Annuity Densities with Application to Tail Development

Daniel R. Corro

Annuity Densities with Application to Tail Development[1]

Dan Corro
National Council on Compensation Insurance, Inc.
May, 2003

Abstract: *This paper considers the task of modeling "pension" claims whose durations may vary, but whose payment pattern is uniform and flat. We derive the aggregate payout pattern from the duration density and discuss and provide examples to show how this idea can be applied to calculating tail development factors.*

It is apparent that the claim duration influences paid loss development. In general a faster (slower) claim closure rate will make paid losses develop faster (slower). While the direct nature of that relationship is apparent, it is not so apparent how to quantify it. This paper quantifies it the case of "pension claims" with a constant periodic payment amount. More precisely, the paper considers continuous payment at a rate that is constant both over time and among claims.

Let $S(t)$ denote a survival function on the time interval $(0,b)$.[2] We regard $S(t)$ as a distribution of closure times and let $F(t) = 1 - S(t)$ be the corresponding cumulative distribution function [CDF]. In effect, all claims are assumed to close on or before time b.

We are interested in a related CDF, which we denote by $\widetilde{F}(t)$ to emphasize its relation with $F(t)$, which models the paid loss development as a function of time. More precisely, $\widetilde{F}(t)$ is the proportion of total loss paid by time t, i.e. the proportion paid out during $(0,t)$ (without any discount adjustment). $\widetilde{F}(t)$ is the reciprocal of the paid to ultimate loss development factor and we will refer to $\widetilde{F}(t)$ as the paid loss development divisor [PLDD]. [3]

We are interested in claims whose payment schedule conforms to two very restrictive assumptions:

- All payments on all claims are of the same amount.
- Payments are made periodically at a common uniform time interval immediately following a common time of loss, $t = 0$, to claim closure.

[1] The author expresses his thanks to Greg Engl, also of NCCI, who reviewed several versions of this paper, pointed out some serious errors, and made many suggestions for improvements.

[2] We are most interested in the case $b < \infty$, although most of what we say applies to the case $b = \infty$. We are, however, rather cavalier about making whatever assumptions are needed to assure that all improper integrals exist and are finite.

[3] Gillam and Couret [2] consider the reciprocal of the loss development factor and call it the loss development divisor.

We refer to these two assumptions in combination as describing the "pension case". We use a continuous model to represent this, translating this into the assumption that for every claim of duration x, the model assumes a continuous and constant payment rate of 1 over the interval $(0,x)$, and 0 elsewhere.

Consider the case when aggregate paid losses are followed over a series of N time units with $N < b$. The usual paid loss development patterns built from these N evaluations will not account for the "tail paid loss development" beyond the final evaluation at time $t = N$. With this notation, observe that this tail factor is just $\lambda = \widetilde{F}(N)^{-1}$.

Workers' Compensation [WC] provides a case in point, as some claims in that long-tail line remain open beyond the reporting window, even though that window has been expanding to near 20 years. It is reasonable to assume that payments beyond some valuation, say after 10 years, will be primarily made on pension like claims. Consequently, a model suited to such pension claims may be helpful in projecting the full payout pattern beyond 10 years. Indeed, suppose you have a collection of PLDD's that cover the portion of the loss "portfolio" that is expected to develop beyond 10 years. That is, for each type of claim you have a PLDD that is deemed appropriate, at least over the time frame beyond 10 years. The paper illustrates how to translate the mix of claims in the loss portfolio into a mixed distribution of those PLDD's (c.f. Corollary 1.3 below). That mixed distribution then provides an estimated tail factor. This approach to deriving a tail factor for WC losses is what motivated this paper.

Notation and Setup

With S, F, $\widetilde{F}$ and b as above, we also let $f(t) = \dfrac{dF}{dt}$ be the probability density function [PDF], $h(t) = \dfrac{f(t)}{S(t)}$ the hazard rate function, $CV = \dfrac{\sigma}{\mu}$ the coefficient of variation of claim duration, and T the random variable that gives the "time" of closure t and has the CDF $F(t)$. We use those same letter symbols and "transparent" notation to specify the relationship between these functions. For example $\widetilde{h}_\alpha(t)$ denotes the hazard rate function of the PLDD $\widetilde{F}_\alpha(t)$ that corresponds to the claim survival function $S_\alpha(t)$ and $\widetilde{T}_\alpha$ the random variable with CDF $\widetilde{F}_\alpha(t)$.

For pension claims, as described here, the entire payment schedule of a claim is completely determined by the claim duration. But for now we consider a somewhat more general situation. We make the assumption that for any time t, $0<t<b$, all claims with duration t have the same pre-determined and differentiable payment pattern. We can capture this mathematically by defining the function

$G(x,t)$=amount paid through time t on a claim, conditional upon claim duration=x.

Then define $g(x,t) = \frac{\partial G}{\partial t}$ = partial derivative of $G(x,t)$ with respect to t. We may interpret $g(x,t)$ as the rate of payment at time t on any claim of duration x. Both $G(x,t)$ and $g(x,t)$ are defined for x,t in $(0,b)$. Note that for $t>x$ we have $g(x,t) = 0$ and $G(x,x) = G(x,t) = G(x,b)$= the ultimate incurred on any claim of duration x. It is convenient to define the claim severity function:

$$\gamma(x) = G(x,x) = G(x,b) \quad x \in (0,b).$$

We consider the cumulative payment for such a claim distribution in which all claims occur at time t=0 and conform to these assumptions (sort of an accident instant, as opposed to an accident year). The only "stochastic" ingredient in this model is claim duration, for which the distribution $F(t)$ is specified. All payments are in effect determined by these assumptions and there is a well defined expected cumulative paid loss per claim $P(t)$ at any time t, from t=0 to ultimate paid at t=b. Indeed, we have:

$$\begin{aligned}
P(t) &= \int_0^t\int_0^b g(x,y)f(x)dxdy = \int_0^b f(x)\int_0^t g(x,y)dydx = \int_0^b f(x)G(x,t)dx \\
&= \int_0^t f(x)G(x,t)dx + \int_t^b f(x)G(x,t)dx \\
&= \int_0^t f(x)\gamma(x)dx + \int_t^b f(x)G(x,t)dx
\end{aligned}$$

