magann/iStock/Thinkstock

Warfare

In "What are the chances of war?" (page 44, April 2016) Pasquale Cirillo and Nassim Taleb attempt to educate us on fat-tailed distributions, but they need not have bothered. The book they attack (but show no signs of having read), Pinker's *The Better Angels of Our Nature*, devotes many pages to the power-law (fat-tailed) distribution of war magnitudes, and many more to the Poisson (temporally random) nature of inter-war intervals. These facts have been known at least since the publication of Lewis Fry Richardson's *Statistics of Deadly Quarrels* in 1960, and have informed quantitative discussions of war ever since.

Contrary to Cirillo and Taleb, we are perfectly aware that a stretch of time without a big war does not imply that a big war cannot happen. The issue is whether the parameters of the processes generating new wars and determining their magnitudes have changed since 1945. This is the era that historians have called "the Long Peace", in which wars between great powers and wars between developed states, common throughout recorded history, essentially disappeared. In the familiar analogy of drawing balls from urns, the idea is not – contrary to Cirillo and Taleb – whether drawing a series of balls with low numbers is taken to suggest that the urn contains no balls with very high numbers; it is whether there is reason to suspect that the urn has been tampered with so as to change the number of balls with numbers of various sizes.

It is true that any test of sample numbers alone is assumption-dependent and may be challenged. That is why all such tests must be interpreted in the light of historical evidence (the equivalent of actually monitoring whether someone has tampered with the urns). In the case of the post-war period, this evidence includes precocious observations by historians of qualitative changes in the international system made decades before the decline in war frequency was apparent, sharp reductions in independent predictors of war such as conscription, length of military service, and military expenditures as a proportion of

GDP, and radical changes in the norms and institutions governing the conduct of states.

For Cirillo and Taleb, no historical evidence can be relevant to the question of whether the risks of war have changed. They fitted a simple model to data on wars spanning two thousand years, failed to spot a trend over these two millennia (as if anyone had claimed there was one), and noted that in their model the probability of a large war in the next century is non-negligible (as if anyone had claimed otherwise). But Cirillo and Taleb failed to perform the critical comparison between the post-World War II period and a comparable one preceding it. More generally, any modelling effort that squeezes two millennia of wildly heterogeneous history into a single distribution, and then affirms the null hypothesis of no overall trend, is ill equipped to shed light on what has happened in the past 70 years.
Michael Spagat, Royal Holloway University of London, and Steven Pinker, Harvard University

Prediction versus time series forecasting

On page 19 of the April 2016 issue, there is a note: "Prediction versus forecasting". The research at bit.ly/1SJxTQz shows an example of the importance of this concept.

A time series forecast can be made for any data item collected on a regular basis. But if you do not have such a time series for every member of a finite population, then a forecast for the estimated total, for such a data item for such a finite population, cannot be made. If you do not want a forecast at all, but an estimate of a total from a current finite population, based on a current sample or a current attempted census with non-response, then one option may
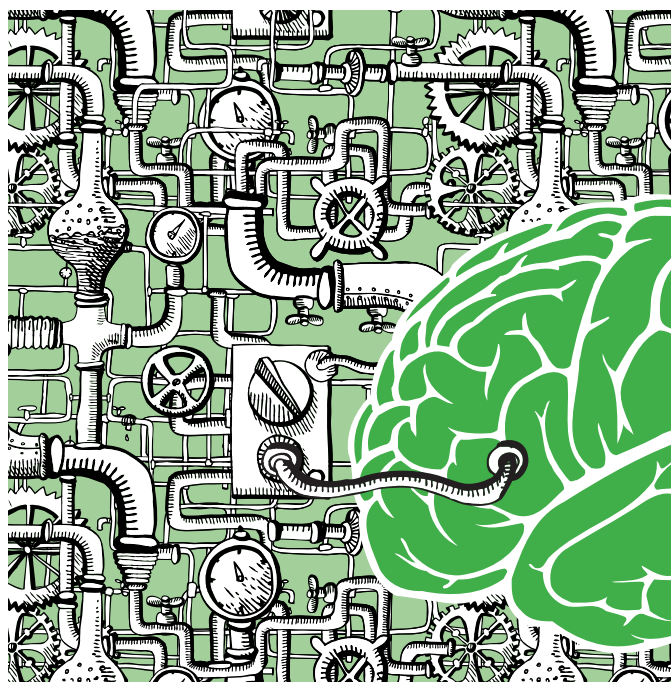
Kevin Reed/CPL

