The Mathematics of Excess Losses

by Leigh J. Halliwell

ABSTRACT

After laying a fairly rigorous foundation for the mathematical treatment of excess losses, this paper shows that the excess-loss function is akin to the probability distribution of its loss. All the moments of the loss can be reclaimed from the excess-loss function, the variance being especially simple. Excess-loss mathematics is a powerful tool for pricing loss layers, as in reinsurance. In some settings it may be more powerful than standard probability techniques. An example featuring the mixed exponential distribution demonstrates this. Two appendices deal with Stieltjes integrals and with proofs of two findings about layered losses that are commonly known among reinsurance actuaries.

KEYWORDS

Excess loss, loss layer, mixed exponential distribution, reinsurance

1. Introduction

The concept of excess loss is widely used and appreciated within the casualty-actuarial world, particularly in reinsurance, but also in retrospective and increased-limits rating. This paper will reveal some of its hidden mysteries. In the next section we will define the excess-loss function, and state its sometime advantages over standard probability concepts. This will prepare us, in the third section, to find new applications of the function, applications extending as far as moment generation. Excess losses naturally imply loss layers, whose probability distributions we will show in the fourth section to be more amenable to an excess-loss treatment than to standard probability theory. Here also we will introduce an example involving the first two moments of a mixed exponential distribution. Then we will, in the fifth section, round out the second moment of the example by considering the covariances among loss layers, and will conclude in the sixth section.

2. The excess-loss function

Let X be a non-negative random variable, i.e., a random variable suitable for representing an amount of loss. Its cumulative distribution function $F_X(a) = Prob[X \le a]$ has the following four properties:

1. If $a < b$, then $F_X(a) \le F_X(b)$	non-decreasing
2. $\lim\limits_{a \to \infty} F_X(a) = 1$	total probability
3. $\lim\limits_{b \to a^+} F_X(b) = F_X(a)$	continuity from the right
4. $\lim\limits_{a \to 0^-} F_X(a) = 0$	non-negative

These properties allow for points of probability mass, since

$$\begin{aligned} Prob[X = a] &= \lim_{b \to a^-} Prob[b < X \le a] \\ &= F_X(a) - \lim_{b \to a^-} F_X(b) \\ &\ge 0. \end{aligned}$$

Of particular note, the probability for X to equal 0 may be positive. Moreover, let G_X be the survival function, the complement of F_X: $G_X(a) = 1 - F_X(a) = Prob[X > a]$. Hence, $dG_X(a) = -dF_X(a)$. G_X is also continuous from the right, and $Prob[X = a] = \lim\limits_{b \to a^-} G_X(b) - G_X(a)$. However, it is non-increasing, its limit at infinity is zero, and $G_X(a) = 1$ for $a < 0$.

For $r \ge 0$, the expected portion of loss X in excess of "retention" r is defined as:

$$\begin{aligned} Excess_X(r) &\equiv \int_{x=-\infty}^{\infty} max(0, x - r)\, dF_X(x) \\ &= \int_{x=r}^{\infty} (x - r)\, dF_X(x). \end{aligned}$$

In particular, $Excess_X(0) = \int_{x=0}^{\infty} x\, dF_X(x) = E[X]$. From integration by parts, we reformulate:

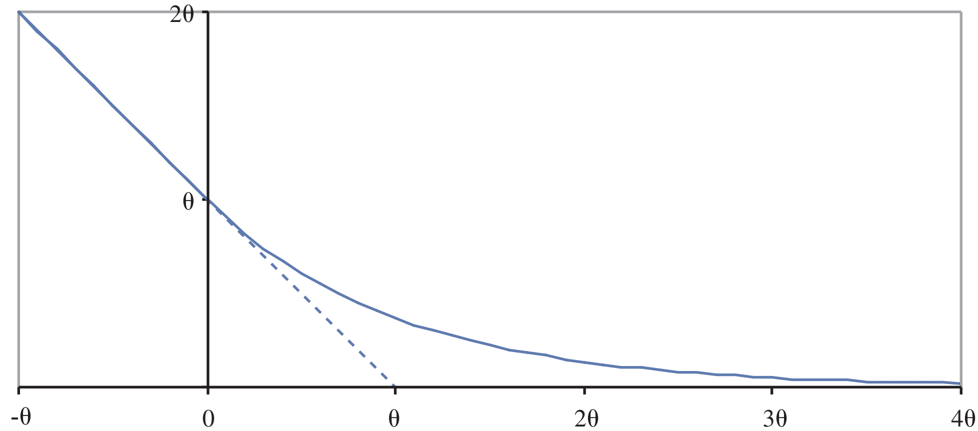
$$\begin{aligned} Excess_X(r) &= \int_{x=r}^{\infty} (x - r)\, dF_X(x) \\ &= -\int_{x=r}^{\infty} (x - r)\, dG_X(x) \\ &= -(x - r)G_X(x)\Big|_r^{\infty} + \int_{x=r}^{\infty} G_X(x)\, d(x - r) \\ &= (x - r)G_X(x)\Big|_{\infty}^{r} + \int_{x=r}^{\infty} G_X(x)\, dx \\ &= 0 - 0 + \int_{x=r}^{\infty} G_X(x)\, dx. \end{aligned}$$

From their studies, casualty actuaries are familiar with this expression.[1]

The excess-loss function is especially useful in reinsurance: if $0 \le a \le b$, the pure premium for the portion of loss X in the layer [a, b] equals $Excess_X(a) - Excess_X(b)$. Note that since G is dimensionless, the unit of the excess-loss function is the unit of dx, which is the unit of the loss amount and the retention. Two appealing properties of the excess-loss function are (1) that it is everywhere continuous, even where the probability density is discrete, and (2) that if it is

[1]It is ably presented and illustrated in Lee (1988), a CAS Exam 9 reading. We interpret all integrals as Stieltjes integrals, the subtleties of which are treated in Appendix A.

Figure 1. Excess-loss function of an Exponential(θ) loss



positive, it strictly decreases. Moreover, its derivative at r, if it exists, equals $-G_X(r)$. Even if it does not exist, at least the left and right derivatives exist, and the difference of the left derivative from the right is the probability mass at r. It is helpful to extend the definition of the excess-loss function to negative retentions, at which $G_X(r) = 1$. For $r < 0$:

$$Excess_X(r) = \int_{x=r}^{\infty} G_X(x)dx = \int_{x=r}^{0} G_X(x)dx + \int_{x=0}^{\infty} G_X(x)dx$$

$$= \int_{x=r}^{0} 1 \cdot dx + Excess_X(0) = -r + E[X].$$

Often useful is the form, valid for all r, $Excess_X(r) = -min(0, r) + \int_{x=max(0,r)}^{\infty} G_X(x)dx$.

It is not difficult to prove two basic theorems. If $c > 0$, then for all r:

$$Excess_{X+c}(r) = Excess_X(r-c)$$

$$Excess_{cX}(r) = cExcess_X(r/c).$$

The first holds true for $c = 0$; the second holds true as well in the limit as $c \to 0^+$, from which it follows that $Excess_0(r) = -min(0, r)$. Furthermore, $Excess_c(r) = -min(0, r - c) = max(0, c - r)$.

