q-Credibility

Olivier Le Courtois *

Abstract

This article extends uniform exposure credibility theory by making quadratic adjustments that take into account the squared values of past observations. This approach amounts to introducing non-linearities in the framework, or to considering higher order cross moments in the computations. We first describe the full parametric approach and, for illustration, we examine the Poisson-gamma and Poisson-Pareto cases. Then, we look at the non-parametric approach where premiums can only be estimated from data and no type of distribution is postulated. Finally, we examine the semi-parametric approach where the conditional distribution is Poisson but the unconditional distribution is unknown. For all of these approaches, the mean square error is, by construction, smaller in the q-credibility framework than in the standard framework.

Keywords

Credibility. Quadratic Approximation. Parametric. Non-Parametric. Semi-Parametric. Poisson-Gamma. Poisson-Pareto. Uniform Exposure.

Mathematics Subject Classification: 62C10, 62C12, 91B30, 97M30
Journal of Economic Literature Classification: C11, G22

*O. Le Courtois, PhD, FSA, CFA, is a Professor of Finance and Insurance at emlyon business school. `lecourtois@em-lyon.com`. Address: 23, Avenue Guy de Collongue, 69134 Ecully Cedex, France. Phone: 33-(0)4-78-33-77-49. Fax: 33-(0)4-78-33-79-28.

Introduction

The origins of credibility theory can be traced back to the papers of Mowbray (1914), Whitney (1918), Bailey (1945, 1950), Longley-Cook (1962), and Mayerson (1964). The core of the theory, as it is known today, is developed in Bühlmann (1967) and in Bühlmann and Straub (1970). See also Hachemeister (1975) for the link with regressions, Zehnwirth (1977) for the link with Bayesian analysis, and Norberg (1979) for the application to ratemaking. General presentations of the theory can be found for instance in Bühlmann (1970), Herzog (1999), Klugman, Panjer, and Willmot (2012), Weishaus (2015), and Norberg (2015). See also the recent broad survey paper by Lai (2012).

In this paper, we construct a quadratic credibility framework where premiums are estimated based on the values of past observations and of past squared observations. In Chapter 7 of Bühlmann and Gisler (2005), it is already mentioned - however without further developments - that credibility estimators are not theoretically restricted to be linear in the observations and that the squares of observations could be used in credibility theory. This paper can be viewed as a first contribution to this research program. See also Chapter 4 of Bühlmann and Gisler (2005) where a maximum likelihood estimator is computed using a logarithmic transformation of the observations, but note that the latter application more provides an appropriate trick for dealing with the Pareto distribution than a non-linear framework *per se*.

We fully compute non-linear, quadratic, credibility estimators in situations that range from parametric to non-parametric settings. The framework that is developed can be useful for the modeler who explicitly wants to deviate from a linear framework and to take into account higher order (cross) moments. For instance, our framework uses the explicit values of the covariance between observations and squared observations, and also the covariance between squared observations. For each of the parametric, non-parametric, and semi-parametric settings explored in this paper, we give illustrations of the reduction of the mean square error gained by going from the classic to the quadratic credibility approach. See Neuhaus (1985) for general results on errors in credibility theory. Note that Norberg (1982), extending De Vylder (1978), uses high order moments and cross moments but in a different context: for the statistical estimation of classic structural parameters. See De Vylder (1985) for a reference on non-linear (in particular exponential) regressions in credibility theory and Hong and Martin (2017) for a flexible non parametric Bayesian approach. See also Taylor (1977) who proposes a Hilbert approach to credibility theory, derives results on sufficient statistics, and constructs an example with non-linear but unbiased statistics of the observations. In the present paper, unbiasedness is obtained by construction. For more details about the Hilbert space approach, see Shiu and Sing (2004).

The paper is organized as follows. The first section develops a parametric quadratic credibility, or q-credibility, approach and provides illustrations of this approach in the Poisson-gamma and Poisson-Pareto settings. Building on the results of the first section, the second section derives a non-parametric approach and the third section concentrates on a semi-parametric approach where the conditional distribution is assumed to be of the Poisson type.

1 Main Results

We consider n random variables $\{X_i\}_{i=1:n}$ that are identically distributed and independent conditionally on a random variable Θ that represents the uncertainty of the system or the parameters of each of the $\{X_i\}_{i=1:n}$. Note that the random variables $\{X_i\}_{i=1:n}$ are not necessarily i.i.d. in full generality. Furthermore, for any strictly positive integer m, we define

$$\mu_m = E(E(X^m|\Theta))$$

and

$$v_m = E(\mathrm{Var}(X^m|\Theta)),$$

and for simplicity we also denote $\mu = \mu_1$ and $v = v_1$. Then, we define

$$a = \mathrm{Var}(\mu(\Theta)),$$

where $\mu(\Theta) = E(X|\Theta)$, and we have:

$$\mathrm{Cov}(X_i, X_k) = a, \qquad \forall i \neq k,$$

and

$$\mathrm{Cov}(X_i, X_i) = \mathrm{Var}(X_i) = a + v. \tag{1}$$

Classic credibility is a method that solves the following program:

$$\min_{\alpha_0, \{\alpha_i\}_{i=1:n}} E\left(\left[\alpha_0 + \sum_{i=1}^{n} \alpha_i X_i - X_{n+1}\right]^2\right)$$

to estimate the future outcome X_{n+1} of a quantity X by using past realizations $\{X_i\}_{i=1:n}$ of this quantity. The solution of this program produces the following estimator of X_{n+1}:

$$\hat{P}_{n+1} = \mu\ (1-z) + z\ \bar{X},$$

where

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

and

$$z = \frac{na}{na+v}.$$

Also note that the mean square error in classic credibility theory is given by

$$\mathrm{MSE}_c = E\left(\left[\hat{P}_{n+1} - X_{n+1}\right]^2\right) = v + a\ (1-z). \tag{2}$$

It is also possible to define

$$\mathrm{MSE}'_c = E\left(\left[\hat{P}_{n+1} - E(X_{n+1}|\Theta)\right]^2\right) = a\ (1-z).$$

In this paper, we introduce q-credibility as a method to estimate the future outcome X_{n+1} of a quantity X by using the past realizations $\{X_i\}_{i=1:n}$ of this quantity but also the past realizations $\{X_i^2\}_{i=1:n}$ of the square of this quantity and by performing a least-squares optimization. Therefore, our goal is to solve the following extended program:

$$\min_{\alpha_{0,q},\{\alpha_i\}_{i=1:n},\{\beta_i\}_{i=1:n}} E\left(\left[\alpha_{0,q} + \sum_{i=1}^{n}\alpha_i X_i + \sum_{i=1}^{n}\beta_i X_i^2 - X_{n+1}\right]^2\right). \tag{3}$$

