Technische Universität Dresden
Institut für Mathematische Stochastik

Dresdner Schriften zur
Versicherungsmathematik
3/2011

**Klaus Th. Hess**

# Maximum–Likelihood and Marginal–Sum Estimation in a Collective Model

**Herausgeber: Die Professoren des Instituts für Mathematische Stochastik**
**ISSN 0946–4727**

# Maximum–Likelihood and Marginal–Sum Estimation in a Collective Model

Klaus Th. Hess

Institut für Mathematik

Universität Rostock

**Abstract**

In the present paper we consider collective models in risk theory and their thinning and decomposition. We give three applications to reinsurance, multiplicative tariffs, and loss reserving. For each of these applications we show how maximum–likelihood and marginal–sum estimation can be used to estimate the parameters.

## 1 Introduction

We consider a portfolio of risks in a particular line of insurance. In order to calculate premiums for future insurance periods (e.g. years) we want to model the total claim amount per insurance period. Based on this model we can predict the total claim amount and hence the premium for a future period.

In actuarial mathematics there are two basic models: the individual model and the collective model.

The *individual model* is a pair

$$\langle n, \{Z_j\}_{j\in\{1,\dots,n\}}\rangle$$

where $n$ is a positive integer and the $Z_j$'s are independent random variables. We interpret $n$ as the number of risks of the portfolio and $Z_j$ as the claim amount of risk $j$ during the next insurance period. The *total claim amount* of the portfolio is given by

$$\sum_{j=1}^{n} Z_j$$

The assumption that the risks are independent is appropriate when there is no interdependence between the risks. However, estimation of the distributions of the individual claim amounts causes a serious problem at least for those risks for which no or only very little claims experience is available from the past.

If we additionally assume that the distribution of the claim amount is identical for all risks in the portfolio, then we get the *individual model for a homogeneous portfolio*. From the statistical point of view, a homogenous portfolio presents an i.i.d. sample of size $n$ and there is a rich theory for this model. Unfortunately, however, insurance portfolios tend to be non–homogeneous such that in most cases the individual model for a homogeneous portfolio is inappropriate.

Because of these reasons, the collective model, which refers to the single claims instead of the single risks, is the more common model in actuarial mathematics. The pair

$$\langle N, \{X_j\}_{j \in \mathbb{N}} \rangle$$

is called a *collective model* if $N$ is a random variable with $P[N \in \mathbb{N}_0] = 1]$ and $\{X_j\}_{j \in \mathbb{N}}$ is a sequence of random variables which is i.i.d. and independent of $N$. We interpret $N$ as the random number of claims caused by the risks of the portfolio during an insurance period and $X_j$ as the claim amount of the $j$th claim. From the statistical point of view, the collective model is an i.i.d. sample with random sample size $N$.

Under the collective point of view, the individual risk causing a particular claim is neglected such that the possible lack of inhomogeneity of the portfolio becomes irrelevant. For this reason, the assumption of i.i.d. claim amounts in the collective model is much more realistic than the assumption of i.i.d. risks in the individual model.

In the collective model the *total claim amount* of the portfolio is given by

$$S \;:=\; \sum_{j=1}^{N} X_j \;=\; \sum_{n=0}^{\infty} \chi_{\{N=n\}} \sum_{j=1}^{n} X_j$$

such that the random sum is defined by a distinction of cases depending on the values of $N$.

Throughout this paper, let $(\Omega, \mathcal{F}, P)$ be the probability space on which all random variables are defined. We denote by $\varphi_Z$ the characteristic function of a random variable $Z$ and by $m_Z$ the probability generating function of a random variable $Z$ with $P[Z \in \mathbb{N}_0] = 1$. We also use the convention $0/0 := 0$.

# 2 The Collective Model

Throughout this section, let

$$\langle N, \{X_j\}_{j\in\mathbb{N}}\rangle$$

be a collective model with total claim amount

$$S \;:=\; \sum_{j=1}^{N} X_j$$

In principle, the distribution of $S$ can be determined by the following results:

**2.1 Theorem.** *The characteristic function of $S$ satisfies $\varphi_S \;=\; m_N \circ \varphi_{X_1}$.*

In the special case where the claim amounts (and hence also the total claim amount) have a discrete distribution on the non–negative integers, the characteristic functions occurring in the previous result can be replaced by probability generating functions:

**2.2 Theorem.** *If $P[X_1 \in \mathbb{N}_0] = 1$, then $P[S \in \mathbb{N}_0] = 1$ and the probability generating function of $S$ is satisfies $m_S \;=\; m_N \circ m_{X_1}$.*

If the characteristic (or probability generating) function of the total claim amount can not be identified as that of a known distribution or if the distributions of the claim number and the claim amounts are unknown except for the first moments, then it is still possible to compute the corresponding first moments of the total claim amount:

**2.3 Theorem (Wald's equations).** *The total claim amount satisfies*

$$\begin{aligned} E[S] &= E[N]\; E[X] \\ var[S] &= E[N]\; var[X] + var[N]\; E[X]^2 \end{aligned}$$

# 3 The Abstract Collective Model

In this section we consider a useful extension of the collective model. The pair

$$\langle N, \{Y_j\}_{j\in\mathbb{N}}\rangle$$

is called an *abstract collective model* if $N$ is a random variable with $P[N \in \mathbb{N}_0] = 1]$ and $\{Y_j\}_{j\in\mathbb{N}}$ is sequence of random variables in $\mathbb{R}^p$ which is i.i.d. and independent of $N$. The (classical) collective model is the special cases with $p = 1$.