As an example, let X be exponentially distributed with mean θ. Hence, $G_X(x) = \begin{cases} e^{-\frac{x}{\theta}} & 0 \le x \\ 1 & x \le 0 \end{cases}.$

So for $r \ge 0$, $Excess_X(r) = \int_{x=r}^{\infty} G_X(x)dx = \int_{x=r}^{\infty} e^{-\frac{x}{\theta}}dx = \theta e^{-\frac{x}{\theta}}\Big|_{\infty}^{r} = \theta e^{-\frac{r}{\theta}}$, and with negative retentions:

$$Excess_X(x) = \begin{cases} \theta e^{-\frac{x}{\theta}} & 0 \le x \\ -x + \theta & x < 0 \end{cases}.$$

Figure 1 graphs this function over the domain $[-\theta, 4\theta]$. In preparation for the next section, we've also extended as a dotted line the negative-retention line, i.e., $f(x) = -x + \theta$.

The straight line itself is the graph of $Excess_\theta(x)$. It is indicative of positive variance that the excess-loss function "pulls up and comes in for a landing" to the right of the dotted line.[2] This is the clue for extracting more information from the excess-loss function.

3. Higher moments and the excess-loss function

The excess-loss function $Excess_X(r) = \int_{x=r}^{\infty} G_X(x)dx$ is not only elegant and useful, it is "full-informational" in the sense that one can derive from it all the moments of X. Above, we saw that the mean of X equals

[2]The excess-loss curve must stay on or above the line. Otherwise, at some value ξ the slope of the excess-loss curve would have to be less than negative one, or equivalently, $G_X(\xi) > 1$, a probability contradiction.

$Excess_X(0)$. That the dimension of the area under the excess-loss function is the square of the loss unit suggests that the area has something to do with the second moment. The following derivation relies on the validity of inverting the order of integration over the region A, which is the part of the first quadrant of the Cartesian plane above the line $y = x$:

$$\int_{x=0}^{\infty} Excess_X(x)\,dx = \int_{x=0}^{\infty}\int_{y=x}^{\infty} G_X(y)\,dy\,dx$$
$$= \iint_A G_X(y)\,dA$$
$$= \int_{y=0}^{\infty}\int_{x=0}^{y} G_X(y)\,dx\,dy$$
$$= \int_{y=0}^{\infty} G_X(y)\left(\int_{x=0}^{y} dx\right)dy$$
$$= \int_{y=0}^{\infty} G_X(y)\,y\,dy$$
$$= \frac{1}{2}\int_{y=0}^{\infty} G_X(y)\,d(y^2)$$
$$= \left.\frac{y^2 G_X(y)}{2}\right|_0^{\infty} - \frac{1}{2}\int_{y=0}^{\infty} y^2\,dG_X(y)$$
$$= 0 - 0 + \frac{1}{2}\int_{y=0}^{\infty} y^2\,dF_X(y)$$
$$= \frac{1}{2}E[X^2].$$

So the area under the excess-loss function in the first quadrant equals half the second moment. Therefore, the area under the excess-loss function but above the dotted line and the x-axis is

$$\frac{1}{2}E[X^2] - \frac{1}{2}E[X]E[X] = \frac{1}{2}\left(E[X^2] - E[X]^2\right)$$
$$= \frac{1}{2}Var[X].$$

Hence, for the excess-loss function to "land" to the right of $E[X]$ indicates a non-trivial variance.

In general, for $h(x)$ continuous on $[0, \infty)$,

$$\int_{x=0}^{\infty} Excess(x)\,dh(x) = \int_{x=0}^{\infty}\int_{y=x}^{\infty} G_X(y)\,dy\,dh(x)$$
$$= \int_{y=0}^{\infty}\int_{x=0}^{y} G_X(y)\,dh(x)\,dy$$
$$= \int_{y=0}^{\infty} G_X(y)\{h(y) - h(0)\}\,dy$$
$$= \int_{y=0}^{\infty} G_X(y)h(y)\,dy - \int_{y=0}^{\infty} G_X(y)h(0)\,dy$$
$$= \int_{y=0}^{\infty} G_X(y)h(y)\,dy - E[X]h(0).$$

Instead of the double integration, we may use integration by parts:

$$\int_{x=0}^{\infty} Excess_X(x)\,dh(x) = Excess_X(x)h(x)\Big|_0^{\infty}$$
$$- \int_{x=0}^{\infty} h(x)\,dExcess_X(x)$$
$$= 0 - Excess_X(0)h(0)$$
$$- \int_{x=0}^{\infty} h(x)(-G_X(x)\,dx)$$
$$= \int_{x=0}^{\infty} G_X(x)h(x)\,dx - E[X]h(0).$$

Now let $H'(x) = h(x)$, or $dH(x) = h(x)dx$. The following derivation relies on a formula from Appendix A, viz., $E[h(X)] = h(0) + \int_{x=0}^{\infty} G_X(x)\,dh(x)$:

$$\int_{x=0}^{\infty} Excess_X(x)\,dh(x) = \int_{x=0}^{\infty} G_X(x)h(x)\,dx - E[X]h(0)$$
$$= \int_{x=0}^{\infty} G_X(x)\,dH(x) - E[X]H'(0)$$
$$= H(0) + \int_{x=0}^{\infty} G_X(x)\,dH(x)$$
$$- H(0) - E[X]H'(0)$$
$$= E[H(X)] - H(0) - E[X]H'(0).$$

The invariance of this formula to the addition of a linear function to H confirms its correctness. For if $H(x) \leftarrow H(x) + cx + d$, then $h(x) \leftarrow h(x) + c$ and

$$\begin{aligned}\int_{x=0}^{\infty} Excess_X(x)d(h(x)+c) &= E[H(X)+cX+d] \\ &\quad -(H(0)+c\cdot 0+d) \\ &\quad -E[X](H'(0)+c) \\ &= E[H(X)]+cE[X] \\ &\quad +d-H(0)-d \\ &\quad -E[X]H'(0)-cE[X] \\ &= \int_{x=0}^{\infty} Excess_X(x)dh(x).\end{aligned}$$

So, for example, if $h(x) = x^k$, where k must be positive for continuity at zero, then $H(x) = \frac{x^{k+1}}{k+1}$ and

$$\begin{aligned}\int_{x=0}^{\infty} Excess_X(x)dx^k &= E\left[\frac{X^{k+1}}{k+1}\right] - \frac{0^{k+1}}{k+1} \\ &\quad - E[X]\cdot 0^k = \frac{1}{k+1}E[X^{k+1}].\end{aligned}$$

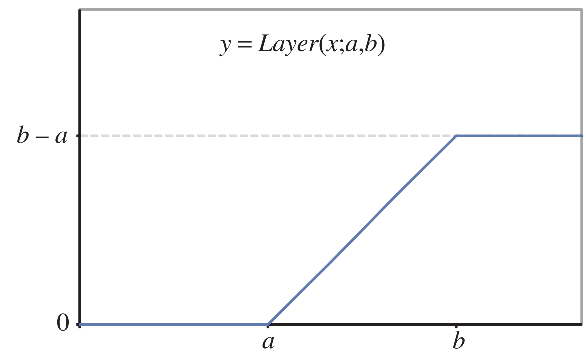
So, the second and higher moments result from integrals of the excess-loss function. Of course, the first moment is just $Excess_X(0)$.

