Zeitschrift: Mitteilungen / Schweizerische Vereinigung der Versicherungsmathematiker = Bulletin / Association Suisse des Actuaires = Bulletin / Swiss Association of Actuaries

Herausgeber: Schweizerische Vereinigung der Versicherungsmathematiker

Band: - (1990)

Heft: 1

Artikel: Restricted minimax credibility : two special cases

Autor: Marazzi, Alfio

DOI: https://doi.org/10.5169/seals-967231

ALFIO MARAZZI, Lausanne

Restricted Minimax Credibility: Two Special Cases

1 Introduction

The simplest credibility formula $\delta(y) = ay + b$, where y is the average claim amount or loss ratio for a contract with risk parameter θ, can be derived within a decision-theoretical framework. Indeed, using a quadratic loss function, $\delta(y)$ is the linear Bayes estimate of $E(y \mid \theta)$ (see *Bühlmann* [3]). Moreover, it is the exact Bayes estimate of $E(y \mid \theta)$ when the density $f(\cdot \mid \theta)$ of y belongs to the single parameter exponential family and the structure function $U(\theta)$ is the corresponding conjugate prior distribution (see *Jewell* [10]).
Within this framework, two sources of error can distort the performances of the credibility estimate: (a) an inappropriate structure function; (b) an unexpected high frequency of very large claims i.e. an inappropriate model $f(\cdot \mid \theta)$. Minimax credibility was suggested by *Bühlmann* [4] and *Marazzi* [11] as a remedy for (a) and data trimming has been used by *Gisler* [6] in order to deal with (b).
We are going to consider the very simple model $y = \theta + e$ as an example and will show how the restricted Bayes and minimax principles proposed by *Hodges/Lehmann* [7] can be applied in order to obtain robust estimates of θ when: (a) the "error" e follows a Gaussian distribution and the structure function is not exactly known; (b) the structure function is Gaussian and the specified error distribution is not accurate. The solutions of the corresponding optimality problems provide a decision-theoretical justification for the well-known data trimming procedures.
The method has been described for the linear model in *Marazzi* [12]. This paper focuses on two simple special cases and indicates possible extensions where the Gaussian distribution is replaced by the exponential family.

1.1 The restricted Bayes and minimax principles

In a decision problem let the unknown parameter θ be a random variable with prior distribution U (structure function). Let $R(\theta, \delta)$ denote the risk

function of a decision procedure δ, for example an estimator of $E(y \mid \theta)$, and let $r(U,\delta) = \int R(\theta,\delta)\, dU(\theta)$ be the mean Bayes risk.
The Hodges & Lehmann approach to the problem of optimal decisions utilizes the available prior information but, at the same time, provides a safeguard in case this information is not correct. It is motivated as follows: the minimax decision does not use the prior information at all and is associated with the smallest possible value m for the maximum of the risk function; but we may be willing to tolerate a somewhat bigger maximum $m + c_0 > m$ if, in case the guess at θ has been a good one, this results in a substantial decrease in the average risk.
This leads to the following problems:

P I: The restricted Bayes problem. Let $c_0 > 0$ be a given number and U_0 a given prior distribution. Minimize $r(U_0, \delta)$ subject to

$$R(\theta,\delta) \leq m + c_0, \qquad \text{for all } \theta .$$

P II: The restricted minimax problem. Let $\varepsilon \in (0,1)$ be a given number, U_0 a given prior distribution, and let

$$\mathscr{P}_\varepsilon = \{U \mid U = (1-\varepsilon)U_0 + \varepsilon H, \quad H \in \mathscr{H}\}.$$

Find δ_ε such that $\sup_{\mathscr{P}_\varepsilon} r(U,\delta_\varepsilon) = \inf_{\mathscr{D}} \sup_{\mathscr{P}_\varepsilon} r(U,\delta)$.

