On the computation of stop-loss premiums

Autor(en): **Gerber, Hans U.**

Objekttyp: **Article**

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

Band (Jahr): **77 (1977)**

PDF erstellt am: **22.07.2024**

Persistenter Link: https://doi.org/10.5169/seals-967010

On the Computation of Stop-Loss Premiums

By Hans U. Gerber

1. Introduction and definitions

The purpose of this note is to carry on some of the arguments that were introduced in [1], and, based on this, to improve the discretization algorithms that are discussed in [2].

If X denotes a random variable with *cdf* $F(x)$, $-\infty < x < \infty$, let

$$P(F, t, a) = \frac{1}{a} \ln E[e^{a(X-t)^+}] \tag{1}$$

for $a > 0$, and

$$P(F, t, 0) = E[(X-t)^+], \tag{2}$$

for $a = 0$.

where $-\infty < t < \infty$. For an arbitrary distribution F, $P(F, t, a)$ is well defined (possibly infinite). In both cases P should be interpreted as the stop-loss premium corresponding to a risk X and a deductible of t. In (2) it is the net premium; in (1) it is the one obtained from the exponential premium calculation principle with parameter $a > 0$. In terms of F, $P(F, t, a)$ equals

$$\frac{1}{a} \ln \{F(t) + \int_t^\infty e^{a(x-t)}\, dF(x)\} \tag{3}$$

$$= \frac{1}{a} \ln \{1 + a \cdot \int_t^\infty e^{a(x-t)} [1 - F(x)]\, dx\}$$

if $a > 0$, and

$$\int_t^\infty (x-t)\, dF(x) = \int_t^\infty [1 - F(x)]\, dx \tag{4}$$

if $a = 0$.

For $a \geq 0$ we define a *partial ordering* among distributions as follows. If G, H are *cdf*, we say that $G \overset{a}{<} H$, if $P(G, t, a) \leq P(H, t, a)$ for $-\infty < t < \infty$. From (3) and (4) we see that an equivalent condition is that

$$\int_t^\infty e^{a(x-t)} [1-G(x)]\, dx \leq \int_t^\infty e^{a(x-t)} [1-H(x)]\, dx \tag{5}$$

for all t.

It is instructive to consider $P(F, t, \infty) = \lim P(F, t, a)$, $a \to \infty$. This limit exists and equals $(r_F - t)+$, if $r_F = \sup\{x \mid F(x) < 1\}$ is finite. In this sense $G \overset{\infty}{<} H$ means that the right hand end point of the range of G is to the left of (or equal to) the right hand end point of the range of H. Thus in the limit we obtain a complete ordering.

Let K_a denote the set of *cdf* F, for which $P(F, t, a)$ is finite. For distributions in K_a we introduce the following *metric*. If $G, H \in K_a$, let

$$d_a(G, H) = \sup_t \left| \int_t^\infty e^{a(x-t)} [G(x) - H(x)]\, dx \right|. \tag{6}$$

From formulas (3) and (4) we see that

$$d_0(G, H) = \sup_t |P(G, t, 0) - P(H, t, 0)| \tag{7}$$

and that

$$d_a(G, H) = \frac{1}{a} \sup_t |e^{aP(G,t,a)} - e^{aP(H,t,a)}| \tag{8}$$

for $a > 0$. Thus the distance (6) is a useful tool for a comparison of stop-loss premiums.

2. Characterization of stop-loss premiums

Let I_a denote the collection of all stop-loss premiums considered functions of the deductible $t(-\infty < t < \infty)$:

$$I_a = \{p \mid p(t) \equiv P(F, t, a) \quad \text{for some} \quad F \in K_a\}. \tag{9}$$

It is easy to see that different *cdf* in K_a lead to different stop-loss premiums in I_a. In fact, given a function $p \in I_a$, the corresponding *cdf* F is given by the formula

$$F(x) = e^{a\,p(x)}(1+p'(x)), \quad -\infty < x < \infty, \tag{10}$$

which has the character of an inversion formula. Note that at points of discontinuity of F, p' should be interpreted as the right-side derivative.
Formula (10) and formulas (3) and (4), respectively, enable us to characterize I_a: A function $p(t)$, $-\infty < t < \infty$, belongs to I_a if and only if the following four conditions are satisfied.