Especially interesting is the integral $\int_{x=0}^{\infty} Excess_X(x)\,dte^{tx}$, for which $h(x) = te^{tx}$ and $H(x) = e^{tx}$:

$$\begin{aligned}\int_{x=0}^{\infty} Excess_X(x)dte^{tx} &= E[e^{tx}] - e^{t\cdot 0} - E[X]\cdot te^{t\cdot 0} \\ &= E[e^{tx}] - 1 - E[X]\cdot t \\ &= M_X(t) - 1 - M'_X(0)\frac{t^1}{1!} \\ &= \sum_{i=2}^{\infty} M_X^{[i]}(0)\frac{t^i}{i!}.\end{aligned}$$

This integral of the excess-loss function reproduces the moment-generating function of X, except for the removal of the first two terms of its Maclaurin series. The whole moment-generating function of X is recovered as $M_X(t) = 1 + Excess_X(0)\cdot t + \int_{x=0}^{\infty} Excess_X(x)dte^{tx}$. This implies that for $t \neq 0$, $\int_{x=0}^{\infty} Excess_X(x)de^{tx} = \sum_{i=2}^{\infty} M_X^{[i]}(0)\frac{t^{i-1}}{i!} = \sum_{i=1}^{\infty} \frac{M_X^{[i+1]}(0)}{i+1}\frac{t^i}{i!}$, which holds true even for $t = 0$. Therefore, $1 + \int_{x=0}^{\infty} Excess_X(x)de^{tx}$ is the moment-generating function of the random variable Y whose moments relate to those of X as $E[Y^i] = \frac{E[X^{i+1}]}{i+1}$. To use a musical analogy, $\int_{x=0}^{\infty} Excess_X(x)de^{tx}$ is a transformation that "plays" X one octave lower.[3]

Figure 2. Layer function



4. Excess-loss functions of layered losses

If X is a non-negative random variable and $0 \leq a \leq b$, then define the portion of loss X in the layer [a, b] as $Layer(X; a, b) \equiv \min(b - a, \max(0, X - a))$. The graph in Figure 2 shows that the layer function is flat except in the interval [a, b], in which $Layer(x; a, b) = x - a$.

Our purpose in this section is to express $Excess_Y$ in terms of $Excess_X$, which requires expressing G_Y in terms of G_X.

Of course, if $r < 0$, then $G_Y(r) = 1$. For $b - a \leq r$, $G_Y(r) = Prob[Y > r] = Prob[Y > b - a] = 0$. And in the middle, for $0 < r < b - a$:

[3]How this relates to the coderived distributions of Corro (2008) we intend to examine in a subsequent paper.

$$G_Y(r) = Prob[Y > r] = Prob[X - a > r]$$
$$= Prob[X > r + a] = G_X(r + a).$$

By continuity from the right, $G_Y(0) = G_Y(0^+) = G_X(0^+ + a) = G_X(a)$. So the final conditional formula is

$$G_Y(y) = \begin{cases} 1 & y < 0 \\ G_X(y + a) & 0 \le y < b - a. \\ 0 & b - a \le y \end{cases}$$

Hence, for $0 \le r < b - a$:

$$Excess_Y(r) = \int_{y=r}^{\infty} G_Y(y)dy$$
$$= \int_{y=r}^{b-a} G_Y(y)dy$$
$$= \int_{y=r}^{b-a} G_X(y+a)dy$$
$$= \int_{y+a=r+a}^{b-a+a} G_X(y+a)d(y+a)$$
$$= \int_{x=r+a}^{b} G_X(x)dx$$
$$= Excess_X(r + a) - Excess_X(b).$$

The form $Excess_Y(r) = Excess_X(\min(b, r + a)) - Excess_X(b)$ is valid for all $r \ge 0$; it also accommodates the three limiting cases $a = 0$, $b = a$, and $b = \infty$.[4] Furthermore, $E[Y] = Excess_Y(0) = Excess_X(a) - Excess_X(b)$, as mentioned in Section 2.

[4]The form $Excess_Y(r) = -min(0, r) + Excess_X(min(b, max(a, r + a))) - Excess_X(b)$ is the most general, valid even for negative r. From it we may derive the following elegant and insightful forms:

$$Excess_Y(r) = -min(0, r) + Excess_X(a + min(b - a, max(0, r)))$$
$$- Excess_X(b)$$
$$= -min(0, r) + Excess_X(a + Layer(r + a; a, b))$$
$$- Excess_X(b)$$
$$= -min(0, r) + Excess_X(a + Layer(r + a; a, b))$$
$$- Excess_X(a + Layer(\infty + a; a, b)).$$

The tables provide an example. First, we mixed four exponential distributions (with weights w and means θ at the top of Table 1).[5] The excess-loss function of the mixed exponential distribution (the "Mxd-Exp Excess" column) is $Excess_{MX}(r) = \sum_{i=1}^{4} w_i \theta_i e^{-\frac{r}{\theta_i}}$. Its values are shown for retentions from zero to 50 million (consider the unit of loss as USD) in steps of one million; the values are also equal to the values of the previous four columns (gray-shaded), weighted according 0.500, 0.250, 0.125, and 0.125. The mean loss is 1,375,000. The final column of Table 1 shows the area under the $Excess_{MX}$ curve from r to infinity. Its formula is

$$\int_{x=r}^{\infty} Excess_{MX}(x)dx = \int_{x=r}^{\infty} \sum_{i=1}^{4} w_i \theta_i e^{-\frac{r}{\theta_i}} dx$$
$$= \sum_{i=1}^{4} w_i \theta_i \int_{x=r}^{\infty} e^{-\frac{r}{\theta_i}} dx = \sum_{i=1}^{4} w_i \theta_i^2 e^{-\frac{r}{\theta_i}}.$$

The total area, which according to Section 3 is $E[MX^2]/2$, is 4.000×10^{12} (USD squared). Therefore the variance of the mixed exponential loss is $2 \times 4.000 \times 10^{12} - 1{,}375{,}000^2$, for a standard deviation of 2,471,715.

