Charting the Path for Workers Compensation Claim Management

Dan Corro

Abstract:

With so much discussion about claim benchmarking, treatment protocols and the like, did you ever wish someone could just point you in the right direction? This analysis of the detailed workers compensation [WC] claim data now becoming available to researchers leads to a picture that resembles a simplified navigational chart. As described in the paper, that map--together with a few rules--provides powerful and potentially valuable guidance in administering WC claims.

Cost analyses are often issue driven. Consequently they tend to be focussed on a single cost liability. Medical costs, wage replacement benefits and loss adjustment expenses are the major categories in WC insurance. The focus is usually on determining their individual, ultimate cost liability. This paper describes some findings based on a new way to model claim costs that puts as much emphasis on their timing and interaction as on the costs themselves. As an illustration, back strain cases are looked at taking note of the mix between medical and lost time benefits. The major finding is hardly a surprise: mix matters. What might surprise you are the prospects for translating esoteric theory into practical guidance.

Introduction

Actuaries, especially life and pension actuaries, have always made use of mortality tables and the stick-man annuity formulas they seem to inspire. Nowadays, that type of analysis is more broadly applied. Engineers, for example, use it to evaluate the mean time to failure of a machine part while medical researchers use it to analyze drug trials and to evaluate treatment protocols. With these applications has come a major facelift. The study of "life contingencies" has been significantly advanced, especially through the incorporation of regression models and statistical theory, and is now called "survival analysis" (see [1] for a succinct, hands-on presentation).

At the same time, advances in data processing have yielded new and different crops of insurance data. Claim information files include a wealth of information never before captured in a readily accessible way. While the driving force was automated claims handling, the information collected may provide researchers with the raw materials needed for more refined statistical analyses. New WC industry-wide claim databases are being built that capture unprecedented detail on individual claims. In some instances there is even the ability to "drill down" to individual payment transactions. The work discussed here is the result of jury-rigging together a methodology to make greater use of that information.

This paper presents some early findings based on this new approach. Back injury cases are studied with an eye toward the interaction of medical and indemnity costs. While the theory is immature and the results only preliminary, hopefully they provide a taste of the fare we expect this new harvest of WC data to bring.

Background

We studied the interaction of medical and indemnity costs for a sample of back strain cases. The claim data used is from the NCCI Detailed Claim Information [DCI] database. The DCI is accurately described as precursor of the newer and more ambitious claim data marts now coming on-line. It is the natural "legacy system" and remains a good test bed for research. The DCI sample used in this study includes lost time claims from the states of Connecticut, New Jersey, New York and Pennsylvania. The study is restricted to injuries from 1983 to 1999 with medical and indemnity benefits each capped at $1,000,000.

Chart 1 groups the claims by indemnity and medical cost quartile, producing 16 [=4x4] separate buckets. Not surprisingly, the saddle shape confirms the strong correlation between indemnity and medical costs, especially at the high and low end cost cases. Because that relationship is so strong, understanding it better should lead to better claims management tools.

Consider, for example, the ongoing debate over the "sports medicine" approach to claims management. Recall the basic argument in its favor: aggressive medical care results in a faster return to work, thereby lowering the wage replacement cost liability. From a simplistic bean counter mentality, the challenge is to identify those cases for which the indemnity savings outweigh the added medical cost. Simplistic as that formulation may be, it poses a difficult question that remains to be resolved. A model that accurately captures the medical-indemnity cost interaction could contribute to that discussion, perhaps leading eventually to a definitive result.

So the goal is to model claims keeping track of the timing, itemization and interaction of claim payments. Individual payment transactions enable us to chart the progress of a claim as a function of time. With a little imagination, we can visualize this as a continuous path. This is a major departure from the traditional way of capturing claim data as a series of discrete snapshots (1^{st} report, 2^{nd} report, ...etc.). Tracing a continuous movement suggests a problem in Newtonian physics. On the other hand, it is more natural to think of a claim as exhibiting survival-oriented behavior, rather than the mindless motion of a "body of mass". This point of view suggests the use of survival analysis techniques, since much of that theory deals with behavioral responses. The model we are investigating is a hybrid, using techniques from survival analysis to organize and process the empirical data and then exploiting some ideas from mathematical physics to do the calculations and derive conclusions.

WC Cartography 101

While it is not really necessary to understand how such a "map" is derived to make use of one, it is helpful to have some basic understanding in order to avoid over-reading and misinterpretation. The discussion here is very general, the mathematical development is presented in [4], albeit without the word "map" (see also [3] and [5]).