(i) p is a continuous, non-increasing function.
(ii) $p(t) \to 0$ for $t \to \infty$, $p(t) \to \infty$ for $t \to -\infty$.
(iii) $e^{a\,p(t)}(1+p'(t))$ is a non-decreasing function.
(iv) $e^{a\,p(t)}(1+p'(t)) \to 0$ for $t \to -\infty$.

The characterization of net stop-loss premiums is particularly simple: A function p belongs to I_0, if and only if it is continuous, non-increasing, concave from above, such that $p(t) \to 0$ for $t \to \infty$ and $p'(t) \to -1$ for $t \to -\infty$. For $a > 0$, these conditions are still necessary (but not sufficient): Conditions (iii) and (iv) are the stronger, the larger a is; thus $I_a \subset I_b$ wherever $0 \leq b < a$.
Using formula (10), we can translate properties of F into properties of $p(t) = P(F, t, a)$, and vice versa. For example, if $F(x)$ is constant over an interval (x_1, x_2), its derivative vanishes and therefore p satisfies the differential equation.

$$p''(t) + a \cdot p'(t) \cdot (1 + p'(t)) = 0 \tag{11}$$

for $t \in (x_1, x_2)$. In terms of the boundary values $p(x_1)$, $p(x_2)$, the solution of this equation is

$$p(t) = \frac{1}{a} \cdot ln \left\{ \frac{(e^{-a x_1} - e^{-a t})\, e^{a\,p(x_2)} + (e^{-a t} - e^{-a x_2})\, e^{a\,p(x_1)}}{e^{-a x_1} - e^{-a x_2}} \right\} \tag{12}$$

$(x_1 \leq t \leq x_2)$, if $a > 0$. If $a = 0$, $p(t)$ is obtained by linear interpolation.

3. A sufficient condition for inequality

Let G, H be *cdf*'s in K_a satisfying the following two conditions:

(A) $\int_{-\infty}^{\infty} e^{a x}[G(x) - H(x)]\, dx \geq 0.$

(B) There is a β, $-\infty \leq \beta < \infty$, such that

$$-G(x) \leq H(x) \quad \text{for} \quad x < \beta$$

$$-G(x) \geq H(x) \quad \text{for} \quad x \geq \beta.$$

Then $G \overset{a}{<} H$.

Proof: We want to verify that inequality (5) holds for all t. If $t \geq \beta$, (5) is satisfied because of condition (A) alone. If $t < \beta$, we first use (B) and then (A) to show that the difference between the right side and the left side in (5) is non-negative:

$$\int_t^\infty e^{a(x-t)}[G(x)-H(x)]\,dx \geq \int_{-\infty}^\infty e^{a(x-t)}[G(x)-H(x)]\,dx \geq 0. \tag{13}$$

q.e.d.

Note that condition (A) is satisfied if $\int e^{ax}\,dG(x) \leq \int e^{ax}\,dH(x)$ (in the case $a > 0$), and if $\int x\,dG(x) \leq \int x\,dH(x)$ (in the case $a = 0$).

4. Examples

In all three examples we assume that F is a *cdf* that is concentrated on a finite interval $[x_1, x_2]$. Let $a \geq 0$.

Example 1. Let G denote the degenerate distribution whose mass is concentrated at the point x_m, where

$$x_m = \begin{cases} \frac{1}{a} \ln \int e^{ax}\,dF(x) & \text{if } a > 0 \\ \int x\,dF(x) & \text{if } a = 0. \end{cases} \tag{14}$$

Then condition (A) is satisfied (with equality holding), and condition (B) is satisfied with $\beta = x_m$. Therefore $G \overset{a}{<} F$.

Example 2. The idea of the previous example was to concentrate the mass of F at the point x_m. Now we look at the other extreme, which is to disperse the mass of F to the endpoints. Let H denote the two-point distribution that is concentrated at x_1 and x_2 such that

$$1-H(x-0) = \begin{cases} \dfrac{\int (e^{ax}-e^{ax_1})\,dF(x)}{e^{ax_2}-e^{ax_1}} & \text{if } a>0 \\[2ex] \dfrac{\int (x-x_1)\,dF(x)}{x_2-x_1} & \text{if } a=0 \end{cases} \tag{15}$$

and $H(x_1) = H(x_2-0)$. Then $F \overset{a}{<} H$, since again conditions (A) (with equality holding) and (B) are satisfied.