Here $\mathscr{H}$ is the set of all prior distributions and $\mathscr{D}$ is a given class of decision functions. The elements of $\mathscr{H}$ are sometimes called *contaminations.*
Under general conditions δ_ε is Bayes for a least favorable (l.f.) distribution U_ε in $\mathscr{P}_\varepsilon$ and $(U_\varepsilon, \delta_\varepsilon)$ is a saddlepoint of the game $(\mathscr{P}_\varepsilon, \mathscr{D}, r)$. Furthermore, the two restricted problems are equivalent in the following sense: if δ_ε is restricted minimax, then δ_ε is a restricted Bayes solution with risk bounded by $\sup_\theta R(\theta, \delta_\varepsilon)$ and the converse also holds.
Our purpose is to apply the restricted Bayes and minimax principles to the problem of estimating θ when $y = \theta + e$ using a quadratic loss $L(\theta,\delta)$.
In Section 2 we assume that e has a normal distribution with a known variance. The exact mathematical solution of the restricted Bayes problem in this case is very messy. However, we show that

$$\text{Minimum Bayes risk } = 1 - I(G)$$

where $I(G)$ denotes the Fisher information for location of the marginal distribution G of y. As G depends on U it follows that the l.f. distribution in $P\,II$ minimizes $I(G)$ over $\mathscr{P}_\varepsilon$. This result is used in order to:

- obtain an approximate analytical solution of the restricted optimal problems;
- obtain accurate numerical approximations of the l.f. distribution and of the corresponding optimal estimate.

In Section 3 we exchange the role of prior and error distribution, i.e. we assume that U is Gaussian and that the error model is in a "neighborhood" of a given distribution F_0, and we modify the restricted Bayes and minimax problems in order to provide a safeguard against deviations from F_0. It turns out that the approximate solution of the corresponding optimality problem is based on data trimming.

2 The case of inaccurate structure function

Let $y = \theta + e$. Suppose that the density of e is $\phi_v(x) = (1/\sqrt{2\pi}\,v)\exp(-x^2/(2v^2))$ (v known) and that θ is distributed according to a structure function U. Let $f(y \mid \theta)$ denote the density of y for given θ and let $g(y) = f \circ U(y)$ be the marginal density of y where $f \circ U(y) = \int f(y \mid \theta)\, dU(\theta)$. The corresponding cumulative distributions are denoted by $F(y \mid \theta)$ and $G(y) = F \circ U(y)$. Let

$$I(G) = \int \left(\frac{d}{dy} \ln g(y)\right)^2 g(y)\, dy$$

be the Fisher information for location of G.
It is desired to estimate θ by an estimate δ using the loss $L(\theta, \delta) = (\theta - \delta)^2$. Without loss of generality, we restrict our attention to estimators of the form $\delta(y) = y + \psi(y)$ where ψ is an absolutely continuous function such that $E_\theta(|\psi'(y)|) < \infty$ and $E_\theta(\cdot)$ denotes the conditional expectation given θ.

Lemma 1.

I) $R(\theta, \delta) = v^2 + v^4 E_\theta(\psi^2(y) + 2\psi'(y))$ for $\delta(y) = y + v^2\psi(y)$.

II) The Bayes estimator of θ is $\delta_U(y) = y + v^2 g'(y)/g(y)$.

III) The minimum Bayes risk is $r(U, \delta_U) = v^2(1 - v^2 I(G))$.

Proof. Consider estimators of the form $\delta_a(y) = y + a\psi(y)$ where a is an arbitrary constant. We obtain:

$$\begin{aligned} R(\theta, \delta_a) &= E_\theta(\delta_a - \theta)^2 \\ &= v^2 + a^2 E_\theta(\psi^2(y)) + 2aE_\theta((y-\theta)\psi(y)). \end{aligned}$$

By partial integration $E_\theta((y-\theta)\psi(y)) = E_\theta(\psi'(y))v^2$ from which I) follows. Moreover:

$$r(U, \delta_a) = v^2 + a^2 E(\psi^2) + 2aE(\psi')v^2 .$$

We minimize first on a, the optimal value being

$$a_0 = -v^2 \frac{E(\psi')}{E(\psi^2)} \quad \text{with} \quad r(U, \delta_{a_0}) = v^2 - v^4 \frac{E(\psi')^2}{E(\psi^2)} .$$

Then we minimize on ψ observing that:

$$\frac{E(\psi')^2}{E(\psi^2)} \leq \int \left(\frac{g'(y)}{g(y)} \right)^2 g(y)\, dy = I(G)$$

by partial integration and Schwarz's inequality. Hence the Bayes estimator of θ is obtained with $a = v^2$ and $\psi = g'/g$. The properties II) and III) follow immediately.

Remark. g can be estimated from available collateral data.

