IN DETAIL

© IWM

The flying bomb and the actuary

Liam P. Shaw and **Luke F. Shaw** follow in the footsteps of R. D. Clarke, a British actuary who sought to determine whether the apparent clustering of V-1 strikes on London during the Second World War was the result of targeting or random chance

In the early morning of 13 June 1944, one week after the D-Day landings, the Nazi regime launched a new weapon at London. The first *Vergeltungswaffe 1* (Vengeance Weapon 1, or V-1) hit a railway bridge in Mile End, killing six people and leaving 200 homeless. Over the following nine months, more than 2,300 "flying bombs" fell on London, killing an estimated 5,500 people.

The V-1 was the first operational cruise missile, capable of delivering an 850 kg warhead from a range of 250 km at any time, day or night. It was certainly a formidable weapon of terror, but if it could be accurately targeted at munitions factories it would also become a strategic threat to the war effort.

As the campaign progressed, there was an impression that the V-1s sometimes fell in clusters. But was this the result of random chance, or of precision guidance? This question was of critical importance. Discovering the answer means going on a journey through bomb maps, statistics textbooks, and even a novel – following in the footsteps of a British actuary.

A discreet actuary

In 1946, R. D. Clarke, a 31-year-old actuary working at the Prudential Assurance Company, published a one-page paper on V-1s.[1] In his paper he discreetly comments that the example he provides arose "in the course of a practical investigation" that he "recently had occasion to make". This is a coded acknowledgement that he had worked in military intelligence during the Second World War, analysing the distribution of V-1s for the Ministry of Aircraft Production (archive records list him as performing "special duties").

In his paper, he writes that, "During the flying-bomb [V-1] attack on London, frequent assertions were made that the points of impact of the bombs tended to be grouped in clusters." Humans are good at seeing patterns and ascribing reasons to them, but statistical analysis can help us decide whether those patterns may be due to chance. So, Clarke writes, "It was accordingly decided to apply a statistical test to discover whether any support could be found for this allegation."

Clarke describes how he first selected a 144 km^2 area where the mean number of V-1s seemed roughly constant. In total, there were 537 V-1 hits within the region. He divided the region up into smaller squares measuring 0.25 km^2 and counted the number of V-1 hits in each individual square. He then counted how many squares had no V-1s, how many squares had one V-1, and so on. Clusters would show up as squares with multiple V-1 hits.

One way to test the "allegation" that the V-1s tended to be clustered is to compare these results with what would arise from a random pattern. But what does "random" mean? A common-sense interpretation might be that every place is just as likely to be hit as everywhere else (what statisticians call "uniformly random"). Some level of clustering would still be expected due to chance, but how much? Nowadays, we might use a computer to simulate 537 random points on the map and see if we get a similar distribution, but Clarke did not have this option. So, how did he decide what "random" would look like? The solution lay in an article he had read two years previously in the *Journal of the Institute of Actuaries* by G. I. Lidstone, which discussed a particular statistical topic: the Poisson distribution.[2]

Finding the right distribution

The Poisson distribution as we know it dates from the nineteenth century and is often applied to thinking about events separated in time. As an example, consider standing in line at the self-checkout tills in a busy supermarket. Here, an "event" is a customer moving to an available till, and the "rate" (λ) is the average number of tills that become available per minute. This scenario can be modelled with the Poisson distribution: events happen randomly (there is no connection between one till becoming available and another till becoming available, as all shoppers are independent) but they happen at a constant average rate. The probability of observing k events in an interval if the average rate is λ is given by a defined function (the "probability mass function"), which we will simply call $P(\lambda, k)$.

The Poisson function $P(\lambda, k)$ can be used to answer questions about the expectation of events, such as "What is the probability that three tills become available in the next minute if the average rate is two per minute?" Answer: compute $P(\lambda = 2, k = 3)$.

Clarke's insight was that a hypothetical uniform but random distribution of V-1s across London meets the criteria for the Poisson distribution. Here an "event" is a V-1 landing in one of his squares (assumed independent of the other V-1s falling), but the average "density" (rather than "rate") of V-1s over the whole area is constant. It is the *same* problem as the Poisson distribution he had read about in Lidstone's article, but with space instead of time: it can be used to answer questions such as "What is the probability that three V-1s land in a square?" Clarke could therefore apply the Poisson function to predict the expected results if the V-1s were falling randomly.