A claim is represented as a trip or path on the map, beginning at the lower left-hand corner. Movement to the right, or due eastward, corresponds to paying medical benefits and movement upward, or due north, to paying indemnity benefits. As the model does not allow for recoveries (negative payments), claims progress in a northeasterly direction with no ability to backtrack.

In conventional survival analysis you observe "lives" and typically only take note of their "births" and "deaths" (and whether they hung around long enough to actually be observed to die). Much of its language has normative content, which can be bothersome. It is usually not good to "die" and often the kinder and gentler terms of "start" and "failure" are used. In our application, however, a life is a WC claim with "birth" corresponding to opening the claim file and "death" to claim closure. In this context, a quick death is not necessarily a bad outcome.

When constructing mortality tables, actuaries make use of the "force of mortality". That term is a bit old-fashioned. Survival analysis uses the more contemporary term "hazard rate function". Either one refers to the (instantaneous) rate of failure, expressed as a positive number. Taking its cue from the older "force" language, a key innovation of the claim model used here is to give hazard both magnitude and direction. Hazard is captured as a vector concept.

To continue with the terminology lesson, the proper name for this mathematical gadget is "vector field". In fact, we visualize the hazard literally as a field (or grassy meadow, if you prefer—the point is that blades of grass look like "vectors", since they have both length and direction). Claims cut out paths through this field from birth in the lower left to their eventual closure[1]—see Chart 2.

Unlike conventional survival analysis, we want to focus on more than just the birth and death of a claim and this is where the physics comes in. We model each observed claim by its entire path through the hazard field. Chart 2 shows two claims, C and D, that both close at the same cost (a,b) in medical and indemnity benefits, respectively. The two claims, however, took different routes in getting to that same end result. Conventional survival analysis is one-dimensional. Think of an infinitesimal bug on a time line. The bug can go fast or slow but not backward in time and has no opportunity to choose the path less traveled by. It is hoped that the use of multi-dimensional models to capture path choices will make all, or at least some of, the difference (with apologies to Robert Frost).

We have discussed two new ideas:

- Modeling a WC claim via its complete payment history and
- Capturing hazard as a vector field.

The two concepts work together: we visualize a claim as a trek over hazardous terrain and we look to our theory for guidance, presented here in the form of a "map".

We will not discuss here the task of constructing the hazard vector field, except to note that this is where 99.9% of the difficulty lies and that this part of the theory remains quite immature. For this study, we used ad hoc regression models to smooth out the discrete survival patterns produced from the empirical data.[2] It is hoped that with further study we will identify some functional forms that provide good analytical representations of WC claim survival data.

The remainder of this paper discusses the implication of a mathematical result known as "Green's Theorem in the Plane", a classical result discussed in most courses on advanced calculus.. It is certainly not necessary to understand Green's theorem to appreciate those

---

[1] In Cartesian coordinates, claims naturally enough originate at the origin (0,0). The x-coordinate tracks the cumulative medical payments while the y-coordinate the cumulative indemnity payments.

[2] For those interested in the methodology, we note that the claim data was fit to a survival vector field, rather than directly to a hazard vector field. More precisely, the steps taken were: (1) produce a lattice of survival vectors from the claim data (2) "invert" that survival lattice into a "gauntlet" hazard vector field and finally (3) use OLS regression models to smooth the gauntlet. (See [3] and [4]). For the last step, the x-component and the log of the y-component of the hazard vector were fit to a list of rational functions in x and y of degree 2 or less ($1, x, y, xy, x^2, y^2, 1/(1+x), 1/(1+y), 1/(1+x^2), 1/(1+y^2), 1/(1+xy)$). Both regressions had R^2 values of 0.95.

implications. Those implications are translated into a simple set of "navigational rules" in the next section. For those readers who are interested, the remainder of this section describes in a non-technical way what Green's theorem says and how it applies here. The truly math-phobic have permission to skip to the next section.

Take a deep breath: Green's Theorem tells us that the difference, *C-D*, in the work done going along two life paths to a common point equals the integral of the "rotation" of the hazard over the area between the paths (whew—and that's the simplest case). That is, the difference can be found by integration over the region *R* in Chart 2. This means that if the rotation is positive (counter-clockwise) on *R*, then more work is accomplished toward claim resolution by taking the lower path *C*. Conversely, if the rotation is negative (clockwise) on *R*, then more work is accomplished by taking the high road *D*. Moreover, while the paths must start and end together, the starting point need not be the origin.