2.1 *Approximate analytical solution of P I and P II*

In order to find a l.f. distribution in $\mathscr{P}_\varepsilon$ one should minimize $I(G)$ on the set

$$\mathscr{R}_\varepsilon = \{G \mid G = (1-\varepsilon)G_0 + \varepsilon K, \quad G_0 = F \circ U_0, \quad K = F \circ H, \quad H \in \mathscr{H}\}.$$

Denote by Θ the support of the l.f. contamination H_ε; let c be a Lagrange multiplier for the condition $\int g(y)\, dy = 1$ and let $\psi = g'/g$. By applying variational methods (as in *Huber* [8], p. 82) one obtains the condition

$$\begin{aligned} c - E_\theta(\psi^2(y) + 2\psi'(y)) &= 0 \qquad \text{for} \quad \theta \in \Theta \\ &\geq 0 \qquad \text{for} \quad \theta \notin \Theta . \end{aligned}$$

We remark, without surprise, that this coincides with the condition $R(\theta, \delta) \leq m + c_0$ with $c_0 = c$ in PI because y is the minimax estimate with $m = v^2$. If K were arbitrary, one would obtain $c - 2(g'/g)' - (g'/g)^2 = 0$ and this differential equation could be solved for g; unfortunately, the condition that K must be a mixture of normal densities makes the problem much harder.
As the function $E_\theta(\psi^2 + 2v^2\psi')$ is analytic in θ, the support Θ is a discrete set. A rigorous proof can be found in *Bickel/Collins* [2]. However, we do not know explicit formulae for the masses of H_ε nor for their abscissae. Therefore, approximate solutions (of approximate optimality problems) are of interest.
We consider the following problem (see also *Berger* [1]):

PI': The approximate restricted Bayes problem. Minimize $r(U_0, \delta)$ for $\delta(y) = y + v^2\psi(y)$ subject to:

$$\psi^2(y) + 2\psi'(y) \leq c_0 \qquad \text{for all} \quad y\,.$$

This condition is clearly motivated by I) in Lemma 1 and is stronger than the condition in PI. On the other hand, we define an *extended game* $(\widehat{\mathscr{R}}_\varepsilon, \mathscr{D}, \widehat{r})$ where

$$\begin{aligned}
\widehat{\mathscr{R}}_\varepsilon &= \{G \mid G = (1-\varepsilon)G_0 + \varepsilon K\,, \quad G_0 = F \circ U_0\,, \\
&\qquad K \text{ is an arbitrary contamination}\} \\
\mathscr{D} &= \{\delta \mid \delta(y) = y + v^2\psi(y)\} \\
\widehat{r}(G, \psi) &= v^2 + v^4 E_G(\psi^2 + 2\psi')
\end{aligned}$$

and $E_G(\cdot)$ denotes expectation using the distribution G. We remark that $\widehat{r}(G, \psi)$ coincides with $r(U, \delta)$ for $\delta \in \mathscr{D}$ and $G = F \circ U$ with $U \in \mathscr{P}_\varepsilon$. Therefore, one can formulate the following problem:

PII': The approximate restricted minimax problem. Let $\varepsilon \in (0,1)$ be a given number. Find $\widehat{\psi}_\varepsilon$ such that

$$\sup_{\widehat{\mathscr{R}}_\varepsilon} \widehat{r}(G, \widehat{\psi}_\varepsilon) = \inf_{\mathscr{D}} \sup_{\widehat{\mathscr{R}}_\varepsilon} \widehat{r}(G, \psi)\,.$$

By standard arguments, PII' leads to minimization of $I(G)$ over $\widehat{\mathscr{R}}_\varepsilon$ i.e. to the minimum condition:

$$c - 2(g'/g)' - (g'/g)^2 \geq 0\,.$$

Therefore, $P\,II'$ is equivalent to $P\,I'$. Moreover, assuming $-\log g_0$ to be convex, the result in *Huber* [8], p. 85 can be used: the l.f. density $\widehat{g}_\varepsilon$ is:

$$\begin{aligned}\widehat{g}_\varepsilon(y) =&(1-\varepsilon)g_0(y_0)e^{d(y-y_0)} && \text{for} \quad y \le y_0\\ &(1-\varepsilon)g_0(y) && \text{for} \quad y_0 < y < y_1\\ &(1-\varepsilon)g_0(y_1)e^{-d(y-y_1)} && \text{for} \quad y_1 < y\end{aligned}$$

where $d = \sqrt{c}$ is related to ε through the condition $\int \widehat{g}_\varepsilon(y)\,dy = 1$ and $y_0 < y_1$ are the endpoints of the interval where $|g_0'/g_0| \le d$. Finally the approximate restricted minimax estimate is

$$\widehat{\delta}_\varepsilon(y) = y + v^2\widehat{\psi}_\varepsilon(y)$$

with $\widehat{\psi}_\varepsilon = \widehat{g}_\varepsilon'/\widehat{g}_\varepsilon$. Clearly $\sup_\theta R(\theta, \widehat{\delta}_\varepsilon) = v^2 + v^4 d^2$.