In the examples given below, the first component of the *claim variable* $Y_j$ is the claim amount of the $j$th claim and the other components contain further information on that claim.

## Transformation

Let us first consider the transformation of a collective model resulting from a transformation of the claim variables. If $g : \mathbb{R}^p \to \mathbb{R}$ is a measurable function, then the pair

$$\langle N, \{g \circ Y_j\}_{j\in\mathbb{N}} \rangle$$

is a collective model since the $\sigma$–algebra generated by $g \circ Y_j$ is contained in that generated by $Y_j$ and since the distribution of $g \circ Y_j$ is a measurable transformation of that of $Y_j$ which is the same for all $j$.

## Thinning

Let $C_1, \dots, C_m$ be a measurable subsets of $\mathbb{R}^p$ such that the *selection probabilities*

$$\eta_k := P[Y_1 \in C_k]$$

satisfy $\eta_k > 0$ for all $k \in \{1, \dots, m\}$ and $\sum_{k=1}^m \eta_k = 1$.

Let us fix $k \in \{1, \dots, m\}$ and consider thinning of the abstract collective model with respect to the set $C_k$. Then the *thinned claim number* $N_k$ is given by

$$N_k \; := \; \sum_{j=1}^{N} \chi_{\{Y_j \in C_k\}}$$

and the *thinned claim variables* are given by

$$Y_{k,j} \; := \; \sum_{h=1}^{\infty} \chi_{\{\nu_{k,j}=h\}} Y_h$$

where the random index $\nu_{k,j}$ is recursively defined by

$$\nu_{k,j} \;\; := \;\; \begin{cases} 0 & \text{if } j = 0 \\ \inf\{h \in \mathbb{N} | \nu_{k,j-1} < h, Y_h \in C_k\} & \text{else} \end{cases}$$

and describes the random position (in the sequence of all claim variables) of the $j$th claim variable taking its value in $C_k$.

**3.1 Theorem.** *For every $k \in \{1, \dots, m\}$, the pair*

$$\langle N_k, \{Y_{k,j}\}_{j\in\mathbb{N}} \rangle$$

*is an abstract collective model. Moreover, the distribution of the thinned claim number is determined by its probability generating function satisfying*

$$m_{N_k}(z) \;\; = \;\; m_N(1 - \eta_k + \eta_k z)$$

*and the distribution of the thinned claim variables satisfies*

$$P[Y_{k,j} \leq y] \;=\; P[Y_1 \leq y |\, Y_1 \in C_k]$$

In this result, the assertion on the probability generating function of $N_k$ is immediate from Theorem 2.2 since $N_k$ is the total claim amount in the collective model

$$\langle N_k, \{\chi_{C_k} \circ Y_j\}_{j\in\mathbb{N}}\rangle$$

and a proof of the second identity can be found in Schmidt (2009; Satz 8.2.9).

Applying Wald's first equation to the collective model $\langle N_k, \{\chi_{C_k} \circ Y_j\}_{j\in\mathbb{N}}\rangle$ we obtain the following result:

**3.2 Lemma.** *For every $k \in \{1, \ldots, m\}$, the expectation of the thinned claim number satisfies*

$$E[N_k] \;=\; \eta_k E[N]$$

**3.3 Example.** Assume that $N$ has the Poisson distribution with parameter $\lambda$. Then the probability generating function of $N$ satisfies $m_N(z) = e^{-\lambda(1-z)}$ and it follows that

$$\begin{aligned} m_{N_k}(z) &= m_N(1-\eta_k+\eta_k z) \\ &= e^{-\lambda(1-(1-\eta_k+\eta_k z))} \\ &= e^{-\lambda\eta_k(1-z)} \end{aligned}$$

Therefore, $N_k$ has the Poisson distribution with parameter $\lambda\eta_k$. $\triangle$

## Decomposition

Thinning with respect to the set $C_1, \ldots, C_m$ yields $m$ collective models. The following result provides information on the joint distribution of the random variables of these models:

**3.4 Theorem.**

(1) *The family of all thinned claim variables $\{Y_{k,j}\}_{k\in\{1,\ldots,m\},j\in\mathbb{N}}$ is independent.*
(2) *The family of all thinned claim variables $\{Y_{k,j}\}_{k\in\{1,\ldots,m\},j\in\mathbb{N}}$ is independent of the family of all thinned claim numbers $\{N_k\}_{k\in\{1,\ldots,m\}}$.*
(3) *The conditional joint distribution of the thinned claim numbers $(N_1, \ldots, N_m)$ is the conditional multinomial distribution with parameters $N$ and $(\eta_1, \ldots, \eta_m)$.*
(4) *The thinned claim numbers and hence the thinned collective models are independent if and only if $N$ has a Poisson distribution.*

A proof of the first assertions can be found in Hess (2000); for the special case of the decomposition into two collective model see also Schmidt (1996). The last assertion follows from the third; see e.g. Hess and Schmidt (2002).

# 4 Excess–of–Loss Reinsurance

Reinsurance allows an insurance company to transform its portfolio into a more homogeneous one by sharing the total claim amount with a reinsurer; see e.g. Mack (2002; Chapter 4) or Hess and Schmidt (2006).