since $G(x,t) = \gamma(x)$ for $x<t$. In particular, the expected ultimate loss per claim is just:

$$P(b) = \int_0^b f(x)\gamma(x)dx.$$

It is convenient to define yet another function of t:

$$\eta(t) = \int_0^t f(x)\gamma(x)dx$$

The (expected) ultimate paid loss development factor from time t is:

$$\lambda(t) = \frac{P(b)}{P(t)}$$

and the inverse provides the PLDD on $(0,b)$ that is the focus of this study:

$$\tilde{F}(t) = \frac{P(t)}{P(b)}.$$

For the PDF of the PLDD, we have, by the fundamental theorem of calculus:

$$\eta(b)\tilde{f}(t) = P(b)\tilde{f}(t) = \frac{d}{dt}\left(\int_0^t\int_0^b g(x,y)f(x)dxdy\right) = \int_0^b g(x,t)f(x)dx = \int_t^b g(x,t)f(x)dx$$

since $g(x,t) = 0$ for $x<t$.

We have established most the following:

Proposition 1: *With the function G(x,t) defined as above, for* $t \in (0,b)$ *:*

i) $\tilde{F}(t) = \dfrac{\eta(t) + \int_t^b f(x)G(x,t)dx}{\eta(b)}$

ii) $\tilde{f}(t) = \dfrac{\int_t^b f(x)g(x,t)dx}{\eta(b)}$

iii) $\tilde{S}(t) = \dfrac{\int_t^b f(x)(\gamma(x) - G(x,t))dx}{\eta(b)}$

Proof: All is clear except perhaps (iii):

$$\tilde{S}(t) = 1 - \tilde{F}(t) = \frac{\eta(b) - \eta(t) - \int_t^b f(x)G(x,t)dx}{\eta(b)} =$$

$$= \frac{\int_0^b f(x)\gamma(x)dx - \int_0^t f(x)\gamma(x)dx - \int_t^b f(x)G(x,t)dx}{\eta(b)}$$

$$= \frac{\int_t^b f(x)\gamma(x)dx - \int_t^b f(x)G(x,t)dx}{\eta(b)}$$

$$= \frac{\int_t^b f(x)(\gamma(x) - G(x,t))dx}{\eta(b)}$$

as required.

In the WC work that motivated this, the focus was on tail development. This in turn led to the consideration of pension cases. Since those cases take longer to resolve, it is natural to try and use that as a way to isolate them. This leads us to consider what happens when there is a delay period prior to closure, i.e. when $f(t) = 0$ for $t \in (0,a)$ where $0 \le a < b$. In that event we have:

Corollary 1.2: *Suppose* $f(t)=0$ *for* $t\in(0,a)$ *where* $0\le a<b$, *then*

$$\widetilde{F}(t)=\frac{\int_a^b f(x)G(x,t)dx}{\eta(b)} \quad \textit{for } t\in(0,a).$$

Proof: This is apparent from Proposition 1 (i), since by our assumption

$$\eta(t)=\int_0^t f(x)\gamma(x)dx=0 \text{ for } t\in(0,a)$$

and the result follows.

This setup is convenient when the distribution of claim durations can be expressed as a mixture of simpler "component" densities. The following Corollary is actually a special case of a more general relationship between mixtures of losses and mixtures of PLDDs that in this pension case is just a simple calculation using Proposition 1:

Corollary 1.3: *Suppose* $F=\sum_{i=1}^{n} w_i F_i$, $\sum_{i=1}^{n} w_i=1$ *is a weighted sum of CDF's on (0,b), then*

$$\widetilde{F}=\sum_{i=1}^{n}\widetilde{w}_i\widetilde{F}_i, \quad \textit{where} \quad \widetilde{w}_i=\frac{w_i\eta_i(b)}{\eta(b)},\ 1\le i\le n.$$

Proof: This is a straightforward application of Proposition 1, noting that the same payment function $G=G_i$ applies to all the claims and so applies to each CDF F_i. More precisely:

$$\begin{aligned}\eta(b)\widetilde{F}(t)&=\eta(t)+\int_t^b f(x)G(x,t)dx=\sum_{i=1}^{n}w_i\eta_i(t)+\int_t^b\sum_{i=1}^{n}w_if_i(x)G(x,t)dx\\&=\sum_{i=1}^{n}w_i\eta_i(t)+\int_t^b\sum_{i=1}^{n}w_if_i(x)G_i(x,t)dx=\sum_{i=1}^{n}w_i\left(\eta_i(t)+\int_t^b f_i(x)G_i(x,t)dx\right)\\&=\sum_{i=1}^{n}w_i\left(\eta_i(b)\widetilde{F}_i(t)\right)\end{aligned}$$

and the result follows.

Recall that μ denotes the mean duration. More generally, define the higher moments of the distributions $F,\widetilde{F}$ as:

With the above notation, the following proposition documents some basic relationships between the duration density and the PLDD density:

Proposition 2: *In the pension case, for* $t \in (0,b)$

i) $\tilde{f}(t) = \dfrac{S(t)}{\mu}$

ii) $\tilde{F}(t) = \dfrac{\eta(t) + tS(t)}{\mu} = \dfrac{\eta(t)}{\mu} + t\tilde{f}(t)$

iii) $\gamma(t) = t$

iv) $\eta(t) = \int_0^t xf(x)dx$

v) $\tilde{h}(t) = \dfrac{S(t)}{\mu\tilde{S}(t)}$

vi) $\tilde{\mu}^{(k)} = \dfrac{\mu^{(k+1)}}{(k+1)\mu}$ *for* $k = 1,2,3,\ldots$

Proof: By the pension case assumptions, we have:

$$g(x,t) = \frac{\partial G}{\partial t} = \begin{cases} 1 & 0 \le t \le x \\ 0 & x \le t \end{cases}$$

$$\Rightarrow G(x,t) = \begin{cases} t & 0 \le t \le x \\ x & x \le t. \end{cases}$$