Example. Let U_0 be the Gaussian distribution with mean μ and variance σ^2; then g_0 is the density of the Gaussian distribution with mean μ and variance $\tau^2 = \sigma^2 + v^2$. We obtain:

$$\widehat{\delta}_\varepsilon(y) = y + v^2 \max[-d, \min[d, (\mu - y)/\tau^2]]$$

or, more explicitly:

$$\begin{aligned}\widehat{\delta}_\varepsilon(y) =&y + v^2 d && \text{for} \quad y < \mu - d\tau^2\\ &y\frac{\sigma^2}{\tau^2} + \mu\frac{v^2}{\tau^2} && \text{for} \quad \mu - d\tau^2 \le y \le \mu + d\tau^2\\ &y - v^2 d && \text{for} \quad y > \mu + d\tau^2\end{aligned}$$

The constants ε and d are related through $2\phi_\tau(d\tau^2)/d + 2\Phi_\tau(d\tau^2) = (2-\varepsilon)/(1-\varepsilon)$ where Φ_τ denotes the Gaussian distribution with mean 0, variance τ^2 and density ϕ_τ. We observe that $\widehat{\delta}_\varepsilon$ coincides with the *limited translation rule* of *Efron/Morris* [5], which follows the Bayes rule as closely as possible subject to the condition $|\delta(y) - y| \le v^2 d$.

2.2 *Numerical approximation of the least favorable distribution*

The discrete nature of Θ suggests the possibility to approximate H_ε numerically. Indeed, for the case $u_0(\theta) = \phi_\sigma(\theta)$ (the Gaussian density with mean 0 and variance σ^2) *Marazzi* [12] minimizes $I(\overline{G})$ over the $2n+2$ parameters $h_1, \ldots, h_n, \theta_1, \ldots, \theta_n, t, b$ of the marginal density

$$\overline{g}(y) = (1-\varepsilon)\phi_\tau(y) + \varepsilon\left[\sum_{j=1}^{n} h_j \alpha(y;\theta_j) + h\int_0^\infty \alpha(y;t+\theta)e^{-b\theta}b\,d\theta\right]$$

where $\alpha(y;\theta) = \phi_v(y+\theta) + \phi_v(y-\theta)$, $\sum h_j + h = 0.5$ and $\tau^2 = v^2 + \sigma^2$. By choosing n sufficiently large, the least favorable marginal density $g_\varepsilon = (1-\varepsilon)\phi_\tau + \varepsilon\phi_v \circ H_\varepsilon$ may be approximated as precisely as desired by functions of this form. Note that $\overline{g}$ has been constructed so that the asymptotic behaviours for large arguments of $\overline{g}'/\overline{g}$ and of the risk function of $\overline{\delta} = y + v^2\overline{g}'/\overline{g}$ coincide with the corresponding behaviours of the analytical approximation of Section 2.1.

In Figure 1 (taken from *Marazzi* [12]) the functions $-\widehat{g}_\varepsilon'/\widehat{g}_\varepsilon$ and $-\overline{g}_\varepsilon'/\overline{g}_\varepsilon$ (obtained by minimizing $I(\overline{G})$) are drawn together with the corresponding risk functions. The numerical approximation mimics the oscillatory behaviour of the optimal rule for low values of y and replaces the oscillations by a simpler curve for those values of y which do not appreciably affect the interesting mean risks $r(U_0, \overline{\delta}_\varepsilon)$ and $r(\overline{H}_\varepsilon, \overline{\delta}_\varepsilon)$.

Some of the numerical results are indicated in Table 1 where the value of $r(\overline{H}_\varepsilon, \overline{\delta}_\varepsilon)$ is an approximation for $\sup_\theta R(\theta, \overline{\delta}_\varepsilon)$. Clearly

$$I(\widehat{G}_\varepsilon) \le \min_{\mathscr{P}_\varepsilon} I(G) \le I(\overline{G}_\varepsilon)$$

and, as the lower bound is numerically close to the upper bound, it can be concluded that the analytical (and the numerical) approximation is nearly optimal.

