Inference about a Population Rate (λ)

JH notes on rates

Sahir Bhatnagar and James Hanley

EPIB 607
Department of Epidemiology, Biostatistics, and Occupational Health
McGill University

sahir.bhatnagar@mcgill.ca
https://sahirbhatnagar.com/EPIB607/

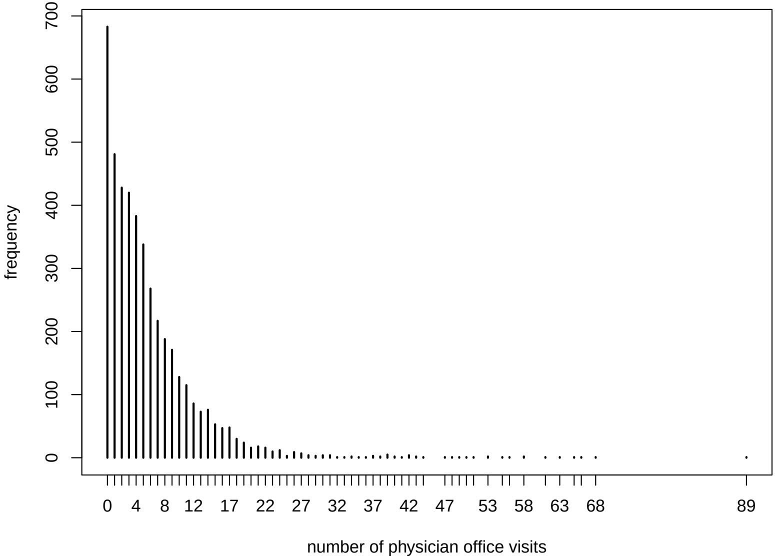
October 22, 2018

Poisson Model for Sampling Variability of a Count in a Given Amount of “Experience”

Motivating example: Demand for medical care

- Data from the US National Medical Expenditure Survey (NMES) for 1987/88
- 4406 individuals, aged 66 and over, who are covered by Medicare, a public insurance program
- The objective of the study was to model the demand for medical care - as captured by <u>the number</u> of physician/non-physician office and hospital outpatient visits - by the covariates available for the patients.

Motivating example: Demand for medical care



Some observations about the previous plot

- Discrete outcome $\rightarrow$ 1, 2, 3, … visits

Some observations about the previous plot

- Discrete outcome $\rightarrow$ 1, 2, 3, ... visits
- There are rare events, e.g. 1 individual with 89 visits

Some observations about the previous plot

- Discrete outcome $\rightarrow$ 1, 2, 3, ... visits
- There are rare events, e.g. 1 individual with 89 visits
- The data are far from normally distributed

Some observations about the previous plot

- Discrete outcome $\rightarrow$ 1, 2, 3, ... visits
- There are rare events, e.g. 1 individual with 89 visits
- The data are far from normally distributed
- Can theoretically go on forever

The Poisson Distribution

- The binomial distribution was derived by starting with an experiment consisting of trials or draws and applying the laws of probability to various outcomes of the experiment.

The Poisson Distribution

- The binomial distribution was derived by starting with an experiment consisting of trials or draws and applying the laws of probability to various outcomes of the experiment.
- There is no simple experiment on which the Poisson distribution is based, although we will shortly describe how it can be obtained by certain limiting operations.

The Poisson Distribution: what it is, and features

- The (infinite number of) probabilities $P_0, P_1, ..., P_y, ...,$ of observing $Y = 0, 1, 2, \ldots, y, \ldots$ events in a given amount of "experience."

The Poisson Distribution: what it is, and features

- The (infinite number of) probabilities $P_0, P_1, ..., P_y, ...,$ of observing $Y = 0, 1, 2, \ldots, y, \ldots$ events in a given amount of "experience."

- These probabilities, $P(Y = k) \rightarrow$ `dpois( )`, are governed by a single parameter, the mean $E[Y] = \mu$ which represents the expected **number** of events in the amount of experience actually studied.

The Poisson Distribution: what it is, and features

- The (infinite number of) probabilities $P_0, P_1, ..., P_y, ...,$ of observing $Y = 0, 1, 2, \ldots, y, \ldots$ events in a given amount of "experience."