Table 2 partitions the support of MX into four non-overlapping layers: [0, 5M], [5M, 10M], [10M, 20M], and [20M, ∞).[6] Let Y_i denote the portion of loss MX in the i^{th} layer. Because of the non-overlapping and complete partition, $MX = \sum_i Y_i$. The excess losses are calculated according to the formula $Excess_Y(r) = Excess_X(\min(b, r+a)) - Excess_X(b)$. The formula for the infinite top layer is simpler: $Excess_Y(r)$

[5]The mixed exponential distribution, which ISO and NCCI have incorporated into their increased limits factors and excess loss factors, was introduced into the actuarial literature by Keatinge (1999). For its use by NCCI, see Corro and Engl (2006). Not only is this distribution versatile (see Keatinge 1999, p. 657); it is also mathematically tractable, hence well-suited as an example.

[6]In reinsurance parlance, the lower bound of an interval is called the retention; the width of an interval is the limit. A layer is usually identified by its retention and limit, e.g., [2M, 5M] would be stated as "3M in excess of 2M," or "3M xs 2M," (our nomenclature in Exhibit 3). Sometimes, especially in Europe, the upper bound is called the plafond (French for ceiling). The reader should note that the example treats of an aggregate loss, as would be reinsured under a stop-loss treaty. More usual is per-claim or per-occurrence layering under a specific-excess treaty.

Table 1. Mixed-exponential excess losses

Wgt (w)	0.500	0.250	0.125	0.125	1.000	
Mean (θ)	500,000	1,000,000	2,000,000	5,000,000	1,375,000 ± 2,471,715	
Retention (r)	Exponential Excess Losses				Mxd-Exp Excess	$\int_r^\infty Excess\,dx$
0	500,000	1,000,000	2,000,000	5,000,000	1,375,000	4.000E+12
1,000,000	67,668	367,879	1,213,061	4,093,654	789,143	2.971E+12
2,000,000	9,158	135,335	735,759	3,351,600	549,333	2.315E+12
3,000,000	1,239	49,787	446,260	2,744,058	411,856	1.839E+12
4,000,000	168	18,316	270,671	2,246,645	319,327	1.476E+12
5,000,000	23	6,738	164,170	1,839,397	252,142	1.192E+12
6,000,000	3	2,479	99,574	1,505,971	201,314	9.667E+11
7,000,000	0	912	60,395	1,232,985	161,901	7.859E+11
8,000,000	0	335	36,631	1,009,483	130,848	6.402E+11
9,000,000	0	123	22,218	826,494	106,120	5.221E+11
10,000,000	0	45	13,476	676,676	86,280	4.263E+11
11,000,000	0	17	8,174	554,016	70,278	3.483E+11
12,000,000	0	6	4,958	453,590	57,320	2.847E+11
13,000,000	0	2	3,007	371,368	46,797	2.329E+11
14,000,000	0	1	1,824	304,050	38,234	1.905E+11
15,000,000	0	0	1,106	248,935	31,255	1.559E+11
16,000,000	0	0	671	203,811	25,560	1.275E+11
17,000,000	0	0	407	166,866	20,909	1.044E+11
18,000,000	0	0	247	136,619	17,108	8.545E+10
19,000,000	0	0	150	111,854	14,000	6.995E+10
20,000,000	0	0	91	91,578	11,459	5.726E+10
21,000,000	0	0	55	74,978	9,379	4.687E+10
22,000,000	0	0	33	61,387	7,678	3.838E+10
23,000,000	0	0	20	50,259	6,285	3.142E+10
24,000,000	0	0	12	41,149	5,145	2.572E+10
25,000,000	0	0	7	33,690	4,212	2.106E+10
26,000,000	0	0	5	27,583	3,448	1.724E+10
27,000,000	0	0	3	22,583	2,823	1.412E+10
28,000,000	0	0	2	18,489	2,311	1.156E+10
29,000,000	0	0	1	15,138	1,892	9.461E+09
30,000,000	0	0	1	12,394	1,549	7.746E+09
31,000,000	0	0	0	10,147	1,268	6.342E+09
32,000,000	0	0	0	8,308	1,039	5.192E+09
33,000,000	0	0	0	6,802	850	4.251E+09
34,000,000	0	0	0	5,569	696	3.481E+09
35,000,000	0	0	0	4,559	570	2.850E+09
36,000,000	0	0	0	3,733	467	2.333E+09
37,000,000	0	0	0	3,056	382	1.910E+09
39,000,000	0	0	0	2,049	256	1.280E+09
40,000,000	0	0	0	1,677	210	1.048E+09
41,000,000	0	0	0	1,373	172	8.583E+08
42,000,000	0	0	0	1,124	141	7.027E+08
43,000,000	0	0	0	921	115	5.753E+08
44,000,000	0	0	0	754	94	4.710E+08
45,000,000	0	0	0	617	77	3.857E+08
46,000,000	0	0	0	505	63	3.157E+08
47,000,000	0	0	0	414	52	2.585E+08
48,000,000	0	0	0	339	42	2.117E+08
49,000,000	0	0	0	277	35	1.733E+08
50,000,000	0	0	0	227	28	1.419E+08

Table 2. Layered losses and moments

	Layers			
	0	5,000,000	10,000,000	20,000,000
	5,000,000	10,000,000	20,000,000	∞
Retention	Excess Losses in Layer			
0	1,122,858	165,861	74,822	11,459
1,000,000	537,001	115,034	58,819	9,379
2,000,000	297,191	75,620	45,861	7,678
3,000,000	159,715	44,568	35,339	6,285
4,000,000	67,185	19,840	26,776	5,145
5,000,000	0	0	19,797	4,212
6,000,000			14,102	3,448
7,000,000			9,451	2,823
8,000,000			5,650	2,311
9,000,000			2,542	1,892
10,000,000			0	1,549
11,000,000				1,268
12,000,000				1,039
13,000,000				850
14,000,000				696
15,000,000				570
16,000,000				467
17,000,000				382
18,000,000				313
19,000,000				256
20,000,000				210
21,000,000				172
22,000,000				141
23,000,000				115
24,000,000				94
25,000,000				77
26,000,000				63
27,000,000				52
28,000,000				42
29,000,000				35
30,000,000				28
E[Y]	1,122,858	165,861	74,822	11,459
Area	1.547E+12	3.347E+11	2.545E+11	5.726E+10
Var[Y]	1.833E+12	6.418E+11	5.033E+11	1.144E+11
Std[Y]	± 1,353,906	± 801,119	± 709,449	± 338,211
CV[Y]	1.21	4.83	9.48	29.52

$= Excess_X(\min(\infty, r + 20M)) - Excess_X(\infty) = Excess_X(r + 20M)$. As expected, the sum of the means of the layered losses, $E[Y_i] = Excess_{Y_i}(0)$, equals 1,375,000; the partitioning conserves the first moment of the loss.