3 The case of inaccurate error distribution

Let $y = \theta + e$. Suppose that the structure function U is the Gaussian distribution with mean μ and variance σ^2 and note the distribution of e by F with density f. We define the *a priori mean squared loss function* of an estimator δ of θ as

$$l(e,\delta) = \int L(\theta, \delta(\theta+e))\,dU(\theta)$$

Table 1. Numerical results

ε	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
(v^2, σ^2)			(1,0)					(1,1)		
$I(\widehat{g}_\varepsilon)$	0.671	0.489	0.354	0.250	0.169	0.336	0.244	0.177	0.125	0.084
$r(U_0, \widehat{\delta}_\varepsilon)$	0.110	0.203	0.293	0.381	0.472	0.555	0.602	0.646	0.691	0.736
$\sup R(\theta, \widehat{\delta}_\varepsilon)$	2.230	1.742	1.469	1.302	1.190	1.650	1.371	1.234	1.151	1.095
$I(\overline{g}_\varepsilon)$	0.697	0.520	0.385	0.279	0.193	0.340	0.250	0.182	0.130	0.088
$r(U_0, \overline{\delta}_\varepsilon)$	0.092	0.176	0.257	0.340	0.425	0.552	0.597	0.640	0.684	0.728
$r(\overline{H}_\varepsilon, \overline{\delta}_\varepsilon)$	2.194	1.698	1.447	1.294	1.190	1.633	1.364	1.231	1.149	1.095
n	3	3	4	4	4	4	5	5	5	5