In excess–of–loss reinsurance the reinsurer assumes, for every claim, liability for the part of the claim amount exceeding a negotiated priority $d$. In actuarial practice, the liability of the reinsurer is usually bounded by a certain limit, but for the sake of simplicity we constrain ourselves to an unlimited excess–of–loss reinsurance contract.

We start with a collective model

$$\langle N, \{X_j\}_{j\in\mathbb{N}}\rangle$$

for the original portfolio and we assume that $0 < P[X_j > d] < 1$. If the claim amount $X_j$ exceeds the priority $d$, then the insurer pays $d$ and the reinsurer pays $X_j - d$; otherwise the insurer pays $X_j$ and the reinsurer pays nothing. Thus, transformations of the collective model yield the collective model

$$\langle N, \{\min\{X_j, d\}_{j\in\mathbb{N}}\rangle$$

for the insurer and the collective model

$$\langle N, \{\max\{X_j - d, 0\}\}_{j\in\mathbb{N}}\rangle$$

for the reinsurer. The total claim amount of the original portfolio is thus split into the sum of the total claim amount of the insurer and the total claim amount of the reinsurer

$$\sum_{j=1}^{N} X_j \;=\; \sum_{j=1}^{N} \min\{X_j, d\} + \sum_{j=1}^{N} \max\{X_j - d, 0\}$$

In general, the random variables occurring in the collective model for the reinsurer are not observable for the reinsurer. It is therefore desirable to construct an equivalent collective model in which all random variables are observable for the reinsurer.

By thinning the collective model $\langle N, \{X_j\}_{j\in\mathbb{N}}\rangle$ of the original portfolio with respect to the set $C_1 := (d, \infty)$, we get the collective model $\langle N_1, \{X_{1,j}\}_{j\in\mathbb{N}}\rangle$ containing only the claims exceeding priority $d$ and the number $N_1$ of those claims; here the selection probability $\eta_1 = P[X_j > d]$ is the probability that a claim exceeds the priority $d$. Since the reinsurer pays only the part of the claim amounts exceeding the priority, a transformation of this collective model leads to the collective model

$$\langle N_1, \{X_{1,j} - d\}_{j\in\mathbb{N}}\rangle$$

of the reinsurer in which indeed all random variable are observable for the reinsurer. The distribution of the claim sizes $X_{1,j} - d$ satisfies

$$P[X_{1,j} - d \leq z] \;=\; P[X_1 \leq d + z |\, X_1 > d]$$

Moreover, we have

$$\sum_{j=1}^{N} \max\{X_j - d, 0\} = \sum_{j=1}^{N_1} (X_{1,j} - d)$$

which means that the total claim amounts of both collective models for the reinsurer are identical; for a proof see Hess (2003).

If the distributions of $N$ and $X_1$ belong to parametric families of distributions, then the reinsurer may estimate the parameters by using truncated data methods. This is of interest with regard to negotiations on future reinsurance contracts with a different priority.

**4.1 Example.** As a very simple example let us consider the collective model $\langle N, \{X_j\}_{j\in\mathbb{N}}\rangle$ where $N$ has the Poisson distribution with parameter $\lambda$ and every $X_j$ has the exponential distribution with expectation $1/\alpha$. Then, in the thinned collective model $\langle N_1, \{X_{1,j} - d\}_{j\in\mathbb{N}}\rangle$, the excess claim number $N_1$ has the Poisson distribution with parameter $\lambda\eta_1$ with

$$\eta_1 \;=\; P[X_j > d] \;=\; e^{-\alpha d}$$

and the excess claim amounts $Z_j := X_{1,j} - d$ have the exponential distribution with expectation $1/\alpha$, due to the fact that the exponential distribution is memoryless. Therefore, the likelihood–function of the thinned collective model satisfies

$$L(\alpha, \lambda | N_1, \{Z_j\}_{j\in\mathbb{N}}) \;=\; e^{-\lambda e^{-\alpha d}} \frac{(\lambda e^{-\alpha d})^{N_1}}{N_1!} \cdot \prod_{j=1}^{N_1} \alpha e^{-\alpha Z_j}$$

Putting the first partial derivatives of the log–likelihood–function equal to zero we obtain the maximum–likelihood estimators

$$\widehat{\alpha}^{\mathrm{ML}} = \left( \frac{1}{N_1} \sum_{j=1}^{N_1} Z_j \right)^{-1}$$

and

$$\widehat{\lambda}^{\mathrm{ML}} = \frac{N_1}{e^{-\widehat{\alpha}^{\mathrm{ML}} d}}$$

Using these estimators, the reinsurer can analyze the effect of changing the priority $d$ in a future reinsurance contract. $\triangle$

# 5 Multiplicative Tariff

In certain lines of business, the risks of the portfolio are classified with respect to a finite number levels of one or more tariff factors. For example, in motor car liability insurance such tariff factors could be the power of the engine or the mileage per year and the levels would be intervals. Once such a classification is given, every risk of the portfolio belongs to a unique tariff cell which is determined by a combination of levels of the different tariff factors. For a discussion see Mack (2002; Chapter 2) and Zocher (2006).

Such a classification of risks is useful, when the full portfolio is inhomogeneous and the subportfolios corresponding to a cell are supposed to be rather homogeneous. In this situation, it would be inappropriate to charge the same premium to all risks in the portfolio and the insurer would like to allow for different premiums for risks belonging to different cells.