Table 2, like Table 1, derives the second moment of each layered loss from the area under its excess-loss curve. But, as one of the advantages of the excess-loss function, it is not necessary to do this anew; it is implicit in Table 1. For algebraically:

$$\frac{E(Y_i^2)}{2} = \int_{y=0}^{\infty} Excess_{Y_i}(y)dy$$

$$= \int_{y=0}^{\infty} \{Excess_{MX}(min(b_i, y + a_i)) - Excess_{MX}(b_i)\}dy$$

$$= \int_{y=0}^{b_i - a_i} \{Excess_{MX}(y + a_i) - Excess_{MX}(b_i)\}dy$$

$$= \int_{y=0}^{b_i - a_i} Excess_{MX}(y + a_i)dy - (b_i - a_i)Excess_{MX}(b_i)$$

$$= \int_{x=a_i}^{b_i} Excess_{MX}(x)dx - (b_i - a_i)Excess_{MX}(b_i)$$

$$= \int_{x=a_i}^{\infty} Excess_{MX}(x)dx - \int_{x=b_i}^{\infty} Excess_{MX}(x)dx - (b_i - a_i)Excess_{MX}(b_i).$$

The values of the last two integrals are those of the last column of Table 1 at retentions a_i and b_i.[7] The "Area" row at the bottom of Table 2 contains these $E[Y_i^2]/2$ values, from which follow the variances and standard deviations. It is well known among reinsurance actuaries that the coefficients of variation, $CV = Std/E$, increase as the layers ascend.[8]

[7]The algebra applies in the limit to the top level as well, since $\lim_{b\to\infty}(b - a)Excess_{MX}(b) = 0$. Otherwise, the variance of MX would be infinite. The relation of the third and higher moments of the excess losses to those of the loss from the ground up is complicated by the presence of powers of x in their integrals; yet they too are implicit. Klugman, Panjer, and Willmot (1998, p. 604) and Patrik (1996, p. 378) treat the second moment in terms of limited expectations of $X \wedge u = min(u, X)$. With the identities (derivable by the reader) $E[X \wedge u] = Excess_X(0) - Excess_X(u)$ and $E[(X \wedge u)^2] = 2\int_{x=0}^{u} Excess_X(x)dx - 2uExcess_X(u)$, their formulations can be converted into ours.

[8]See Appendix B for a proof.

The sum of the four areas, 2.193×10^{12}, does not equal the 4.000×10^{12} of the MX area; nor is the sum of the four variances, 3.093×10^{12}, equal to the variance of MX, 6.109×10^{12}. What is lacking in the conservation of the second moment is the covariance among the layered losses, to which we now turn.

5. Covariances among non-overlapping layers

Since $Cov[Y_i, Y_j] = E[Y_iY_j] - E[Y_i]E[Y_i]$, and the means are known, we need a formula for the product moments $E[Y_iY_j]$, where $i \neq j$. The actual formula, based on the loss variable of which they are layers, is $E[Y_iY_j] = \int_{x=-\infty}^{\infty} y_i(x)y_j(x)dF(x)$. A formal derivation is not necessary; the following argument will suffice. Since the layers are different but non-overlapping, one is above the other. The range of the integration may be restricted to the range over which the integrand $y_i(x)y_j(x)$ is non-zero, which range is the intersection of the non-zero ranges of $y_i(x)$ and $y_j(x)$ separately. However, due to the non-overlapping layering, the range of the higher layer must be a subset of that of the lower. Therefore, the range of integration may be restricted to the range over which the higher layer is non-zero. But over this range the lower layer is exhausted, or equal to its width. Hence, the product moment of two different layers equals the product of the width of the lower layer and the mean of the higher.[9]

The 4×4 blue-shaded block at the bottom of Table 3 contains the product moments. Down its diagonal are $E[Y_i^2]$, or twice the values of the "Area" row of Table 2. Off the diagonal are the lower-width-and-higher-mean products. The augmenting margin (unshaded) pertains to the loss from ground up, or MX;

[9]Similarly for multinomial products, the expectation depends on the highest layer, the other layers being determined by their exhaustion. Moreover, since only the highest layer can be infinite, the problem of multiplying by an infinite width will never arise.

Table 3. Two-moment summary

			Ground Up	5M xs 0	5M xs 5M	10M xs 10M	∞ xs 20M
Loss Layer	Mean	Std			Variance		
Ground Up	1,375,000	± 2,471,715	6.109E+12	2.811E+12	1.702E+12	1.269E+12	3.27 9E+11
5M xs 0	1,122,858	± 1,353,906	2.811E+12	1.833E+12	6.431E+11	2.901E+11	4.443E+10
5M xs 5M	165,861	± 801,119	1.702E+12	6.431E+11	6.418E+11	3.617E+11	5.539E+10
10M xs 10M	74,822	± 709,449	1.269E+12	2.901E+11	3.617E+11	5.033E+11	1.137E+11
∞ xs 20M	11,459	± 338,211	3.279E+11	4.443E+10	5.539E+10	1.137E+11	1.144E+11
					Correlation		
			100%	84%	86%	72%	39%
			84%	100%	59%	30%	10%
			86%	59%	100%	64%	20%
			72%	30%	64%	100%	47%
			39%	10%	20%	47%	100%
					E[Z] E[Z]′		
			1.89E+12	1.54E+12	2.28E+11	1.03E+11	1.58E+10
			1.54E+12	1.26E+12	1.86E+11	8.40E+10	1.29E+10
			2.28E+11	1.86E+11	2.75E+10	1.24E+10	1.90E+09
			1.03E+11	8.40E+10	1.24E+10	5.60E+09	8.57E+08
			1.58E+10	1.29E+10	1.90E+09	8.57E+08	1.31E+08
					E[ZZ′]		
			8.000E+12	4.355E+12	1.930E+12	1.372E+12	3.437E+11
			4.355E+12	3.094E+12	8.293E+11	3.741E+11	5.729E+10
			1.930E+12	8.293E+11	6.693E+11	3.741E+11	5.729E+10
			1.372E+12	3.741E+11	3.741E+11	5.089E+11	1.146E+11
			3.437E+11	5.729E+10	5.729E+10	1.146E+11	1.145E+11

it contains row and column sums of the blue-shaded block, since

$$E\left[Y_i MX\right] = E\left[Y_i\left(\sum_{j=1}^{4} Y_j\right)\right] = E\left[\sum_{j=1}^{4} Y_i Y_j\right] = \sum_{j=1}^{4} E\left[Y_i Y_j\right]$$

$$E\left[MX^2\right] = E\left[\left(\sum_{i=1}^{4} Y_i\right)MX\right] = E\left[\left(\sum_{i=1}^{4} Y_i MX\right)\right]$$

$$= \sum_{i=1}^{4} E\left[Y_i MX\right] = \sum_{i=1}^{4}\sum_{i=1}^{4} E\left[Y_i Y_j\right].$$

The soundness of our logic is confirmed inasmuch as $E\left[MX^2\right] = 8.000\times 10^{12}$, which is twice the area under the $Excess_{MX}$ curve, according to Table 1.