Example 3. Let H denote the mixture between the uniform distribution over (x_1, x_2) and the degenerate distribution with mass concentrated at either x_1 or x_2, for which

$$\int e^{ax}[F(x)-H(x)]\,dx = 0. \tag{16}$$

Thus condition (A) is satisfied. If we make the additional assumption that F is unimodal, condition (B) is also satisfied and $F \overset{a}{<} H$. This example is due to Verbeek (see [3] for $a = 0$).

5. Properties

The following result shows that the larger the parameter a is, the more distributions are comparable.

Proposition 1. If $0 \leq a < b$, $G \overset{a}{<} H$ implies that $G \overset{b}{<} H$.

Proof. Suppose that $H \in K_b$ (otherwise the statement is trivial). Then we integrate by parts to rewrite

$$\begin{aligned} &\int_t^\infty e^{b(x-t)}[1-H(x)]\,dx = \\ &-\int_t^\infty e^{(b-a)(x-t)} \frac{d}{dx} \int_x^\infty e^{a(y-t)}[1-H(y)]\,dy\,dx \end{aligned} \tag{17}$$

as

$$\begin{aligned} &\int_t^\infty e^{a(y-t)}[1-H(y)]\,dy + \\ &(b-a)\cdot \int_t^\infty e^{(b-a)(x-t)} \int_t^\infty e^{a(y-t)}[1-H(y)]\,dy\,dx. \end{aligned} \tag{18}$$

Suppose now that $G \overset{a}{<} H$, i.e., that inequality (5) holds for all t. Thus if we replace H by G in expression (18) we obtain a lower bound for any t, which in turn means that $G \overset{b}{<} H$.

q.e.d.

As an illustration let us consider G, the degenerate distribution with mass concentrated at 2, and H, the exponential distribution with parameter 1. Then $P(G, t, a) = (2-t)+$ and

$$P(H, t, a) = \begin{cases} \frac{1}{a} \cdot ln \frac{1}{1-a} - t & \text{if } t < 0 \\ \\ \frac{1}{a} \cdot ln \left\{ 1 + \frac{a}{1-a} e^{-t} \right\} & \text{if } t \geq 0 \end{cases} \tag{19}$$

if $0 < a < 1$ and infinite if $a \geq 1$.

Thus $G \overset{a}{<} H$ whenever $P(G, 0, a) \leq P(H, 0, a)$, i.e., whenever $(1-a)\, e^{2a} \leq 1$, or $a \geq .7968...$.

Proposition 2. If $G_i \overset{a}{<} H_i$ $(i = 1, 2, ...)$ and $\{p_i\}$ is a sequence of probability weights, then

a) $\sum_i p_i G_i \overset{a}{<} \sum_i p_i H_i$

b) $G_1 * \ldots * G_n \overset{a}{<} H_1 * \ldots * H_n$ for all n.

The proof is similar to the one given in [1] (for $a = 0$) and is omitted. The following result shows how mixing and convoluting affects the metric.

Proposition 3. Let F, G, H, G_i, H_i be *cdf*'s in K_a, and let $\{p_i\}$ be a sequence of probability weights. Then

a) $d_a\left(\sum_i p_i G_i, \sum_i p_i H_i\right) \leq \sum_i p_i\, d_a(G_i, H_i)$

b) $d_a(F * G, F * H) \leq d_a(G, H)$

c) $d_a(G^{*n}, H^{*n}) \leq n \cdot d_a(G, H)$.

Proof. a) Easy. b) for all t

$$\left|\int_t^\infty e^{a(x-t)}[F*G(x)-F*H(x)]\,dx\right|$$

$$=\left|\int\int_t^\infty e^{a(x-t)}[G(x-z)-H(x-z)]\,dx\,dF(z)\right| \tag{20}$$

$$\le\int\left|\int_t^\infty e^{a(x-t)}[G(x-z)-H(x-z)\,dx\right|dF(z)$$

$$\le d_a(G,H).$$

Thus, taking the supremum of the left side, we obtain the desired inequality. c) First we use the triangle inequality and then b) to get the estimate

$$d_a(G^{*n},H^{*n}) \tag{21}$$

$$\le\sum_{k=0}^{n-1} d_a(G*G^{*n-1-k}*H^{*k},H*G^{*n-1-k}*H^{*k})$$

$$\le n\cdot d_a(G,H).$$

q.e.d.