In a multiplicative tariff, the net premiums for the different cells determined in such a way that, for every cell, the net premium is proportional to certain parameters presenting the levels of the tariff factors.

For the sake of simplicity we consider the construction of the multiplicative tariff in the case of two tariff factors with respective levels $1, \dots, I$ and $1, \dots, K$ and tariff cells $(i, k) \in \{1, \dots, I\} \times \{1, \dots, K\}$.

To this end we consider the abstract collective model

$$\langle N, \{(X_j, U_j, T_j)\}_{j \in \mathbb{N}} \rangle$$

We interpret $N$ the number of claims in the portfolio, $X_j$ as the claim amount of the $j$th claim, and $(U_j, T_j)$ as the tariff cell of the risk which causes the $j$th claim.

We assume that $X_j$ and $(U_j, T_j)$ are independent. Furthermore, we assume that there exist known volume measures $v_{i,k} > 0$ (e.g. the number of risks in cell $(i, k)$) and unknown parameters $\alpha_i > 0$ and $\beta_k > 0$ such that

$$P[(U_j, T_j) = (i, k)] \;=\; \alpha_i \, \beta_k \, v_{i,k}$$

holds for all $(i, k) \in \{1, \dots, I\} \times \{1, \dots, K\}$.

Then we have

$$\sum_{i=1}^{I} \sum_{k=1}^{K} \alpha_i \, \beta_k \, v_{i,k} \;=\; 1$$

We want to decompose the collective model $\langle N, \{(X_j, U_j, T_j)\}_{j \in \mathbb{N}} \rangle$ with respect to the tariff cells. Define $C_{i,k} := \mathbb{R} \times \{i\} \times \{k\}$. Then the selection probabilities are $\eta_{i,k} = \alpha_i \, \beta_k \, v_{i,k}$ and we get the thinned collective models

$$\langle N_{i,k}, \{(X_{i,k;j}, U_{i,k;j}, T_{i,k;j})\}_{j \in \mathbb{N}} \rangle$$

By the construction of the thinned collective models we have $U_{i,k;j} = i$ and $T_{i,k;j} = k$. Thus, the second and third component of the thinned claim variables are redundant and we get the collective models

$$\langle N_{i,k}, \{X_{i,k;j}\}_{j \in \mathbb{N}} \rangle$$

From the statistical point of view this is a poststratification with respect to the tariff cells.

Since $X_j$ and $(U_j, T_j)$ are independent, we get

$$\begin{aligned} P[X_{i,k;j} \leq x] &= P[X_{i,k;j} \leq x, U_{i,k;j} = i, T_{i,k;j} = k] \\ &= P[X_j \leq x, U_j = i, T_j = k | \, U_j = i, T_j = k] \\ &= P[X_j \leq x] \end{aligned}$$

this means that the claim amounts $X_{i,k;j}$ have the same distribution as the original claim amounts $X_j$. Therefore the distribution or moments of the claim amounts of any cell may be estimated from the full sample by usual methods.

In order to interpret the parameters, we consider the expectation of the total claim amount of the thinned collective models. Applying Lemma 3.2, we get

$$E[N_{i,k}] = \alpha_i \, \beta_k \, v_{i,k} \, E[N]$$

and using this identity together with $E[S] = E[N] \, E[X_1]$ we obtain

$$\begin{aligned} E[S_{i,k}] &= E[N_{i,k}] \, E[X_{i,k;1}] \\ &= \alpha_i \, \beta_k \, v_{i,k} \, E[N] \, E[X_1] \\ &= \alpha_i \, \beta_k \, v_{i,k} \, E[S] \end{aligned}$$

The ratio $S_{i,k}/v_{i,k}$ is total claim amount per risk in cell $(i,k)$ and its expectation

$$E\left[\frac{S_{i,k}}{v_{i,k}}\right] = \alpha_i \, \beta_k \, E[S]$$

is the *net premium* per risk in cell $(i,k)$. Up to the estimation of the parameters, we have thus constructed a multiplicative tariff.

We want to estimate the unknown parameters by the maximum–likelihood method. Applying Theorem 3.4 (1), we get the joint distribution of the thinned claim numbers

$$\begin{aligned} P\left[\bigcap_{i=1}^{I} \bigcap_{k=1}^{K} \{N_{i,k} = n_{i,k}\}\right] &= P\left[\bigcap_{i=1}^{I} \bigcap_{k=1}^{K} \{N_{i,k} = n_{i,k}\} \,\middle|\, N = n\right] \cdot P[N = n] \\ &= \frac{n!}{\prod_{i=1}^{I} \prod_{k=1}^{K} n_{i,k}!} \prod_{i=1}^{I} \prod_{k=1}^{K} \eta_{i,k}^{n_{i,k}} \cdot p(n) \end{aligned}$$

with $n = \sum_{i=1}^{I} \sum_{k=1}^{K} n_{i,k}$ and $p(n) := P[N = n]$. Then the likelihood function satisfies

$$\begin{aligned} & L(\alpha_1, \ldots, \alpha_I, \beta_1, \ldots, \beta_K | \{N_{i,k}\}_{i \in \{1,\ldots,I\}, k \in \{1,\ldots,K\}}) \\ &= \frac{N!}{\prod_{i=1}^{I} \prod_{k=1}^{K} N_{i,k}!} \prod_{i=1}^{I} \prod_{k=1}^{K} (\alpha_i \, \beta_k \, v_{i,k})^{N_{i,k}} \cdot p(N) \end{aligned}$$