Table 3 gives the label ‘Z’ to the 5×1 vector whose first element is MX and remaining four are the Y_is. So the augmented product-moment matrix is $E[ZZ']$, as labeled. The box above it is the outer product of the means, $E[Z]E[Z]'$. The vector variance, $Var[Z] = E[ZZ'] - E[Z]E[Z]'$, is shown under the heading “Variance.” The “Std” column contains the square roots of the diagonal elements of the variance matrix, which values agree with those of Tables 1 and 2. When covariance is taken into account, the second moment of the loss is conserved. Finally, removing the standard-deviation scale from the variance matrix results in the “Correlation” matrix. It bears out something else well known to reinsurance actuaries, namely, that layered losses are positively correlated, although the correlation diminishes as the distance between the layers increases.[10]

6. Conclusion

The mathematics of excess losses is not only beautiful; it is powerful. The excess-loss function impounds all the information of the probability distribution of its loss. Therefore, although from the beginning actuaries and underwriters have found it convenient for the calculation of the pure premiums of layered losses, it is just as serviceable for the calculation of higher moments, whether the integrals involved be calculated analytically (as done in our example) or approximated numerically. The versatile mixed exponential distribution lessens the difficulty of such calculations.

References

Corro, D., "Grouping Loss Distributions by Tail Behavior," Casualty Actuarial Society *E-Forum* (Fall 2008), pp. 448–541, www.casact.org/pubs/forum/08fforum.

Corro, D., and G. J. Engl, "The 2004 NCCI Excess Loss Factors," Casualty Actuarial Society *Forum* (Fall 2006), pp. 513–571, www.casact.org/pubs/forum/06fforum/517.pdf.

Keatinge, C. L., "Modeling Losses with the Mixed Exponential Distribution," *Proceedings of the Casualty Actuarial Society* 86, 1999, pp. 654–698, www.casact.org/pubs/proceed/proceed99/99578.pdf.

Klugman, S. A., H. H. Panjer, and G. E. Willmot, *Loss Models: From Data to Decisions,* New York: Wiley, 1998.

Lee, Y.-S., "The Mathematics of Excess of Loss Coverage and Retrospective Rating—A Graphical Approach," *Proceedings of the Casualty Actuarial Society* 75, 1988, pp. 49–77, www.casact.org/pubs/proceed/proceed88/88049.pdf.

Patrik, G. S., "Reinsurance," chapter 6 in *Foundations of Casualty Actuarial Science* (3rd ed.), Casualty Actuarial Society, 1996, pp. 283–380.

[10]Again, as per footnote 8, see Appendix B for a proof.

Appendix A

Stieltjes integrals: Watch your step!

The expectation of $h(X)$ is $\sum_{i=1}^{\infty} h(x_i) \cdot Prob[X = x_i]$, if X is discreet, and $\int_{x=-\infty}^{\infty} h(x) f_X(x) dx$, if X is continuous. But random variables may be mixed, i.e., continuous with discrete steps.[11] For the sake of generality we can employ the Stieltjes integral [1, p. 12]: $E[h(X)] = \int_{x=-\infty}^{\infty} h(x) dF_X(x)$, where $F_X(x)$ is the cumulative distribution function of X. Of course, if F is differentiable, $dF_X(x) = f_X(x)dx$, and the Stieltjes integral reverts back to the familiar (Cauchy-Riemann) integral $\int_{x=-\infty}^{\infty} h(x) f_X(x) dx$. The Stieltjes integral is defined as

$$\int_{x=a}^{b} h(x)dF(x) = \lim_{\max\{\Delta x_i\}\to 0} \sum_{i=1}^{n} h(\xi_i)\Delta F_i,$$

where the interval is partitioned as $a = x_0 < x_1 < \cdots < x_n = b$, $\Delta x_i = x_i - x_{i-1}$, and $\Delta F_i = F(x_i) - F(x_{i-1})$. Each ξ_i is arbitrarily chosen from the subinterval $x_i - 1 \le \xi_i \le x_i$.

If u is continuous over the interval, nothing is problematic about this definition. Now for our purposes X is a non-negative random variable; hence, for $x < 0$, $dF_X(x) = 0$, and $E[h(X)] = \lim_{\varepsilon\to 0^-} \int_{x=\varepsilon}^{\infty} h(x)dF_X(x)$. One is tempted to simplify this to $\int_{x=0}^{\infty} h(x)dF_X(x)$ would miss any discrete step at zero, since the cumulative distribution function is continuous from the right. Therefore, the Stieltjes integral $\int_{x=a}^{b} h(x)dF_x(x)$ counts probability mass at the upper limit, but not at the lower. This asymmetry ensures that $\int_{x=a}^{b} h(x)dF_x(x) + \int_{x=b}^{c} h(x)dF_x(x) = \int_{x=a}^{c} h(x) dF_x(x)$; otherwise, probability mass at the endpoints might be either ignored or double-counted.

Therefore, the correct formulation for a non-negative random variable X is $E[h(X)] = h(0)\, Prob[X = 0] + \int_{x=0}^{\infty} h(x)dF_X(x)$. From integration by parts, we derive the form for the survival function $G_X(x) = 1 - F_X(x)$, which also is continuous from the right:

$$\begin{aligned}
E[h(X)] &= h(0)Prob[X=0] + \int_{x=0}^{\infty} h(x)dF_X(x) \\
&= h(0)Prob[X=0] + \int_{x=\infty}^{\infty} h(x)dG_X(x) \\
&= h(0)Prob[X=0] + h(x)dG_X(x)\Big|_{\infty}^{0} \\
&\quad - \int_{x=\infty}^{\infty} G_X(x)dh(x) \\
&= h(0)Prob[X=0] + h(0)G_X(0) - 0 \\
&\quad + \int_{x=0}^{\infty} G_X(x)dh(x). \\
&= h(0)Prob[X \ge 0] + \int_{x=0}^{\infty} G_X(x)dh(x) \\
&= h(0) + \int_{x=0}^{\infty} G_X(x)dh(x).
\end{aligned}$$

For this reason we subtitled this appendix "Watch your step!" For the $[0, \infty]$ Stieltjes integrals in this paper do not count probability mass at zero. For two reasons it is easy to overlook this subtlety. First, if $Prob[X = 0] = 0$, $E[h(X)] = \int_{x=0}^{\infty} h(x)dF_X(x)$, although the other form is still $E[h(X)] = h(0) + \int_{x=0}^{\infty} G_X(x)dh(x)$. And second, it is common for $h(x)$ to be a positive power of x, in which case $h(0) = 0$.