be "prediction" (using regression). The regression sought is between this current sample or near-census, and regressor data (which may possibly be an older census of the same data item). Prediction is one option for obtaining an estimated current total for a data item (such as propane production), for a finite population (say, in a given geographic region), not a forecast of it based only on previous data. Yet you can either forecast or predict for an individual respondent, depending upon what other data you possess.

When we want to know something about current economic conditions – say, to report official statistics – we are not so interested in forecasts based only on old data. Imputing for some missing data using time series data for individual respondents, when available, might often do well, for a forecast, but when it is most important – when, say, there has been a sudden change in a market – a time series cannot tell you about current conditions. They have to assume old patterns are still in effect.
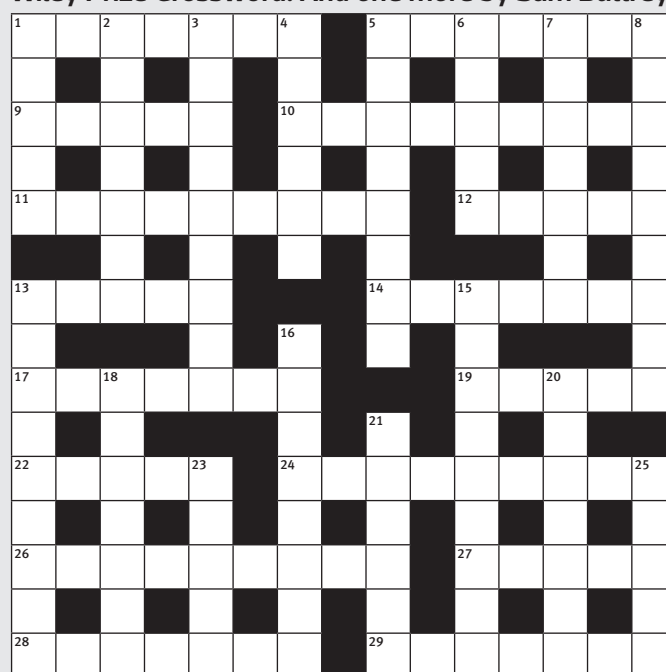
Thus it is very important to distinguish between time series forecasting and prediction. Time series forecasting extrapolates from patterns. Prediction seems a misnomer. With prediction you are modelling current data, using regression, to employ that relationship to estimate ("predict") for the unobserved data in a current finite population.

James R. Knaub, Jr.
Reston, Virginia, USA

Stopping rule revisited

Michael Wininger's analysis of the potential for shortened (American) college-level football games using early-stopping rules (page 30, December 2015) seems to have included a tautological query. In evaluating data from 2544 games over a three-year span (2012–14), he looked for the minimum score differential at each interim quarter that would ensure correct identification of the game's victor at least 99% of the time, with a minimum *n* of 30 games. These thresholds were 22, 21, and 15 points at the end of the first quarter (Q1), half-time (Q2), and the third quarter (Q3), respectively. He then tests the values (derived from this data set) against the same data set to determine "how accurate ▶

Wiley Prize Crossword: *And one more* by Sam Buttrey

Send your solution to: *Significance* Crossword Competition, Royal Statistical Society, 12 Errol Street, London, EC1Y 8LX or scan it and email to significance@rss.org. uk. The competition is sponsored by Wiley (wiley.com/statistics), who will give the winner £100 or $150 to spend on Wiley books. Closing date: 16 July 2016. The winner will be chosen randomly from the correct entries, and the correct solution published in a future issue. Photocopies are acceptable.