For this purpose, we first introduce four new structural parameters b, g, c, and h, defined as follows:

$$\mathrm{Cov}(X_i^2, X_k) = b, \qquad \forall i \neq k, \tag{4}$$

and

$$\mathrm{Cov}(X_i^2, X_i) = b + g, \tag{5}$$

and also

$$\mathrm{Cov}(X_i^2, X_k^2) = c, \qquad \forall i \neq k, \tag{6}$$

and

$$\mathrm{Cov}(X_i^2, X_i^2) = \mathrm{Var}(X_i^2) = c + h. \tag{7}$$

We can easily check that $b = \mathrm{Cov}(E(X^2|\Theta), E(X|\Theta))$, $g = E(\mathrm{Cov}(X^2, X|\Theta))$, $c = \mathrm{Var}(E(X^2|\Theta))$, and $h = E(\mathrm{Var}(X^2|\Theta))$. We can now state the main result of this section.

Proposition 1.1 (q-credibility)**.** *The q-credibility premium* $\hat{P}^q_{n+1}$ *that solves the program* (3) *and that gives the best quadratic estimator of* X_{n+1} *can be expressed as a function of the empirical mean* $\bar{X}$ *of the past values, of the empirical mean* $\overline{X^2} = \frac{1}{n}\sum_{i=1}^n X_i^2$ *of the past squared values, and of the high-order co-moments defined in Eqs* (4) *to* (7)*, as follows:*

$$\hat{P}^q_{n+1} = \alpha^*_{0,q} + z_q\ \bar{X} + y_q\ \overline{X^2}, \tag{8}$$

where

$$\alpha^*_{0,q} = \mu\ (1 - z_q) - y_q\ (\mu^2 + a + v), \tag{9}$$

and

$$z_q = \frac{n\,[a(nc+h) - b(nb+g)]}{(na+v)(nc+h) - (nb+g)^2}, \tag{10}$$

and

$$y_q = \frac{n(bv - ag)}{(na+v)(nc+h) - (nb+g)^2}, \tag{11}$$

where $\lim_{n\to+\infty} z_q(n) = 1$ *and* $\lim_{n\to+\infty} y_q(n) = 0$*, so where the best estimator in the presence of infinite experience is simply the empirical mean.*

In this proposition, we assume that the denominators of Eqs (10) and (11) are non-null. Nothing prevents the credibility factor z_q to be negative when the experience is limited, so when n is small. The last two illustrations of the paper will show situations where this is the case, even when all the structural parameters have positive estimated values. Indeed, q-credibility theory only provides the outcome of a least-squares optimization. As long as we do not expect more from the framework than what it can provide, it is not inconsistent that the best estimator of a future claim or claim number negatively depends on the empirical mean of past values, as long as a correction by the empirical mean of past squared values is applied. Let us now make four important remarks.

Remark 1.2. Similar to the classic estimator $\hat{P}_{n+1}$, the quadratic estimator $\hat{P}^q_{n+1}$ is by construction unbiased. Indeed, we have that $E(\hat{P}^q_{n+1}) = E(X_{n+1})$ from Eq. (27) in the proof of Proposition 1.1.

Remark 1.3. We note that $\hat{P}^q_{n+1}$ can also be written as

$$\hat{P}^q_{n+1} = \mu\ (1 - z_q) + \bar{X}\ z_q + y_q\ (\overline{X^2} - \mu_2),$$

or as

$$\hat{P}^q_{n+1} = \mu + z_q\ (\bar{X} - \mu) + y_q\ (\overline{X^2} - \mu_2),$$

so that q-credibility amounts to correcting credibility premiums by a proportion of the difference between the empirical and the theoretical non-centered second-order moments.

Remark 1.4. When $b = g = 0$, and without imposing any constraint on c or h, then $z_q = \frac{na}{na+v} = z$, $y_q = 0$ and $\alpha_{0,q} = \mu\ (1 - z_q) = \mu\ (1 - z)$, so we recover the classic credibility case.

Remark 1.5. It can be easily checked that the solution to Eq. (3) also solves

$$\min_{\alpha_{0,q},\{\alpha_i\},\{\beta_i\}} E\left(\left[\alpha_{0,q} + \sum_{i=1}^{n} \alpha_i X_i + \sum_{i=1}^{n} \beta_i X_i^2 - E\left(X_{n+1} \mid \Theta\right)\right]^2\right).$$

To measure the gain reached by going from the credibility to the q-credibility framework, we derive the following proposition, in the spirit of Neuhaus (1985).

Proposition 1.6 (Mean Square Error). *In the q-credibility framework, the mean square error - or quadratic loss - is equal to*

$$MSE_q = E\left(\left[\hat{P}^q_{n+1} - X_{n+1}\right]^2\right) = v + a\ (1 - z_q) - y_q\ b. \qquad (12)$$

We also define

$$MSE'_q = E\left(\left[\hat{P}^q_{n+1} - E(X_{n+1}|\Theta)\right]^2\right) = a\ (1 - z_q) - y_q\ b.$$

It is in fact possible to relate MSE_c and MSE_q as follows.

Remark 1.7. Because the space of the combinations of the $\{X_i\}_{i=1:n}$ is a subspace of the combinations of the $\{X_i\}_{i=1:n}$ and of the $\{X_i^2\}_{i=1:n}$, which is itself a subspace of the space of the square integrable random variables, we have the following Pythagorean result:

$$E\left(\left[X_{n+1} - \hat{P}_{n+1}\right]^2\right) = E\left(\left[X_{n+1} - \hat{P}^2_{n+1}\right]^2\right) + E\left(\left[\hat{P}_{n+1} - \hat{P}^2_{n+1}\right]^2\right),$$

which expresses that the projection on a subsubspace is the projection on a subspace plus the square distance between the two projections. Therefore, we have:

$$\text{MSE}_c = \text{MSE}_q + \text{Var}\left(\hat{P}_{n+1} - \hat{P}^2_{n+1}\right).$$

We can also define

$$\Delta\text{MSE} = \text{MSE}_c - \text{MSE}_q = a\ (z_q - z) + b\ y_q.$$

Note that relative gains will be measured by the quantity:

$$\kappa = \frac{\Delta\text{MSE}}{\text{MSE}_c} = \frac{a\ (z_q - z) + b\ y_q}{v + a\ (1 - z)},$$

or by

$$\kappa' = \frac{\Delta\text{MSE}'}{\text{MSE}'_c} = \frac{a\ (z_q - z) + b\ y_q}{a\ (1 - z)}.$$

We can make the following additional remarks:

Remark 1.8. The introduction of the β_i's in the optimization program allows us to reach a smaller quadratic distance between the estimator and X_{n+1}. Therefore, we always have $\text{MSE}_q \leq \text{MSE}_c$.

and

Remark 1.9. When $b = g = 0$, and without imposing any constraint on c or h, then $\text{MSE}_q = \text{MSE}_c$.

and also

Remark 1.10. Eqs (9), (10), (11), (2), and (12) are valid in all this text.