- These probabilities, $P(Y = k) \rightarrow$ `dpois()`, are governed by a single parameter, the mean $E[Y] = \mu$ which represents the expected **number** of events in the amount of experience actually studied.

- We say that a random variable $Y \sim \text{Poisson}(\mu)$ distribution if

$$P(Y = k) = \frac{\mu^k}{k!} e^{-\mu}, \quad k = 0, 1, 2, \ldots$$

The Poisson Distribution: what it is, and features

- The (infinite number of) probabilities $P_0, P_1, ..., P_y, ...,$ of observing $Y = 0, 1, 2, \ldots, y, \ldots$ events in a given amount of "experience."

- These probabilities, $P(Y = k) \to$ `dpois()`, are governed by a single parameter, the mean $E[Y] = \mu$ which represents the expected **number** of events in the amount of experience actually studied.

- We say that a random variable $Y \sim \mathrm{Poisson}(\mu)$ distribution if

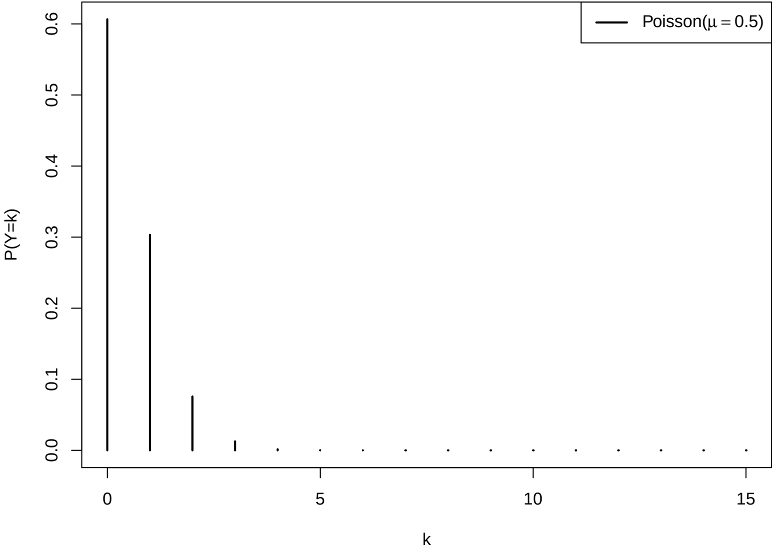
$$P(Y = k) = \frac{\mu^k}{k!} e^{-\mu}, \quad k = 0, 1, 2, \ldots$$

- Note: in `dpois()` μ is referred to as `lambda`

- Note the distinction between μ and λ
  - μ: expected **number** of events
  - λ: **rate** parameter

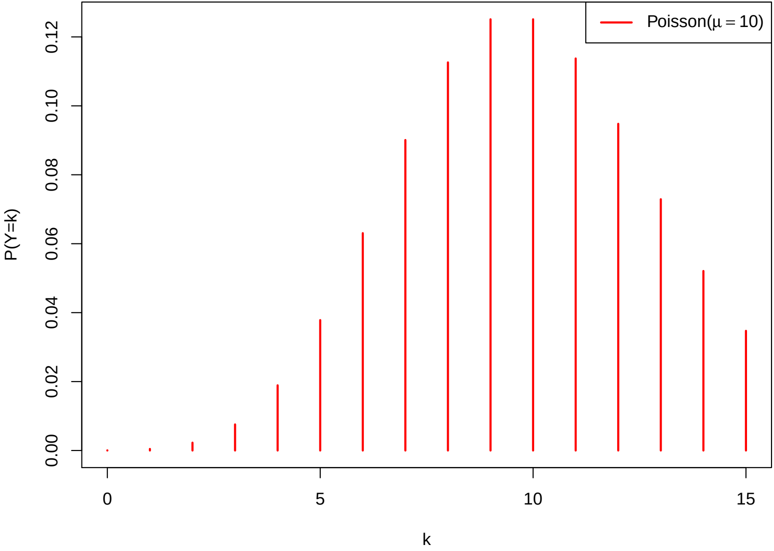
The probability mass function for $\mu = 0.5$

```r
dpois(x = 0:15, lambda = 0.5)
```