6. Application: Discretization

In this section we discuss numerical procedures for evaluating $P(F,t,a)$ for some $a\ge0$, if F (the distribution of aggregate claims) is of the form

$$F(x)=\sum_{i=0}^{\infty} p_i B^{*i}(x). \tag{22}$$

Here $p_i=\text{prob}(N=i)$, where N denotes the number of claims, and B denotes the *cdf* of individual claim amounts. It is assumed that both B and F are in K_a (in the special case where the claim number distribution is Poisson, $B\in K_a$ implies that $F\in K_a$).

For a given $d>0$, let K_{ad} denote the distributions in K_a that are arithmetic with span d, i.e., whose probability mass is concentrated at the points 0, $\pm d$, $\pm 2d$, ..., and let I_{ad} denote the corresponding subset of I_a.

The general idea is to replace the original claim amount distribution by a distribution $B_{\#} \in K_{ad}$, to do the calculations for

$$F_{\#}(x) = \sum_{i=0}^{\infty} p_i B_{\#}^{*i}(x), \tag{23}$$

and then to get information about $P(F, t, a)$ from $P(F_{\#}, t, a)$. Specifically, three methods are suggested:

I. Estimation of the difference

For an arbitrary $B_{\#} \in K_{ad}$ we can use the inequality

$$d_a(F, F_{\#}) \leq E[N] \cdot d_a(B, B_{\#}), \tag{24}$$

which follows from Proposition 3 (Parts a) and c)), to estimate the difference between $P(F, t, a)$ and $P(F_{\#}, t, a)$.

II. Upper bounds (method of dispersal)

If we select a $B_{\#} \in K_{ad}$ for which $B \overset{a}{<} B_{\#}$, Proposition 2 tells us that $F \overset{a}{<} F_{\#}$. This rases the question whether among the distributions in K_{ad} that dominate B there is a smallest one (in the sense of $\overset{a}{<}$). The answer is yes; we shall construct a $B_u \in K_{ad}$ and then verify that it satisfies these properties.
First let us write B as a mixture of conditional distributions:

$$B = \sum_{q_i \neq 0} q_i B_i, \tag{25}$$

where $q_i = B((i+1)d) - B(id)$ and

$$B_i(x) = \begin{cases} 0 & \text{if } \; x < id \\ [B(x) - B(id)]/q_i & \text{if } \; id \leq x < (i+1)d \\ 1 & \text{if } \; x \geq (i+1)d \end{cases} \tag{26}$$

Following Example 2, we replace each of these B_i's by the corresponding two-point distribution that is concentrated at the adjacent points id and $(i+1)d$. The mixture of these two-point distributions is a distribution $B_u \in K_{ad}$ for which $B \overset{a}{<} B_u$ (reason: Part a) of Proposition 2). Observe that the atoms $b_i = B_u(id) - B_u(id-0)$ of B_u are given by the following formula:

$$b_i = \frac{\int\limits_{(i-1)d}^{id} (e^{ax} - e^{a(i-1)d})\, dB(x)}{e^{aid} - e^{a(i-1)d}} + \frac{\int\limits_{id}^{(i+1)d} (e^{a(i+1)d} - e^{ax})\, dB(x)}{e^{a(i+1)d} - e^{aid}} \tag{27}$$

if $a > 0$. The expression for $a = 0$ can be obtained as an obvious limit and has already appeared elsewhere (formula (22) in [2]).

A short calculation shows that $P(B, t, a) = P(B_u, t, a)$ whenever t is a multiple of the span d. Let us now consider an arbitrary $B_{\#} \in K_{ad}$ for which $B \overset{a}{<} B_{\#}$. Thus $P(B_u, t, a) \leq P(B_{\#}, t, a)$ whenever t is a multiple of d. But, since the interpolation formula (12) is monotone in the boundary conditions, this means that this inequality holds for all t, i.e., that $B_u \overset{a}{<} B_{\#}$. So B_u is indeed the least element of K_{ad} that dominates B.