(note that $N = \sum_{i=1}^{I} \sum_{k=1}^{K} N_{i,k}$). Positive maximizers $\widehat{\alpha}_1^{\mathrm{ML}}, \dots, \widehat{\alpha}_I^{\mathrm{ML}}$ and $\widehat{\beta}_1^{\mathrm{ML}}, \dots, \widehat{\beta}_K^{\mathrm{ML}}$ of the likelihood function fulfilling the constraint

$$\sum_{i=1}^{I} \sum_{k=1}^{K} \widehat{\alpha}_i^{\mathrm{ML}} \, \widehat{\beta}_k^{\mathrm{ML}} \, v_{i,k} \;=\; 1$$

are *maximum–likelihood estimators* of the parameters $\alpha_1, \dots, \alpha_I$ and $\beta_1, \dots, \beta_K$. As proved in Hess (2009) the maximum–likelihood estimators are solutions of the *marginal–sum equations*

$$\begin{aligned} \sum_{i=1}^{I} \widehat{\alpha}_i \, \widehat{\beta}_k \, v_{i,k} &= \sum_{i=1}^{I} \frac{N_{i,k}}{N} & \qquad & k \in \{1, \dots, K\} \\ \sum_{k=1}^{K} \widehat{\alpha}_i \, \widehat{\beta}_k \, v_{i,k} &= \sum_{k=1}^{K} \frac{N_{i,k}}{N} & \qquad & i \in \{1, \dots, I\} \end{aligned}$$

under the constraint

$$\sum_{i=1}^{I} \sum_{k=1}^{K} \widehat{\alpha}_i \, \widehat{\beta}_k \, v_{i,k} \;=\; 1$$

Since $v_{i,k} > 0$ these equations have a solution which is unique up to scaling and can be obtained by iteration (see Dietze, Riedrich and Schmidt (2006)).

As an alternative to maximum–likelihood estimation, the marginal–sum equations can also be deduced from the identities for the expected claim numbers: We have

$$E[N_{i,k}] = \alpha_i \, \beta_k \, v_{i,k} \, E[N]$$

and summation yields

$$\begin{aligned} \sum_{i=1}^{I} \frac{E[N_{i,k}]}{E[N]} &= \sum_{i=1}^{I} \alpha_i \, \beta_k \, v_{i,k} & \qquad & k \in \{1, \dots, K\} \\ \sum_{k=1}^{K} \frac{E[N_{i,k}]}{E[N]} &= \sum_{k=1}^{K} \alpha_i \, \beta_k \, v_{i,k} & \qquad & i \in \{1, \dots, I\} \end{aligned}$$

Replacing the expectations by the corresponding random variables and the unknown parameters by their estimators then leads to the marginal–sum equations. Also, applying the same principle to the identities

$$E[S_{i,k}] = \alpha_i \, \beta_k \, v_{i,k} \, E[S]$$

yields a different set of marginal equations, using the total claim amounts of the cells instead of the claim numbers.

Therefore, marginal–sum estimation may be regarded as a principle of estimation in its own right.

# 6 Loss Reserving

The problem of loss reserving results from the fact that by the end of an insurance year certain claims are not yet settled since either the claim size is not yet known or the claim has not yet been reported. Since all claims of a given accident year have to be paid from the premium income of the same accident year, the insurer has to determine a reserve for claims payments in future development years. In certain lines of business like liability insurance, the amount of all future payments may be huge when compared with the annual premium income. In actuarial mathematics there is a rich literature on loss reserving; see also Schmidt (2011) in the present volume.

In the construction of models for loss reserving it is usually assumed that every claim can be settled either in the accident year of with a delay of at most $I$ years. Accordingly, it is usually assumed that certain data from $I+1$ accident years are available.

We consider $I+1$ independent abstract collective models

$$\langle N^{(i)}, \{(X_j^{(i)}, D_j^{(i)})\}_{j\in\mathbb{N}}\rangle$$

corresponding to the accident years $i \in \{0,1,\ldots,I\}$ and we assume for all $i \in \{0,1,\ldots,I\}$ and $j \in \mathbb{N}$ that

- $N^{(i)}$ has the Hofmann distribution $\pi_{a_i,p_i,c_i}$ with parameters $a_i \in \mathbb{R}_+$ and $p_i, c_i \in (0,\infty)$ (see Appendix A for details),
- $X_j^{(i)}$ and $D_j^{(i)}$ are independent,
- $X_j^{(i)}$ is a real–valued random variable, and
- the distribution of $D_j^{(i)}$ is given by $P[D_j^{(i)} = k] = \vartheta_k$ with $k \in \{0,1,\ldots,I\}$ and parameters $\vartheta_0, \vartheta_1, \ldots, \vartheta_I \in (0,\infty)$ with $\sum_{k=0}^I \vartheta_k = 1$.

For every accident year $i$, we interpret $N^{(i)}$ as the number of claims, $X_j^{(i)}$ as the claim amount of the $j$th claim, and $D_j^{(i)}$ as the delay until settlement of the $j$th claim (such that there is exactly one payment per claim). We assume that the claim variables $(X_j^{(i)}, D_j^{(i)})$ are observable if $i + D_j^{(i)} \le I$ and that they are non–observable if $i + D_j^{(i)} \ge I+1$.