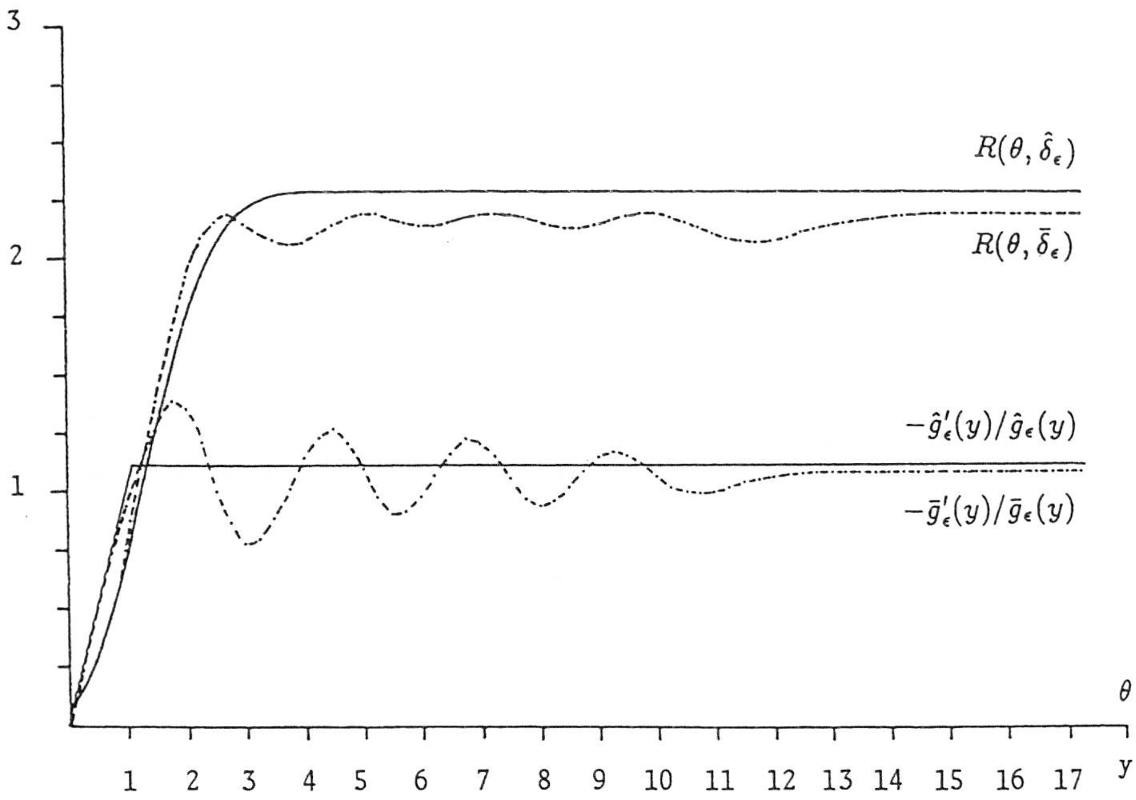

Figure 1. Minimax functions $-\overline{g}'_\varepsilon/\overline{g}_\varepsilon$, $-\widehat{g}_\varepsilon/\widehat{g}_\varepsilon$ and corresponding risk functions (for $\varepsilon = 0.1$, $v^2 = 1$, $\sigma^2 = 0$, $n = 3$)

and use $L(\theta, \delta) = (\theta - \delta)^2$. The mean (Bayes) risk is then

$$r(F, \delta) = \int l(e, \delta)\, dF(e).$$

In this section it will be convenient to consider the mean risk as a function of δ and F. As in Section 2 let $g(y) = f \circ U(y) = \int f(y - \theta)\, dU(\theta)$ and $I(G)$ be the Fisher information for location of G.

Lemma 2.

I) $l(e, \delta) = \sigma^2 + \sigma^4 \int (\psi^2(y) + 2\psi'(y))\, dU(\theta)$ for $\delta(y) = \mu - \sigma^2 \psi(y)$.

II) The Bayes estimator of θ is $\delta_F(y) = \mu - \sigma^2 g'(y)/g(y)$.

III) The minimum Bayes risk is $r(F, \delta_F) = \sigma^2(1 - \sigma^2 I(G))$.

Proof. Similar to the proof of Lemma 1 in Section 2.

3.1 *Modified restricted optimality principles*

In Section 2 the cause of an "outlying" value of y was assigned to an outlying value of θ and it was appropriate to obtain robustness by bounding $R(\theta, \delta)$. It is clearly impossible to define a sensible robust procedure based on the single observation y without assuming one of its two components, θ or e, to be correct. Yet, it does not seem unreasonable to treat the two components of y in a similar way. Therefore, if the cause of an outlying value of y is now assigned to a bad value of e it may be appropriate to obtain robustness by bounding $l(e, \delta)$. The following approach simply paraphrases the previous section by exchanging the roles of the structure function and the error distribution.

Let F_0 be a given error distribution. The goal is to find an optimal estimator of θ that utilizes the information contained in F_0 but, at the same time, provide a safeguard in case this information is incorrect. Consider $\delta^0(y) \equiv \mu$, the prior mean of θ; this estimator does not use F_0 at all and its maximum a priori mean squared loss $\sup_e l(e, \delta^0) \equiv l(e, \delta^0) \equiv \sigma^2$ is the smallest possible value for the maximum of the a priori mean squared loss function. But we might be willing to tolerate a somewhat larger maximum a priori mean loss if there results a substantial decrease in the mean risk when F_0 is correct. This leads to the following problem:

P i: The modified restricted Bayes problem. Minimize $r(F_0, \delta)$ subject to the condition

$$l(e,\delta) \le \sigma^2 + c_0 \qquad \text{for all} \quad e \qquad (c_0 > 0)\,.$$

The Hodges & Lehmann theory for $P\,I$ and $P\,II$ can obviously be applied by exchanging the prior and error distributions. In particular one needs to consider the sets

$$\mathscr{P}_\varepsilon = \{F \mid F = (1-\varepsilon)F_0 + \varepsilon H, \quad H \in \mathscr{H}\}$$

where $\mathscr{H} = \{\text{all distributions}\}$ and

$$\mathscr{R}_\varepsilon = \{G \mid G = (1-\varepsilon)G_0 + \varepsilon K, \quad G_0 = F_0 \circ U, \quad K = H \circ U, \quad H \in \mathscr{H}\}$$

where $\varepsilon \in (0,1)$. There is an equivalent minimax problem:

P ii: The modified restricted minimax problem. Minimize the maximum Bayes risk over all error distributions $F \in \mathscr{P}_\varepsilon$.