The navigation map is just a plot of where the rotation is positive and negative. To express this in familiar terms, areas where the rotation is positive are called "land" and areas where the rotation is negative are called "water". Boundaries, where the rotation is zero, are (you guessed it) "coastlines". The navigation map produced in the back strain case study is shown in Chart 3. A coastline is "eastern" ("western", "southern", "northern", etc.) when you move east to reach the coast from inland. New York City, for example, is on the eastern US coastline, irrespective of whether it happens to fall on the left or right hand side of any particular map you are reading. In Chart 3, for example, the coastline on the left is an eastern coastline while the land area on the right is bordered by a western and by a southern coastline.

Rules to Die For

It is easy to use a claim navigation map like Chart 3, prepared from the back strain case study, provided you keep a few simple rules in mind. These rules apply when you have pre-allocated amounts of medical and indemnity dollars to spend. This is because the life paths must start and end together in order for what Green's theorem says about work to work. Remember that this simple model does not provide for subrogation or other recoveries, and so you can only go north or east. There are four cases, depending upon your current circumstances.

- *You are on water with no land in sight.* Head north then east to make more progress toward resolving the claim.

- *You are on land with no water in sight.* Head east then north to make more progress toward resolving the claim.

- *You are near a western coastline.* Avoid the coast to make more progress toward resolving the claim. (Western coastlines are paths of least resistance and so following them minimizes the work accomplished toward closing the claim).

- *You are near an eastern coastline.* Follow the coastline to make more progress toward resolving the claim. (Eastern coastlines are paths of maximum resistance).

As in so many adventure novels, it all comes down to finding the right map.

It is important to understand that the map and rules discussed here do not reveal any "best" course toward resolving a claim, they are only helpful in deciding between two ways of getting to the same place. It is clear from the theory that questions about the existence, uniqueness and determination of optimal paths are much harder. See [4] for an illustration of how the theory rhymes with fixed asset allocation and benefit cost minimization (What, too many syllables?).

Of course, this simple, two-color map can be refined into a "contour map" that warns of particularly rough terrain and especially turbulent waters. Also, while the bean counters would certainly urge you to shorten your trips, distance traveled ($\sqrt{\Delta x^2 + \Delta y^2}$) does not equal the money paid getting there ($\Delta x + \Delta y$); suggesting maybe using an alternative scale. Hopefully, advances will be made on these and on related issues as the theory is applied.

There are two basic problems to be addressed by a mature theory:

- First, assess the "work" remaining to resolve a claim
- And then, determine an efficient path for completing that work.

The next section presents a case study with more and less efficient paths and so the path choice does matter. The extent to which these problems can be solved remains to be found

Back Strains: A Case Study

We are finally at the fun part. Refer to Chart 3 which shows the map for resolving back strain claims. When there is no rotation, the path does not matter[3]. The basic finding of this study is that timing matters and that there is both positive and negative rotation out there influencing the resolution of the claim sample.

For example, what does the map suggest as regards the sports medicine debate? First, note that we are only considering **dollars** of medical and indemnity benefits. Other such

---

[3] Vector fields with rotation identically 0 are called *conservative*. These are the vector field that have a potential function and are characterized by the fact that the amount of work done moving from one point to another is independent of the path taken. For example, the potential energy a rock loses when moved from the top to the bottom of a hill will be the same whether you throw it, kick it or carry it in your shoe. The astute reader will note that the map discussion conveniently ignores the possibility of "conservative coastal areas". While perhaps politically odd, such areas can occur. The smoothing functions used here reduce them to (lower dimensional, measure zero) subsets that can be ignored. In any event, where the rotation vanishes identically, progress toward claim resolution is independent of the path and the only guidance Green has to offer is to the limits of indifference. Short form: it would have messed up the rules without adding anything.

models may incorporate better medical utilization metrics; here we make do with medical dollar costs as a surrogate for medical utilization.

With that limitation in mind, though, the experience of the lower cost cases (near the origin) does not support the sports medicine approach. There, it is best to follow the eastern coastline which allocates the lion's share of dollars to replacing lost income. This has some common sense appeal. Note that this observation applies only so far because further north that coastline veers west along an impossible track

Look next toward the right but still along the bottom. That part of the map pictures a danger inherent in the sports medicine approach. There you run the risk of becoming trapped within an inlet and being forced aground on a western shore. Recall the rule to avoid this because expending resources along a western coastline achieves minimal progress toward resolving claims.