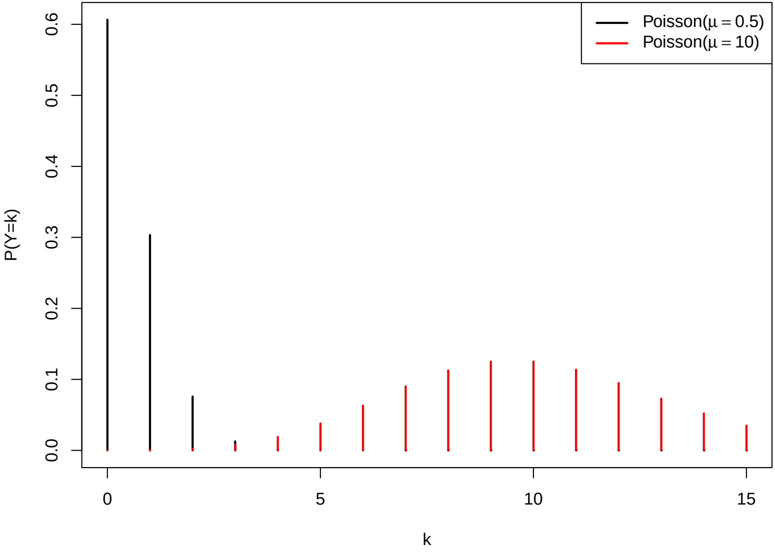


The probability mass function for $\mu = 10$

```r
dpois(x = 0:15, lambda = 10)
```



The probability mass function



The Poisson Distribution: what it is, and features

- $\sigma_Y^2 = \mu \ \rightarrow \ \ \sigma_Y = \sqrt{\mu}.$

The Poisson Distribution: what it is, and features

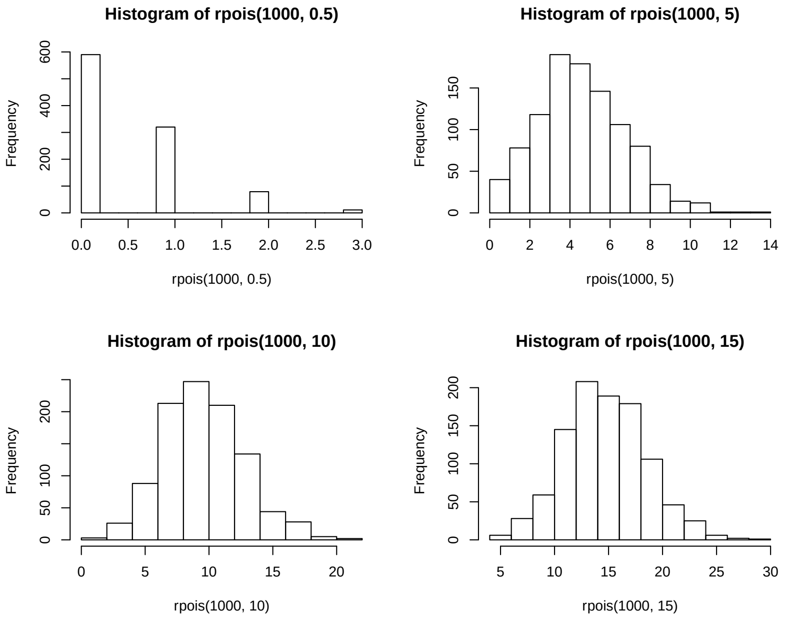
- $\sigma_Y^2 = \mu \ \rightarrow \ \ \sigma_Y = \sqrt{\mu}$.

- Approximated by $\mathcal{N}(\mu, \sqrt{\mu})$ when $\mu >> 10$

The Poisson Distribution: what it is, and features

- $\sigma_Y^2 = \mu \ \rightarrow \ \ \sigma_Y = \sqrt{\mu}.$
- Approximated by $\mathcal{N}(\mu, \sqrt{\mu})$ when $\mu >> 10$
- Open-ended (unlike Binomial), but in practice, has finite range.
- Poisson data sometimes called ”numerator only”: (unlike Binomial) may not “see” or count “non-events”

Normal approximation to Poisson is the CLT in action



How it arises

- Count of events or items that occur randomly, with low homogeneous intensity, in time, space, or ‘item’-time (e.g. person–time).

