Fluctuations in the numbers of deaths may be represented as the outcome of a Poisson process

In many cases, the number of deaths fluctuates ‘randomly’ from period to period (e.g., from week to week, or from year to year). In such cases, the number of deaths can be represented as the outcome of a ‘Poisson process’. This is a process in which events occur at random, with the probability of an event occurring depending upon the underlying rate of occurrence (not upon how long it has been since a previous event, nor upon the number of events that have occurred in a recent period). For example, suppose that there are, on average, 10 such events per period, and no overall trend. Then, the underlying rate of occurrence is ten events per period, but random fluctuations could lead to the number of events in a series of seven periods being as follows: 10, 14, 8, 10, 7, 12 and 9: a total of 70 events in 7 periods, and therefore an average of 10 per period, but with the actual numbers in different periods varying randomly between 7 and 14.

In such cases, statistical theory can be used to calculate a likely range of random period-to-period variation in the figures, such as what statisticians call a ‘95% confidence interval’. Based on statistical theory, one would expect only about one period in twenty, on average, to have a figure outwith that range.

For the purpose of calculating such a range, it is assumed that the underlying rate of occurrence of deaths is represented by an appropriate average (e.g. the overall average number of deaths per period, or a moving average of the figures for, say, five consecutive periods - whatever seems reasonable in the particular case).

A characteristic of a Poisson distribution is that the values of the mean and the (statistical) variance are the same. The estimated underlying rate of occurrence can therefore be used as the estimate of the variance of the number of deaths, and its square root as the estimate of the standard deviation of the number of deaths.

A likely range of random variation around the expected number of deaths in a period can then be estimated using the estimated underlying rate for that period (which is the expected number) and the estimated standard deviation. The range can be calculated in a similar way to a 95% confidence interval - i.e.:

- if the underlying rate of deaths for the period is 100 or more, the Normal approximation is used - i.e. the range is simply the expected number of deaths (the underlying rate for the period) plus or minus twice the estimated standard deviation (i.e., +/- twice the square root of the underlying rate).
- if the underlying rate of deaths for the period is less than 100, the range (like an exact confidence interval) can be calculated using the inverse Chi-squared distribution, as a result of which:

    - the range is not symmetric about the expected number of casualties; and
    - in cases where the numbers are small, it is not possible for the lower limit of the range to have a value of less than zero.

Of course, deaths do not occur at random. For example, death rates are generally higher in bad winters, and the number of deaths in a given period for a particular part of Scotland could include several deaths which were caused by (say) a bad road accident, a fire in a care home or an outbreak of an infectious disease. Therefore, the number of deaths per period may fluctuate more than would be expected from a Poisson process, and it may well turn out to be outwith the ‘likely range’ in more than the expected 5% of periods.

NB: because it is easier to use the Normal approximation than to calculate exact confidence intervals, the formula ‘underlying rate +/- twice its square root’ is sometimes used to provide a rough indication of the likely range of random period-to-period variation in cases where the underlying rate is less than 100 deaths per period. The table below shows that the smaller the underlying rate is, the greater the extent (in proportionate terms) to which such a rough indication of the range will differ from an exact confidence interval calculated using the inverse Chi-square distribution. For example, when the underlying rate is just one death per period, the ‘rough’ lower limit is clearly unrealistic (as it is -1), whereas when the underlying rate is 100 deaths per period, the ‘rough’ limits differ from the exact limits by only 1-2%.

‘Likely ranges’ - ‘95% confidence intervals’
for numbers of events in a given period

Assuming that the number of events is the result of a Poisson process,
with the specified underlying rate per period
‘Rough’ limits: subtract / add twice the square root of the rate
‘Exact’ limits calculated using the inverse of the Chi-square distribution

Underlying rate per period	Rough lower limit	Rough upper limit	Exact lower limit (rounded)	Exact upper limit (rounded)
1	-1	3	0	6
4	0	8	1	10
9	3	15	4	17
16	8	24	9	26
25	15	35	16	37
30	19	41	20	43
40	27	53	29	54
50	36	64	37	66
60	45	75	46	77
70	53	87	55	88
80	62	98	63	100
90	71	109	72	111
100	80	120	81	122