Next, we obtain general expressions for the structural parameters.

Proposition 1.11 (Parameters in the general case)**.** *We recall that* $v = E[Var(X|\Theta)]$ *and* $a = Var[E(X|\Theta)]$ *where,* $\forall i \neq k$*,* X_k *and* X_i *are identically distributed and can be replaced, when there is no ambiguity, by a representative random variable* X*. Further,* X_i^2 *and* X_k *are assumed independent conditionally on* Θ*. The quantities* b*,* g*,* c*, and* h *defined in Eqs* (4) *to* (7) *can be expressed as functions of* Θ *as follows:*

$$b = Cov[E(X^2|\Theta), E(X|\Theta)] = E[E(X^2|\Theta)\ E(X|\Theta)] - E[E(X^2|\Theta)]E[E(X|\Theta)], \tag{13}$$

and

$$g = E[Cov(X^2, X|\Theta)] = E[E(X^3|\Theta)] - E[E(X^2|\Theta)\ E(X|\Theta)], \tag{14}$$

and also

$$c = Var[E(X^2|\Theta)], \tag{15}$$

and

$$h = E[Var(X^2|\Theta)] = E[E(X^4|\Theta) - E(X^2|\Theta)^2]. \tag{16}$$

2 The Parametric Poisson Case

We now examine a few cases where parametric expressions are postulated for Θ and for X given Θ. In this context, we first derive expressions for the structural parameters in the conditional Poisson setting.

Proposition 2.1 (Parameters in the conditional Poisson case)**.** *We assume that X conditional on Θ is Poisson distributed. We recall that $\mu = v = E(\Theta)$ and $a = Var(\Theta)$ in classic credibility theory. The quantities b, g, c, and h can be written as functions of the moments of Θ as follows:*

$$b = a + E(\Theta^3) - E(\Theta^2)\ E(\Theta),$$

and

$$g = E(\Theta) + 2E(\Theta^2), \tag{17}$$

and also

$$c = 2b - a + Var(\Theta^2),$$

and

$$h = E(\Theta) + 6E(\Theta^2) + 4E(\Theta^3). \tag{18}$$

When the distribution of Θ is gamma, q-credibility reduces to standard credibility. Indeed, the classic credibility premium coincides with the Bayesian premium in the Poisson-gamma case. This means that it is not possible to further reduce the mean square error and therefore the q-credibility predictor can only be equal to the classic credibility predictor. We are in a situation of *exact q-credibility* that we summarize in the next proposition.

Proposition 2.2 (q-Credibility in the Poisson-gamma case)**.** *In the Poisson-gamma case, q-credibility reduces to classic credibility, with*

$$y_q = 0,$$

$$z_q = z,$$

and

$$\alpha_{0,q} = \alpha_0 = \mu\ (1 - z),$$

and the q-credibility predictor, similar to the credibility predictor, is equal to the Bayesian predictor.

When we assume that the distribution of Θ is Pareto, so when we use $E(\Theta^k) = \frac{\eta\chi^k}{\eta-k}$ that is valid as long as $\eta > k$, we obtain:

Proposition 2.3 (Parameters in the Pareto case)**.** *Let Θ be Pareto-distributed*[1] *with parameters (η, χ), under the restriction $\eta > 4$. In this context, it is already known that $\mu = v = \frac{\eta\chi}{\eta-1}$ and $a = \frac{\eta\chi^2}{\eta-2} - \left(\frac{\eta\chi}{\eta-1}\right)^2$. The quantities b, g, c, and h can be written as follows:*

$$b = a + \frac{\eta\chi^3}{\eta-3} - \left(\frac{\eta\chi^2}{\eta-2}\right)\left(\frac{\eta\chi}{\eta-1}\right), \tag{19}$$

and

$$g = \frac{\eta\chi}{\eta-1} + 2\frac{\eta\chi^2}{\eta-2}, \tag{20}$$

and also

$$c = 2b - a + \frac{\eta\chi^4}{\eta-4} - \left(\frac{\eta\chi^2}{\eta-2}\right)^2, \tag{21}$$

and

$$h = \frac{\eta\chi}{\eta-1} + 6\frac{\eta\chi^2}{\eta-2} + 4\frac{\eta\chi^3}{\eta-3}. \tag{22}$$

We now construct an example where the parameters of the Pareto distribution are $\eta = 5$ and $\chi = 4$ and we assume that 5 claims have been observed in the past $n = 2$ years. In this example, the average number of observations is $\bar{X} = \frac{5}{2} = 2.5$.

To compute the quantity $\overline{X^2}$, we need to know how the 5 claims were distributed between the 2 years. There are three possible scenarios: 3 claims in one year and 2 claims in the other year, 4 claims in one year and 1 claim in the other year, and 5 claims in one year and 0 claim in the other year. The order in which the numbers of claims are observed is not relevant. In the first case, $\overline{X^2} = \frac{3^2+2^2}{2} = 6.5$. In the second case, $\overline{X^2} = \frac{4^2+1^2}{2} = 8.5$. Finally, in the third case, $\overline{X^2} = \frac{5^2+0^2}{2} = 12.5$.

According to the classic credibility theory,

$$\mu = v = 5,$$

1. We use the following density for the Pareto distribution:

$$f_\Theta(x) = \frac{\eta\,\chi^\eta}{x^{\eta+1}}\,\mathbb{1}_{x>\chi}.$$

and

$$a = \frac{5}{3}.$$

Therefore,

$$k = \frac{v}{a} = 3,$$

and

$$z = \frac{n}{n+k} = \frac{2}{2+3} = \frac{2}{5}.$$

The expected number of claims for the coming period is given by

$$\hat{P} = z\ \bar{X} + (1-z)\ \mu = \frac{2}{5}\ \frac{5}{2} + \frac{3}{5}\ 5 = 4.$$

Not surprisingly, this value is comprised between the empirical mean $\bar{X} = 2.5$ and the theoretical mean $\mu = 5$. The mean square error is in the classic setting:

$$\text{MSE}'_c = 1.$$

To compute the q-credibility estimator, we start by computing Eqs (19) to (22). We obtain:

$$g = \frac{175}{3} \approx 58.33, \qquad b = \frac{85}{3} \approx 28.33, \qquad c = \frac{5615}{9} \approx 623.88, \qquad h = 805.$$

Then, we have:

$$z_q = \frac{11}{131} \approx 0.083969,$$

and

$$y_q = \frac{3}{131} \approx 0.022901,$$

so,

$$\alpha^*_{0,q} = \frac{505}{131} \approx 3.85496.$$

In Table 1, we show $\hat{P}_q = \alpha^*_{0,q} + z_q\ \bar{X} + y_q\ \overline{X^2}$ for each of the three possible scenarios for $\overline{X^2}$.

