Journal of Risk Finance, Vol. 9, No. 4, 2008.

EDITORIAL

The Sequential Sawyer –

A Tale of Frequentist Fright

Michael R. Powers*

Some years ago, the mythical small town of Two Bits, Texas was terrorized by a series of five gruesome (and unsolved) murders. At the time, the sheriff and sole law-enforcement officer of Two Bits was a man named R.A. While the use of initials as first names is quite popular in the Lone Star State, the residents of Two Bits had taken this tradition to something of an extreme by forgoing surnames entirely, and identifying themselves exclusively by initial pairs. So R.A. was just R.A. – not "Ronald Aylmer," or anything of that sort; and in fact, at the time of the events I am about to relate, the town's most recent census showed a population of exactly twenty-five censees, all denominated in like manner:

A.C.	B.A.	C.B.	D.A.	E.R.
A.D.	B.B.	C.E.	D.C.	M.R.
A.I.	B.E.	C.O.	D.J.	R.A.
A.K.	B.O.	C.V.	D.P.	R.B.
A.Y.	B.S.	C.W.	D.T.	T.Y.

Two Bits was a quiet town in which nothing of much significance ever seemed to occur – until the aforesaid murders eliminated twenty percent of its population in a matter of just a few weeks. The first sign of trouble arose during the quiet hours of a Saturday evening in early October. Long before the sun had set, but well after the time when all the good citizens had retired to bed, the stillness was shattered by the sudden roar of a chainsaw. Although bothered by the unholy sound, the townsfolk were content to bury their heads in their pillows until it had

subsided. The next morning, however, they quickly discovered that the previous evening's disruption had left a chilling memento: the lifeless and awkwardly contorted body of A.C. was found in his home, having been sawed into two (bits) at the waist.

The following Saturday, when the unmistakable chords of an angry chainsaw again filled the air, the residents of Two Bits were quicker to respond. Jumping from bed, Sheriff R.A. immediately formed a posse to search for the noisy killer. Regrettably, his efforts were of no avail. Upon checking every domicile in town, the posse was too late to save resident B.E. from a bifurcated ending.

After a week of grave uneasiness, a similar drama was repeated the next Saturday, at which time the crimson entrails of C.E. announced the killer's latest victim. Then, following another week of wondering who would be the next to feel the concatenated cutter's greasy steel, the residents' relived their terror of previous weeks with the discovery of D.C.'s divided corpse.

This time, however, there was a minor enhancement of the killer's *modus operandi*. Neatly tucked into the victim's jacket pocket, so that only Sheriff R.A. would find it, was the following typewritten note:

A.C., B.E.,

C.E., D.C. –

In Two Bits saw each letter pair.

Four heads, or feet,

A line then pleat,

To point to Sawyer's Hallows' fare.

– The Sequential Sawyer

Now, while R.A. may have been the simple sheriff of a mythical small town, he was not to be underestimated intellectually. Prior to entering the field of law enforcement, he had worked for many years as a cryptanalyst for the U.S. government's National Security Agency; and before that, he had earned a Ph.D. in statistics from a prestigious west-coast university. Given these impressive qualifications, he quite readily saw the killer's pattern: The first initial of the first victim was an A; the first initial of the second victim was a B; the first initial of the third victim was a C; and the first initial of the fourth victim was a D. Thus, the killer was selecting his victims alphabetically – and so next week's target would be E.R., the only resident of Two Bits whose first name began with an E.

However, R.A. was not only a deep thinker, but also a conscientious scientist. Although he believed he had solved the problem, he wanted to make sure that he had left no stone unturned in his investigation – and that required a complete understanding of the killer's cryptic poem. Apparently, there were several steps necessary for the analysis:

- The third line indicated that each of the four pairs of initials – A.C., B.E., C.E., and D.C. – was to be divided into two separate letters.
- The fourth line suggested, rather graphically, that these individual letters were to be segregated into two groups, "heads" and "feet". Apparently, the first initials – A, B, C, and D – were the "heads", and the second initials – C, E, E, and C – were the "feet".
- According to the remainder of the poem, either the four "heads" initials or the four "feet" initials would form a "line" to "point" to the intended victim of the following Saturday, which happened to fall on October 31 (Halloween). In other words, one of the two sequences,

  (I) A, B, C, D, … or

(II) C, E, E, C, …,

should direct the sheriff to the killer's fifth target.

