to the outliers) in this static case: the only difference from the situation without missing observations will appear in the different numbers $n_1, \dots, n_k$ of observations $y_{it}$ in particular groups.

## 4. Dynamic models

The dynamic models allow for time-varying parameters (see e.g. [5], [14], [16], [19], [20], [21], [25]). One can extend the static models from Section 2 to the dynamic ones e.g. replacing (2.1) by the system

$$y_{it} = x'_{it}\beta_{it} + \varepsilon_{it}, \tag{4.1}$$

$$\beta_{it} = T\beta_{i,t-1} + \nu_{it} \tag{4.2}$$

($i = 1, \dots, k$; $t = 1, \dots, n$), where the models (4.1) fulfill the same assumptions as in Section 2, i.e., in particular,

$$\varepsilon_{it} \sim N(0, \sigma^2), \tag{4.3}$$

$T$ is a known matrix and $\nu_{it}$ are independent $p$-dimensional errors that are independent of errors $\varepsilon_{it}$ in (4.1) and fulfill

$$\nu_{it} \sim N(0, \sigma^2\Lambda) \tag{4.4}$$

(the parameter matrix $\Lambda$ is also unknown). The interpretation of (4.1) and (4.2) in the actuarial context is the same as in Section 2: one considers $n$ claims $y_{it}$ with dynamic development in time ($t = 1, \dots, n$) observed in $k$ insurance groups ($i = 1, \dots, k$).

One can treat the dynamic models of this type recursively by means of the Kalman filter. If one writes $\beta_{it}|y_{i,t-1}, y_{i,t-2}, \dots \sim N(\hat\beta_{it}^{t-1}, \sigma^2 P_{it}^{t-1})$ and $\beta_{it}|y_{it}, y_{i,t-1}, \dots \sim N(\hat\beta_{it}^{t}, \sigma^2 P_{it}^{t})$ then the Kalman filter provides

$$\hat\beta_{it}^{t} = \hat\beta_{it}^{t-1} + P_{it}^{t-1} x_{it} \frac{y_{it} - x'_{it}\hat\beta_{it}^{t-1}}{x'_{it}P_{it}^{t-1}x_{it} + 1}, \tag{4.5}$$

$$P_{it}^{t} = P_{it}^{t-1} - \frac{P_{it}^{t-1}x_{it}x'_{it}P_{it}^{t-1}}{x'_{it}P_{it}^{t-1}x_{it} + 1}, \tag{4.6}$$

where the predictive values $\hat\beta_{it}^{t-1}$ and $P_{it}^{t-1}$ are given by

$$\hat\beta_{it}^{t-1} = T\hat\beta_{i,t-1}^{t-1}, \tag{4.7}$$

$$P_{it}^{t-1} = TP_{i,t-1}^{t-1}T' + \Lambda. \tag{4.8}$$

It is not difficult to show that (4.5) can be obtained by the minimization procedure

$$\begin{aligned} \hat\beta_{it}^{t} &= \operatorname{argmin}\left\{\frac{1}{2}(\hat\beta_{it}^{t-1} - \beta_{it})'(\sigma^2 P_{it}^{t-1})^{-1}(\hat\beta_{it}^{t-1} - \beta_{it}) + \frac{1}{2\sigma^2}(y_{it} - x'_{it}\beta_{it})^2\right\} \\ &= \operatorname{argmin}\{(\hat\beta_{it}^{t-1} - \beta_{it})'(P_{it}^{t-1})^{-1}(\hat\beta_{it}^{t-1} - \beta_{it}) + (y_{it} - x'_{it}\beta_{it})^2\}, \end{aligned} \tag{4.9}$$

where argmin is taken over $\beta_{it} \in R^p$. The estimation of $\sigma^2$ and $\Lambda$ can be based on the maximum likelihood principle giving in the case of a non-informative prior (see e.g. [19])

