Zeitschrift: Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries

Herausgeber: Vereinigung Schweizerischer Versicherungsmathematiker

Band: 78 (1978)

Artikel: A computer algorithm for the cumul model

Autor: Berger, Gottfried

DOI: https://doi.org/10.5169/seals-555074

A Computer Algorithm for the Cumul Model

By Gottfried Berger

Abstract

The first two sections of this paper describe a risk theory model which was introduced recently by Tellenbach. The model involves Laplace-Stieltjes transforms which pose severe computational difficulties.
The remainder of this paper describes a simple algorithm which appears to work in the special case where the time span considered is reasonably small. The latter condition is typically met if the model refers to cumulative claims. A thorough mathematical treatment of the subject is not even attempted. Rather, the emphasis is on the computational aspect of the problem.

1. A Claims Process

Let t_j denote the interclaim time between the $j-th$ and the next following claim ($j = 0, 1, 2\ldots$). Suppose the probability distribution functions $P_j(t)$ for the stochastic variables t_j are known. We assume $t_j \geq 0$ and thus $P_j(t) = 0$ for $t < 0$. Our objective is to find the probabilities $p_k(t)$ for k claims in the time interval $(0, t)$.
Clearly, $p_0(t) = 1 - P_0(t)$. The rest is not as easy, since we have to perform convolutions. This can be done, at least in theory, by means of Laplace-Stieltjes transforms. See for instance Seal (1969), Appendix A, or Tellenbach (1977).
We thus define the following Laplace-Stieltjes transforms:

$$\begin{aligned} \Phi_j &= \Phi_j(s) = L\{P_j(t)\} = \int_{0-}^{\infty} e^{-st}\, dP_j(t), (j = 0, 1, 2\ldots). \\ \Psi_k &= \Psi_k(s) = L\{p_k(t)\} = \int_{0-}^{\infty} e^{-st}\, dp_k(t), (k = 0, 1, 2\ldots). \end{aligned} \tag{1}$$

We now introduce two simplifications. First, we require that all interclaim times t_j (except possibly t_0) are independently and identically distributed, i.e., $\Phi_j = \Phi_1$ for $j \geq 1$. The theory of the renewal process then shows that:

$$
\begin{aligned}
L\{1-p_0(t)\} &\equiv \Phi_0 \\
L\{p_1(t)\} &\equiv \Psi_1 = \Phi_0(1-\Phi_1) \\
L\{p_k(t)\} &\equiv \Psi_k = \Psi_{k-1}\,\Phi_1 = \Psi_1\Phi_1{}^{k-1} \text{ for } k > 1
\end{aligned}
\tag{2}
$$

Second, we stipulate a stationary claims process:

$$\mu_1(t) = \sum_{1}^{\infty} k \cdot p_k(t) = \text{const} \cdot t = t/\alpha_1 \tag{3}$$

If (3) holds, $\mu_1(t)$, the expected claims number in $(0, t)$, is proportional to t, and inversely proportional to α_1 = average interclaim time = mean of $P_1(t)$. We obtain from (2) and (3):

$$L\{\mu_1(t)\} = \Psi_1(1+2\Phi_1+3\Phi_1{}^2+\ldots) = \Psi_1/(1-\Phi_1)^2 = \Phi_0/(1-\Phi_1) = 1/\alpha_1 s$$

Thus, we can rewrite equations (2) as follows:

$$
\begin{aligned}
L\{1-p_0(t)\} &\equiv \Phi_0 = (1-\Phi_1)/\alpha_1 s \\
L\{p_1(t)\} &\equiv \Psi_1 = \Phi_0(1-\Phi_1) \\
L\{p_k(t)\} &\equiv \Psi_k = \Psi_{k-1}\,\Phi_1 \text{ for } k > 1
\end{aligned}
\tag{4}
$$

Equations (4) confirm that the stationary claims process is completely determined by the function $\Phi_1 = \Phi_1(s)$, the Laplace-Stieltjes transform of $P_1(t)$, the d.f. of the interclaim time.