How it arises

- Count of events or items that occur randomly, with low homogeneous intensity, in time, space, or ‘item’-time (e.g. person–time).

- Binomial(n, π) when $n \to \infty$ and $\pi \to 0$, but $n \times \pi = \mu$ is finite.

How it arises

- Count of events or items that occur randomly, with low homogeneous intensity, in time, space, or ‘item’-time (e.g. person–time).
- Binomial(n, π) when $n \rightarrow \infty$ and $\pi \rightarrow 0$, but $n \times \pi = \mu$ is finite.
- $Y \sim Poisson(\mu_Y)$ if time (T) between events follows an $T \sim \mathrm{Exponential}(\mu_T = 1/\mu_Y)$. http://www.epi.mcgill.ca/hanley/bios601/Intensity-Rate/Randomness_poisson.pdf

How it arises

- Count of events or items that occur randomly, with low homogeneous intensity, in time, space, or 'item'-time (e.g. person–time).
- Binomial(n, π) when $n \to \infty$ and $\pi \to 0$, but $n \times \pi = \mu$ is finite.
- $Y \sim Poisson(\mu_Y)$ if time (T) between events follows an $T \sim \text{Exponential}(\mu_T = 1/\mu_Y)$. http://www.epi.mcgill.ca/hanley/bios601/Intensity-Rate/Randomness_poisson.pdf
- As sum of ≥ 2 *independent* Poisson random variables, with same **or different** μ's: $Y_1 \sim \text{Poisson}(\mu_1)$ $Y_2 \sim \text{Poisson}(\mu_2) \Rightarrow Y = Y_1 + Y_2 \sim \text{Poisson}(\mu_1 + \mu_2)$.

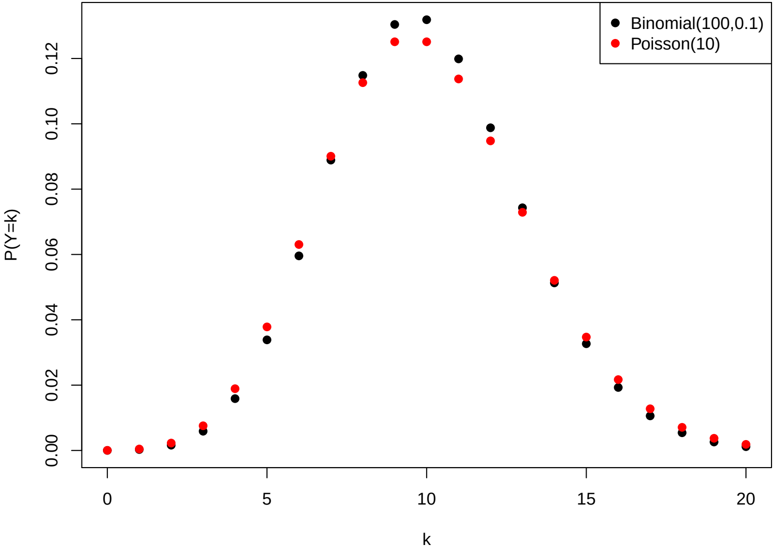
Poisson distribution as a limit

The rationale for using the Poisson distribution in many situations is provided by the following proposition.

Proposition 1 (Limit of a binomial is Poisson)

Suppose that $Y \sim Binomial(n, \pi)$. If we let $\pi = \mu/n$, then as $n \to \infty$, $Binomial(n, \pi) \to Poisson(\mu)$. Another way of saying this: for large n and small π, we can approximate the $Binomial(n, \pi)$ probability by the $Poisson(\mu = n\pi)$.

Poisson approximation to the Binomial



Examples

- numbers of asbestos fibres
- deaths from horse kicks*
- needle-stick or other percutaneous injuries
- bus-driver accidents*
- twin-pairs*
- radioactive disintegrations*
- flying-bomb hits*
- white blood cells
- typographical errors
- cell occupants – in a given volume, area, line-length, population-time, time, etc. [1]