Liam P. Shaw is a computational biologist in the Nuffield Department of Medicine, University of Oxford.

Luke F. Shaw is a data scientist in the Office for National Statistics Data Science Campus.

A textbook example

Assuming the V-1s were falling at random, Clarke estimated the average density as:

$$\lambda = \frac{N_{\text{V-1s}}}{N_{\text{squares}}} = \frac{537}{576} = 0.932$$

He then used the Poisson function to calculate the expected number of squares with $k = 0, 1, 2, 3, 4$, and 5 or more V-1 hits, and compared these expected values to his observed values. Even by eye it is apparent that the two sets of numbers are uncannily close (Table 1).

Clarke next performed a chi-squared test to determine how well the observed values fit those that would be expected if his specified null hypothesis was true – that is, if V-1s did indeed follow the Poisson distribution. The chi-squared test assigns a probability p (between 0 and 1) of seeing the observed results, or those more extreme, conditional on the null hypothesis being true. Clarke's test value was large ($p = 0.88$), so he concluded that the observed data were described very well by the Poisson distribution.

In the years that followed, Clarke's paper became widely known. In 1950 Professor W. Feller of Cornell University included it in a list of applications of the Poisson distribution in a textbook on probability, noting that the fit was "surprisingly good".[3]

However, Clarke's discretion meant he only provided the summary table of his results, rather than the original high-resolution map (perhaps because it was classified information at the time). Therefore, nobody has been able to directly repeat his test of randomness. Did V-1s falling on London really follow a Poisson distribution? We tried to find out.

The bomb maps

During the Second World War, the London County Council (LCC) kept bomb damage maps, meticulously recording all damage to buildings within the London County boundary.[4] The maps were regularly updated by hand, with buildings coloured according to bomb damage. The location of every V-1 hit was marked with an inked circle. These maps were republished in 2015 by the London Metropolitan Archives using stunning photographs of the originals, which are a sobering depiction of the catastrophic consequences of aerial bombing.

To analyse the data, we had to manually extract the V-1 locations. One of us (LPS) went through all 190 A3 pages of maps in the book and added every V-1 hit to a Google Maps layer (tiny.cc/london-V1). In total, the map contains 887 V-1s, including the first to hit London (coded V1.124 on the map) and one that destroyed the Institute of Actuaries hall on 24 August 1944 (V1.157).

As Londoners observed at the time, there are multiple apparent clusters (Figure 1 shows one such example). We downloaded the data set of V-1 locations for further analysis. What happens if we now apply Clarke's original test?

Repeating the randomness test 75 years on

It is apparent from looking at the distribution of V-1 strikes (Figure 2) that they were not completely uniform over the London County region. We cannot be sure, but we believe that the LCC maps do not include the entire area that Clarke used; the only information he gives is that it was in "south London". However, like him, we can pick a representative region where the density is "very nearly constant",[1] and repeat his randomness test.

Following Clarke, we choose a rectangular region of 144 km^2 of south London containing 532 total V-1 hits (compared to 537 in his paper) and divide it into 576 squares measuring

TABLE 1 Comparing the expected results if V-1s followed the Poisson distribution to the actual results over a region of south London. The expected number of squares is given by $P(\lambda, k) \times 576$ with $\lambda = 537/576$ and $k = 0, 1, 2, \ldots$. Clarke's reported p-value of 0.88 is for a chi-squared test with 4 degrees of freedom (DOF).[1] There are 6 classes in the table and we have estimated one parameter (λ), so DOF = (6 – 1) – 1 = 4.

Number of V-1s in square (k)	Expected number of squares (Poisson)	Observed number of squares
0	226.74	229
1	211.39	211
2	98.54	93
3	30.62	35
4	7.14	7
5 and over	1.57	1
	576	576

The original version of this article incorrectly stated that the degrees of freedom in Clarke's chi-squared test was 5. In fact it is 4, as Clarke wrote originally, because the value of λ is estimated from the observations. The authors thank Erik Holst for this correction.

