Timeline Simulation: Theory and Practice

Rodney Kreps
Intuition Quantified, LLC
Steve White
Guy Carpenter & Co, LLC

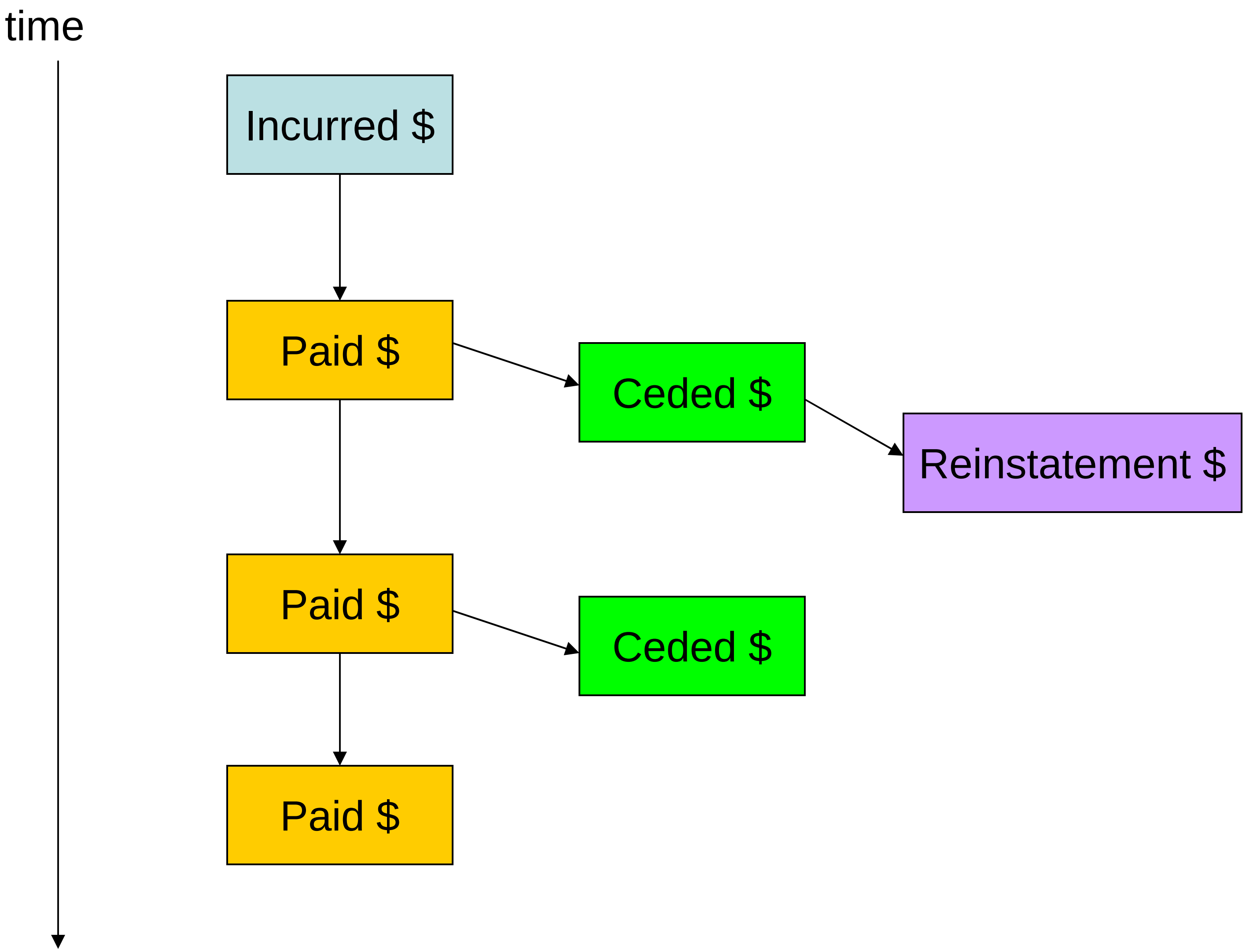
What is Timeline Simulation?

- Current simulations
  - Collective Risk model
  - Choose a time interval, and ask “how many events in the interval?”
- Timeline simulation
  - Associate a time with all events – everything is on a time line.
  - Ask “how long to the next event?”

Motivation: we want

- Transparency – complete audit trail for each realization
- Causality during each realization
- Reality – be able to model closer to how things actually happen
- Intuitive modeling.
- Reproducibility of current procedures

Transparency – Event Cascade



Event Cascade

- Every event occurs at a time.
- Every event is connected with causal links back to its ultimate source.
- We can pick up any event and see the chain that led there.
- Events can influence the generation of other events.