For higher cost claims, the upper right region tells yet another story. There we see a western coastline just below the line y = x and nearly parallel with it. There is an identifiable path of least resistance along which medical benefits and income replacement benefits continue to be paid out at about equal rates. The spine and especially the "saddle horn" in Chart 1 suggest that this is a popular route. Since western coastlines are to be avoided, this advises against such a middle of the road course. While it is not clear which is better in any given case, the suggestion here is to either adopt or clearly reject the sports medicine model in any given case. And further, sticking by that decision whenever possible. It warns of maintaining a level of palliative care inadequate to bring the injured worker back to work. Of course, in practice there may be little recourse away from that track.

Combining the map (Chart 2) with the claim distribution (Chart 1) highlights the value of making a determination early on and breaking away from the pack. This observation again has some common sense appeal. At this stage, the map offers little but an "I told you so" in the event of a bad call. Consider how much more valuable the theory would become if it could lead to identifying the "correct" choice on a case by case basis. By investigating how certain claim characteristics impact the geography, the approach provides a blueprint for resolving the debate over the sports medicine model.

The skeptic may view the upper right of the chart as just a graphical representation of a known and rather obvious pitfall to avoid when managing a back strain injury claim. Nevertheless, this picture was drawn from "hard" empirical data, not anecdotes. At such an early stage, the theory is unable to assess the degree that this picture is the reflection of intelligent versus blind choices.

Suppose you are confronted with a fairly serious back strain injury. You recognize that there is much "work" needed to resolve this case and so you look toward the upper right as your likely final destination. You believe you would do better ending up on land and so you decide to use the sports medicine approach. You must make an important strategic decision and decide upon a landing point along that dreaded western coastline.

You also need to avoid being sucked along the coast, as that path offers less resistance to having claim payments just continue on. That western coast is especially dangerous since it offers an optimal course for those seeking to maximize their take from the WC system. Naively, then, the map suggests landing on the south shore, since that avoids being drawn into the "y=x" pitfall for at least the near term. Nevertheless, you must still be wary of medical costs looming near due east along that same coastline. However, more specific information would be needed together with some number crunching to determine whether that is actually what the model indicates in any particular case scenario.

Summary

A confluence of factors has combined to produce a new generation of computerized WC insurance claim information. This paper describes, in a mostly non-technical way, a new mathematical model for WC claims. The model was developed to take advantage of that claim data. It combines ingredients from contemporary survival analysis with classical physics. A case study of back strains was done to determine whether the theory could be applied to real world data, and if so whether anything of interest would come from it. To illustrate the potential applications, the theory is used to construct a "map" to help navigate the resolution of WC claims. That simple picture is a "surface map" in more than one sense. Hopefully it represents only the surface of what this theory may potentially yield. There is the chance we may strike gold by digging deeper into the theory and mining the data.

References

[1] Allison, Paul D., *Survival Analysis Using the SAS® System*: A Practical Guide, The SAS Institute, Inc., 1995.

[2] Corro, Dan, *Determining the Change in Mean Duration Due to a Shift in the Hazard Rate Function*, CAS Forum, Winter 2001.

[3] Corro, Dan, *Modeling Loss Development with Micro Data*, CAS Forum, Fall 2000.

[4] Corro, Dan, *Modeling Multi-Dimensional Survival with Hazard Vector Fields*, CAS Forum, Winter 2001.

[5] Corro, Dan, *A Note on the Inverse Relationship of Hazard with Life Expectancy*, submitted to CAS Forum.

[6] Spivak, Michael, *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus*, Perseus Books, 1965.