FIGURE 1 A cluster of three V-1s near a railway line in Lewisham, south London. Patterns like these motivated Clarke to test whether their distribution was random. (a) A contemporary aerial photograph of the V-1 sites, showing the large radius of destruction of houses around the bomb sites (© IWM, catalogue no. CH 15109). (b) The same region from the LCC bomb damage maps[4] (p. 162), showing the three V-1 hits (black circles). Damage to houses is coded by colour, with black indicating total destruction. (c) The same region in the Google Maps layer (hit locations V1.509–V1.511).

© 2019 The Royal Statistical Society

0.25 km^2, counting up the squares with k V-1 hits. Using $\lambda = 532/576 = 0.923$ (almost identical to Clarke's value for λ), we see a very good fit to the Poisson distribution. The test gives $p = 0.70$ – not quite as clear-cut as Clarke's example, but still convincing.

We also have the option of using more modern computational methods. First, rather than choosing an arbitrary rectangular area for the test, we can sample from the whole area covered by the map. Second, we can sample with replacement as many times as we like (the basis of an approach known as "bootstrapping") rather than using fixed squares.

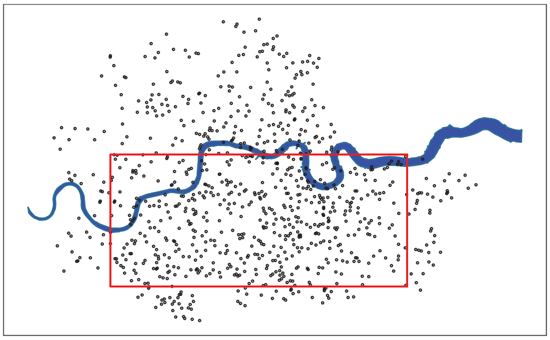
One issue is that we need to know the boundary of the London County: the LCC maps only record the 887 V-1s which fell within this boundary, so we do not have the complete distribution of the approximately 2,300 V-1s which fell over London as a whole. We do not want to accidentally sample from an area without data and incorrectly count it as an area where no bombs fell.

The London County no longer exists as an administrative region. But while we do not know its boundary exactly, we can recreate it from the set of 887 V-1 hits using a "hull", which defines the border around the points like the hull of a ship. We can then use this to check if a simulated circle is within the border and, if not, exclude it and try again.

Using this approach (Figure 3), the chi-squared test gives $p = 0.02$. This is much lower than either Clarke's or our original test result, assigning only a 2% probability to obtaining these values (or a worse agreement) if we assume that V-1s follow the Poisson distribution, meaning the data is a poor fit. So, did Clarke get it wrong?

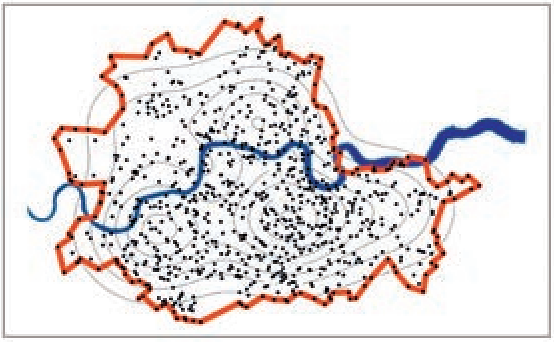
The problem is that the assumption of constant density required for the Poisson distribution does not now hold true. Overall, the distribution of V-1s was *not* uniformly random over London but was skewed to the south. We have found that Clarke's analysis is correct, but only within this southern region. He would have been well aware of this. During late 1944, British military intelligence had noticed that the V-1s tended to fall short of the centre of London, on less densely populated suburbs. To save lives, they therefore attempted to increase this shortfall using misinformation (bit.ly/2beLbvT).

While this fascinating story of deception is beyond the scope of this article, it is not the only instance of fiction related to the V-1. We turn now to a very unexpected application of the Poisson distribution.