Each of the abstract collective models can be decomposed with respect to the values of the delays $D_j^{(i)}$ by choosing $C_k := \mathbb{R} \times \{k\}$ for all $k \in \{0,1,\ldots,I\}$. Then $\vartheta_k$ is the selection probability for the set $C_k$ (delay of $k$ years) and we get the abstract collective models

$$\langle N_k^{(i)}, \{(X_{k,j}^{(i)}, D_{k,j}^{(i)})\}_{j\in\mathbb{N}}\rangle$$

for all $i \in \{0,\ldots,I\}$. Due to the choice of the sets $C_k$, we have $D_{k,j}^{(i)} = k$.

Letting $N_{i,k} := N_k^{(i)}$ and $X_{i,k;j} := X_{k,j}^{(i)}$, we thus obtain the collective models

$$\langle N_{i,k}, \{X_{i,k;j}\}_{j\in\mathbb{N}}\rangle$$

which are observable if $i+k \leq I$ and non–observable if $i+k \geq I+1$. These collective models are independent for different $i$'s since we started with independent models for the accident years. Therefore all claim amounts $X_{i,k;j}$ are independent. For a given accident year $i$ we can use all observable claim amounts for estimating their common distribution or moments.

For every accident year $i$, the joint distribution of the thinned claim numbers $N_{i,k}$ can be obtained from Theorem 3.4 (3) and the representation of a Hofmann–distribution as a mixed Poisson–distribution (see Proposition A.1 with $t = 1$). We thus obtain

$$\begin{aligned}
&P\left[\bigcap_{k=0}^{I}\{N_{i,k} = n_{i,k}\}\right] \\
&= P\left[\bigcap_{k=0}^{I}\{N_{i,k} = n_{i,k}\}\,\middle|\, N^{(i)} = \sum_{k=0}^{I} n_{i,k}\right] \cdot P\left[N^{(i)} = \sum_{k=0}^{I} n_{i,k}\right] \\
&= \frac{\left(\sum_{k=0}^{I} n_{i,k}\right)!}{\prod_{k=0}^{I} n_{i,k}!}\prod_{k=0}^{I}\vartheta_k^{n_{i,k}} \cdot \int_{\mathbb{R}} e^{-\lambda}\frac{\lambda^{\sum_{k=0}^{I} n_{i,k}}}{\left(\sum_{k=0}^{I} n_{i,k}\right)!}\, dQ_{a_i,p_i,c_i}(\lambda) \\
&= \int_{\mathbb{R}}\prod_{k=0}^{I} e^{-\lambda\vartheta_k}\frac{(\lambda\vartheta_k)^{n_{i,k}}}{n_{i,k}!}\, dQ_{a_i,p_i,c_i}(\lambda)
\end{aligned}$$

Since we started with independent collective models for the different accident years, the joint distribution of all thinned claim numbers is given by

$$P\left[\bigcap_{i=0}^{I}\bigcap_{k=0}^{I}\{N_{i,k} = n_{i,k}\}\right] = \prod_{i=0}^{I}\int_{\mathbb{R}}\prod_{k=0}^{I} e^{-\lambda\vartheta_k}\frac{(\lambda\vartheta_k)^{n_{i,k}}}{n_{i,k}!}\, dQ_{a_i,p_i,c_i}(\lambda)$$

This model was considered by Schmidt and Zocher (2005; Section 7) who showed that the joint distribution of all observable claim numbers satisfies

$$\begin{aligned}
&P\left[\bigcap_{i=0}^{I}\bigcap_{k=0}^{I-i}\{N_{i,k} = n_{i,k}\}\right] \\
&= \prod_{i=0}^{I}\left(\frac{\left(\sum_{k=0}^{I-i} n_{i,k}\right)!}{\prod_{k=0}^{I-i} n_{i,k}!}\prod_{k=0}^{I-i}\left(\frac{\vartheta_k}{\sum_{l=0}^{I-i}\vartheta_l}\right)^{n_{i,k}} \cdot \pi_{a_i,p_i\sum_{k=0}^{I-i}\vartheta_k,c_i\sum_{k=0}^{I-i}\vartheta_k}\left(\sum_{k=0}^{I-i} n_{i,k}\right)\right)
\end{aligned}$$

Positive maximizers $\widehat{p}_0^{\mathrm{ML}}, \widehat{p}_1^{\mathrm{ML}}, \ldots, \widehat{p}_I^{\mathrm{ML}}$ and $\widehat{\vartheta}_0^{\mathrm{ML}}, \widehat{\vartheta}_1^{\mathrm{ML}}, \ldots, \widehat{\vartheta}_I^{\mathrm{ML}}$ of the corresponding likelihood function fulfilling the constraint

$$\sum_{k=0}^{I}\widehat{\vartheta}_k^{\mathrm{ML}} = 1$$

are *maximum–likelihood estimators* of the parameters $p_0, p_1, \ldots, p_I$ and $\vartheta_0, \vartheta_1, \ldots, \vartheta_I$. Schmidt and Zocher (2005) have shown that the maximum–likelihood estimators fulfill the marginal–sum equations

$$\begin{aligned}
\sum_{i=0}^{I-k} \widehat{p}_i \, \widehat{\vartheta}_k &= \sum_{i=0}^{I-k} N_{i,k} && k \in \{0,1,\ldots,I\} \\
\sum_{k=0}^{I-i} \widehat{p}_i \, \widehat{\vartheta}_k &= \sum_{k=0}^{I-i} N_{i,k} && i \in \{0,1,\ldots,I\}
\end{aligned}$$

under the constraint

$$\sum_{k=0}^{I} \widehat{\vartheta}_k = 1$$

It is well known that these equations have a unique and explicit solution (see e.g. Schmidt and Wünsche (1998)).