We then find that:

$$\gamma(t) = G(t,t) = t$$

$$\Rightarrow \eta(t) = \int_0^t \gamma(x)f(x)dx = \int_0^t xf(x)dx \Rightarrow \eta(b) = \mu$$

which establishes (iii) and (iv). Note that from Proposition 1:

$$\tilde{F}(t) = \frac{\eta(t) + \int_t^b f(x)G(x,t)dx}{\eta(b)} = \frac{\eta(t) + t\int_t^b f(x)dx}{\mu} = \frac{\eta(t) + tS(t)}{\mu}$$

$$\tilde{f}(t) = \frac{\int_t^b f(x)g(x,t)dx}{\eta(b)} = \frac{\int_t^b f(x)dx}{\mu} = \frac{S(t)}{\mu}$$

proving (i) and (ii). For (v), observe that:

$$\tilde{h}(t)=\frac{\tilde{f}(t)}{\tilde{S}(t)}=\frac{S(t)}{\mu\tilde{S}(t)}$$

and for (vi), integration by parts also gives:

$$E(T^{k+1})=(k+1)\int_0^b x^k S(x)dx$$

and we have:

$$\tilde{\mu}^{(k)}=\int_0^b x^k\tilde{f}(x)dx=\int_0^b x^k\left(\frac{S(x)}{\mu}\right)dx=\frac{k+1}{(k+1)\mu}\int_0^b x^k S(x)dx=\frac{E(T^{k+1})}{(k+1)\mu}=\frac{\mu^{(k+1)}}{(k+1)\mu}$$

This completes the proof of Proposition 2.

Now we clearly have that the PDF $\tilde{f}(t)$ is decreasing, indeed $\frac{d\tilde{f}}{dt}=-\frac{f(t)}{\mu}\leq 0$ and so the mode of the PLDD $\tilde{F}(t)$ is 0. From the following Corollary, we see how the shift from $F(t)$ to $\tilde{F}(t)$ impacts the mean, in particular, we find that the shift increases the mean exactly when $\sigma>\mu$.

Corollary 2.1: $2\tilde{\mu}=\mu+\sigma CV$

Proof: From Proposition 2 (vi):

$$2\tilde{\mu}=\frac{\mu^{(2)}}{\mu}=\frac{\mu^2+\sigma^2}{\mu}=\mu+\sigma CV.$$

Corollary 2.2: *Suppose* $f(t)=0$ *for* $t\in(0,a)$ *where* $0\leq a<b$, *then*

$$\tilde{F}(t)=\frac{t}{\mu}\quad \text{for } t\in(0,a).$$

Proof: Under these assumptions, Corollary 1.2 implies that for $t<a$:

$$\tilde{F}(t)=\frac{\int_a^b f(x)G(x,t)dx}{\eta(b)}=\frac{\int_a^b f(x)tdx}{\mu}=\frac{t\int_a^b f(x)dx}{\mu}=\frac{t\int_0^b f(x)dx}{\mu}=\frac{t}{\mu}$$

as claimed.

Probably the most useful family of distributions defined on a finite interval is the class of Beta densities on (0,1). Recall that the Beta distribution is a two-parameter, α,β, distribution that is usually defined in terms of its PDF:

$$f(\alpha,\beta;x)=\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha,\beta)}=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}\quad x\in(0,1),\alpha>0,\beta>0$$

where B and Γ denote the usual Beta and Gamma functions (c.f.[1], [3]). The CDF of the Beta density is:

$$\mathrm{B}(\alpha,\beta;t)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^t x^{\alpha-1}(1-x)^{\beta-1}dx \quad t\in(0,1),\alpha>0,\beta>0$$

and we have:

Corollary 2.3: *For* $\alpha>0,\beta>0$ *let* $f(t)=f(\alpha,\beta;t)$, $F(t)=\mathrm{B}(\alpha,\beta;t)$ *be the PDF and CDF of a beta density on* $(0,1)$, *as just defined, then:*

$$\tilde{f}(t)=\frac{\alpha+\beta}{\alpha}(1-\mathrm{B}(\alpha,\beta;t))=\frac{\alpha+\beta}{\alpha}(\mathrm{B}(\beta,\alpha;1-t))=\frac{\mathrm{B}(\beta,\alpha;1-t)}{\mu}$$

$$\tilde{F}(t)=\tilde{\mathrm{B}}(\alpha,\beta;t)=\mathrm{B}(\alpha+1,\beta;t)+\frac{(\alpha+\beta)t}{\alpha}\mathrm{B}(\beta,\alpha;1-t) \quad 0<t<1.$$

Proof: The proof is a straightforward application of Proposition 2. For the PDF, note that:

$$\tilde{f}(t)=\frac{1-F(t)}{\mu}=\frac{1-\mathrm{B}(\alpha,\beta;t)}{\mu}=\frac{1-\mathrm{B}(\alpha,\beta;t)}{\left(\frac{\alpha}{\alpha+\beta}\right)}=\frac{(\alpha+\beta)\mathrm{B}(\beta,\alpha;1-t)}{\alpha}=\frac{\mathrm{B}(\beta,\alpha;1-t)}{\mu}.$$

Note too that:

$$\begin{aligned}\eta(t)&=\int_0^t xf(x)dx=\left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right)\int_0^t xx^{\alpha-1}(1-x)^{\beta-1}dx\\&=\left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right)\int_0^t x^{(\alpha+1)-1}(1-x)^{\beta-1}dx\\&=\left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right)\left(\frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1+\beta)}\right)\left(\frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)}\right)\int_0^t x^{(\alpha+1)-1}(1-x)^{\beta-1}dx\\&=\left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+1)}\right)\left(\frac{\Gamma(\alpha+1)}{\Gamma(\alpha)}\right)\mathrm{B}(\alpha+1,\beta;t)\\&=\left(\frac{\alpha}{\alpha+\beta}\right)\mathrm{B}(\alpha+1,\beta;t)=\mu\mathrm{B}(\alpha+1,\beta;t).\end{aligned}$$

And so for the CDF:

$$\tilde{F}(t)=\frac{\eta(t)+tS(t)}{\mu}=\mathrm{B}(\alpha+1,\beta;t)+\frac{(\alpha+\beta)t\mathrm{B}(\beta,\alpha;1-t)}{\alpha}$$

as claimed.