Number of V-1s in square (k)	Expected number of squares (Poisson)	Observed number of squares
0	228.72	237
1	211.25	189
2	97.56	115
3	30.03	28
4	6.94	6
5 and over	1.50	1
	576	576

FIGURE 2 Repeating a version of Clarke's randomness test with the bomb damage map data. Black points indicate V-1 hits. White regions outside the central cluster of points are not included in the maps. Following Clarke, we therefore pick a rectangular region of 144 km^2 of south London and divide it into areas of 0.25 km^2, counting up the squares with k V-1 hits. The expected distribution from the Poisson function shows a good fit to the observed values (chi-squared test: $p = 0.70$).



Number of V-1s in circle (k)	Expected number of circles (Poisson)	Observed number of circles
0	469.07	508
1	355.09	306
2	134.40	121
3	33.91	51
4	6.42	14
5 and over	0.97	0
	1,000	1,000

FIGURE 3 Using the whole bomb map to apply Clarke's randomness test. We fit a boundary (red line) around the V-1 hits (black points) and sample 1,000 circles of area 0.25 km^2 each from within it. The results do not fit a Poisson distribution well with $\lambda = 0.757$ (lower than Clarke's value) because the density of the V-1s is not constant. Fitting the density of the V-1s using a technique called a Gaussian mixture model – which estimates the number of centred distributions which must be added together to explain the observed distribution – suggests three main centres of density (overlaid contours in grey). However, the data is limited by the boundaries of the London County map.

From the flying bomb to the rocket

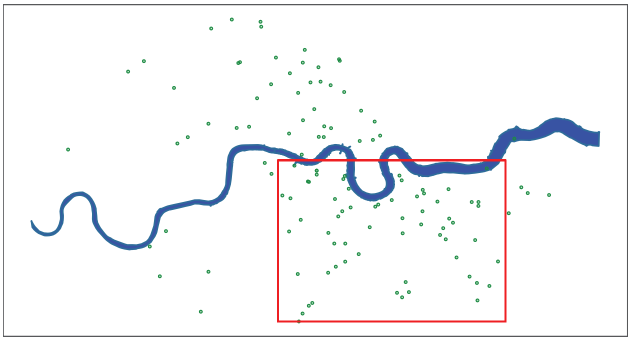
> The rockets are distributing about London just as Poisson's equation in the textbooks predicts … [the] little bureau is dominated now by a glimmering map … ruled off into 576 squares, a quarter square kilometer each. Rocket strikes are represented by red circles. The Poisson equation will tell, for a number of total hits arbitrarily chosen, how many squares will get none, how many one, two, three, and so on.[5]

Clarke anticipated his wartime statistics being included in textbooks, but probably not that they would be used in a postmodern novel. In *Gravity's Rainbow*,[5] the 1973 novel by Thomas Pynchon, the Poisson distribution is central to the plot. V-weapons are falling on London, and the statistician Roger Mexico realises that they are following a Poisson distribution. But there's a problem with the history: the V-weapons in *Gravity's Rainbow* are not V-1s – they are V-2s.

The V-2 or "rocket" was the second Vengeance Weapon: a ballistic missile which fell at supersonic speed, first fired at London in September 1944. Although neither Clarke's paper nor Feller's textbook mention V-1s and talk only of "flying bombs", it is unlikely that Pynchon simply made a mistake. More likely, he chose the V-2 because it suited his story better as a powerful image of random chance, falling without any warning whatsoever.

Pynchon is notoriously reclusive (he has appeared on *The Simpsons*, but only with a cartoon paper bag over his head), so we cannot ask him for more details. However, we can tell that he knew about Clarke's paper: the giveaway is those "576 squares, a quarter square kilometer each". Our guess is that he encountered the example as part of an introductory statistics course when he was studying physics and engineering at Cornell University in 1954. It seems likely that Feller's textbook would have been on the syllabus.

We doubt that Pynchon actually analysed the distribution of V-2s. However, the LCC maps do show the locations of V-2 hits, and we added these onto our Google Maps layer. We can now apply Clarke's test to V-2s, to see if Pynchon's literary conceit was correct. We have to use a slightly different region, as the density of the V-2s is clearly highly uneven over south London (Figure 4). Furthermore, V-2s are far rarer than V-1s (with $\lambda = 64/408 = 0.156$). Despite these problems, the test shows that the data fit the Poisson distribution ($p = 0.19$) – although not nearly as well as for Clarke's test of V-1s.