Number of claims distrib.	(3,2)	(4,1)	(5,0)
$\overline{X^2}$	6.5	8.5	12.5
$\hat{P}_q$	4.2137	4.2595	4.3511

Table 1 – q-Credibility Estimates

We observe from the table that the more the number of claims is irregular with years, the greater $\overline{X^2}$ and the greater the correction to classic credibility theory made by q-credibility. Also note that the repartition of claims along years, *ceteris paribus*, is a feature that cannot be taken into account by classic credibility theory while quadratic credibility measures this effect in the premiums it produces. However, although q-credibility can capture irregularities, it cannot capture trends.

The mean square error is in the quadratic setting:

$$\text{MSE}'_q = \frac{115}{131} \approx 0.8779.$$

Therefore, the following relative reduction in the error is observed in this experiment:

$$\kappa' = \frac{16}{131} \approx 12.21\%.$$

Let us now examine what q-credibility means in a semi-parametric case.

3 The Semi-Parametric Case

The semi-parametric approach to credibility corresponds to a situation where the distribution of a number of claims X conditionally on Θ is known. However, neither the distribution of Θ nor the unconditional distribution of X are known.

Assume we observed M insured during a particular year. During that year, X_i is the number of insureds for which i claims occurred. We can estimate the average number of claims as follows:

$$\hat{\mu} = \sum_{i=0}^{+\infty} i \, \frac{X_i}{\sum\limits_{j=0}^{+\infty} X_j} = \frac{1}{M} \sum_{i=0}^{+\infty} i \; X_i,$$

where we note that $M = \sum\limits_{j=0}^{+\infty} X_j$.

Because we are in a conditional Poisson setting, we readily have:

$$v = \mu$$

by taking the expectation of

$$\Theta = E(X|\Theta) = \text{Var}(X|\Theta).$$

Using the unbiased estimator of the variance of X, which is equal to $\hat{a}+\hat{v}$, we can write the classic formula for the estimator of a:

$$\hat{a} = \frac{\sum_{i=0}^{+\infty}(i-\hat{\mu})^2\ X_i}{\sum_{i=0}^{+\infty} X_i - 1} - \hat{v} = \frac{1}{M-1}\sum_{i=0}^{+\infty}(i-\hat{\mu})^2\ X_i - \hat{v}.$$

The next proposition gives the q-credibility semi-parametric estimators.

Proposition 3.1 (Semi-parametric estimators)**.** *Under the conditional Poisson assumption, we have:*

$$\hat{g} = \frac{1}{M}\sum_{i=0}^{+\infty}(2i^2-i)\ X_i,$$

and

$$\hat{b} = \frac{1}{M-1}\sum_{i=0}^{+\infty}\left(i^2 - \frac{1}{M}\sum_{j=0}^{+\infty} j^2\ X_j\right)\left(i - \frac{1}{M}\sum_{j=0}^{+\infty} j\ X_j\right)\ X_i - \hat{g},$$

and also

$$\hat{h} = \frac{1}{M}\sum_{i=0}^{+\infty}(4i^3-6i^2+3i)\ X_i,$$

and

$$\hat{c} = \frac{1}{M-1}\sum_{i=0}^{+\infty}\left(i^2 - \frac{1}{M}\sum_{j=0}^{+\infty} j^2 X_j\right)^2\ X_i - \hat{h}.$$

Let us now come to an illustration of these results. Assume we observed the data given in Table 2. This table expresses that 560 insureds incurred no claim in the past period, 134 insureds incurred one claim in the past period, and so on. We want to compute the expected future number of claims for an insured who incurred i claims in the past period.

i	0	1	2	3
X_i	560	134	14	2

Table 2 – Dataset

In this example, no insured incurred more that three claims. We have:

$$M = \sum_{i=0}^{3} X_i = 710.$$

We can compute

$$\hat{\mu} = \hat{v} = 0.2366, \qquad \hat{a} = 0.0006834,$$

Using the classic credibility formulas (for $n = 1$ year of observations), we have:

$$k = \frac{\hat{v}}{\hat{a}} = 346.26, \qquad z = \frac{1}{1+k} = 0.0029.$$

According to classic credibility theory, we can compute the expected future number of claims for an insured who incurred i claims as follows:

$$\hat{P}(i) = z\ i + (1-z)\ \hat{\mu} \qquad 0 \leq i \leq 3,$$

which yields

$$\hat{P} = [0.2359 \quad 0.2388 \quad 0.2417 \quad 0.2446],$$

where the observation and prediction periods are of the same length.

We now compute the q-credibility estimators given in Proposition 3.1. We obtain

$$\hat{b} = 0.0044, \qquad \hat{g} = 0.3493, \qquad \hat{c} = 0.0052, \qquad \hat{h} = 0.6423.$$

Then, we have, using Eqs (9) to (11):

$$z_q = -0.0393, \qquad y_q = 0.0283, \qquad \alpha^*_{0,q} = 0.2376.$$

Based on Eq. (8), we compute

$$\hat{P}_q(i) = \alpha^*_{0,q} + z_q\ i + y_q\ i^2 \qquad 0 \leq i \leq 3,$$

because i and i^2 represent the first- and second-order empirical non-centered moments over one period for each line of the dataset considered. We obtain the q-credibility estimates

$$\hat{P}_q = [0.2376 \quad 0.2266 \quad 0.2722 \quad 0.3743].$$

Using the formulas of Proposition 1.6, we find that the mean square error is in the classic setting:

$$\mathrm{MSE}'_c = 0.000681,$$

while we have in the quadratic setting:

$$\text{MSE}'_q = 0.000585.$$

Therefore, the following relative reduction in the error is observed in this experiment:

$$\kappa' = 14.1\%.$$

Note that the illustration of the semi-parametric case that we conduct here, where we compare the quadratic situation to the classic situation well-known of SOA and CAS actuaries (see for instance the book of Klugman et al. (2012)) is not devoid of drawbacks. For instance, the grouping of policies per number of claims measured per year may lead to the construction of inconsistent classes. We leave to another publication the development of other illustrations of the semi-parametric framework.