Answers to the eleven italicised clues need to be modified in a similar way to fit into their spaces. For those clues, enumerations are withheld, but in the spirit of fairness, two-word and hyphenated answers are indicated. Other clues are normal. Solvers are invited to identify the sequence of letters missing from the grid.

Across

1 *Small red car breaks inside, leading to declaration of hostilities* (2 wds)
5 *Jefferson, initially, touches oil to anoint place with three arms* (2 wds, hyph.)
9 *Criticism might lead to small chills*
10 School I cut after leaving home with sister, trailing, curving back (9)
11 Tragic heart wail is who's left when will is lost (4-2-3)
12 *Source of heat to confuse, irk papa*
13 Atria redesigned on cool side of India (5)
14 Suffered flagellation with alarm (3, 4)
17 *Remain agitated after flipped-over vehicle ends up underwater*
19 Viennese piazza holds old picture (5)
22 *Canal builder is without truncated records*
24 "Aye, aye," and so on; secretly, I covet veto as time passes (5, 4)
26 Knock over vehicle, split the stuff that's easier to get to (4, 5)

27 *Modifies entrances outside the county*
28 Old Spanish coin I purloined after second of April (7)
29 *Voters run amok, causing politician replacements*

Down

1 Size, with depth in the interior (5)
2 Tell chronicler, albatross (in idyll) is captured (7)
3 Range of time during African voyage leading to Jamaican movement (9)
4 Fords tops of dammed stream covered in slippery fish (6)
5 *With resistance, unfortunately, Raj was not in place for stories of Paris*
6 Some of stench implicates hairy fellow (5)
7 Establish a place for the spirits? (7)
8 Lost again, bewildered, in "Remembrance of Things Past"? (9)
13 Old iron, moved higher, heated up quickly (7, 2)
15 Creeds end, unconventionally re-thinking part of p's and q's (9)
16 *June revels confused far-out writer* (2 wds)
18 Two religious groups have crosses? (7)
20 Using ends of paddle, grab broken motor and push (7)
21 *Nautical preserver – albeit of destruction*
23 Thus was written a thorough defeat of a sadist (5)
25 Cuts down on trouble from eastern fool who grabs spouse's behind (5)

Solution to February issue's crossword:
Turn up the heat by Goujeers

Answers to italicised clues are types of chilli (or chile) peppers. The ordering is by increasing order of heat on the Scoville scale.

Across: 1 FRES(h) NO, **4** anag, **9** 2 defs, **10** POT in CHILE, **12** anag, **13** BAN ANA, **15** CAR in REIN NATES, **18** anag, **21** (fortissi) MO RITA(rdando), **22** ALAN in anag, popper, **24** EYE = pheasant's brood, **25** Linear A, Linear B ancient Greek writing systems, **26** IN in SISTER, heraldic left, **27** (n)UN HOL(l)Y.

Down: 1 FURT(ive) HERS, **2** CLIP in CITE rev, **3** anag, **5** (moust)ACHE, **6** anag, **7** PARTNE(r) rev, **8** NE (neon, rare gas) in OMAN, **11** R in MIDI FF, **14** YEN in CANE, **16** A BAN in HERO, **17** anag, **19** AM, BASE with E moving to the top EBAS, **20** hidden, **23** homophone of bite.

F	R	E	S	N	O		J	A	L	A	P	E	N	O
U		C		E			C		L		N		N	
R	O	L	L	U	P		C	H	I	P	O	T	L	E
T		I		R		M		E		H		R		M
H	Y	P	N	O	T	I	C		B	A	N	A	N	A
E		T		S		D		C		B		P		N
R	E	I	N	C	A	R	N	A	T	E	S			
S		C		I		I		Y		T		H		B
			B	E	E	F	T	E	R	I	Y	A	K	I
A		D		N		F		N		S		B		T
M	O	R	I	T	A		I	N	H	A	L	A	N	T
E		Y		I		B		E		T		N		E
B	I	R	D	S	E	Y	E		L	I	N	E	A	R
A		U		T		T			O		R		L	
S	I	N	I	S	T	E	R		U	N	H	O	L	Y

Winner, April: Patrick Hanson, North Berwick, UK

▶ [we would] be in identifying the winner in games that were [as] lopsided in these interim assessments". The answer, of course, is, 99% (1134 of 1142 games that would have been stopped would have been correctly forecast)!