Events

- Events are basically anything of interest that you want to model.
- E.g. premiums, losses, ceded losses, loss adjustment expense, other expenses.
- Could include indexes, unemployment rates, economic activity, ...

Causality

Events can influence the generation of other events.

- Inflation hits claim payments and index clauses
- Successful suits engender more of the same
- Management rules affect writings or rates
- Economy can influence frequency and severity. Mortgage derivative insurance, anyone?

Reality

- Events *do* happen at points in time. Later events *can* be causally influenced by earlier events.
- Discounting and index clauses can be done exactly.
- Frequency and severity need not be independent.
- Seasonality can be easily done.

Intuitive

- Event generation is separated from reporting: there is one timeline and AY, AQ, RY, RQ etc are just different views.
- We can do new models simply. E.g. “big claims pay out later” is easy if you make the distribution of the time lag to payment depend on the realized severity.

Reproducibility

- Collective risk modeling generally amounts to assuming all losses happen in the middle of the period. This is a special case.
- Arbitrary frequency distributions can be used. Poisson, negative binomial, and more generally any mixed Poisson distributions are natural and easy.

Some Problems Avoided:

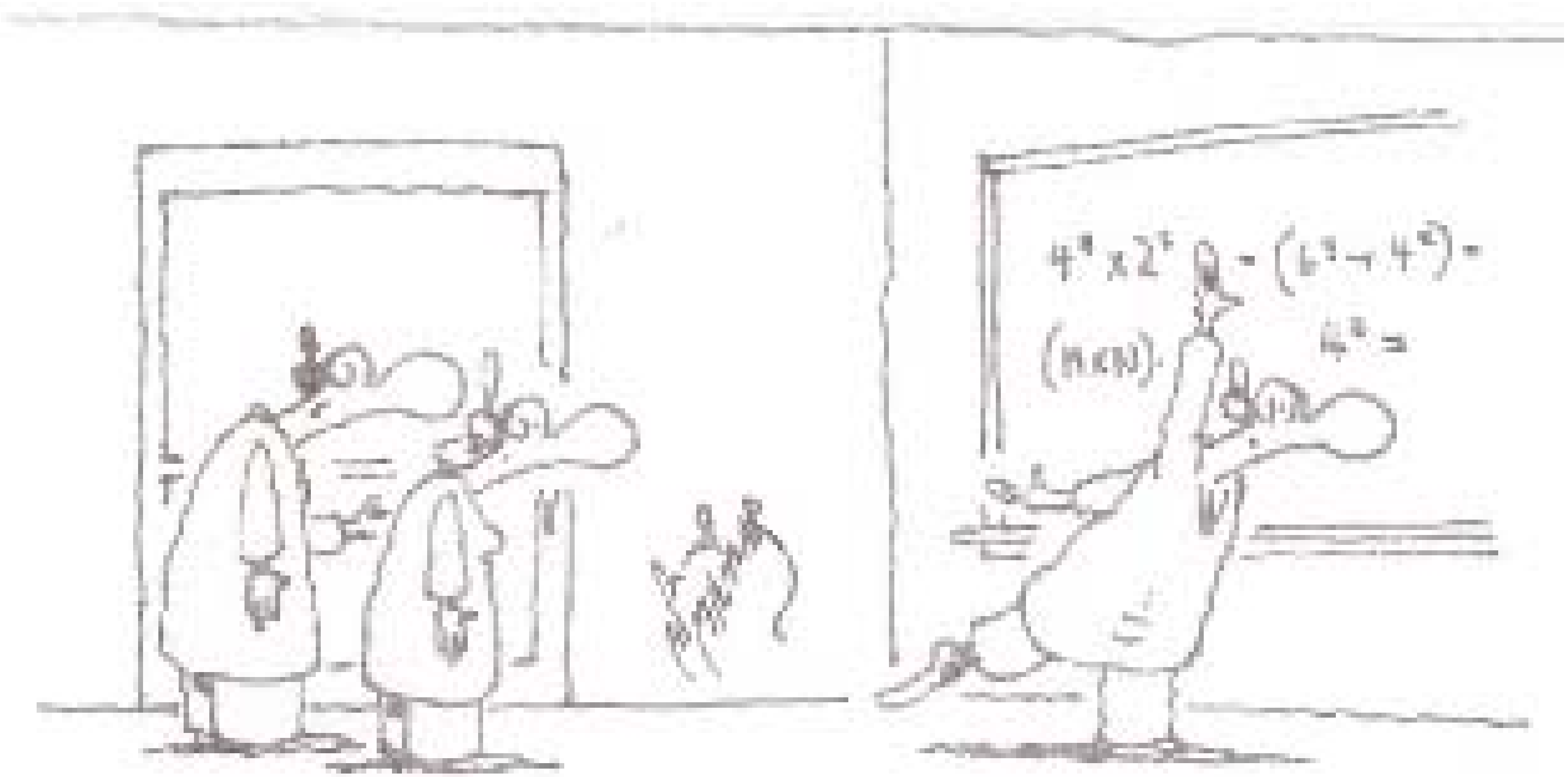
- Sparse matrices: with weekly periods most entries are zero, but still need housekeeping.
- Event generation dependent on reporting.
- Inappropriate allocation of deductibles.

