Does the ODP Really Agree with Chain Ladder?

The ODP here is defined as the distribution of aggregate claims where frequency is Poisson in λ and severity is a constant θ. This distribution may also be called the Poisson-constant severity, or PCS. Some authors define ODP more broadly as any distribution in the exponential family in which the variance is proportional to the mean. By certain uniqueness properties of exponential families it is not difficult to show that the PCS is the only distribution with this mean-variance relationship.

If X is the total loss random variable, X/θ is Poisson in $\lambda = EX/\theta = \mu/\theta$. Thus $Pr(X/\theta = n) = e^{-\mu/\theta}(\mu/\theta)^n/n!$. For $x = \theta n$, $Pr(X=x) = e^{-\mu/\theta}(\mu/\theta)^{x/\theta}/(x/\theta)!$. Now if μ is modeled by some covariates and parameters, say $\mu_{w,d} = m(U_w, g_d, h_{w+d})$, but θ is fixed, then an observation of $X_{w,d}$, say $q_{w,d}$, has $Pr(X_{w,d} = q_{w,d}) = e^{-\mu_{w,d}/\theta}(\mu_{w,d}/\theta)^{q_{w,d}/\theta}/(q_{w,d}/\theta)!$.

This is all fine, and you can show that in this case if $\mu_{w,d} = U_w g_d$, then the estimates of U_w and g_d by MLE agree with chain ladder.

A more complete specification of the probability function is $p(x) = Pr(X=x) = e^{-\mu/\theta}(\mu/\theta)^{x/\theta}/(x/\theta)!$ if x/θ is a non-negative integer and is zero otherwise. But it would be nice to be able to use p(x) as a density function on the positive reals, extending the factorial by the gamma function, so it is still defined and agrees at integer values, i.e., defining a! as $\Gamma(1+a)$. But then is p(x) really a density, that is, does it integrate to 1? Numerical testing finds that it is not. But Thomas Mack (????) has developed a workable alternative, which might be called the zero-modified continuous scaled Poisson but here will just be called Mack's Poisson.

The distribution is defined by setting the density $f(x;\mu,\theta) = e^{-\mu/\theta}(\mu/\theta)^{x/\theta}/[\theta(x/\theta)!]$ for $x > 0$ and adding a point mass at $x = 0$ to make the total probability 1. To see how much probability is needed at 0, define the function $pois(x,\lambda) = \lambda^x e^{-\lambda}/x!$ and the function

$zm(\lambda) = 1 - \int_0^\infty pois(x,\lambda)dx$. Then with a change of variable in f(x) to $y = x/\theta$ and defining $\lambda = \mu/\theta$, it is easy to see that the integral of $f(x;\mu,\theta)$ is $1 - zm(\lambda)$. Thus the point mass needed at zero is $zm(\mu/\theta)$.

Comparing f to p reveals that there is an extra θ in the denominator of f. But that will not affect the MLE of μ nor of the elements of μ if μ is a function of covariates. Thus the estimates of U_w and g_d will be those given by the chain ladder (as long as there are not any zero observations).

But there is still a need for caution. The expected value of X is no longer exactly μ. Integrating xf(x) shows that the mean is actually:

$$EX = \mu[1 - zm(\mu/\theta) + \int_{-1}^{0} pois(x,\mu/\theta)dx].$$

This is greater than μ, but not much so unless λ is small, as the table shows.

μ/θ	**zm(μ/θ)**	**EX/μ – 1**
0.2	.48628	.33861
1	.16619	.03291
5	.00216	9.43e-05
25	3.19e-12	1.96e-14

In a recent application with a fairly noisy runoff triangle, μ/θ was only less than 2 for one observation and less than 5 for five observations, out of 55. Thus only a few small observations would have fitted means noticeably different from the chain ladder's. Also the point masses are reasonable, but are a bit below those from a pure Poisson.

In the case where $\mu_{w,d} = U_w g_d$, the loglikelihood function is

$$l = \sum \left(\frac{q_{w,d}}{\theta} \ln \frac{U_w g_d}{\theta} - \frac{U_w g_d}{\theta} - \ln\left(\frac{q_{w,d}}{\theta}! \right) - \ln(\theta) \right)$$

The derivatives wrt U_w and g_d are the same as for the PCS, so the parameter distributions from the information matrix are also the same. Without the last term in the sum, l is an increasing function of θ, so θ would not be estimable by MLE, but with this term it should have an MLE estimate. The reserve estimate for small cells would be affected by θ, so the parameter variance of the reserve from the delta method would be affected as well. Also the process variance is slightly lower than for the PCS for small cells.