We next consider some more specific examples:

Example 1: Consider the case when $S_1(t)=\frac{b-t}{b}$, $b<\infty$, is a DeMoivre survival curve. In this case:

$$F_1(b;t) = F_1(t) = \frac{t}{b} \quad f_1(t) = \frac{1}{b}.$$

Note that

$$\eta_1(t) = \int_0^t x f_1(x)dx = \frac{1}{b}\int_0^t x dx = \frac{1}{b}\left[\frac{x^2}{2}\right]_0^t = \frac{t^2}{2b} \Rightarrow \mu_1 = \eta_1(b) = \frac{b}{2}$$

from which we find:

$$\tilde{f}_1(t) = \frac{S_1(t)}{\mu_1} = \left(\frac{b-t}{b}\right)\left(\frac{2}{b}\right) = \frac{2(b-t)}{b^2}$$

and so the density of this PLDD decreases linearly with time. Whence:

$$\tilde{F}_1(t) = \frac{\eta_1(t) + tS_1(t)}{\mu_1} = \frac{2}{b}\left(\frac{t^2}{2b} + t\left(\frac{b-t}{b}\right)\right) = \frac{2}{b}\left(t - \frac{t^2}{2b}\right) = \frac{t}{b}\left(2 - \frac{t}{b}\right)$$

and finally:

$$\begin{aligned}\tilde{S}_1(t) &= 1 - \tilde{F}_1(t) = 1 - \frac{t}{b}\left(2 - \frac{t}{b}\right) \\ &= 1 - 2\left(\frac{t}{b}\right) + \left(\frac{t}{b}\right)^2 = \left(1 - \frac{t}{b}\right)^2 = \left(\frac{b-t}{b}\right)^2 = S_1(t)^2.\end{aligned}$$

Example 2: It is easy to generalize the first example, for $\varphi > 0$ let the claim closure have the CDF:

$$F_2(b;t) = F_2(t) = \begin{cases}\left(\frac{t}{b}\right)^{\varphi} & t \le b \\ 1 & t \ge b\end{cases}$$

then we have:

$$f_2(t) = \begin{cases}\frac{\varphi t^{\varphi-1}}{b^{\varphi}} & t \le b \\ 0 & t \ge b\end{cases}$$

$$\eta_2(t) = \int_0^t x f_2(x)dx = \frac{\varphi}{b^{\varphi}}\int_0^t x^{\varphi}dx = \frac{\varphi}{b^{\varphi}}\left[\frac{x^{\varphi+1}}{\varphi+1}\right]_0^t = \frac{\varphi t^{\varphi+1}}{(\varphi+1)b^{\varphi}}.$$

In particular, we have:

$$\mu_2 = \eta_2(b) = \frac{\varphi b}{\varphi + 1}$$
$$\widetilde{f}_2(t) = \frac{S_2(t)}{\mu_2} = \frac{(\varphi+1)}{\varphi b}\left(1 - \left(\frac{t}{b}\right)^{\varphi}\right)$$
$$\widetilde{F}_2(t) = \frac{\eta_2(t) + tS_2(t)}{\mu_2} = \frac{t}{\varphi b}\left(\varphi + 1 - \left(\frac{t}{b}\right)^{\varphi}\right).$$

Example 3: Consider the case when fewer claims close over time according to a linear pattern (like the PLDD of Example 1):

$$f_3(b;t) = f_3(t) = \begin{cases} \dfrac{2(b-t)}{b^2} & t \le b \\ 0 & b \le t. \end{cases}$$

By Example 1, $f_3(t)$ is indeed a PDF on $[0,b]$ and we have:

$$\eta_3(t) = \int_0^t x f_3(x)dx = \frac{2}{b^2}\int_0^t x(b-x)dx = \frac{2}{b^2}\left[\frac{bx^2}{2} - \frac{x^3}{3}\right]_0^t = \frac{2}{b^2}\left(\frac{bt^2}{2} - \frac{t^3}{3}\right) \Rightarrow \mu_3 = \frac{b}{3}.$$

But then from Example 1:

$$S_3(t) = \left(\frac{b-t}{b}\right)^2$$
$$\Rightarrow \widetilde{f}_3(t) = \frac{S_3(t)}{\mu_3} = \frac{3}{b^3}(t-b)^2$$
$$\Rightarrow \widetilde{S}_3(t) = \int_t^b \widetilde{f}_3(x)dx = \left(\frac{b-t}{b}\right)^3$$
$$\Rightarrow \widetilde{F}_3(b;t) = \widetilde{F}_3(t) = 1 - \left(\frac{b-t}{b}\right)^3.$$

In the WC work that motivated this, we seek to find a 19th to ultimate paid loss development factor. One idea that we considered is to use a weighted sum (mixture) of PLDDs of the form $w\widetilde{F}_3(b_1;t) + (1-w)\widetilde{F}_3(b_2;t)$ (c.f. Corollary 1.3, extend to a common interval by setting the density of the shorter interval to vanish outside it's natural domain). This means that we are assuming all claims close after $Max(b_1, b_2)$ years and one part of the loss "portfolio" close by time $t = b_1$ and the complement by b_2. Empirical loss development factor data is used to fit a non-linear model in which the mixing weight variable w is a "parameter". When these simple functions are used with b_1, b_2 as selected constants, it is straightforward to set up the calculation so as to assure a closed form solution for the value of w that gives the best fit to the data.

Our experience to date of comparing this approach to alternative methods, suggests that the use of linear survival models for the claim duration distribution, while pedagogically and theoretically helpful, may be too simplistic for practical application. Although payment duration is effectively limited by the beneficiary's life-span, there may be no applicable limit to the incurred loss, especially when long term medical care may be covered. While we are primarily interested in the case of finite support, it may be useful to consider a couple of examples when $b = \infty$.