It is instructive to visualize this result graphically. For simplicity let us consider the case $a = 0$. A function $p_{\#} \in I_0$ is in I_{0d}, if and only if it is piecewise linear, the discontinuities of the first derivative being at the multiples of d. Given a $p \in I_0$, we find the smallest element in I_{0d} that dominates p by connecting successive points $(id, p(id))$ with linear line segments (reason: the graph of p is concave from above). Formula (10) with $a = 0$ tells us that the weights of p_u are obtained as the differences of successive slopes:

$$b_i = p_u'(id+0) - p_u'(id-0). \tag{28}$$

In terms of p, this means that

$$b_i = \frac{p((i+1)d) - 2p(id) + p((i-1)d)}{d} \tag{29}$$

or, since $p(t) = \int\limits_t^{\infty} [1 - B(x)]\, dx$,

$$b_i = \left\{ \int\limits_{(i-1)d}^{id} [1 - B(x)]\, dx - \int\limits_{id}^{(i+1)d} [1 - B(x)]\, dx \right\} / d. \tag{30}$$

Integrating by parts we see that this expression is consistent with formula (27).

III. Lower bounds

The idea is to chose a $B_{\#} \in K_{ad}$ for which $B_{\#} \overset{a}{<} B$. Then, because of Proposition 2, $F_{\#} \overset{a}{<} F$. In general, there is no biggest lower bound for B in K_{ad}; so it is not obvious which $B_{\#}$ should be chosen. We shall discuss two methods for this. In each case we first write B as a mixture of distributions, such that the equivalent point masses, according to formula (14), are located at multiples of the span d. Then, following Example 1, we replace these distributions by the corresponding degenerate distributions. By virtue of Part a) of Proposition 2 we obtain a lower bound for B.
In both cases we assume that $B(x) = 0$ for $x<0$, and that the probability for a claim of size zero is "sufficiently large". The latter condition is less artificial than it appears at first sight (see the remark below).
For the *truncation method* we write

$$B = \sum_{\substack{i=1 \\ q_i \neq 0}}^{\infty} (q_i + r_i) B_i + \left(1 - \sum (q_i + r_i)\right) C, \tag{31}$$

where $q_i = B((i+1)d-0) - B(id-0)$; $r_i \geq 0$ is defined by the equation

$$(q_i + r_i)\, e^{aid} = r_i + \int_{id}^{(i+1)d} e^{ax}\, dB(x) \tag{32}$$

(if $a>0$) and B_i from

$$(r_i + q_i) \cdot B_i(x) = \begin{cases} 0 & \text{if } \; x<0 \\ r_i & \text{if } \; 0 \leq x < id \\ r_i + B(x) - B(id-0) & \text{if } \; id \leq x < (i+1)d \\ r_i + q_i & \text{if } \; x \geq (i+1)d \end{cases} \tag{33}$$

Finally, the remainder C is a distribution that is concentrated on $[0, d)$. In the expression (31) we replace B_i by the degenerate distribution with mass concentrated at id, and C by the one with mass concentrated at 0, to get a lower bound B_t for B. ("Sufficiently large" means here that $\sum (q_i + r_i) \leq 1$).)
The basis of the truncation method was to combine the probability mass of an interval $[id, (i+1)d)$ with a point mass at zero of a suitable size. Alternatively, we may want to combine only masses of single intervals. This is accomplished by the *partition method*. We assume that B has a finite range; let x_0 denote its right hand end point. For simplicity let us assume that $B(x)$ is

continuous for $x \neq 0$. Then we determine sequentially $x_0 > y_0 > x_1 > y_1 > \dots$ as follows: Let y_i be the largest multiple of d less than x_i. Then let x_{i+1} be the smallest solution of

$$q_i e^{a y_i} = \int_{x_{i+1}}^{x_i} e^{ax}\, dB(x) \tag{34}$$

(if $a > 0$), where $q_i = B(x_i) - B(x_{i+1})$. This construction stops when $x_n \in (0, d)$ or when $x_n = 0$ (in the latter case we have to use part of the point mass at zero). Thus B can be written as the following mixture:

$$B = \sum_{i=0}^{n-1} q_i B_i + (1 - \sum q_i)\, C, \tag{35}$$

where

$$q_i B_i(x) = \begin{cases} 0 & \text{if } x < 0 \\ q_i - B(x_i) + B(x) & \text{if } x_{i+1} \leq x < x_i \\ q_i & \text{if } x \geq x_i \end{cases} \tag{36}$$

and the remainder C is a distribution that is concentrated on $(0, d)$. Now we replace B_i by the degenerate distribution with mass concentrated at y_i, and C by the degenerate distribution with mass concentrated at zero, to get a lower bound $B_l \in K_{ad}$ for B.