A more convincing analysis would have been to derive the stopping rule from a random sample of the games, then apply it to the remaining cases from the data set. Alternatively, one could apply the stopping rule to a different year's game results. Aside from this point, the only other comment I would make is that far more (American) high school football games are played each year than college games – several orders of magnitude, with arguably less medical expertise and poorer facilities available for athletes on-site. Wininger's premise that an early-stopping rule would be a simple way to reduce the incidence of traumatic injury to players is certainly worth serious consideration by athletic conferences and federations. Of course, for early game mismatches, the parents of marching band and associated performers would still insist on having the half-time show go on!
David Morse
Mississippi State University

The author responds: I agree with David Morse's valid remark regarding the tautological nature of the analysis as presented. In response, I have chosen to cross-validate the published rule, built from 2012–14 data (a so-called "training set"), against game play data from 2009–11 (the "testing set"); data were extracted from the same resource, and processed identically.

The data from the training set and testing set were similar in terms of basic descriptors:

	Training set (2012–14)	Testing set (2009–11)
Sample size (*N* games)	2544	2426
End-of-game score differential (median, IQR)	16 (7, 28) points	15 (6, 27) points
Number of games > 20 pts	1016 (39.9%)	944 (38.9%)
Number of games > 40 pts	246 (9.7%)	225 (9.3%)

In simulated stopping of the testing set following the rule derived from the training set, we see a similar stopping rate:

	Training set (2012–14)	Testing set (2009–11)
Games stopped after Q1	46 (1.8%)	29 (1.2%)
Games stopped after Q2	396 (15.6%)	386 (15.9%)
Games stopped after Q3	700 (27.5%)	626 (25.8%)
Total games stopped	1142 (44.9%)	1041 (42.9%)

The error rate in the testing set exceeds the 99% target: 8 out of 1041 games (99.2%) stopped in simulation yielded a win ($n = 6$) or tie-to-win in real world play ($n = 2$).

8213erika/iStock/Thinkstock

I believe that a simplistic stopping rule, such as the one described in *Significance*, is ideal for implementation at the high school level: simply provide the referees with the point-differential threshold at Q1, Q2, and Q3; if the score exceeds the threshold, stop the game. A more sophisticated (and ultimately more powerful rule) would be to provide minute-by-minute stopping rules that could be consulted after each scoring play. And, of course, if the game is stopped before half-time, the band could still be summoned to perform a flourish in celebration of statistically sound, data-driven prevention of risk. Huzzah!
Michael Wininger
Connecticut, USA

Profiling the past

Google Scholar is a powerful, free resource for researchers. Not only is it a search engine dedicated to the scholarly literature, but also it provides researchers with their own editable web page listing their publications and metrics related to how often they have been cited. This latter information can be made public, and it is then searchable by name or research topic.

Google Scholar is clearly very useful, and all researchers should consider signing up to it. But it is particularly informative about famous past scientists, providing a convenient and comprehensive record of their output. For example, the eminent American statisticians Leo Breiman (1928–2005), William G. Cochran (1909–1980) and Frederick Mosteller (1916–2006) all have their own Scholar profiles. However, someone else must have set them up, as these individuals all died before Google Scholar became available.

Sadly, there appear to be no such profiles for deceased British statisticians. We propose that the Royal Statistical Society should take on the role of creating Scholar profiles for eminent past British statisticians, as a service to the statistical community. How about it, RSS?
Tim Cole and Mario Cortina Borja
University College London Institute of Child Health