4 The Non-Parametric Case

We first give the main results obtained in a non-parametric setting for the structural parameters and then we provide an illustration of these results.

In classic credibility theory, the estimator of expected hypothetical means is

$$\hat{\mu} = \frac{1}{rn} \sum_{i=1}^{r} \sum_{j=1}^{n} X_{ij},$$

where the claims X_{ij} are doubly indexed to reflect the fact that we now consider r policyholders over n periods. The estimator of expected process variance is

$$\hat{v} = \frac{1}{r(n-1)} \sum_{i=1}^{r} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2,$$

where $\bar{X}_i = \frac{1}{n} \sum\limits_{j=1}^{n} X_{ij}$ is the empirical mean of past observations for insured i. Then, the estimator of the variance of hypothetical means is

$$\hat{a} = \frac{1}{r-1} \sum_{i=1}^{r} (\bar{X}_i - \bar{X})^2 - \frac{\hat{v}}{n},$$

where $\bar{X}$ is the empirical mean of past observations for all insureds, which is equal to $\hat{\mu}$. We obtain similar estimators for q-credibility parameters in the following proposition.

Proposition 4.1 (Non-parametric estimators)**.** *The non-parametric estimators for the quantities h, c, g, and b are given as follows.*

$$\hat{h} = \frac{1}{r(n-1)} \sum_{i=1}^{r} \sum_{j=1}^{n} \left(X_{ij}^2 - \overline{X_i^2} \right)^2, \tag{23}$$

where $\overline{X_i^2} = \frac{1}{n} \sum_{j=1}^{n} X_{ij}^2$ *is the empirical mean of past squared observations for a given insured i. Then,*

$$\hat{c} = \frac{1}{r-1} \sum_{i=1}^{r} \left(\overline{X_i^2} - \overline{X^2} \right)^2 - \frac{\hat{h}}{n}, \tag{24}$$

where

$$\overline{X^2} = \frac{1}{rn} \sum_{i=1}^{r} \sum_{j=1}^{n} X_{ij}^2$$

is the empirical mean of past squared observations for all insureds. Next,

$$\hat{g} = \frac{1}{r(n-1)} \sum_{i=1}^{r} \sum_{j=1}^{n} (X_{ij}^2 - \overline{X_i^2})(X_{ij} - \bar{X}_i), \tag{25}$$

and

$$\hat{b} = \frac{1}{r-1} \sum_{i=1}^{r} (\overline{X_i^2} - \overline{X^2})(\bar{X}_i - \bar{X}) - \frac{\hat{g}}{n}. \tag{26}$$

Let us now study the use of these estimators via a simple example. Assume $r = n = 3$ and we have the following data:

$$\mathbf{X} = \begin{pmatrix} 1 & 2 & 6 \\ 1 & 10 & 13 \\ 1 & 1 & 1 \end{pmatrix},$$

where each line is for one insured and gives three consecutive numbers of observations.

We compute $\left(\bar{X}_1 = 3, \bar{X}_2 = 8, \bar{X}_3 = 1\right)$ and $\hat{\mu} = \bar{X} = 4$. Then, according to classic credibility theory, $\hat{v} = \frac{46}{3}$ and $\hat{a} = \frac{71}{9}$. We deduce $k = \frac{138}{71}$ and $z = \frac{3}{3+k} = \frac{213}{351}$.

The expected number of claims in the next period for the first insured is, according to classic credibility theory,

$$\hat{P}_1 = z\ \bar{X}_1 + (1-z)\ \mu = \frac{1,191}{351} \approx 3.3932.$$

Similarly, we have for the second insured:

$$\hat{P}_2 = z\ \bar{X}_2 + (1-z)\ \mu = \frac{2,256}{351} \approx 6.4274,$$

and for the third insured:

$$\hat{P}_3 = z\ \bar{X}_3 + (1-z)\ \mu = \frac{765}{351} \approx 2.1795.$$

Next, we turn to the q-credibility approach. For simplicity, we denote $X^2 = X \circ X$ the element-wise product of X with itself. We have:

$$\mathbf{X^2} = \begin{pmatrix} 1 & 4 & 36 \\ 1 & 100 & 169 \\ 1 & 1 & 1 \end{pmatrix}.$$

Therefore, $\left(\overline{X_1^2} = \frac{41}{3}, \overline{X_2^2} = 90, \overline{X_3^2} = 1\right)$ and $\overline{X^2} = \frac{314}{9}$. The q-credibility parameters are estimated using Eqs (23) to (26). We obtain

$$\hat{h} = \frac{22,522}{9}, \qquad \hat{c} = \frac{13,355}{9}, \qquad \hat{g} = 190, \qquad \hat{b} = \frac{325}{3}.$$

Then, we have, using Eqs (9) to (11):

$$z_q = -\frac{18,862}{40,401}, \qquad y_q = \frac{365}{4,489}, \qquad \alpha^*_{0,q} = \frac{12,023}{4,489}.$$

The expected number of claims in the next period for the first insured is, according to q-credibility theory,

$$\hat{P}_{q,1} = \alpha^*_{0,q} + z_q\ \bar{X}_1 + y_q\ \overline{X_1^2} = \frac{10,724}{4,489} \approx 2.3890.$$

Similarly, we have for the second insured:

$$\hat{P}_{q,2} = \alpha^*_{0,q} + z_q\ \bar{X}_2 + y_q\ \overline{X_2^2} = \frac{252,961}{40,401} \approx 6.2613,$$

and for the third insured:

$$\hat{P}_{q,3} = \alpha^*_{0,q} + z_q\ \bar{X}_3 + y_q\ \overline{X_3^2} = \frac{92,630}{40,401} \approx 2.2928.$$

The relative changes $\left(\frac{\hat{P}_{q,i}-\hat{P}_i}{\hat{P}_i}\right)_{i=1:3}$ induced by the quadratic correction are respectively -29.6%, -2.58%, and 5.2%. They are not negligible and can be of any sign.

In the classic setting, we find that the mean square error is

$$\mathrm{MSE}'_c = 3.1016,$$

while in the quadratic setting we have:

$$\mathrm{MSE}'_q = 2.7634.$$

Therefore, the following relative reduction in the error observed in this experiment is

$$\kappa' = 10.9\%.$$

Conclusion

In this article, we have examined the effect of adding a quadratic correction to credibility theory. We have shown how the parametric, semi-parametric and non-parametric settings can be extended to incorporate this correction. The three settings have been illustrated and we have found a decent decrease of about 10% in the mean square error in each case.