Using the maximum–likelihood estimators, we can now predict the non–observable claim numbers In fact, using Lemma 3.2 and Proposition A.3 we obtain

$$E[N_{i,k}] = p_i \vartheta_k$$

This identity suggest to use the observable random variable

$$\widehat{N}_{i,k} := \widehat{p}_i^{\mathrm{ML}} \, \widehat{\vartheta}_k^{\mathrm{ML}}$$

to predict the non–observable claim numbers $N_{i,k}$ with $i+k \geq I+1$ (or to estimate their expectations). Since the total claim amounts $S_{i,k}$ satisfy

$$\begin{aligned}
E[S_{i,k}] &= E[N_{i,k}] \, E[X_{i,k;j}] \\
&= E[N_{i,k}] \, E[X_1]
\end{aligned}$$

one may use the observable random variable

$$\widehat{S}_{i,k} := \widehat{p}_i^{\mathrm{ML}} \, \widehat{\vartheta}_k^{\mathrm{ML}} \, \widehat{\mu}$$

with an estimator $\widehat{\mu}$ of the expectation $\mu := E[X_1]$ to predict the non–observable total claim amounts $S_{i,k}$ with $i+k \geq I+1$ (or to estimate its expectation).

As in the model considered in the previous section, the marginal–sum equations can also be deduced from the expectations of the observable claim numbers: As noted before, we have $E[N_{i,k}] = p_i \vartheta_k$, and summation yields

$$\begin{aligned}
\sum_{i=0}^{I-k} E[N_{i,k}] &= \sum_{i=0}^{I-k} p_i \, \vartheta_k && k \in \{0,1,\ldots,I\} \\
\sum_{k=0}^{I-i} E[N_{i,k}] &= \sum_{k=0}^{I-i} p_i \, \vartheta_k && i \in \{0,1,\ldots,I\}
\end{aligned}$$

Replacing the expectations by the corresponding random variables and the unknown parameters by their estimators then leads to the marginal–sum equations.

# 7 Comparison of the Marginal–Sum Equations

In the case of a multiplicative tariff and also in the case of loss reserving, maximum–likelihood estimation in a collective model leads to marginal–sum equations. Although the structure of the marginal–sum equations in both cases is similar, there are nevertheless important differences:

In the case of a multiplicative tariff the claim number is observable and this makes maximum–likelihood estimation of the parameters independent of the distribution of the claim number. By contrast, in the case of loss reserving certain claim numbers are non–observable such that we cannot use Theorem 3.4 (3) to obtain the joint distribution of the decomposed claim numbers; in this case, the assumption of Hofmann distributed claim numbers allows us to determine the joint distribution of the thinned claim numbers.

But also the marginal–sum equations are quite different: In the case of a multiplicative tariff they have a rectangular structure and cannot solved explicitly; for a discussion on their solvability, see Dietze et al. (2006). In the case of loss reserving they have a triangular structure and can solved explicitly; see e.g. Schmidt and Wünsche (1998).

# A Hofmann family and Hofmann distribution

The Hofmann family was introduced by Hofmann (1955). We use the presentation by Schmidt and Zocher (2005).

We consider a family $\{Q_t\}_{t\in\mathbb{R}}$ of distributions $Q_t : \mathcal{P}(\mathbb{N}_0) \to [0,1]$ and a sequence $\{\Pi_k\}_{k\in\mathbb{N}_0}$ of functions $\Pi_k : \mathbb{R}_+ \to [0,1]$ such that the identity

$$Q_t[\{k\}] \;=\; \Pi_k(t)$$

holds for all $t \in \mathbb{R}_+$ and $k \in \mathbb{N}_0$.

The family $\{Q_t\}_{t\in\mathbb{R}}$ is said to be the *Hofmann family* $\mathbf{H}(a,p,c)$ with parameters $a \in \mathbb{R}_+$ and $p, c \in (0,\infty)$ if there exists a differentiable function $\nu_{a,p,c} : \mathbb{R}_+ \to \mathbb{R}$ such that

$$\begin{aligned}
\nu_{a,p,c}(0) &= 0 \\
\frac{d\nu_{a,p,c}}{dt}(t) &= \frac{p}{(1+ct)^a} \\
\Pi_0(t) &= e^{-\nu_{a,p,c}(t)} \\
\Pi_k(t) &= \frac{(-t)^k}{k!}\frac{d^k\Pi_0}{dt^k}(t)
\end{aligned}$$

hold for all $t \in \mathbb{R}_+$ and $k \in \mathbb{N}$.