Example 4: Consider the case of a single parameter Pareto (c.f. [3], p 584):

$$F_4(\alpha,\theta;t) = F_4(t) = 1 - \left(\frac{\theta}{t}\right)^{\alpha} \qquad f_4(\alpha,\theta;t) = f_4(t) = \frac{\alpha\theta^{\alpha}}{t^{\alpha+1}} \quad \text{for } t > \theta.$$

It is natural to extend the definition of the PDF to assign $f_4(t) = 0$ for $t \in (0,\theta]$. Assume that $\alpha > 1$. In this case we have:

$$\mu_4 = \frac{\alpha\theta}{\alpha - 1} \quad \Rightarrow \tilde{f}_4(t) = \frac{S_4(t)}{\mu_4} = \begin{cases} \dfrac{\alpha - 1}{\alpha\theta} & t \le \theta \\[2ex] \dfrac{\left(\frac{\theta}{t}\right)^{\alpha}}{\frac{\alpha\theta}{\alpha - 1}} = \left(\dfrac{\alpha - 1}{\alpha\theta}\right)\left(\dfrac{\theta}{t}\right)^{\alpha} & t \ge \theta \end{cases}$$

and we find that:

$$\eta_4(t) = \int_0^t x f_4(x)dx = \begin{cases} 0 & t \le \theta \\ \alpha\theta^{\alpha}\displaystyle\int_\theta^t \frac{dx}{x^{\alpha}} = \left(\frac{\alpha\theta}{\alpha - 1}\right)\left(1 - \left(\frac{\theta}{t}\right)^{\alpha-1}\right) & t \ge \theta \end{cases}$$

$$\Rightarrow \tilde{F}_4(t) = \frac{\eta_4(t) + tS_4(t)}{\mu_4} = \begin{cases} \left(\dfrac{\alpha - 1}{\alpha\theta}\right)t & t \le \theta \\[2ex] 1 - \dfrac{\left(\frac{\theta}{t}\right)^{\alpha-1}}{\alpha} & t \ge \theta \end{cases}$$

$$\Rightarrow \tilde{S}_4(t) = \frac{\left(\frac{\theta}{t}\right)^{\alpha-1}}{\alpha} = \frac{S_4(t)^{\frac{\alpha-1}{\alpha}}}{\alpha} \quad \text{for } t \ge \theta.$$

Example 5: Consider the case when claim closures follow an exponential density, so here again $b = \infty$. In this case, we have:

$$F_5(\theta;t) = F_5(t) = 1 - e^{-\frac{t}{\theta}}$$

$$f_5(\theta;t) = f_5(t) = \frac{e^{-\frac{t}{\theta}}}{\theta}.$$

Then from Proposition 2 we have:

$$\tilde{f}_5(\theta;t) = \tilde{f}_5(t) = \frac{S_5(t)}{\mu_5} = \frac{e^{-\frac{t}{\theta}}}{\theta} = f_5(t)$$
$$\Rightarrow \tilde{F}_5(t) = F_5(t)$$

and we find that, in the pension case, an exponentially distributed duration has an exponentially distributed PLDD, with the same mean. This suggests that the use of an exponential density, or a mixed exponential (c.f. Corollary 1.3), to fit the PLDD may be quite reasonable when analyzing tail behavior of coverages for which the payments on long term claims become pension like.

Example 6: It is tempting to generalize Example 5, so consider the case when claim closures follow a Weibull density, and so here again $b = \infty$. In this case, we have:

$$F_6(\theta,\tau;t) = F_6(t) = 1 - e^{-\left(\frac{t}{\theta}\right)^\tau}$$

$$f_6(\theta,\tau;t) = f_6(t) = \frac{\tau\left(\frac{t}{\theta}\right)^\tau e^{-\left(\frac{t}{\theta}\right)^\tau}}{t}.$$

$$\mu_6(\theta,\tau;t) = \mu_6 = \theta \cdot \Gamma\left(1 + \frac{1}{\tau}\right).$$

Then from Proposition 2 we have:

$$\tilde{f}_6(\theta,\tau;t) = \tilde{f}_6(t) = \frac{S_6(t)}{\mu_6} = \frac{e^{-\left(\frac{t}{\theta}\right)^\tau}}{\theta \cdot \Gamma\left(1 + \frac{1}{\tau}\right)}$$

and we find that, in the pension case, a Weibull distributed duration has a "transformed exponential" as PLDD.

Example 7: Finally, suppose a PLDD $\widetilde{F}_7(t)$ took the form of a Weibull density:

$$\widetilde{F}_7(\theta,\tau;t) = \widetilde{F}_7(t) = 1 - e^{-\left(\frac{t}{\theta}\right)^\tau}$$

$$\widetilde{f}_7(\theta,\tau;t) = \widetilde{f}_7(t) = \frac{\tau\left(\frac{t}{\theta}\right)^\tau e^{-\left(\frac{t}{\theta}\right)^\tau}}{t}.$$

Then from Proposition 2 we would have:

$$S_7(t) = \mu_7 \cdot \widetilde{f}_6(t) = \frac{\mu_7 \cdot \tau\left(\frac{t}{\theta}\right)^\tau e^{-\left(\frac{t}{\theta}\right)^\tau}}{t}.$$

But then by L'Hôpital:

$$1 = \underset{t\to 0}{Lim}\, S_7(t) = \mu_7 \cdot \tau\, \underset{t\to 0}{Lim} \frac{\left(\frac{t}{\theta}\right)^\tau e^{-\left(\frac{t}{\theta}\right)^\tau}}{t}$$

$$= \mu_7 \cdot \tau\, \underset{t\to 0}{Lim} \frac{\left(\frac{t}{\theta}\right)^\tau e^{-\left(\frac{t}{\theta}\right)^\tau}}{t}$$

$$= \mu_7 \cdot \tau\, \underset{t\to 0}{Lim} \frac{\left(\frac{t}{\theta}\right)^\tau e^{-\left(\frac{t}{\theta}\right)^\tau}\left(-\left(\frac{\tau}{\theta}\right)\left(\frac{t}{\theta}\right)^{\tau-1}\right) + e^{-\left(\frac{t}{\theta}\right)^\tau}\left(\left(\frac{\tau}{\theta}\right)\left(\frac{t}{\theta}\right)^{\tau-1}\right)}{1}$$

$$= \frac{\mu_7 \cdot \tau^2}{\theta} \underset{t\to 0}{Lim} \frac{e^{-\left(\frac{t}{\theta}\right)^\tau}\left(\frac{t}{\theta}\right)^{\tau-1}\left(1-\left(\frac{t}{\theta}\right)^\tau\right)}{1} = \begin{cases} 0 & \tau > 1 \\ \frac{\mu_7}{\theta} & \tau = 1 \\ \infty & 0 < \tau < 1 \end{cases}$$

And it follows that the only Weibull density that can be a PLDD in the pension case is the exponential.