Watching one's step at zero also consistently handles a constant shift in $h(x)$:

[11]The number of discrete steps may be infinite, but it must be countable, or denumerable. To prove this, let $p(x) = Prob[X = x] = F_X(x) - \lim_{y\to x^-} F_X(y)$, and $M = \{x \in \Re : p(x) > 0\}$. M is the set of points of probability mass. To prove that M must be countable, partition it into the countable set of subsets $M = \bigcup_{i=1}^{\infty} M_i$, where $M_i = \{x \in M : 2^{-i} < p(x) \le 2^{-i+1}\}$. Each element of M is in one and only one of the subsets. If M_i were infinite, any sum of 2^i of its elements would exceed one, since all its elements are greater than 2^{-i}. But this would lead to total probability in excess of one; in fact, total probability would be infinite. Hence, each subset is finite. But a countable union of finite sets is countable. In symbols, $|M| = \left|\bigcup_{i=1}^{\infty} M_i\right| \le \aleph_0 \cdot \aleph_0 = \aleph_0$, where $|M|$ means the cardinality of M, and $\aleph_0$ is the cardinality of the natural numbers.

$$
\begin{aligned}
E[h(X)+c] &= E[h(X)]+c \\
&= h(0)Prob[X=0] + \int_{x=0}^{\infty} h(x)\,dF_X(x) + c \\
&= h(0)Prob[X=0] + \int_{x=0}^{\infty} h(x)\,dF_X(x) \\
&\quad + c\left\{Prob[X=0] + \int_{x=0}^{\infty} dF_X(x)\right\} \\
&= (h(0)+c)Prob[X=0] \\
&\quad + \int_{x=0}^{\infty} (h(x)+c)\,dF_X(x).
\end{aligned}
$$

And

$$
\begin{aligned}
E[h(X)+c] &= E[h(X)]+c \\
&= h(0) + \int_{x=0}^{\infty} G_X(x)\,dh(x) + c \\
&= (h(0)+c) + \int_{x=0}^{\infty} G_X(x)\,d(h(x)+c).
\end{aligned}
$$

Appendix B

Two theorems about reinsurance layers

Here we will give proofs of the two facts which this paper claims to be "well known to reinsurance actuaries":

1. that the coefficients of variation, $CV = Std/E$, increase as the layers ascend, and
2. that layered losses are positively correlated, although the correlation diminishes as the distance between the layers increases.

Our proofs will begin with "differential" layers, i.e., to layers whose width is dx. But, as we shall show, one can integrate such layers into layers of any width.

Let X be a non-negative random variable, whose survival function (the complement of the cumulative distribution function) is G_X. The probability that $X > x$ is $G_X(x)$; therefore, the probability of a non-zero loss in the interval $[x, x + \Delta x]$ is $G_X(x)$. In the limit, as $\Delta x \to 0^+$, the expected loss in the layer, $E[Layer(X; x, x + \Delta x)]$, approaches $G_X(x)\Delta x$. Defining $dY(x)$ as the portion of X in the differential layer $[x, x + dx]$, we may say that $dY(x) \sim Bernoulli(G_X(x)) \cdot dx$. Accordingly, $E[dY(x)] = G_X(x)dx$ and $E\left[(dY(x))^2\right] = G_X(x)(dx)^2$. Arguing as we did in Section 5, we have $E[dY(x_1)dY(x_2)] = min(G_X(x_1), G_X(x_2))dx_1dx_2$, of which $E[(dY(x))^2] = G_X(x)(dx)^2$ is a special instance in which $x_1 = x_2 = x$.[12]

Before we prove the two theorems, it will be instructive to see how a layer can be integrated from differential layers. If Y is the portion of X in layer $[a,b]$, then $Y = Layer(X;a,b) = \int_{x=a}^{b} dY(x)$. Hence, $E[Y] = E\left[\int_{x=a}^{b} dY(x)\right] = \int_{x=a}^{b} E[dY(x)] = \int_{x=a}^{b} G_X(x)\,dx$.

[12]One may no more object to the differential random variable $dY(x)$ than to the $dX(t)$ of the Wiener process. However, unlike the Wiener process, in which the $dX(t)$ are independent, the survival or filtration process imposes a co-moment structure upon the $dY(x)$.

Moreover, the second moment is

$$E[Y^2] = E\left[\left(\int_{x=a}^{b} dY(x)\right)^2\right]$$

$$= E\left[\int_{x_1=a}^{b} dY(x_1) \int_{x_2=a}^{b} dY(x_2)\right]$$

$$= E\left[\int_{x_1=a}^{b} \int_{x_2=a}^{b} dY(x_2)\,dY(x_1)\right]$$

$$= \int_{x_1=a}^{b} \int_{x_2=a}^{b} E[dY(x_2)\,dY(x_1)]$$

$$= \int_{x_1=a}^{b} \int_{x_2=a}^{b} min(G_X(x_1), G_X(x_2))\,dx_2\,dx_1$$

$$= \int_{x_1=a}^{b} \left\{ \int_{x_2=a}^{x_1} min(G_X(x_1), G_X(x_2))\,dx_2 \right.$$

$$\left. + \int_{x_2=x_1}^{b} min(G_X(x_1), G_X(x_2))\,dx_2 \right\} dx_1$$

$$= \int_{x_1=a}^{b} \left\{ \int_{x_2=a}^{x_1} G_X(x_1)\,dx_2 + \int_{x_2=x_1}^{b} G_X(x_2)\,dx_2 \right\} dx_1$$

$$= \int_{x_1=a}^{b} \int_{x_2=a}^{x_1} G_X(x_1)\,dx_2\,dx_1 + \int_{x_1=a}^{b} \int_{x_2=x_1}^{b} G_X(x_2)\,dx_2\,dx_1$$

$$= \int_{x_1=a}^{b} \int_{x_2=a}^{x_1} G_X(x_1)\,dx_2\,dx_1 + \int_{x_2=a}^{b} \int_{x_1=a}^{x_2} G_X(x_2)\,dx_1\,dx_2$$

$$= 2\int_{x_1=a}^{b} \int_{x_2=a}^{x_1} G_X(x_1)\,dx_2\,dx_1 .$$

$$= 2\int_{x=a}^{b} G_X(x)(x-a)\,dx$$

$$= \int_{x=a}^{b} G_X(x)\,d(x-a)^2 .$$

For ease of understanding, the derivation proceeded in many small steps; nevertheless, line seven deserves an explanation. Since G_X is non-increasing, $min(G_X(x_1), G_X(x_2)) = G_X(max(x_1, x_2))$. So by dividing the inner integral into the two regions, we can identify the minimum.[13] The reproduction of the moments of Y confirms the legitimacy of the formula $Y = \int_{x=a}^{b} dY(x)$.

One more notion is required for our proofs, a notion which we will call the *coefficient of covariance.* It is the covariance between two random variables whose means have been normalized to unity, i.e.,

$$CoefCov[X,Y] = Cov\left[\frac{X}{E[X]}, \frac{Y}{E[Y]}\right]$$
$$= \frac{Cov[X,Y]}{E[X]E[Y]} = \frac{E[XY]-E[X]E[Y]}{E[X]E[Y]}$$
$$= \frac{E[XY]}{E[X]E[Y]} - 1.$$

Of course, $CoefCov[X, X] = CV^2[X]$. The coefficient of covariance between two differential layers is

$$CoefCov[dY(x_1), dY(x_2)] = \frac{E[dY(x_1)dY(x_2)]}{E[dY(x_1)]E[dY(x_2)]} - 1$$
$$= \frac{min(G_X(x_1), G_X(x_2))dx_1dx_2}{G_X(x_1)dx_1G_X(x_2)dx_2} - 1$$
$$= \frac{min(G_X(x_1), G_X(x_2))}{G_X(x_1)G_X(x_2)} - 1$$
$$= \frac{min(G_X(x_1), G_X(x_2))}{min(G_X(x_1), G_X(x_2)) \cdot max(G_X(x_1), G_X(x_2))} - 1$$