To study the problem formally, R.A. decided that it would be best to replace each of the letters in models (I) and (II) with the integer corresponding to its place in the alphabet (i.e., to substitute 1 for A, 2 for B, 3 for C, etc.). This yielded the following number sequences:

(I) 1, 2, 3, 4, …

(II) 3, 5, 5, 3, ….

Clearly, the first sequence was just the counting numbers, and so the fifth term should be 5. However, despite R.A.'s extensive analytical background, he couldn't immediately see any obvious mathematical structure underlying the second sequence.

Turning the words of the killer's poem over in his mind, he finally decided that the answer must lie in the phrases "a line then pleat" and "to point to." Just as the first four counting numbers, 1, 2, 3, 4, formed a line in two dimensions when plotted against equally spaced intervals, so the other four integers, 3, 5, 5, 3, also must form a line when similarly plotted. However, in the latter case, the line would have to be a statistical *line of best fit*, since no ordinary line would pass through those four particular numbers. Fortunately, the fitted locus was easy to envision: it was simply the horizontal line passing through the integers 4, 4, 4, 4. Consequently, just as the line through 1, 2, 3, 4 "pointed to" the number 5, so the line through 4, 4, 4, 4 "pointed to" the number 4.

In other words, model (I) predicted that the first ("heads") initial of the next victim would be an E (i.e., that the killer's target was E.R.), whereas model (II) predicted that the second ("feet") initial of the next victim would be a D (i.e., that the killer's target was A.D.).

R.A. was pleased to see that his original guess (E.R.) had survived this analysis, but he still needed to confirm the earlier conclusion by showing that model (I) was more appropriate than model (II). Fortunately, as a skillful *frequentist* statistician, he had the powerful method of *maximum likelihood* (ML) at his disposal. Under the ML principle, all that was necessary was to determine which of the models, (I) or (II), was most consistent with the observed data (i.e., with the identities of the first four murder victims). This could be accomplished by computing the probability of obtaining the particular sequence of murder victims assuming that each of the two models, in turn, was the correct one. The model that yielded the greatest such probability, or "likelihood", would be the one he should use to identify the killer's next target.

To this end, R.A. made the following calculations:

$$\text{(I)}\quad \Pr\{\text{A.C.,B.E.,C.E.,D.C.}|\text{Model (I)}\} = (1/5)(1/5)(1/5)(1/5) = 1/625$$

$$\text{(II)}\quad \Pr\{\text{A.C.,B.E.,C.E.,D.C.}|\text{Model (II)}\} = (1/2)(1/2)(1)(1) = 1/4\,.[^1]$$

Rather remarkably, he found that the likelihood associated with model (II) was greater than that associated with model (I) by a factor of more than 150. In other words, the model that pointed to A.D. as the next victim was vastly more "likely" than the model that pointed to E.R.! So how could R.A. have been so wrong in his original assessment?

"Ah," he thought to himself, "Now I see the problem. Just like one of those simpletons who thinks he see faces on the surface of Mars, I was tricked by the familiar! When I saw the pattern 1, 2, 3, 4, I gave too much weight to my *prior belief* that the sequence would be simple and recognizable. Now I know that it's A.D., not E.R., who needs protection on Halloween."

Well, Halloween came and went, and appropriately enough, our hero was tricked – but not by the familiar. As is clear to anyone with common sense, there are reasons why the familiar is familiar. In this case, it was because the sequence of counting numbers was just as familiar to

the killer as it was to R.A.; and that was much more relevant to the problem than the fact that so many of Two Bits' residents had names beginning with A, B, C, and D (thereby lowering the likelihood of model (I)). Consequently, while R.A. was wasting his efforts tending A.D., the killer was effortlessly rending the waist of E.R.

---

* Editor, *Journal of Risk Finance*; Professor of Risk Management and Insurance, Fox School of Business, Temple University; Distinguished Visiting Professor of Finance, School of Economics and Management, Tsinghua University; e-mail: michael.powers@temple.edu.

[1] In the model (I) likelihood calculation, the factors 1/5, 1/5, 1/5, and 1/5 are used because five residents have first initial A, five residents have first initial B, five residents have first initial C, five residents have first initial D, and each of these letters appears only once. In the model (II) likelihood calculation, the factors 1/2, 1/2, 1, and 1 are used because two residents have second initial B, two residents have second initial E, and both letters B and E appear twice (so that after the first one appears, the probability of the second one increases from 1/2 to 1).