Number of V-2s in square (k)	Expected number of squares (Poisson)	Observed number of squares
0	348.77	356
1	54.71	41
2	4.29	10
3	0.22	1
4	0.01	0
5 and over	0.00	0
	408	408

FIGURE 4 Applying Clarke's randomness test to the V-2. We pick a rectangular region of 102 km^2 of south London, over which the density of V-2 hits is roughly constant, and divide it into squares measuring 0.25 km^2, counting up the squares with k V-2 hits. The chi-squared test gives $p = 0.19$.

How does this compare to the distribution in *Gravity's Rainbow*? At one point, Roger Mexico counts his squares: "A chance of 0.37 that … a given square on his map will have suffered only one hit, 0.17 that it will suffer two" (p. 55). From this second value for $k = 2$, we can tell from Table 1 that instead of Clarke's actual data (93/576 = 0.16 to two decimal places), Pynchon has used the theoretical Poisson function values (98.54/576 = 0.17). In *Gravity's Rainbow*, if not in real life, the distribution of the bombs is perfectly Poissonian.

Conclusion

> All of us in this Institute know that we are debtors to our profession. We are also debtors to those who have gone before.[6]

In our journey through the history of Clarke's textbook example, we have discovered many fascinating connections. But we should not lose sight of the real consequences of aerial bombing in the Second World War. In total on all sides, hundreds of thousands of civilians were killed. As humans, we see patterns; we want events to have explanations or narratives. The knowledge that who lived and who died was driven by random chance is hard to bear, no matter how conclusive the statistical tests.

In the Second World War, British statisticians like Clarke were not only analysing the distributions of bombs falling on London. Many were actively involved in optimising mass bombing campaigns against German cities like Hamburg and Dresden. Examining the bomb damage maps here should give us pause to contemplate the utter destruction caused by bombing civilians – not just in London in 1944, but in cities around the world, then and now. In studying the history of statistics, we are studying the history of humanity: our best, and our worst. ■

© IWM

Acknowledgements

We are extremely grateful to Laurence Ward and others at the London Metropolitan Archives for leading the publication of the LCC bomb damage maps for 1939–1945, including those who originally compiled them at the LCC; to Jeremy Ottevanger for the suggestion of searching the Imperial War Museum Archives to find suitable photographs; to John Porter, the group archivist at Prudential plc, for finding invaluable biographical information about R. D. Clarke; and to Quentin Shaw and Claire Hall for helpful feedback on a draft version of this article. We also thank Jeremy Bernstein and Freeman Dyson for comments. Finally, we thank our grandfather, Philip Shaw, for sharing his personal memories of the V-1 and V-2.

About this article

Shaw and Shaw's article is the winner of our 2019 Statistical Excellence Award for Early-Career Writing, awarded in partnership with the Young Statisticians Section of the Royal Statistical Society. Judges found the story to be gripping and well told – at times almost like a mystery novel in its approach. It wraps a history lesson around an introduction to probability, while showcasing the tools available to the modern data scientist and data analyst.

Congratulations to Liam and Luke, and congratulations also to our runners-up: Marco Antonio Andrade Barrera of the National Autonomous University of Mexico, for his article, "A story about a tiny bot", and Maximilian Aigner of the University of Lausanne for "Trouble in paradise: polarisation and the popular vote in Switzerland". These articles will be published in the coming weeks on our website, significancemagazine.com.

Details of our 2020 writing competition and award will be announced in February 2020.

References

1. Clarke, R. D. (1946) An application of the Poisson distribution. *Journal of the Institute of Actuaries*, **72**, 481.

2. Lidstone, G. I. (1942) Notes on the Poisson frequency distribution. *Journal of the Institute of Actuaries*, **71**, 284–291.

3. Feller, W. (1950) *An Introduction to Probability Theory and its Applications*, Volume I. New York: John Wiley & Sons.

4. Ward, L. (2015) *The London County Council Bomb Damage Maps 1939–1945*. London: Thames & Hudson.

5. Pynchon, T. (1973) *Gravity's Rainbow*. London: Cape.

6. Presentation of an Institute Silver Medal to Mr. Roland David Clarke (1981). *Journal of the Institute of Actuaries*, **108**, 7–8.