Some examples where a timeline approach is useful:

- Success of one claim engenders others – think toxic mold.
- A change in exposure affects premiums, frequency of individual losses, and severity of aggregated losses.
- Indexation clauses
- Probability of two large hurricanes within two weeks of each other.

Theory

- Simple in principle: after each event ask for the time to the next, rather than how many events there are in a given time.
- The essential results are
  - Simulations are easy.
  - We can reproduce how we currently do things.
  - We can get new modes of thinking and models.



“At some point his theory becomes so abstract it can only be conveyed using interpretative dance”

-NOT

Theory

- Assumptions:
  - In an arbitrarily short time interval Δt there is at most one event. Clearly true for insurance work.
  - The probability of an event in Δt is proportional to Δt. The proportionality is the *instantaneous frequency*.

Instantaneous Frequency λ

- Intuitively, the propensity for an event to occur.
- It can depend on time, number of events, or anything else in the problem, such as previous events.
- As a formula, the probability of an event in Δt is

$$\Pr = \lambda(t, n, \ldots)\,\Delta t$$

Fundamental relation

- To have n events at $t+\Delta t$ you either already had n at t and did not get another in Δt, or you had n-1 and did get another.
- So: the probability of having exactly n events at time $t+\Delta t$ is the sum of the probability of n events at time *t* times the probability of no events between *t* and $t+\Delta t$ plus the probability of n-1 events at time *t* times the probability of one event between *t* and $t+\Delta t$.

Fundamental relation (2)

- Easier to see as a formula on probabilities:

$$P_n(t+\Delta t) = P_n(t)\left[1-\lambda(t,n)\Delta t\right] + P_{n-1}(t)\left[\lambda(t,n-1)\Delta t\right]$$

Had n

Didn't get one

Had n-1

Did get one

Differential Equations

- Go to limit as $\Delta t \to 0$:

$$\frac{d}{dt}P_n(t) = -\lambda(t,n)P_n(t) + \lambda(t,n-1)P_{n-1}(t)$$

- Boundary condition: No claims at t=0.

$$P_0(0) = 1, \;\; P_{n>0}(0) = 0$$

Probability of no events

- This is ultimately used in simulation.
- In general,

$$\frac{d}{dt} P_0(t) = -\lambda(t,0) P_0(t)$$

- The solution with the boundary condition is

$$P_0(t) = \exp\left\{ -\int_0^t \lambda(\tau,0)\, d\tau \right\}$$

Waiting Time Distribution

- The distribution of waiting time T from time t=0 is

$$F(T) = 1 - P_0(T) = 1 - \exp\left\{ -\int_0^T \lambda(\tau, 0)\, d\tau \right\}$$

- The substitution of a random uniform deviate for F(T) will generate a random time T to the next event. We need to solve for it, of course.

Poisson Process

- Defined by λ being constant.
- Can solve the differential equations and get

$$P_n(t) = \frac{(\lambda t)^n}{\Gamma(n+1)} e^{-\lambda t}$$

- This is recognizably the form usually used for a Poisson, although for arbitrary time.

More Poisson

- Exponential waiting time distribution from t

$$F\left(T,t\right)=1-\mathrm{e}^{-\lambda\left(T-t\right)}$$

- Random wait time given as

$$T-t=-\frac{1}{\lambda}\ln\left(uniform\,random\right)$$

- Timeline simulation is basically piecewise Poisson, with the exception of trend and seasonality which use their explicit time dependence.

Mean Count and Frequency

- The rate of change of the mean count is the frequency averaged over events:

$$\frac{d}{dt} mean(t) = \sum_{n=0}^{\infty} \lambda(t,n) P_n(t)$$

- If the frequency does not depend on count, the frequency *is* the derivative.

$$\lambda(t) = \frac{d}{dt} mean(t)$$

- For a Poisson, the mean increases linearly.