Application to Tail Development

We seek to fit Workers Compensation (WC) age-to-age paid LDFs to a PLDD distribution $\widetilde{F}(t)$. With an eye on Corollary 1.3 and recognizing that pension claims represent a small minority among all WC losses, we decide to consider models of the form:

$$\widetilde{F}(t) = w\widetilde{F}_\alpha(t) + (1-w)\widetilde{F}_\beta(t), \quad 0 \le w \le 1.$$

where $\widetilde{F}_\alpha(t)$ and $\widetilde{F}_\beta(t)$ are PLDD's of "known type." Because we focus on the tail, we only look at development beyond an 11-th report. Throughout this section, we consider as given a set of age-to-age paid loss development factors from an 11th to a 19th report:

$$\begin{aligned}
\lambda_1 &= 11^{\text{th}} \text{ to } 12^{\text{th}} \text{ paid loss LDF} \\
\lambda_2 &= 12^{\text{th}} \text{ to } 13^{\text{th}} \text{ paid loss LDF} \\
&\vdots \\
\lambda_8 &= 18^{\text{th}} \text{ to } 19^{\text{th}} \text{ paid loss LDF.}
\end{aligned}$$

If we knew the true "tail factor" = 19th to ultimate paid loss LDF, we could readily combine that information with the λ_i to determine several "actual" values of $\widetilde{F}(t)$, namely for t=19, 18, …,11. More precisely, let v^{-1} = tail factor, then the "true" PLDD would equal:

$$\begin{array}{ll}
v & \text{at } t = 19 \\
v\lambda_8^{-1} & \text{at } t = 18 \\
\vdots & \\
v\prod_{i=k}^{8} \lambda_i^{-1} & \text{at } t = 10 + k \\
\vdots & \\
v\prod_{i=1}^{8} \lambda_i^{-1} & \text{at } t = 11.
\end{array}$$

So defining $G(10+k) = \prod_{i=k}^{8} \lambda_i^{-1}$, we want:

$$vG(k) \approx \widetilde{F}(k) \quad k = 11, 12, \ldots, 19.$$

More precisely, we seek values of the "parameters" w and v, which minimize the weighted sum of squared differences:

$$D(w,v) = \sum_{k=11}^{19} (k-10)\left(\widetilde{F}(k) - vG(k)\right)^2 = \sum_{k=11}^{19} (k-10)\left(w\widetilde{F}_\alpha(k) + (1-w)\widetilde{F}_\beta(k) - vG(k)\right)^2.$$

Since the focus is on the tail, we opt to weight the sum heavier with increasing k. Setting the two partial derivatives $\frac{\partial D}{\partial w}$ and $\frac{\partial D}{\partial v}$ to 0 gives two equations in the two "unknowns" w and v, which are readily solved:

$$
\begin{aligned}
0 = \frac{\partial D}{\partial w} &= \sum_{k=11}^{19} 2(k-10)\left(\widetilde{F}_\beta(k) + w\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right) - vG(k)\right)\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right) \\
\Rightarrow 0 = & \sum_{k=11}^{19} (k-10)\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right)\widetilde{F}_\beta(k) \\
& + w\sum_{k=11}^{19} (k-10)\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right)^2 \\
& - v\sum_{k=11}^{19} (k-10)\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right)G(k) \\
&= a_0 + a_1 w - a_2 v
\end{aligned}
$$

$$
\begin{aligned}
0 = \frac{\partial D}{\partial v} &= \sum_{k=11}^{19} 2(k-10)\left(\widetilde{F}_\beta(k) + w\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right) - vG(k)\right)\left(-G(k)\right) \\
\Rightarrow 0 = & \sum_{k=11}^{19} (k-10)\left(\widetilde{F}_\beta(k) G(k)\right) \\
& + w\sum_{k=11}^{19} (k-10)\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right)G(k) \\
& - v\sum_{k=11}^{19} (k-10)G(k)^2 \\
&= a_3 + a_2 w - a_4 v
\end{aligned}
$$

where:

$$
\begin{aligned}
a_0 &= \sum_{k=11}^{19} (k-10)\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right)\widetilde{F}_\beta(k) \\
a_1 &= \sum_{k=11}^{19} (k-10)\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right)^2 \\
a_2 &= \sum_{k=11}^{19} (k-10)\left(\widetilde{F}_\alpha(k) - \widetilde{F}_\beta(k)\right)G(k)
\end{aligned}
$$

$$
\begin{aligned}
a_3 &= \sum_{k=11}^{19} (k-10)\left(\widetilde{F}_\beta(k) G(k)\right) \\
a_4 &= \sum_{k=11}^{19} (k-10)G(k)^2
\end{aligned}
$$

and which lead to the solution:

$$
\begin{aligned}
a_0 &= -a_1 w + a_2 v \\
-a_3 &= a_2 w - a_4 v \\
a_0 a_2 &= -a_1 a_2 w + a_2^{\,2} v \\
-a_1 a_3 &= a_1 a_2 w - a_1 a_4 v \\
a_0 a_2 - a_1 a_3 &= \left(a_2^{\,2} - a_1 a_4\right) v \\
\Rightarrow v &= \frac{a_0 a_2 - a_1 a_3}{a_2^{\,2} - a_1 a_4}
\end{aligned}
$$

$$
\begin{aligned}
a_0 a_4 &= -a_1 a_4 w + a_2 a_4 v \\
-a_2 a_3 &= a_2^{\,2} w - a_2 a_4 v \\
a_0 a_4 - a_2 a_3 &= \left(a_2^{\,2} - a_1 a_4\right) w \\
\Rightarrow w &= \frac{a_0 a_4 - a_2 a_3}{a_2^{\,2} - a_1 a_4}
\end{aligned}
$$

If this solution falls outside the square $[0,1]\times[0,1]$, it is necessary to inspect the edges and corners to determine the optimal choice for w and v.