From Lemma 2 it follows that the l.f. distribution H_ε in $P\,ii$ minimizes $I(G)$ over $\mathscr{P}_\varepsilon$ and $\mathscr{R}_\varepsilon$. With the aid of a Lagrange multiplier for $\int g(y)\,dy = 1$ we obtain the minimum condition

$$c - \int (\psi^2(y) + 2\psi'(y)) f(y-\theta)\, dU(\theta) = 0 \qquad \text{for} \quad y \in \Gamma$$

where $\psi = g'/g$ and Γ is the support of H_ε. The equality sign is replaced by $\ge$ for $y \notin \Gamma$.
A numerical procedure similar to the one described in Section 2.2 is applicable to the determination of H_ε but we may be satisfied with the l.f. marginal distribution $\widehat{G}_\varepsilon$ in the extended class

$$\widehat{\mathscr{R}}_\varepsilon = \{G \mid G = (1-\varepsilon)G_0 + \varepsilon K, \quad G_0 = F_0 \circ U, \quad K \text{ arbitrary}\}.$$

The form of $\widehat{g}_\varepsilon$ is the same as in Section 2.1 and the approximate restricted minimax estimate is

$$\widehat{\delta}_\varepsilon(y) = \mu - \sigma^2 \widehat{\psi}_\varepsilon(y)$$

where $\widehat{\psi}_\varepsilon = \widehat{g}_\varepsilon'/\widehat{g}_\varepsilon$. Clearly $\sup_e l(e, \widehat{\delta}_\varepsilon) = \sigma^2 + \sigma^4 d^2$.

4 Extensions and open problems

The method has been extended by *Marazzi* [12] to the linear model

$$y = X\vartheta + e$$

where y is an n-vector of observations, ϑ a p-vector of parameters, X an $n \times p$ matrix of constants and e an n-vector of errors. Two cases have been considered: (a) the distribution of e is the n-variate Gaussian distribution and the p-variate structure function is not exactly known; (b) the structure function is a p-variate Gaussian distribution and the specified n-variate error distribution F_0 is affected by contamination.

For example, in case (b) with $p = 1$ and $X = (1, \dots, 1)^T$, if F_0 is the n-variate Gaussian distribution with mean vector O and covariance matrix $v^2 I$ one obtains an approximate restricted minimax estimate of the form

$$\widehat{\delta}_\varepsilon(y) = \mu - \sigma^2 \max[-d_n, \min[d_n, (\mu - \bar{y})/\tau^2]]$$

where $\bar{y}$ is the arithmetic mean of the components of y, d_n is an appropriate constant, $\mu = E(\vartheta)$, $\sigma^2 = \mathrm{Var}\,(\vartheta)$ and $\tau^2 = \sigma^2 + v^2/n$.

The estimate $\widehat{\delta}_\varepsilon(y)$ is based on the assumption that the components of e are independent and identically distributed with probability $(1 - \varepsilon)$; however, e comes from an arbitrary multivariate contamination with probability ε. A different model assumes that the distribution of $e = (e_1, \dots, e_n)^T$ is of the form

$$F(e) = F_1(e_1) \cdot F_2(e_2) \cdot \cdots \cdot F_n(e_n)$$

where each factor F_i is a mixture of a given univariate distribution and an arbitrary univariate contamination. The application of the restricted Bayes and minimax principles to this situation is still an open problem.

The crucial identity of Section 2 allowing to relate the minimum Bayes risk to the Fisher information is $E_\theta((y - \theta)\psi(y)) = v^2 E(\psi'(y))$. This identity can be generalized to the continuous exponential family

$$f(y \mid \theta) = \exp(\theta y - \gamma(\theta))\beta(y)$$

with support $\mathbb{R} = (-\infty, \infty)$. If the support is a bounded interval we need a supplementary condition (see *Hudson* [9]). Indeed, with $s(y) = -\beta'(y)/\beta(y)$, we obtain

$$E_\theta((s(y) - \theta)\psi(y)) = E_\theta(\psi'(y))$$

for any absolutely continuous function ψ on $\mathbb{R}$ such that $E_\theta(|\psi'(y)|) < \infty$. The Bayes estimator of θ with respect to a prior distribution U is $\delta_U(y) = s(y) + g'(y)/g(y)$ and the minimum Bayes risk is $r(\delta_U, U) = E_G(s'(y)) - I(G)$. A similar extension of the results of Section 3 is also possible.
Therefore, in order to find a l.f. distribution in a (weakly compact and convex) set of prior distributions $\mathscr{P}$ one has to minimize the functional $J(G) = I(G) - E_G(s')$ on $\mathscr{R} = \{G \mid G(y) = \int F(y \mid \theta)\, dU(\theta), \quad U \in \mathscr{P}\}$. Again we may consider $\mathscr{P} = \mathscr{P}_\varepsilon$ in which case we note $\mathscr{R}$ by $\mathscr{R}_\varepsilon$. Moreover, if we allow G to belong to the extended set $\widehat{\mathscr{R}}_\varepsilon$, we obtain the condition