At this stage, three types of extensions could be devised. First, it could be possible to conduct a study on exact q-credibility by introducing quadratic exponential functions, so by enlarging the linex paradigm. Then, for all of our analysis, we have considered uniform exposures: it could be interesting to develop q-credibility in a non-uniform setting in a distinct paper. Finally, adding parameters in a system does not go without increasing the uncertainty of this system. It could be interesting to examine in a further study the tradeoff between the precision added by moving to q-credibility and the cost of estimating an increased number of structural parameters. For this purpose, simulations could be conducted and the comparison of training and test MSEs could be performed (see, e.g., James et alii (2017)).

Acknowledgments and a final remark

The author wishes to thank referees and Daniel Bauer, Michel Dacorogna, Abdou Kelani, Zinoviy Landsman, Ragnar Norberg, Li Shen, and Xia Xu for their insightful comments. The proof of the first proposition of this article is long *on purpose*, to show to a broad readership how the q-credibility estimator is constructed. It is of course possible to derive an existence and uniqueness result in a shorter way, using a Hilbert space Pythagorean lemma as in Remark 1.7 dealing with the mean square error.

References

Bailey, A. (1945): "A Generalized Theory of Credibility," *Proceedings of the Casualty Actuarial Society*, 37, 13–20.

——— (1950): "Credibility Procedures," *Proceedings of the Casualty Actuarial Society*, 37, 7–23, 94–115.

Bühlmann, H. (1967): "Experience Rating and Credibility," *ASTIN Bulletin*, 4(3), 199–207.

——— (1970): *Mathematical Methods in Risk Theory*. Springer, Berlin.

Bühlmann, H., and A. Gisler (2005): *A Course in Credibility Theory and its Applications*. Universitext, Springer, Berlin.

Bühlmann, H., and E. Straub (1970): "Glaubwürdigkeit für Schadensätze (Credibility for Loss Ratios)," *Mitteilungen der Vereiningun Schweizerischer Versicherungs - Mathematiker*, 70, 111–133.

De Vylder, F. (1978): "Parameter Estimation in Credibility Theory," *ASTIN Bulletin*, 10, 99–112.

——— (1985): "Non-Linear Regression in Credibility Theory," *Insurance: Mathematics and Economics*, 4, 163–172.

Hachemeister, C. A. (1975): "Credibility for Regression Models with Application to Trend," in *Credibility: Theory and Applications*, ed. by P. M. Kahn, pp. 129–169. Academic Press, New York.

Herzog, T. N. (1999): *Credibility Theory*. ACTEX Publications, Third Edition.

Hong, L., and R. Martin (2017): "A Flexible Bayesian Nonparametric Model for Predicting Future Insurance Claims," *North American Actuarial Journal*, 21(2), 228–241.

James, G., D. Witten, T. Hastie, and R. Tibshirani (2017): *An Introduction to Statistical Learning with Applications in R.* Springer.

Klugman, S. A., H. H. Panjer, and G. E. Willmot (2012): *Loss Models. From Data to Decisions.* Wiley, Fourth Edition.

Lai, T. L. (2012): "Credit Portfolios, Credibility Theory, and Dynamic Empirical Bayes," *International Scholarly Research Network: Probability and Statistics.*

Longley-Cook, L. (1962): "An Introduction to Credibility Theory," *Proceedings of the Casualty Actuarial Society*, 49, 194–221.

Mayerson, A. L. (1964): "The Uses of Credibility in Property Insurance Ratemaking," *Giornale dell'Istituto Italiano degli Attuari*, 27, 197–218.

Mowbray, A. H. (1914): "How Extensive a Payroll Exposure Is Necessary to Give a Dependable Pure Premium," *Proceedings of the Casualty Actuarial Society*, 1, 24–30.

Neuhaus, W. (1985): "Choice of Statistics in Linear Bayes Estimation," *Scandinavian Actuarial Journal*, pp. 1–26.

Norberg, R. (1979): "The Credibility Approach to Ratemaking," *Scandinavian Actuarial Journal*, pp. 181–221.

——— (1982): "On Optimal Parameter Estimation in Credibility," *Insurance: Mathematics and Economics*, 1, 73–89.

——— (2015): "Credibility Theory," *Wiley StatsRef: Statistics Reference Online.*

Shiu, E. S., and F. Y. Sing (2004): "Credibility Theory and Geometry," *Journal of Actuarial Practice*, 11, 197–216.

Taylor, G. C. (1977): "Abstract Credibility," *Scandinavian Actuarial Journal*, pp. 149–168.

Weishaus, A. (2015): *ASM Study Manual Exam C/Exam 4, 17th Edition.* Actex Publications.

Whitney, A. W. (1918): "The Theory of Experience Rating," *Proceedings of the Casualty Actuarial Society*, 4, 274–292.

Zehnwirth, B. (1977): "The Mean Credibility Formula Is a Bayes Rule," *Scandinavian Actuarial Journal*, pp. 212–216.

Appendix

Proof of Proposition 1.1

The goal is to minimize the Mean Square Error function:

$$f = E\left(\left[\alpha_{0,q} + \sum_{i=1}^{n}\alpha_i X_i + \sum_{i=1}^{n}\beta_i X_i^2 - X_{n+1}\right]^2\right).$$

We first set the derivative of f with respect to $\alpha_{0,q}$ equal to 0:

$$\frac{\partial f}{\partial \alpha_{0,q}} = 0 = 2\cdot E\left(\alpha_{0,q}^* + \sum_{i=1}^{n}\alpha_i^* X_i + \sum_{i=1}^{n}\beta_i^* X_i^2 - X_{n+1}\right), \tag{27}$$

and we obtain

$$\alpha_{0,q}^* + \sum_{i=1}^{n}\alpha_i^* E(X_i) + \sum_{i=1}^{n}\beta_i^* E(X_i^2) = E(X_{n+1}). \tag{28}$$

Then, we set the derivative of f with respect to each α_k equal to 0:

$$\frac{\partial f}{\partial \alpha_k} = 0 = 2\cdot E\left(X_k\left[\alpha_{0,q}^* + \sum_{i=1}^{n}\alpha_i^* X_i + \sum_{i=1}^{n}\beta_i^* X_i^2 - X_{n+1}\right]\right), \qquad \forall k = 1:n,$$

and we obtain

$$\alpha_{0,q}^* E(X_k) + \sum_{i=1}^{n}\alpha_i^* E(X_i X_k) + \sum_{i=1}^{n}\beta_i^* E(X_i^2 X_k) = E(X_{n+1}X_k), \qquad \forall k = 1:n. \tag{29}$$