2. The Tellenbach Model

Tellenbach (1977) considered the following choice of $P_1(t)$:

$$P_1(t) = h \cdot P(t) + (1-h) \cdot u(t) \tag{5}$$

Here, $P(t)$ is an arbitrary d.f. which applies to the claims process with the probability h $(0 < h \leq 1)$. Multiple claims may occur with the probability $(1-h)$, since

$$u(t) = \begin{cases} 0 \text{ if } t < 0, \\ 1 \text{ if } t \geq 0. \end{cases}$$

Intuitively, $P(t)$ controls the number of events, while $P_1(t)$ determines the number of claims. The smaller the parameter h is chosen, the more claims one associates with each claim event.
Let $\Phi = \Phi(s)$ denote the Laplace-Stieltjes transform of $P(t)$, and let α be the mean of $P(t)$. If $a < \infty$ we obtain from (5):

$$L\{P_1(t)\} = \Phi_1(s) = (1-h)+h\cdot\Phi(s). \tag{6}$$

Tellenbach applied this model to actual data on auto insurance, published by Thyrion (1961). In doing so, he determined t such that $\mu_1(t)$ matches the empirical mean $\hat{\mu}_1$. That is,

$$t = \alpha_1\mu_1(t) = \alpha h\mu_1(t) = \alpha h\hat{\mu}_1. \tag{7}$$

This leaves the free parameter h which may be chosen so that an appropriate error measure is minimized.

Tellenbach considered for $P(t)$ the negative exponential as well as the one-sided Gauss distribution. The results were superior in the latter case which, however, involves computational difficulties. Tellenbach solved the problem by Monte-Carlo techniques.

3. An Algorithm

We shall now assume that the distribution function $P(t)$ can be expressed as a power series of t which converges reasonably fast. This assumption will hold if t is small enough. The data of Thyrion, for instance, require according to equation (7) for the one-sided Gauss distribution:

$$t = \alpha h\mu = \sqrt{2/\pi}\cdot 0.214\cdot h = 0.171\cdot h.$$

Hoping for convergence we thus develop:

$$P'(t) = \alpha\, e^{-t^2/2} = \alpha(1-t^2/2+t^4/8-\ldots),\ \alpha = \sqrt{2/\pi}$$

Hence,

$$\Phi(s) = \alpha(1/s-1/s^3+3/s^5-\ldots),$$

and from (6):

$$\Phi_1(s) = (1-h)+h\alpha(1/s-1/s^3+3/s^5-\ldots).$$

The first equation of (4) yields:

$$\Phi_0(s) = 1/\alpha s - 1/s^2 + 1/s^4 - 3/s^6 - \ldots$$

$$1 - p_0(t) = t/\alpha - t^2/2! + t^4/4! - 3t^6/6! + \ldots$$

An algorithm to calculate the probabilities $p_k(t)$ for $k = 0, 1, 2\ldots$ and for the special value t obtained from condition (7) is described below. Please note that the three steps refer to the three equations (4).

Step 1 – Define: $\Phi_1(s) = z_0 + z_1/s + z_2/s^2 + \ldots$

Calculate: $\Phi_0(s) = (1 - \Phi_1)/\alpha_1 s = y_1/s + y_2/s^2 + \ldots$

Calculate: $p_0(t) = 1 - y_1 t - y_2(t^2/2!) - \ldots$

Step 2 – Calculate: $\Psi_1 = \Phi_0(1 - \Phi_1) = \big((1 - z_0)y_1\big)/s + \big((1 - z_0)y_2 - z_1 y_1\big)/s^2 + \ldots$

Redefine: $\Psi_1 = y_1/s + y_2/s^2 + \ldots$

Calculate: $p_1(t) = y_1 t + y_2 \cdot t^2/2! + \ldots$

Set: $k = 1$

Step 3 – Calculate: $\Psi_{k+1} = \mu_k \Phi_1 = z_0 y_2/s + (z_0 y_2 + z_1 y_1)/s^2 + \ldots$

Redefine: $\Psi_{k+1} = y_1/s + y_2/s^2 + \ldots$

Calculate: $p_{k+1}(t) = y_1 t + y_2 t^2/z! + \ldots$

Set: $k = k+1$, return to Step 3.

In the computer language APL, Step 3 would read:

$$W \leftarrow W,\ T + \cdot \times Y \leftarrow Z + \cdot \times Y. \qquad (8)$$

Please note the APL statements are evaluated from the right to the left. The symbols contained in (8) have the following meanings:

Y is a vector which holds the first n coefficients y_j of Ψ_k, to be replaced by the coefficients y_j of Ψ_{k+1}.

Z is a matrix built from the first n coefficients of Φ_1. For $n = 3$, Z looks like:

$$\begin{matrix} z_0 & 0 & 0 \\ z_1 & z_0 & 0 \\ z_2 & z_1 & z_0 \end{matrix}$$

T is a vector holding the values $t, t^2/2!, \ldots, t^n/n!$
W is a vector holding the values $p_k(t)$. Each time Step 3 is traversed, the value $p_{k+1}(t)$ is appended to W.