$$\hat{\sigma}^2 = \frac{1}{(n-p)k} \sum_{i=1}^{k} \sum_{t=p+1}^{n} \frac{(y_{it} - x'_{it}\hat{\beta}_{it}^{t-1})^2}{x'_{it}P_{it}^{t-1}x_{it} + 1} \tag{4.10}$$

and the concentrated log-likelihood function for $\Lambda$

$$\hat{\Lambda} = \operatorname{argmax}\Big\{\text{constant } - \frac{(n-p)k}{2}\log\hat{\sigma}^2 - \frac{1}{2}\sum_{i=1}^{k}\sum_{t=p+1}^{n}\log(x'_{it}P_{it}^{t-1}x_{it} + 1)\Big\}. \tag{4.11}$$

The shrinkage model of the dynamic credibility approach has the form

$$\beta_{in} = \beta_n + e_{in} \tag{4.12}$$

$(i = 1, \ldots, k)$, where $e_{in}$ are independent $p$-dimensional errors fulfilling

$$e_{in} \sim N(0, \sigma^2 B_n) \tag{4.13}$$

(the parameters $\beta_n$ and $B_n$ of appropriate dimensions are unknown). It is sufficient to formulate the shrinkage model only for the time $t = n$ since the claim forecasts should be based on parameter estimates from the current time $n$. The corresponding credibility estimates $b_{in}$ of $\beta_{in}$ including the estimates $\hat{\beta}_n$ and $\hat{B}_n$ of $\beta_n$ and $B_n$ can be obtained as $b_i$, $\hat{\beta}$ and $B$ in Section 2 (see also Section 5).

As a special case of the described dynamic models one obtains the Gerber–Jones evolutionary model for $x_{it} = 1$ and $T = 1$: this model is a random walk

$$\beta_{it} = \beta_{i,t-1} + \nu_{it} \tag{4.14}$$

observed with a random error

$$y_{it} = \beta_{it} + \varepsilon_{it} \tag{4.15}$$

(see [9]). As practical applications are concerned the dynamic models include the cases of dynamic trends, dynamic seasonality or general structural models (see e.g. [12]).

## 5. Dynamic models with outliers and missing observations

The Kalman filter robustified by means of the $M$-estimation principle (see [4], [23]) seems to be a suitable instrument from the practical point of view for treating outliers and missing observations in the dynamic models from Section 4.

If outliers appear among observations $y_{it}$ which can be modelled by contaminating the normal distributions of the errors $\varepsilon_{it}$ in (4.1) by heavy-tailed distributions then one replaces the minimization procedure (4.9) by

$$\hat{\beta}_{it} = \operatorname{argmin}\Big\{\frac{1}{2}(\hat{\beta}_{it}^{t-1} - \beta_{it})'(\sigma^2 P_{it}^{t-1})^{-1}(\hat{\beta}_{it}^{t-1} - \beta_{it}) + \varrho\Big(\frac{y_{it} - x_{it}'\beta_{it}}{\sigma}\Big)\Big\}, \tag{5.1}$$

where $\varrho$ is a robustifying loss function with psi-function $\psi = \varrho'$ (the special choice $\varrho(z) = z^2/2$ reduces (5.1) to (4.9)).

In the most common case of Huber's psi-function $\psi_H$ the explicit solution of (5.1) can be found as

$$\hat{\beta}_{it}^{t} = \hat{\beta}_{it}^{t-1} + P_{it}^{t-1} x_{it}\sigma\psi_H\Big(\frac{y_{it} - x_{it}'\hat{\beta}_{it}^{t-1}}{\sigma(x_{it}'P_{it}^{t-1}x_{it} + 1)}\Big) \tag{5.2}$$

with $P_{it}^{t}$ calculated for simplicity according to (4.6). For a general psi-function $\psi$ one can use (5.2) (with $\psi$ instead of $\psi_H$) as an approximation, or the IWLS scheme can be applied similarly as in Section 3.