[13]Equivalently, one may argue from symmetry. Let x_1 be the larger variable and integrate $\int_{x_1=a}^{b}\int_{x_2=a}^{x_1}$. But this is half the value of the full integral. This extends to higher moments. For example,

$$E[Y^3] = \int_{x_1=a}^{b}\int_{x_2=a}^{b}\int_{x_3=a}^{b} G_X(max(x_1, x_2, x_3))dx_3dx_2dx_1$$
$$= 3!\int_{x_1=a}^{b}\int_{x_2=a}^{x_1}\int_{x_3=a}^{x_1} G_X(x_1)dx_3dx_2dx_1 = 3\int_{x=a}^{b} G_X(x)(x-a)^2\,dx$$

In general, $E[Y^k] = k\int_{x=a}^{b} G_X(x)(x-a)^{k-1}\,dx = \int_{x=a}^{b} G_X(x)d(x-a)^k$, which is the layer-appropriate version of the formula of Section 2 and Appendix A, $E[h(X)] = h(0) + \int_{x=0}^{\infty} G_X(x)dh(x)$, for $h(x) = (x-0)^k$.

$$= \frac{1}{max(G_X(x_1), G_X(x_2))} - 1$$
$$= \frac{1}{G_X(min(x_1, x_2))} - 1.$$

This coefficient is well defined when $G_X(x_1)$ and $G_X(x_2)$ are non-zero; loss in the differential layers must be possible. Furthermore, $CV^2[dY(x)] = CoefCov[dY(x), dY(x)] = \frac{1}{G_X(x)} - 1$. Due to the properties of G_X, $CV^2[dY(x)]$ is a non-decreasing function in x.

With this preparation, the first fact is easily proven. If $Y = Layer(X; a,b) = \int_{x=a}^{b} dY(x)$ and $G_X(a) \geq G_X(b) > 0$, then

$$CV^2[Y] = Var[Y]/E[Y]^2$$
$$= \frac{\int_{x_1=a}^{b}\int_{x_2=a}^{b} Cov[dY(x_1), dY(x_2)]}{\left(\int_{x=a}^{b} E[dY(x)]\right)^2}$$
$$= \frac{\int_{x_1=a}^{b}\int_{x_2=a}^{b} E[dY(x_1)]E[dY(x_2)]CoefCov[dY(x_1), dY(x_2)]}{\int_{x_1=a}^{b}\int_{x_2=a}^{b} E[dY(x_1)]E[dY(x_2)]}.$$

Hence, $CV^2[Y]$ is a weighted average (weighted over two dimensions) of the coefficients of the layer's covariances. And since the weights are non-negative, the weighted average must be bounded by the minimum and maximum coefficients, which are at the endpoints:

$$CV^2[dY(a)] = CoefCov[dY(a), dY(a)] \leq CV^2[Y]$$
$$\leq CoefCov[dY(b), dY(b)]$$
$$= CV^2[dY(b)].$$

Therefore, of two layers, $[a, b]$ and $[c, d]$, where $a < b \leq c < d$, the CV^2 of the lower will be less than or equal to that of the higher. And if probability is consumed anywhere in these layers,[14] the inequality

[14]We say that "probability is consumed," if $G_X(a) > \lim_{x \to d^-} G_X(x)$.

will be strict. Even if the two layers overlap (i.e., $a \le c < b$ and $b \le d$, but not both $a = c$ and $b = d$), one can consider three intervals, the middle interval being the overlap. Then, as above, $CV^2(A) \le CV^2(B) \le CV^2(C)$. Because the unions involve weighted-averaging, $CV^2(A) \le CV^2(A \cup B) \le CV^2(B) \le CV^2(B \cup C) \le CV^2(C)$. Therefore, the CV^2 of a higher layer is greater than or equal to that of a lower layer, even if there is some overlap; the inequality is strict, if probability is consumed. Finally, since $CV \ge 0$, the inequalities are as valid for CV as for CV^2. Note that the widths of the layers do not need to be equal.

Second, as to correlation, let $Y_1 = Layer(X;a,b) = \int_{x=a}^{b} dY(x)$ and $Y_2 = Layer(X;c,d) = \int_{x=c}^{d} dY(x)$, for $a < b \le c < d$. Therefore, we know that the minimum G_X will be in the higher interval $[c, d]$.

Under these conditions,

$$
\begin{aligned}
Cov[Y_1, Y_2] &= E[Y_1 Y_2] - E[Y_1]E[Y_2] \\
&= \int_{x_1=a}^{b} \int_{x_2=c}^{d} E[dY(x_1)dY(x_2)] - E[Y_1]E[Y_2] \\
&= \int_{x_1=a}^{b} \int_{x_2=c}^{d} min(G_X(x_1), G_X(x_2))dx_2\, dx_1 - E[Y_1]E[Y_2] \\
&= \int_{x_1=a}^{b} \int_{x_2=c}^{d} G_X(x_2)dx_2\, dx_1 - E[Y_1]E[Y_2]. \\
&= (b-a)\int_{x_2=c}^{d} G_X(x_2)dx_2 - E[Y_1]E[Y_2] \\
&= (b-a)E[Y_2] - E[Y_1]E[Y_2] \\
&= (b-a-E[Y_1])E[Y_2].
\end{aligned}
$$

Because $E[Y_1] = \int_{x=a}^{b} G_X(x)dx \le \int_{x=a}^{b} 1 \cdot dx = b - a$, $b - a - E[Y_1] \ge 0$; hence, $Cov[Y_1, Y_2] \ge 0$. So the correlation coefficient between the portions of X in the two layers is

$$
\begin{aligned}
Corr[Y_1, Y_2] &= \frac{Cov[Y_1, Y_2]}{\sigma_{Y_1}\sigma_{Y_2}} \\
&= \frac{b-a-E[Y_1]}{\sigma_{Y_1}} \cdot \frac{E[Y_2]}{\sigma_{Y_2}} \\
&= \frac{\left(\dfrac{b-a-E[Y_1]}{\sigma_{Y_1}}\right)}{CV[Y_2]}.
\end{aligned}
$$

Now consider shifting $[c, d]$ to the right, i.e., to $[c + \xi, d + \xi]$, where $\xi \ge 0$. And let $Y_2(\xi) = Layer(X; c + \xi, d + \xi) = \int_{x=c+\xi}^{d+\xi} dY(x)$. From the first proof we know that $CV[Y_2(\xi)]$ is non-decreasing. Since the numerator of $Corr$ is constant, $Corr[Y_1, Y_2(\xi)]$ is non-increasing; strictly decreasing if probability is consumed. Therefore, as the retention of the upper layer so moves away from that of the lower as to consume probability, the correlation decreases. This implies that in the absence of compensating risk premiums, a reinsurer should not underwrite neighboring layers.