Subtracting $E(X_k)$ times Eq. (28) from Eq. (29), we have:

$$\sum_{i=1}^{n}\alpha_i^*\left[E(X_i X_k) - E(X_i)E(X_k)\right] + \sum_{i=1}^{n}\beta_i^*\left[E(X_i^2 X_k) - E(X_i^2)E(X_k)\right]$$
$$= E(X_{n+1}X_k) - E(X_{n+1})E(X_k), \qquad \forall k = 1:n,$$

or

$$\sum_{i=1}^{n}\alpha_i^* \operatorname{Cov}(X_i, X_k) + \sum_{i=1}^{n}\beta_i^* \operatorname{Cov}(X_i^2, X_k) = \operatorname{Cov}(X_{n+1}, X_k), \qquad \forall k = 1:n. \tag{30}$$

Finally, we set the derivative of f with respect to each β_k equal to 0:

$$\frac{\partial f}{\partial \beta_k} = 0 = 2 \cdot E\left(X_k^2 \left[\alpha_{0,q}^* + \sum_{i=1}^{n} \alpha_i^* X_i + \sum_{i=1}^{n} \beta_i^* X_i^2 - X_{n+1} \right] \right), \qquad \forall k = 1:n,$$

and we obtain

$$\alpha_{0,q}^* E(X_k^2) + \sum_{i=1}^{n} \alpha_i^* E(X_i X_k^2) + \sum_{i=1}^{n} \beta_i^* E(X_i^2 X_k^2) = E(X_{n+1} X_k^2), \qquad \forall k = 1:n. \tag{31}$$

Subtracting $E(X_k^2)$ times Eq. (28) from Eq. (31), we have:

$$\sum_{i=1}^{n} \alpha_i^* \left[E(X_i X_k^2) - E(X_i) E(X_k^2) \right] + \sum_{i=1}^{n} \beta_i^* \left[E(X_i^2 X_k^2) - E(X_i^2) E(X_k^2) \right]$$
$$= E(X_{n+1} X_k^2) - E(X_{n+1}) E(X_k^2), \qquad \forall k = 1:n,$$

or

$$\sum_{i=1}^{n} \alpha_i^* \operatorname{Cov}(X_i, X_k^2) + \sum_{i=1}^{n} \beta_i^* \operatorname{Cov}(X_i^2, X_k^2) = \operatorname{Cov}(X_{n+1}, X_k^2), \qquad \forall k = 1:n. \tag{32}$$

Recall that the X_i's are assumed identically distributed. Therefore, the index 'i' does not play any role and we have: $\forall i = 1:n,\ \alpha_i^* = \alpha^*$ and $\forall i = 1:n,\ \beta_i^* = \beta^*$. Then, Eqs (30) and (32) become

$$\alpha^*(na + v) + \beta^*(nb + g) = a, \tag{33}$$

and

$$\alpha^*(nb + g) + \beta^*(nc + h) = b. \tag{34}$$

Solving Eqs (33) and (34), we obtain

$$\alpha^* = \frac{a(nc + h) - b(nb + g)}{(na + v)(nc + h) - (nb + g)^2},$$

and

$$\beta^* = \frac{bv - ag}{(na + v)(nc + h) - (nb + g)^2}.$$

Next, Eq. (28) can be rewritten as follows:

$$\alpha_{0,q}^* + n\alpha^* \mu + n\beta^*(\mu^2 + a + v) = \mu.$$

because

$$E\left(X_i^2\right) = E\left(X_i\right)^2 + \text{Var}(X_i) = \mu^2 + a + v$$

thanks to Eq. (1).

Introducing

$$z_q = n\alpha^*,$$

and

$$y_q = n\beta^*,$$

we obtain

$$\alpha^*_{0,q} = \mu(1 - z_q) - y_q(\mu^2 + a + v).$$

These are Eqs (9) to (11). Then, we want to estimate X_{n+1} using the past realizations $\{\hat{X}_i\}_{i=1:n}$ of the $\{X_i\}_{i=1:n}$. Using

$$\alpha^*_{0,q} + \sum_{i=1}^{n} \alpha^*_i \hat{X}_i + \sum_{i=1}^{n} \beta^*_i \hat{X}_i^2$$

that we rewrite as

$$\alpha^*_{0,q} + n\alpha^* \, \frac{\sum_{i=1}^{n} \hat{X}_i}{n} + n\beta^* \, \frac{\sum_{i=1}^{n} \hat{X}_i^2}{n},$$

we obtain Eq. (8).

Proof of Proposition 1.6

We need to compute

$$\text{MSE}_q = E\left(\left[\hat{P}^q_{n+1} - X_{n+1}\right]^2\right) = \text{Var}\left(\hat{P}^q_{n+1} - X_{n+1}\right)$$

or

$$\text{MSE}_q = \text{Var}\left(\hat{P}^q_{n+1}\right) + \text{Var}\left(X_{n+1}\right) - 2\ \text{Covar}\left(\hat{P}^q_{n+1}, X_{n+1}\right).$$

We already know that

$$\text{Var}\left(X_{n+1}\right) = v + a$$

and we compute

$$\text{Var}\left(\hat{P}^q_{n+1}\right) = z_q^2\ \text{Var}\left(\overline{X}\right) + y_q^2\ \text{Var}\left(\overline{X^2}\right) + 2\ z_q\ y_q\ \text{Covar}\ \left(\overline{X}, \overline{X^2}\right).$$

Elementary computations yield

$$\text{Var}\left(\overline{X}\right) = a + \frac{v}{n},$$

$$\text{Var}\left(\overline{X^2}\right) = c + \frac{h}{n},$$

and

$$\text{Covar}\ \left(\overline{X}, \overline{X^2}\right) = b + \frac{g}{n},$$

so that

$$\text{Var}\left(\hat{P}^q_{n+1}\right) = z_q^2\ \left(a + \frac{v}{n}\right) + y_q^2\ \left(c + \frac{h}{n}\right) + 2\ z_q\ y_q\ \left(b + \frac{g}{n}\right).$$

Because Eqs (33) and (34) can be rewritten as follows:

$$z_q\ \left(a + \frac{v}{n}\right) + y_q\ (b + \frac{g}{n}) = a$$

and

$$z_q\ \left(b + \frac{g}{n}\right) + y_q\ (c + \frac{h}{n}) = b,$$

we have:

$$\text{Var}\left(\hat{P}^q_{n+1}\right) = z_q\ a + y_q\ b.$$

Finally, we can compute:

$$\text{Covar}\left(\hat{P}^q_{n+1}, X_{n+1}\right) = \text{Covar}\left(z_q\ \overline{X} + y_q\ \overline{X^2}, X_{n+1}\right) = z_q\ a + y_q\ b.$$

Recombining the above results, we obtain the result of the proposition.