Moreover, if in addition to outliers some observations $y_{it}$ are missing the following iterative procedure is suggested using the previous results:

$$\hat{\beta}_{it}^{t-1(m)} = T\hat{\beta}_{i,t-1}^{t-1(m)}, \tag{5.3}$$

$$P_{it}^{t-1(m)} = TP_{i,t-1}^{t-1(m)}T' + \hat{\Lambda}^{(m)}, \tag{5.4}$$

$$\hat{\beta}_{it}^{t(m)} = \begin{cases} \hat{\beta}_{it}^{t-1(m)} + P_{it}^{t-1(m)}\hat{\sigma}^{(m-1)}\psi_H\Big(\dfrac{y_t - x_{it}'\hat{\beta}_{it}^{t-1(m)}}{\hat{\sigma}^{(m-1)}[x_{it}'P_{it}^{t-1(m)}x_{it} + 1]}\Big), \\ \hat{\beta}_{it}^{t-1(m)}, \end{cases} \tag{5.5}$$

$$P_{it}^{t(m)} = \begin{cases} P_{it}^{t-1(m)} - \dfrac{P_{it}^{t-1(m)}x_{it}x_{it}'P_{it}^{t-1(m)}}{x_{it}'P_{it}^{t-1(m)}x_{it} + 1}, \\ P_{it}^{t-1(m)} \end{cases} \tag{5.6}$$

(the first variant in (5.5) and (5.6) is for the observed $y_{it}$ and the second one for the missing $y_{it}$),

$$[\hat{\sigma}^{(m)}]^2 = \frac{[\hat{\sigma}^{(m-1)}]^2}{0.7785(S - kp)}\sum_i\sum_t \psi_H^2\Big(\frac{y_{it} - x_{it}'\hat{\beta}_{it}^{t-1(m)}}{\hat{\sigma}^{(m-1)}[x_{it}'P_{it}^{t-1(m)}x_{it} + 1]^{1/2}}\Big), \tag{5.7}$$

where $S$ is the number of all observed $y_{it}$, the sums are taken over all observed $y_{it}$ and the value $\hat{\Lambda}^{(m)}$ necessary for recursive calculation of (5.3)–(5.7) is obtained by minimizing

$$\text{(5.8)} \qquad \hat{\Lambda}^{(m)} = \operatorname{argmin}\Big\{(S-kp)\log[\hat{\sigma}^{(m)}]^2 + \sum_i \sum_t \log[x_{it}' P_{it}^{t-1(m)} x_{it} + 1]\Big\}.$$

Now one can continue the iterative procedure in a similar way as in Section 3:

$$\text{(5.9)} \qquad Z_{in}^{(m)} = \hat{B}_n^{(m-1)}(\hat{B}_n^{(m-1)} + P_{in}^{n(m-1)})^{-1},$$

$$\text{(5.10)} \qquad \hat{\beta}_n^{(m)} = \Big(\sum_i Z_{in}^{(m)}\Big)^{-1} \sum_i Z_{in}^{(m)} \hat{\beta}_{in}^{(m)},$$

$$\text{(5.11)} \qquad H_n^{(m)} = \frac{1}{k-1} \sum_i Z_{in}^{(m)} (\hat{\beta}_{in}^{(m)} - \hat{\beta}_n^{(m)})(\hat{\beta}_{in}^{(m)} - \hat{\beta}_n^{(m)})',$$

$$\text{(5.12)} \qquad B_n^{(m)} = \frac{1}{2[\hat{\sigma}^{(m)}]^2}(H_n^{(m)} + H_n^{(m)\prime}),$$

$$\text{(5.13)} \qquad b_{in}^{(m)} = Z_{in}^{(m)} \hat{\beta}_{in}^{(m)} + (I - Z_{in}^{(m)}) \hat{\beta}_n^{(m)}.$$