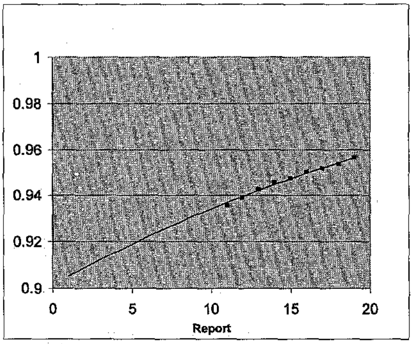
In this application, we break down $\widetilde{F}(t) = w\widetilde{F}_\alpha(t) + (1-w)\widetilde{F}_\beta(t)$ by selecting as one subset of claims those claims that close prior to an eleventh report. So we have:

$$\widetilde{F}_\alpha(k) = 1, \quad k = 11, 12, \ldots, 19.$$

There remains the selection of $\widetilde{F}_\beta$. Numeric examples are given which illustrate for $\widetilde{F}_\beta = \widetilde{F}_3$ and $\widetilde{F}_\beta = \widetilde{F}_6$, perhaps the two most attractive from the examples defined above. In the numeric examples below, we use the following age-to-age LDFs:

i	λ_i	k=10+i	$G(k)$
1	1.004808	11	.9779
2	1.003861	12	.9817
3	1.002915	13	.9855
4	1.001947	14	.9883
5	1.002930	15	.9903
6	1.001957	16	.9932
7	1.001961	17	.9951
8	1.002950	18	.9971
		19	1

Numeric Application 1: Since workers rarely start work much younger than age 20 and live beyond age 100, we select $\widetilde{F}_\beta(t) = \widetilde{F}_3(80;t) = 1 - \left(\dfrac{80-t}{80}\right)^3$. In the above notation, the solution occurs at w= 0.9016 and v=0.9565. We suggest that the tail factor be selected as $\frac{1}{\widetilde{F}(19)} = 1.046$. Charting the points $vG(k)$ together with the graph of the function $\widetilde{F}(t)$, the picture is:



In the next example we use an exponential. For that, it is convenient to note that the WC financial calls include open and closed indemnity claim counts. This provides the ability to estimate the conditional probability p_k of closure from report k to k+1, assuming a claim is open at report k. So suppose we have the eight probabilities $p_{11}, p_{12}, \ldots p_{18}$. We want to use that information to estimate the parameter $\hat{\theta}$ of $\widetilde{F}_5(\hat{\theta};t)$. For simplicity, suppose there were 100 claims open at report 11, then we would expect the following closure pattern:

$$c_1 = 100 p_{11} \quad \text{would close for some t} \in [11,12]$$
$$c_2 = 100 p_{12}(1-p_{11}) \quad \text{would close for some t} \in [12,13]$$
$$c_3 = 100 p_{13}(1-p_{11})(1-p_{12}) \quad \text{would close for some t} \in [13,14]$$
$$\vdots$$
$$c_k = 100 p_k \prod_{j=11}^{k-1}(1-p_j) \quad \text{would close for some t} \in [\text{k}, \text{k}+1]$$
$$\vdots$$
$$c_8 = 100 p_{18} \prod_{j=11}^{17}(1-p_j) \quad \text{would close for some t} \in [18,19]$$

and $\text{d} = 100 - \sum_{j=11}^{18} \text{c}_\text{j} = 100 \prod_{j=11}^{18}(1-p_j)$ would remain open at report 19.

It is convenient to simplify this still further and assume that the c_k claims all close at the midpoint of the time interval $= t_k = k + 10½$. Then it is easy to write out the maximum likelihood function for $\theta = \hat{\theta} + 11$. Indeed, we have c_k observed "failures" at $t = t_k$ and d observations "censored" at $t = 19 = s$.

$$L(\theta) = \prod_{i=1}^{8} \left(\frac{e^{-\frac{t_i}{\theta}}}{\theta} \right)^{c_i} \left(e^{-\frac{s}{\theta}} \right)^{d} = \theta^{-\sum_{i=1}^{8} c_i} e^{-\frac{\sum_{i=1}^{8} c_i t_i + ds}{\theta}}.$$

and the log-likelihood function is:

$$LL(\theta) = \log(L(\theta)) = \left(-\sum_{i=1}^{8} c_i \right) \log(\theta) - \frac{\sum_{i=1}^{8} c_i t_i + ds}{\theta}$$

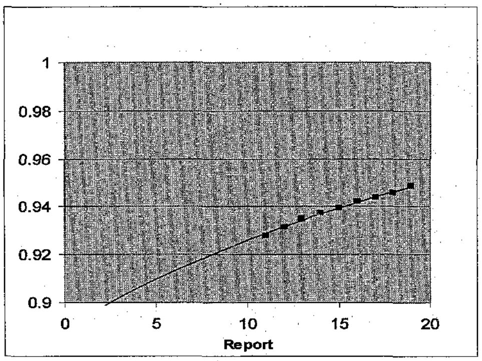
Setting $\frac{dLL}{d\theta} = 0$ to obtain the MLE estimator for θ. We obtain:

$$\theta = \frac{\sum_{i=1}^{8} c_i t_i + ds}{\sum_{i=1}^{8} c_i}$$

which is a useful rule of thumb for determining mean duration with such censored data—take the ratio that defines the weighted mean duration except only include the weight of non-censored observations in the denominator. For our purposes, this provides a simple way to estimate θ from the available WC data and so to specify $\widetilde{F}_\beta(t) = \widetilde{F}_5(\hat{\theta}; t)$. For a specific numeric example, suppose:

k	p_k	c_k	t_k	*Censored?*
1	0.070	7	11.5	No
2	0.075	7	12.5	No
3	0.081	7	13.5	No
4	0.076	6	14.5	No
5	0.082	6	15.5	No
6	0.075	5	16.5	No
7	0.081	5	17.5	No
8	0.070	4	18.5	No
		53	19	Yes