The function $\nu_{a,p,c}$ has an explicit representation:

$$\nu_{a,p,c} \;=\; \begin{cases} pt & \text{if } a = 0 \\ \dfrac{p}{c}\log(1+ct) & \text{if } a = 1 \\ \dfrac{p}{c}\dfrac{(1+ct)^{1-a}-1}{1-a} & \text{if } a \in (0,1)\cup(1,\infty) \end{cases}$$

The special cases $a = 0$ and $a = 1$ are the Poisson case and the negativebinomial case, respectively.

Applying the Bernstein–Widder Theorem to the completely monotone function $\Pi_0$, we see that the Hofmann family has a mixed Poisson representation; see e.g. Berg et al. (1984):

**A.1 Proposition.** *There exists a probability measure* $Q_{a,p,c} : \mathcal{B}(\mathbb{R}) \to [0,1]$ *with* $Q_{a,p,c}[\mathbb{R}_+] = 1$ *and such that*

$$Q_t[\{k\}] \;=\; \int_{\mathbb{R}_+} e^{-\lambda t}\frac{(\lambda t)^k}{k!}\, dQ_{a,p,c}(\lambda)$$

*holds for all* $t \in \mathbb{R}_+$ *and* $k \in \mathbb{N}_0$.

Using this Proposition, we get the following result on the probability generating function:

**A.2 Proposition.** *The probability generation function of* $Q_t$ *satisfies*

$$m_{Q_t}(z) \;=\; e^{-\nu_{a,p,c}(t-tz)}$$

If $\{Q_t\}_{t\in\mathbb{R}_+}$ is the Hofmann family $\mathbf{H}(a,p,c)$, then the discrete distribution $\pi_{a,p,c}$ on the non–negative integers satisfying

$$\pi_{a,p,c}[\{k\}] \;=\; Q_1[\{k\}]$$

for all $k \in \mathbb{N}_0$ is called the *Hofmann distribution* with parameters $a \in \mathbb{R}_+$ and $p, c \in (0,\infty)$. Proposition A.1 shows that every Hofmann distribution is a mixed Poisson distribution. The probability generating function of a Hofmann distribution is obtained by letting $t = 1$ in Proposition A.2 and yields its first and second order moments:

**A.3 Proposition.** *If* $N$ *is a random variable with distribution* $\pi_{a,p,c}$*, then*

$$\begin{aligned} E[N] &= p \\ var[N] &= p(1+ac) \end{aligned}$$

Therefore we have $E[N] \le var[N]$ which is true for every mixed Poisson distribution and is typical for the moments of real data.

# Acknowledgement

The author would like to thank Klaus D. Schmidt for many years of collaboration and he would also like to thank Tobias Franke, Markus Hübel, Martin Jahnke, Wolfgang Macht and Mathias Zocher for their contributions on the topics of this paper.


# References


**Berg, C.; Christensen, J. P. R.; Ressel, P. (1984):** *Harmonic Analysis on Semigroups.* Berlin – Heidelber – New York: Springer.

**Dietze, S.; Riedrich, Th.; Schmidt, K. D. (2006):** *On the Solution of Marginal–Sum Equations.* Dresdner Schriften zur Versicherungsmathematik 1/2006.

**Hofmann, M. (1955):** *Über zusammengesetzte Poisson–Prozesse und ihre Anwendungen in der Unfallversicherung.* Mitteilungen VSVM **55**, 499–575.

**Hess, K. Th. (2000):** *Random Partitions of Samples.* Dresdner Schriften zur Versicherungsmathematik 1/2000.

**Hess, K. Th. (2003):** *Das kollektive Modell der Risikotheorie in der Schadenexzedenten–Rückversicherung.* Allg. Stat. Archiv **87**, 309–320.

**Hess, K. Th. (2009):** *Marginal–sum and maximum–likelihood estimation in a multiplicative tariff.* AStA Advances **93**, 221–233.

**Hess, K. Th.; Schmidt, K. D. (2002):** *A comparison of models for the chain–ladder method.* Insurance: Math. Econom. **31**, 351–364.

**Hess, K. Th.; Schmidt, K. D. (2006):** *Risikoteilung und Rückversicherung.* Wiss. Z. TU Dresden **55**, 19–23.

**Mack, Th. (2002):** *Schadenversicherungsmathematik.* Karlsruhe: Verlag Versicherungswirtschaft, 2$^{\mathrm{nd}}$ ed.

**Schmidt, K. D. (1996):** *Lectures on Risk Theory.* Stuttgart: Teubner.

**Schmidt, K. D. (2009):** *Versicherungsmathematik.* Berlin: Springer, 3$^{\mathrm{rd}}$ ed.

**Schmidt, K. D. (2011):** *Loss Prediction Based on Run–Off Triangles.* AStA Advances **XX**, XXX–XXX.

**Schmidt, K. D.; Wünsche, A. (1998):** *Chain ladder, marginal sum and maximum likelihood estimation.* Blätter DGVM **23**, 267–277.

**Schmidt, K. D.; Zocher, M. (2005):** *Loss reserving and Hofmann distributions.* Mitteilungen SAV 127–162.

**Zocher, M. (2006):** *Risikoadäquate Tarifierung in der Kraftfahrthaftpflichtversicherung.* Wiss. Z. TU Dresden **55**, 131–135.

Klaus Th. Hess
Institut für Mathematik
Universität Rostock
D–18051 Rostock
Germany

e–mail: klaus-thomas.hess@uni-rostock.de

internet: http://www.math.uni-rostock.de/~hess/

April 13, 2021