## 6. Numerical example

The recursive relations (5.3)–(5.8) that form the substantial part of the suggested iterative estimation procedure for dynamic models with outliers were applied to data from [22] simulated by means of the model (4.14) and (4.15) for $k = 1$ and $n = 31$:

$$\text{(6.1)} \qquad y_t = \beta_t + \varepsilon_t, \quad \varepsilon_t \sim \text{iid}\, N(0,4),$$

$$\text{(6.2)} \qquad \beta_t = \beta_{t-1} + \nu_t, \quad \nu_t \sim \text{iid}\, N(0,1)$$

($t = 1, \ldots, 31$), i.e. $\sigma^2 = 2^2$, $\sigma^2\Lambda = 1$. The simulated values $y_t$ and $\beta_t$ are given in Table 1. Obviously, the value $y_{20} = 35.00$ is an outlier. Here in difference from the previous treatment of these data (see [4], [22]) the variances $\sigma^2$ and $\sigma^2\Lambda$ are taken as unknown.

We used Huber's psi-function (3.3) with $c = u_{0.95} = 1.645$ and the initial values $\hat{\beta}_1^{1(m)} = y_1$, $P_1^{1(m)} = \text{var}(\hat{\beta}_1^{1(m)} - \beta_1)/\sigma^2 = \text{var}(y_1 - \beta_1)/\sigma^2 = 1$ for each $m$ and $[\hat{\sigma}^{(0)}]^2 = \sum(y_t - \overline{y})^2/(n-1)$ (sample variance).

After $m = 20$ iterations of (5.3)–(5.8) we obtained $\hat{\sigma}^2 = 2.78^2$, $\hat{\sigma}^2\hat{\Lambda} = 0.85$ and the estimates $\hat{\beta}_t$ ($t = 1, \ldots, 31$) given in Table 1. The results for $\hat{\beta}_t$ are comparable with those in [4], [22] in spite of the unknown variances in this case.

Table 1. Data and estimates after $m = 20$ iterations using (5.3)–(5.8)

| $t$ | $y_t$ | $\beta_t$ | $\hat{\beta}_t$ | $t$ | $y_t$ | $\beta_t$ | $\hat{\beta}_t$ |
|---|---|---|---|---|---|---|---|
| 1 | 8.65 | 10.00 | 8.65 | 17 | 7.29 | 3.69 | 7.17 |
| 2 | 7.28 | 9.83 | 7.93 | 18 | 5.94 | 3.37 | 6.83 |
| 3 | 7.44 | 9.98 | 7.74 | 19 | 1.96 | 3.25 | 5.48 |
| 4 | 11.13 | 8.99 | 8.86 | 20 | 35.00 | 2.81 | 7.23 |
| 5 | 11.18 | 9.36 | 9.56 | 21 | −0.62 | 2.36 | 5.48 |
| 6 | 5.45 | 8.50 | 8.37 | 22 | 4.13 | 2.46 | 5.11 |
| 7 | 6.17 | 8.90 | 7.74 | 23 | −0.84 | 0.82 | 3.46 |
| 8 | 3.92 | 8.20 | 6.67 | 24 | 2.78 | 0.24 | 3.27 |
| 9 | 12.32 | 8.47 | 8.25 | 25 | 1.93 | 1.62 | 2.90 |
| 10 | 6.95 | 7.46 | 7.89 | 26 | 0.45 | 1.46 | 2.22 |
| 11 | 10.46 | 6.49 | 8.60 | 27 | 2.54 | 1.96 | 2.31 |
| 12 | 9.54 | 7.34 | 8.86 | 28 | −0.95 | 2.62 | 1.41 |
| 13 | 7.07 | 7.82 | 8.36 | 29 | 2.69 | 2.95 | 1.76 |
| 14 | 8.17 | 7.06 | 8.31 | 30 | −0.89 | 1.40 | 1.03 |
| 15 | 5.59 | 6.85 | 7.56 | 31 | 2.83 | 2.84 | 1.53 |
| 16 | 5.99 | 5.67 | 7.12 | | | | |

### *References*


[1] *D. Birkes, Y. Dodge*: Alternative Methods of Regression. Wiley, New York, 1993.