In this example $\theta = 36$ and $\hat{\theta} = 25$. Then w=0.8895 and v=0.9485, $\frac{1}{\widetilde{F}(19)} = 1.054$ and the picture is:



We close with one more numeric application. The idea here is to use a model like $\widetilde{F}_3$ for $\widetilde{F}_\beta$ but to match the mean duration as we did with the exponential $\widetilde{F}_5$. Suppose, as before, we have used financial data to establish the mean duration $\hat{\theta}$ of claim conditional upon the claim being open at an 11th report. Unlike with an exponential, it is not so convenient to relate the conditional duration $\hat{\theta}$ for $t>11$ with the unconditional duration μ. Instead, we will get around this by generalizing $F_3(b;t)$ to allow for a deferral period of $a=11$ (c.f. Corollary 2.2).

So first generalize $F_3(b;t)$ to $F_3(a,b;t)$ as follows:

$$f_3(a,b;t) = f_3(t) = \begin{cases} 0 & t \le a \\ \dfrac{2(b-t)}{(b-a)^2} & a \le t \le b \\ 0 & b \le t \end{cases}$$

$$F_3(a,b;t) = F_3(t) = \begin{cases} 0 & t \le a \\ 1-\left(\dfrac{b-t}{b-a}\right)^2 & a \le t \le b \\ 1 & b \le t. \end{cases}$$

We find that for $a \le t \le b$:

$$\eta_3(a,b;t) = \int_0^t x f_3(x)dx = \frac{2}{(b-a)^2}\int_a^t x(b-x)dx$$

$$= \frac{2}{(b-a)^2}\left[\frac{bx^2}{2} - \frac{x^3}{3}\right]_a^t = \frac{2}{(b-a)^2}\left(\frac{b(t^2-a^2)}{2} - \frac{t^3-a^3}{3}\right)$$

$$=\frac{2(t-a)}{(b-a)^2}\left(\frac{b(t+a)}{2}-\frac{t^2+at+a^2}{3}\right)$$

$$=\frac{(t-a)}{(b-a)^2}\left(b(t+a)-\frac{2}{3}\left(t^2+at+a^2\right)\right)$$

and so when t=b:

$$\mu_3=\mu_3(a,b)=\eta_3(b)=\frac{1}{b-a}\left(b(b+a)-\frac{2}{3}\left(b^2+ab+a^2\right)\right)$$

$$=\frac{1}{3(b-a)}\left(3b^2+3ab-2b^2-2ab-2a^2\right)=\frac{1}{3(b-a)}\left(b^2+ab-2a^2\right)$$

$$=\frac{(b-a)(b+2a)}{3(b-a)}=\frac{(b+2a)}{3}=a+\frac{b-a}{3}$$

which we could also have arrived at by recalling from the earlier example that the mean for the one parameter $F_3(b-a;t)$ case is just $\mu_3(b-a)=\dfrac{b-a}{3}$ and so:

$$\mu_3=\mu_3(a,b)=a+\mu_3(b-a)=a+\frac{b-a}{3}.$$

In our numeric example we would calculate b via:

$$25=\hat{\theta}=\frac{b-a}{3}=\frac{b-11}{3}\Rightarrow b=75+11=86$$

which suggests using 86 as the maximum time duration, a selection that is at least consistent with earlier considerations. Finally, we have:

$$\widetilde{F}_3(a,b;t)=\frac{\eta_3(a,b;t)+tS_3(a,b;t)}{\mu_3(a,b)}=\frac{3}{b+2a}\left(\eta_3(a,b;t)+tS_3(a,b;t)\right)$$

$$=\frac{3}{b+2a}\left(\frac{(t-a)}{(b-a)^2}\left(b(t+a)-\frac{2}{3}\left(t^2+at+a^2\right)\right)+t\left(\frac{b-t}{b-a}\right)^2\right)$$

$$=\frac{3}{(b+2a)(b-a)^2}\left(b\left(t^2-a^2\right)-\frac{2}{3}\left(t^3-a^3\right)+t(t-b)^2\right)$$

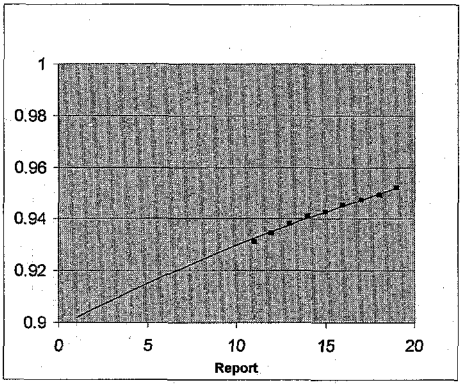
$$=\frac{3\left(b\left(t^2-a^2\right)+t(t-b)^2\right)-2\left(t^3-a^3\right)}{(b+2a)(b-a)^2}$$

$$=\frac{3\left(t^3-bt^2+b^2t-a^2b\right)-2\left(t^3-a^3\right)}{(b+2a)(b-a)^2}=\frac{t^3-3b\left(t^2-bt\right)+a^2(2a-3b)}{(b+2a)(b-a)^2}$$

When a=11, b=86 and $11\le t\le 86$, we have the PLDD determined as the cubic polynomial:

$$\widetilde{F}_\beta(t) = \widetilde{F}_3(11{,}86;t) = \frac{t^3 - 258t^2 + 22188t - 28556}{607500}$$

Applying this to the above example, the best fit is for w=0.99028 and v=0.9520, $\frac{1}{\widetilde{F}(19)} = 1.051$ and the picture is:



References:

[1] Corro, Dan, *Fitting Beta Densities to Loss Data*, CAS Forum, Summer 2002.

[2] Gillam, William R.; Couret, Jose R. *Retrospective Rating: 1997 Excess Loss Factors*, PCAS LXXXIV, 1997, including discussion: Mahler, H. PCAS, LXXXV, 1998.

[3] Klugman, Stuart A.; Panjer Harry H.; and Willmot, Gordon E., *Loss Models*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., 1998.