The partition method has an interesting geometric interpretation. For simplicity let us look at in the case $a = 0$. Given $p(t) = P(B, t, 0)$, one can construct $p_l(t) = P(B_l, t, 0)$ as follows: First $p_l(t) = 0$ for $t \geq y_0$. Then we draw the tangent line from the point $(y_0, 0)$ to the graph of p. The point of contact is $(x_1, p(x_1))$. Then we extend this tangent line to get to the point $(y_1, p_l(y_1))$. From this point we draw again the tangent line to the graph of p. The point of contact is now $(x_2, p(x_2))$, etc.

Remark. If the claim number distribution is mixed Poisson,

$$p_n = \int_0^\infty e^{-\theta} \frac{\theta^n}{n!}\, dU(\theta), \tag{37}$$

we can always generate claims of size zero (of a probability arbitrarily close to one) by a transformation of the structure function. If the new claim amount distribution should be $\tilde{B} = (1 - w)\, I + wB$ for some $0 < w < 1$ (where $I(x) = 0$

for $x<0$ and 1 for $x \geq 0$), the transformed structure function $\tilde{U}$ is given by the equation $\tilde{U}(\theta) = U(w\theta)$, $\theta \geq 0$. The distribution F remains invariant under these transformations. *Examples:* 1) If the claim number distribution is Poisson with parameter $\lambda > 0$, the transformed claim number distribution is again Poisson, namely with parameter $\tilde{\lambda} = \lambda/w$. 2) If the claim number distribution is negative binomial (i.e., if U is Gamma), say with parameters $\alpha > 0$ and $0 < p < 1$, the transformed claim number distribution is again negative binomial, namely, with parameters $\tilde{\alpha} = \alpha$, and

$$\tilde{p} = \frac{p}{1-q(1-w)} \tag{38}$$

where $q = 1-p$.

References

[1] *H. Bühlmann, B. Gagliardi, H. Gerber, E. Straub:* Some inequalities for stop-loss premiums ASTIN Bulletin (1977), Vol. IX, 75–83.

[2] *H. Gerber, D. Jones:* Some practical considerations in connection with the calculation of stop-loss premiums. Transactions of the Society of Actuaries (1976), Vol. XXVIII, 215–235.

[3] *H. Verbeek:* A stop-loss inequality for compound Poisson processes with unimodal claim size distribution. ASTIN Bulletin (1977), Vol. IX, 247–256.

Zusammenfassung

Eine Familie von Abständen und Halbordnungen unter Verteilungen wird eingeführt, und ihre Eigenschaften unter Mischung und Faltung werden untersucht. Diese Begriffe werden sodann angewendet zur Abschätzung des Fehlers, der entsteht, wenn die Schadenhöhenverteilung diskretisiert wird bei der Berechnung einer Stop-Loss-Prämie.

Résumé

L'auteur introduit une famille d'écarts et d'ordres partiaux parmi plusieurs distributions et étudie leurs propriétés sous mixtures et convolutions. Certaines sont employées ensuite pour évaluer l'erreur qui se produit quand la distribution des sommes de sinistre est discrétisée pour le calcul numérique.

Riassunto

Si introduce una famiglia di distanze e ordini parziali fra distribuzioni e si esaminano le loro qualità sotto mescolanza e convoluzione. Questi concetti si applicano poi per valutare l'errore che risulta discretizzando la distribuzione dell'ammontare dei sinistri nel calcolo del premio stop loss.

Summary

A family of distances and partial orderings among distributions is introduced, and their properties under mixing and convolution are discussed. These concepts are then applied to estimate the error that results when the claim amount distribution is discretizised for calculating a stop-loss premium.