An Example of Count Dependence

- When $\lambda(t,n) = a + bn$ with $b>0$, the resulting distribution is negative binomial at any fixed time.
- Its mean increases exponentially in time

$$mean = \frac{\lambda}{b}\left(e^{bt} - 1\right)$$

Frequency mixing of Poisson distributions is

- necessary because of parameter uncertainty.
- also useful because a negative binomial can be represented as a gamma-distributed mix of Poissons.
- done at the start of a realization, and intuitively corresponds to choosing which world will be represented.

Frequency Mixed Poisson

- Formula is

$$P_n(t) = \int_0^\infty \frac{(\lambda t)^n e^{-\lambda t}}{\Gamma(n+1)} f(\lambda)\, d\lambda$$

- Mean number of events is

$$E(n) = t\int_0^\infty \lambda f(\lambda) d\lambda \equiv \mu t$$

- Variance to mean ratio is

$$\left[\text{var}/\text{mean}\right]_{count} = 1 + t\left[\text{var}/\text{mean}\right]_{mixing}$$

Negative Binomial as Gamma mix

- $f(\lambda) = \dfrac{\lambda^{\alpha-1} e^{-\lambda/\theta}}{\theta^{\alpha}\Gamma(\alpha)}$ The mixing mean is $\alpha\theta$ and the variance to mean is θ.
- The count distribution is

$$P_n(t) = \frac{\left(\frac{\theta t}{1+\theta t}\right)^n \left(\frac{1}{1+\theta t}\right)^{\alpha} \Gamma(n+\alpha)}{\Gamma(n+1)\Gamma(\alpha)}$$

- The mean is $\alpha\theta$t and the variance to mean is 1+θt.

Uniform mix

- $f(\lambda) = \dfrac{1}{b-a}$ for $a \le \lambda \le b$ and zero otherwise.
- The count distribution is

$$P_n(t) = \frac{G(bt, n+1) - G(at, n+1)}{(b-a)t}$$

- where

$$G(\lambda, n) = 1 - e^{-\lambda}\left\{1 + \lambda + \frac{\lambda^2}{2} + \ldots + \frac{\lambda^{n-1}}{\Gamma(n)}\right\}$$

Given the probabilities

- We can in principle always find a mixing function, but it may not be a distribution.
- E.g. if there is exactly one event, then the probability for no events is zero. That is,

$$0 = P_0\left(t\right) = \int_0^\infty f\left(\lambda\right) d\lambda$$

- This is not possible unless $f\left(\lambda\right) < 0$ somewhere, and probability densities are inherently positive.
- We can still generate timeline events, though.

Simulation

- At the start, and after each event, get the time to the next event.
- Events may be randomly realized, created in response to earlier events, or scheduled.
- Respectively, examples could be losses, reinsurance, and premium payments – or a lot of other possibilities.

How to do it?

- One way is to let time increase by intervals of Δt and in each interval see if there is an event using the current frequencies.
- This is a lot of realizations, and not necessary.
- We can use the waiting time distributions to find the next event. This is exactly equivalent to looking in each Δt, since the probability is still $\lambda\Delta t$.

So, A Better Way

- Take the problem as Poisson at any point in time, frequencies fixed until the next event.
- The time to the next event is

$$T - t = -\frac{1}{\lambda}\ln\left(uniform\,random\right)$$

- With many processes, we can evaluate the time for each of them and choose the earliest.

A Better Way (2)

- A sum of Poissons is Poisson. Having processes with frequencies $\lambda 1$, $\lambda 2$, … create the sum $\lambda = \lambda 1 + \lambda 2 + ...$
- Get the waiting time for this, and then choose which process pro-rata on the frequencies.
- This gives the same result, since the probability for an event in Δt for process n is $(\lambda \Delta t)(\lambda_n/\lambda) = \lambda_n \Delta t$

A Better Way (3)

- Having the next random event, compare to the next scheduled and use the earlier.
- After the event, poll the frequencies and get the next event.
- This allows arbitrary interdependence between events.
- The entire realization history is available to affect process parameters – and anything else that is modeled.

An Improved Better Way

- Let each generator keep track of its own next event time, rather than just frequency.
- Recalculate the time when an affecting event happens.
- With this, solely time-dependent frequencies can be done exactly without having to approximate them as steps.
- E.g. trend only requires solving a quadratic.

Spreadsheet examples

- Pure Poisson-Pareto with an XS cover.
- Pure negative binomial as a gamma mix, with a variable number of payments.
- Projection and parameter uncertainty for a negative binomial, and a variable number, timing, and amount of payments.
- Exposure-driven premium, large, and aggregated losses.
- More in the spreadsheet.

And finally

Now that you have seen the model T,

take a look at the Ferrari.