Proof of Proposition 1.11

For $X \neq Y$,

$$b = \text{Cov}(X^2, Y) = E[\text{Cov}(X^2, Y|\Theta)] + \text{Cov}[E(X^2|\Theta), E(Y|\Theta)].$$

Then, by conditional independence of X^2 and Y,

$$b = \text{Cov}[E(X^2|\Theta), E(Y|\Theta)] = E[E(X^2|\Theta)\ E(Y|\Theta)] - E[E(X^2|\Theta)]E[E(Y|\Theta)],$$

and then, because X and Y are identically distributed,

$$b = \text{Cov}[E(X^2|\Theta), E(X|\Theta)] = E[E(X^2|\Theta)\ E(X|\Theta)] - E[E(X^2|\Theta)]E[E(X|\Theta)].$$

In

$$\text{Cov}(X^2, X) = E[\text{Cov}(X^2, X|\Theta)] + \text{Cov}[E(X^2|\Theta), E(X|\Theta)],$$

we can identify

$$g = E[\text{Cov}(X^2, X|\Theta)] = E[E(X^3|\Theta)] - E[E(X^2|\Theta)\ E(X|\Theta)].$$

The derivation of the expressions of c and h is similar to that of a and v, replacing X by X^2.

Proof of Proposition 2.1

We use the fact that $E(X|\Theta) = \Theta$, $E(X^2|\Theta) = \Theta + \Theta^2$, $E(X^3|\Theta) = \Theta + 3\Theta^2 + \Theta^3$, and $E(X^4|\Theta) = \Theta + 7\Theta^2 + 6\Theta^3 + \Theta^4$. Note that

$$b = E[E(X^2|\Theta)\ E(X|\Theta)] - E[E(X^2|\Theta)]E[E(X|\Theta)] = E[(\Theta+\Theta^2)\ \Theta] - E(\Theta+\Theta^2)E(\Theta),$$

and

$$b = E(\Theta^2) + E(\Theta^3) - E(\Theta)^2 - E(\Theta^2)E(\Theta) = \text{Var}(\Theta) + E(\Theta^3) - E(\Theta^2)\ E(\Theta).$$

Then, we use

$$g = E[E(X^3|\Theta)] - E[E(X^2|\Theta)\ E(X|\Theta)] = E(\Theta + 3\Theta^2 + \Theta^3) - E((\Theta + \Theta^2)\Theta).$$

Next, we have:

$$c = \text{Var}[E(X^2|\Theta)] = \text{Var}(\Theta + \Theta^2) = \text{Var}(\Theta) + \text{Var}(\Theta^2) + 2\ \text{Cov}(\Theta, \Theta^2),$$

and

$$c = \text{Var}(\Theta) + \text{Var}(\Theta^2) + 2\ \left[E(\Theta^3) - E(\Theta^2)\ E(\Theta)\right] = b + \text{Var}(\Theta^2) + E(\Theta^3) - E(\Theta^2)\ E(\Theta),$$

and finally

$$h = E[E(X^4|\Theta) - E(X^2|\Theta)^2] = E(\Theta + 7\Theta^2 + 6\Theta^3 + \Theta^4 - (\Theta + \Theta^2)^2).$$

Proof of Proposition 2.2

Because

$$bv - ag = \chi gv - ag = g(\chi v - a) = g(\chi^2\eta - \chi^2\eta) = 0$$

we have from Eq. (11) that $y_q = 0$. A few lines of computation give $z_q = \frac{n\chi}{n\chi+1} = z$. The expression of $\alpha_{0,q}$ follows.

Proof of Proposition 3.1

We have:

$$E(X) = E(E(X|\Theta)) = E(\Theta)$$

Then,

$$\mathrm{Var}(X) = E(X^2) - E(X)^2 = E(\mathrm{Var}(X|\Theta)) + \mathrm{Var}(E(X|\Theta)) = E(\Theta) + \mathrm{Var}(\Theta)$$

yields

$$E(X^2) - E(X)^2 = E(\Theta) + E(\Theta^2) - E(\Theta)^2$$

so that

$$E(\Theta^2) = E(X^2) - E(X). \tag{35}$$

Using similar arguments, we obtain:

$$E(\Theta^3) = E(X^3) - 3E(X^2) + 2E(X). \tag{36}$$

Replacing Eqs (35) and (36) in Eqs (17) and (18) and then taking the unbiased estimates of the non-centered moments gives the expressions of $\hat{g}$ and $\hat{h}$. For the computation of $\hat{b}$, we recall that

$$b + g = \mathrm{Cov}(X^2, X),$$

and we take the unbiased estimator of this covariance term. Similarly for the computation of $\hat{c}$, we have:

$$c + h = \mathrm{Var}(X^2),$$

and we take the unbiased estimate of this variance term.

Proof of Proposition 4.1

The estimators $\hat{h}$ and $\hat{c}$ can be derived in the same way as $\hat{v}$ and $\hat{a}$, replacing observations by squared observations. From Proposition 1.11, we have:

$$g = E[\mathrm{Cov}(X^2, X|\Theta)],$$

The inner sum in Eq. (25) is obtained by taking the unbiased estimator of the covariance in the above formula for g. Then, the outer sum in (25) is derived using the unbiased estimator of the mean in the expression of g. Next, to compute $\hat{b}$, we start by computing

$$\mathrm{Cov}\left(\overline{X_i^2}, \overline{X_i}\right) = \mathrm{Cov}\left(E(X_{i,j}^2|\Theta_i), E(X_{i,j}|\Theta_i)\right) + \frac{1}{n}E\left(\mathrm{Cov}\left(X_{i,j}^2, X_{i,j}|\Theta_i\right)\right)$$

where we have used the conditional independence of $X_{i,k}^2$ and $X_{i,j}$ knowing Θ_i. Then, we recognize in the above formula:

$$\mathrm{Cov}\left(\overline{X_i^2}, \overline{X_i}\right) = \hat{b} + \frac{\hat{g}}{n}$$

and the expression of $\hat{b}$ follows by computing the unbiased estimator of $\mathrm{Cov}\left(\overline{X_i^2}, \overline{X_i}\right)$.