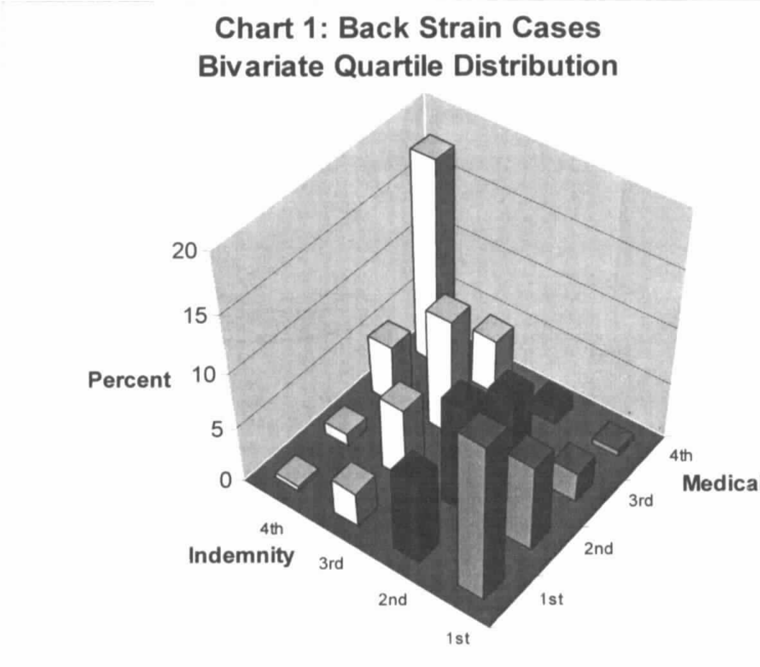
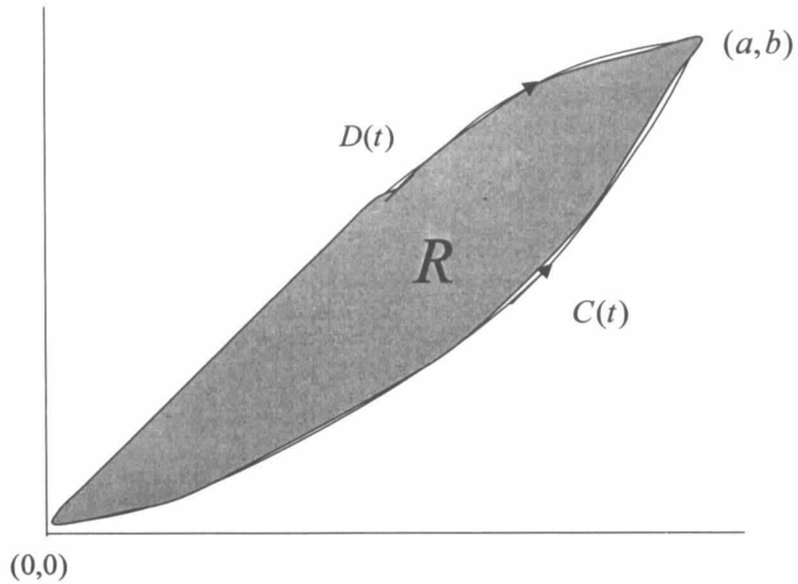


Chart 1: Back Strain Cases
Bivariate Quartile Distribution
Percent
20
15
10
5
0
Indemnity
4th
3rd
2nd
1st
Medical
4th
3rd
2nd
1st

Chart 2: Claim Paths



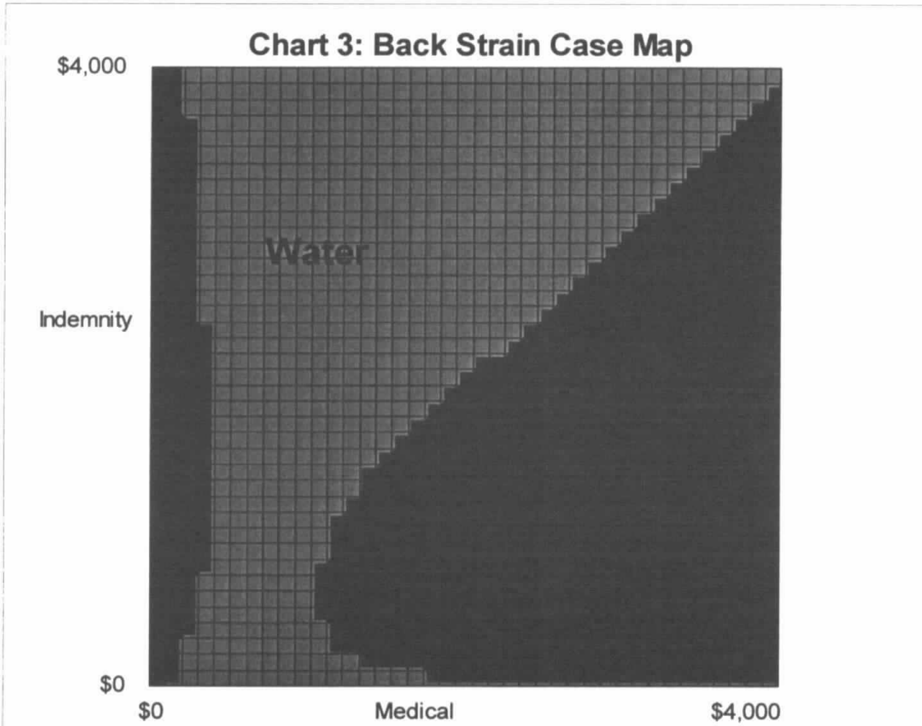


Chart 3: Back Strain Case Map
$4,000
Indemnity
$0
$0
Medical
$4,000
Water