$$c - s' - 2(g'/g)' - (g'/g)^2 = 0$$

on the set of y-values where g can be freely varied. Writing $z(y) = \sqrt{g(y)}$ the equation becomes

$$z''(y) + \frac{1}{4}[s'(y) - c]z(y) = 0\,.$$

Beyond this point the problem remains open.

Alfio Marazzi
Institut universitaire de médecine sociale et préventive
Bugnon 17,
1004 Lausanne

Bibliography

[1] *Berger, J. O.* (1982): Estimation in continuous exponential families: Bayesian estimation subject to risk restrictions and inadmissibility results. In: S. S. Gupta and J. O. Berger (Eds.), Statistical decision theory and related topics III. Academic Press, New York.

[2] *Bickel, P. J./Collins, J. R.* (1983): Minimizing Fisher information over mixtures of distributions. Sankhyã: The Indian Journal of Statistics *45*, Series A, Pt. 1, 1–19.

[3] *Bühlmann, H.* (1967): Experience rating and credibility. ASTIN Bulletin *4*, 199–207.

[4] *Bühlmann, H.* (1975): Minimax credibility. In: P. M. Kahn (Ed.), Credibility theory and applications. Academic Press, New York.

[5] *Efron, B./Morris, C.* (1971): Limiting the risk of Bayes and empirical Bayes estimators – Part I: the Bayes case. Journal Amer. Statist. Assoc. *66*, 807–815.

[6] *Gisler, A.* (1980): Optimum trimming of data in the credibility model. Bulletin of the Association of Swiss Actuaries *80*, 313–325.

[7] *Hodges, J. L. Jr./Lehmann, E. L.* (1952): The use of previous experience in reaching statistical decisions. Ann. Math. Statist. *23*, 396–407.

[8] *Huber, P. J.* (1980): Robust statistic. Wiley, New York.

[9] *Hudson, H. M.* (1978): A natural identity for exponential families with applications in multiparameter estimation. Ann. of Statist. *6*, 473–484.

[10] *Jewell, W. S.* (1974): Credible means are exact Bayesian for exponential families. ASTIN Bulletin *8*, 77–90.

[11] *Marazzi, A.* (1976): Minimax credibility. Bulletin of the Association of Swiss Actuaries *76*, 219–229.

[12] *Marazzi, A.* (1985): On constrained minimization of the Bayes risk for the linear model, Statistics & Decisions *3*, 277–296.

Summary

The restricted Bayes and minimax principles are used in order to derive robust credibility estimates of a risk parameter θ in the very simple case where the claim is the sum of θ and an "error term". Two examples are considered: (a) the error distribution is Gaussian and the structure function is not exactly known; (b) the structure function is Gaussian and the error model is not precise. Approximate analytical and numerical solutions as well as possible extensions are discussed.

Zusammenfassung

Eingeschränkte Bayes- und Minimaxprinzipien werden angewandt, um robuste Credibility Schätzungen eines Risikoparameters θ abzuleiten, dies im ganz einfachen Falle, wo der Schadenbetrag der Summe von θ und einem "Fehlerwert" gleich gesetzt ist. Zwei Beispiele werden betrachtet: (a) die Fehlerverteilung ist nach Gauss und die Strukturfunktion nicht genau bekannt; (b) die Strukturfunktion ist nach Gauss und das Fehlermodell ungenau. Annähernde analytische und numerische Lösungen sowie mögliche Erweiterungen werden beschrieben.

Résumé

Les critères restreints de Bayes et minimax sont utilisés pour développer des estimateurs de crédibilité robustes d'un paramètre de risque θ dans le cas simple où le sinistre est la somme de θ et d'une "erreur". Deux exemples sont considérés: (a) la distribution de l'erreur est gaussienne et la fonction de structure n'est connue qu'approximativement; (b) la fonction de structure est gaussienne et le modèle d'erreur est imprécis. Des solutions approximatives analytiques et numériques ainsi que des extensions possibles sont décrites.