4. Numerical Results

The APL program described above runs fast even on a micro-computer. The main practical difficulty is that we do not know beforehand how far to extend the vector Z. The length of n of Z should be determined by the condition that

$$t^n/n! = (\alpha h \hat{\mu}_1)^n/n!$$

becomes insignificantly small. However, it is advisable to make additional control runs with increased values of n. Of course, the required length n of Z may exceed the computer space; then the method suggested in this paper has to be abandoned.

For the data considered by Tellenbach, final results were already achieved for $n = 5$.

Attached are copies of two sample runs; namely, for $h = 0.86$ and 0.83. The former run yields the lowest sum of error-squares; this is the measure used by Tellenbach. The latter run (with $h = 0.83$) approximates the actual second moment $\mu_2(t) = \sum k^2 p_k(t)$. The printouts apply the terminology of Tellenbach which differs from the terminology used in this paper as follows:

Tellenbach	This paper	Comments
$Q(s)$	$P_0(t)$	1 − Error Function
$P_1(s)$	$P(t)$	Gauss
p	h	

The printouts display:

$$\varepsilon = t^n/n!$$

$$P, T = h, t$$

$$\text{MOM} = \sum k^j p_k(t) \quad \text{for } j = 0, 1 \text{ and } 2$$

$$\text{ERR} = \sum (p_k(t) - \hat{p}_k)^j \quad \text{for } j = 1 \text{ and } 2$$

The columns show:

$$N = k$$
$$P[N] = p_K(t)$$
$$P-\underline{P} = p_K(t)-\hat{p}_K$$
$$(P-\underline{P})^{*}2 = (p_K(t)-\hat{p}_K)^2$$

Appendix 1 compares the actual number of claims (namely, $\hat{p}_K \cdot 9{,}461$) with the corresponding figures from Tellenbach and the runs for $h = 0.86$ and $h = 0.83$, respectively.
The run $h = 0.86$ comes reasonably close to the Monte-Carlo results of Tellenbach. This may justify the "naive" approach suggested in this paper.
The run $h = 0.83$ fits better to the tail than the run $h = 0.86$, but is less accurate for $k = 0$ through 4.

Appendix 1

		Number of Policies		
Number of Claims	Actual	Tellenbach (Monte-Carlo)	Computer-Algorithm h = .86	h = .83
0	7,840	7,831	7,819.1	7,872.9
1	1,317	1,311	1,327.6	1,242.1
2	239	255	255.6	271.8
3	42	54	47.9	58.5
4	14	9	8.8	12.4
5	4	1.3	1.6	2.6
6	4	0	0.3	0.5
7	1	0	0	0.1
	9,461	9,461.3	9,460.9	9,460.9
$\hat{\mu}_1$ =	0.214	0.214	0.214	0.214
$\hat{\mu}_2$ =	0.335	0.316	0.315	0.333
$\sum(p_n-\hat{p}_n)^2$ =		566	907	8,058

Note: Computer printout figures are multiplied by 9,461. For example: $0.826453 \cdot 9{,}461 = 7{,}819.1$.

0.83
*** Q=1-ERF *** P1=GAUSS ***

ε~4.8036E¯7
P,T = 0.83 0.14195
MOM = 1 0.21435 0.33249
ERR = ¯2.4438E¯8 9.0058E¯5

0	.832145	.003480	.000012
1	.131286	¯.007917	.000063
2	.028732	.003471	.000012
3	.006180	.001740	.000003
4	.001310	¯.000169	.000000
5	.000275	¯.000148	.000000
6	.000057	¯.000366	.000000
7	.000012	¯.000094	.000000
8	.000002	.000002	.000000
9	.000000	.000000	.000000
10	.000000	.000000	.000000
N	P[N]	P-P	(P-P)*2

ε~6.4503E¯14
P,T = 0.83 0.14195
MOM = 1 0.21435 0.3325
ERR = ¯1.0693E¯14 9.006E¯5

0	.832145	.003480	.000012
1	.131286	¯.007917	.000063
2	.028732	.003471	.000012
3	.006180	.001740	.000003
4	.001310	¯.000169	.000000
5	.000275	¯.000148	.000000
6	.000057	¯.000366	.000000
7	.000012	¯.000094	.000000
8	.000002	.000002	.000000
9	.000000	.000000	.000000
10	.000000	.000000	.000000
11	.000000	.000000	.000000
12	.000000	.000000	.000000
13	.000000	.000000	.000000
14	.000000	.000000	.000000
15	.000000	.000000	.000000
16	.000000	.000000	.000000
17	.000000	.000000	.000000
18	.000000	.000000	.000000
19	.000000	.000000	.000000
N	P[N]	P-P	(P-P)*2

