

1 What do extra-poisson and less-than-poisson mean?

As JH explains in the follow-up video, extra-Poisson means ‘more than Poisson’, resulting in an SD that is (much) larger than the $\sqrt{\mu}$ that characterizes the Poisson distribution.

Less-than-Poisson means ‘less than Poisson’, resulting in an SD that is (much) smaller than the $\sqrt{\mu}$ that characterizes the Poisson distribution. An example might be the numbers of olives on pizzas, especially if the company insists that it be close to (say) 6 per pizza.

‘Close to Poisson’ would be if raisins were well mixed in the dough to make cookies. Extra would be if the raisin were clumped, and there were hot spots and empty parts of the dough. Less than would be if the raisins were added to each cookie in a systematic way.

2 How to find the confidence interval for counts without the canned functions?

Use the table provided in the Notes (it goes up to 25!)

If count is large enough for Normal approximation, then use the z-based interval using $SE = \sqrt{count}$.

If not, use the ‘hand’ method in Rothman, or the Notes.

3 What do you mean when you say that the Poisson (or binomial) tails have different distributions?

The Binomial distribution has a different shape at different π values along the (0,1) axis. If (when counts are small) you use 2 of these to arrive at lower and upper limits, they will result in a different ‘+’ than ‘-’ for the margins of error.

4 How to identify which is a Poisson distribution and which is not?

By thinking about it hard, and asking if there are extrinsic factors like weather or human behaviour or the like that would add variation.

(Even if population stays stable) the year to year variation in deaths is greater than Poisson because some years are bad flu years, etc.

The ideal is a radioactive source that produces disintegrations, or sampling from a liquid where the cells are very well mixed.

The 3 examples of ‘not’ covered in the Notes and video are daily births, ‘accidents’ in bus drivers and accidents in different Mondays in the Spring.

If you have data, such as those shown in the notes, you can check the SD vs. the mean, and the fitted and observed frequencies.

5 What are the assumptions for the Poisson distribution?

- Discrete events
- Greater than or equal to 0
- Can theoretically go to infinity

6 I’m uncertain based on the last examples when would Poisson be a good fit for the data? What are the steps we should walk through to check the fit?

See above. (1) logic and reflection (2) SD vs. \sqrt{mean} and (3) comparison of observed and expected frequencies.

7 In what cases would events be overestimated by a Poisson?

Not sure what you mean by ‘cases’. Do you mean ‘contexts’?

We don’t estimate events; we estimate rates, and do so via counts in known amounts of population-time.

Just like any fallible measurements, we can over-measure (include cases that don’t fit definitions) or under-measure (miss cases).

This is different from sampling variation, and its attendant overshooting or undershooting of the target.