[2] *H. Bühlmann, A. Gisler, W. Jewell*: Excess claims and data trimming in the context of credibility rating procedures. Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker *82* (1982), 117–147.

[3] *H. Bühlmann, E. Straub*: Glaubwürdigkeit für Schadensätze. Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker *70* (1970), 11–113.

[4] *T. Cipra, R. Romera*: Robust Kalman filter and its application in time series analysis. Kybernetika *27* (1991), 481–494.

[5] *P. deJong, B. Zehnwirth*: Credibility theory and the Kalman filter. Insurance: Mathematics and Economics *2* (1983), 281–286.

[6] *F. deVylder*: Practical credibility theory with emphasis on optimal parameter estimation. ASTIN Bulletin *12* (1981), 115–131.

[7] *F. deVylder*: Practical models in credibility theory, including parameter estimation. Premium Calculation in Insurance (F. deVylder, M. Goovaerts, J Haezendonck, eds.). Reidel, 1984.

[8] *R. Dutter*: Robust regression: Different approaches to numerical solution and algorithms. Research Report No. 6 (1975). ETH Zürich.

[9] *H. Gerber, D. Jones*: Credibility formulas of the updating type. Transactions of the Society of Actuaries *27* (1975), 31–46.

[10] *A. Gisler, P. Reinhard*: Robust credibility. ASTIN Bulletin *23* (1993), 117–143.

[11] *C. Hachemeister*: Credibility for regression models with applications to trend. Credibility: Theory and Applications (P.M. Kahn, ed.). Academic Press, New York, 1975.

[12] *A. Harvey, C. Fernandes*: Time series models for insurance claims. Journal of Institute of Actuaries *116* (1989), 513–528.

[13] *P. Huber*: Robust Statistics. Wiley, New York, 1981.
[14] *E. Kremer*: Credibility theory for some evolutionary models. Scandinavian Actuarial Journal (1982), 129–142.
[15] *E. Kremer*: Large claims in credibility. Blätter der Deutschen Gesellschaft für Versicherungsmathematik *20* (1991), 123–150.
[16] *E. Kremer*: Empirical Kalman-credibility. Blätter der Deutschen Gesellschaft für Versicherungsmathematik *22* (1995), 17–28.
[17] *E. Kremer*: Robust credibility via robust Kalman filtering. ASTIN Bulletin *24* (1994), 221–233.
[18] *H.R. Künsch*: Robust methods for credibility. ASTIN Bulletin *22* (1992), 33–49.
[19] *J. Ledolter, S. Klugman, C.S. Lee*: Credibility models with time-varying trend components. ASTIN Bulletin *21* (1991), 73–91.
[20] *J. Ledolter, C.S. Lee*: Analysis of many short time sequences: Forecast improvements achieved by shrinkage. Journal of Forecasting *12* (1993), 1–11.
[21] *W. Neuhaus*: Early warning. Scandinavian Actuarial Journal (1987), 128–156.
[22] *D. Peña, J. Guttman*: Optimal collapsing of mixture distributions in robust recursive estimation. Commun. Statist. - Theory Meth. *18* (1989), 817–833.
[23] *R. Romera, T. Cipra*: Practical implementation of robust Kalman filtering. Commun. Statist. - Simulation and Computation *24* (1995), 461–488.
[24] *K.D. Schmidt*: Stochastische Modellierung in der Erfahrungstarifierung. Blätter der Deutschen Gesellschaft für Versicherungsmathematik *20* (1992), 441–455.
[25] *B. Sundt*: Recursive credibility estimation. Scandinavian Actuarial Journal (1981), 3–21.

*Author's address*: *Tomáš Cipra*, Matematicko-fyzikální fakulta UK (Faculty of Mathematics and Physics, Charles University), Sokolovská 83, 186 00 Praha 8, Czech Republic.