0.86
*** Q=1-ERF *** P1=GAUSS ***

ε~5.7367E¯7
P,T = 0.86 0.14709
MOM = 1 0.21435 0.31451
ERR = ¯3.2897E¯8 1.0159E¯5

0	.826453	¯.002212	.000005
1	.140327	.001124	.000001
2	.027016	.001755	.000003
3	.005067	.000628	.000000
4	.000931	¯.000549	.000000
5	.000168	¯.000254	.000000
6	.000030	¯.000393	.000000
7	.000005	¯.000100	.000000
8	.000001	.000001	.000000
9	.000000	.000000	.000000
N	P[N]	P-P	(P-P)*2

ε~8.879E¯14
P,T = 0.86 0.14709
MOM = 1 0.21435 0.31452
ERR = ¯3.3131E¯15 1.0158E¯5

0	.826453	¯.002212	.000005
1	.140327	.001124	.000001
2	.027016	.001755	.000003
3	.005067	.000628	.000000
4	.000931	¯.000549	.000000
5	.000168	¯.000254	.000000
6	.000030	¯.000393	.000000
7	.000005	¯.000100	.000000
8	.000001	.000001	.000000
9	.000000	.000000	.000000
10	.000000	.000000	.000000
11	.000000	.000000	.000000
12	.000000	.000000	.000000
13	.000000	.000000	.000000
14	.000000	.000000	.000000
15	.000000	.000000	.000000
16	.000000	.000000	.000000
17	.000000	.000000	.000000
18	.000000	.000000	.000000
N	P[N]	P-P	(P-P)*2

References

Seal, H. (1969): Stochastic Theory of a Risk Business, John Wiley & Sons

Tellenbach, U. (1977): Berechnung der Verteilung der Schadenzahlen bei bekannter Verteilung der Wartzeiten, MSV 77.1, 35–46.

Thyrion, P. (1961): Contribution à l'étude du bonus pour un sinistre en assurance automobile, ASTIN Bull. 1, 142–162.

Gottfried Berger
Cologne Life Reinsurance
Company
POB 300
Stamford, Conn. 06904
USA

Zusammenfassung

In der Arbeit von Tellenbach (Mitteilungen 77/1) geht es darum, aus der Verteilung der Zwischenschadenzeiten auf die Anzahl Schäden im Intervall (0,t) zu schliessen. Dieses Problem lässt sich mittels der Laplacetransformierten theoretisch lösen. Insbesondere hat Tellenbach Zwischenankunftszeiten, deren Verteilung ein Atom im Nullpunkt hat, für die Beschreibung von Kumuleffekten verwendet. Numerisch hat er die auftretende Laplacetransformierte aber nicht invertiert, sondern sich mit Monte-Carlo-Methoden beholfen. *Berger* gibt nun einen konkreten Algorithmus zur Inversion der auftretenden Laplacetransformierten, welche er in eine Reihe entwickelt.

Résumé

Au début l'auteur présente un modèle introduit récemment par Tellenbach qui permet de tirer des conclusions sur le nombre des sinistres en (0, t) si on connaît la distribution entre deux sinistres consécutifs. Ce modèle demande des transformations du type Laplace-Stieltjes qui ne sont pas faciles à calculer numériquement.
Dans la deuxième partie on trouve un algorithme simple qui semble bien fonctionner dans le cas où l'intervalle de temps est suffisamment petit. Cet algorithme permet l'inversion de la transformation Laplace; l'article traite surtout des problèmes de calcul et moins des rapports mathématiques.

Riassunto

All'inizio l'autore presenta un modello, introdotto recentemente da Tellenbach, che permette delle conclusioni sul numero dei danni nell'intervallo (0, t) quando si conosce la distribuzione del tempo tra due danni consecutivi. Questo metodo richiede l'applicazione delle integrali Laplace-Stieltjes che non sono facili da calcolare. Poi viene proposto un algoritmo semplice, per l'inversione di trasformazioni di Laplace, che sembra funzionare bene nel caso dove (0, t) è piccolo.