[1] * included in
http://www.epi.mcgill.ca/hanley/bios601/Intensity-Rate/

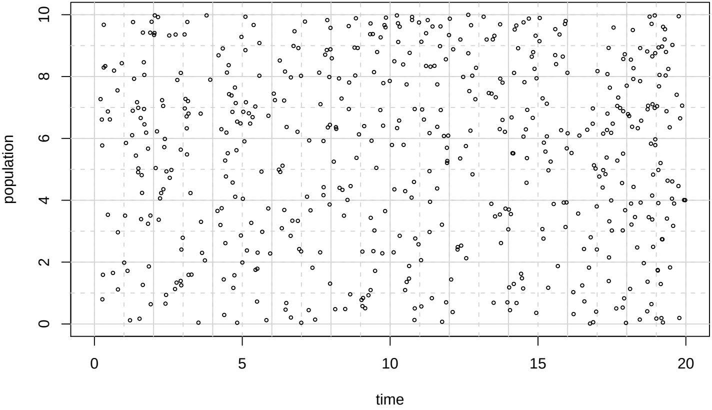


Fig.: Events in Population-Time randomly generated from intensities that are constant within (2 squares high by 2 squares wide) ‘panels’, but vary between such panels. In Epidemiology, each square might represent a number of units of population-time, and each dot an event.

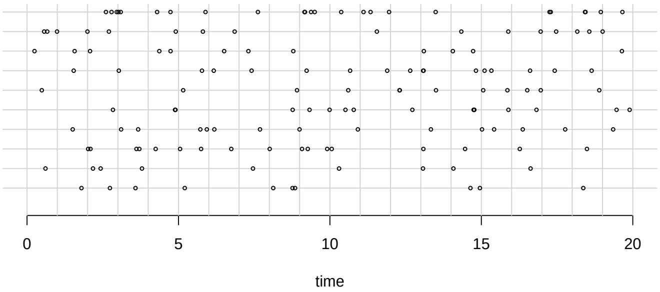


Fig.: Events in Time: 10 examples, randomly generated from constant over time intensities. Simulated with 1000 Bernoulli($_{\text{small}}\ \pi$)'s per time unit.

Does the Poisson Distribution apply to.. ?

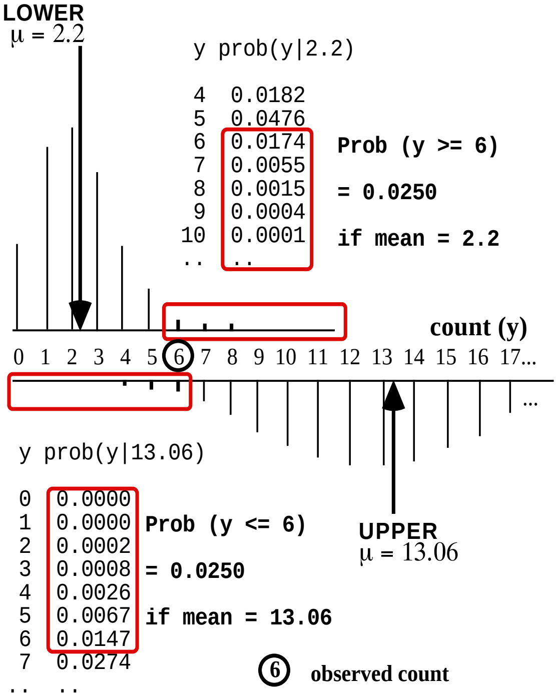
1. Yearly variations in numbers of persons killed in plane crashes
2. Daily variations in numbers of births
3. Weekly variations in numbers of births
4. Daily variations in numbers of deaths
5. Daily variations in numbers of traffic accidents
6. Variations across cookies/pizzas in numbers of chocolate chips/olives

Inference regarding μ, based on observed count y

Confidence interval for μ

- Instead of the usual "point-estimate $\pm$ some (z or t) multiple of standard error," a *first-principles* $100(1-\alpha)\%$ CI is the pair $(\mu_{LOWER},\ \mu_{UPPER})$ such that $P(Y \geq y \mid \mu_{LOWER}) = \alpha/2 \ \text{ and } \ P(Y \leq y \mid \mu_{UPPER}) = \alpha/2.$

- For example, as is shown in the Figure on the next slide, the 95% CI for μ, based on $y = 6$, is $\{\underline{2.20}, \underline{13.06}\}$.


LOWER
μ = 2.2
y prob(y|2.2)
4 0.0182
5 0.0476
6 0.0174
7 0.0055
8 0.0015
9 0.0004
10 0.0001
.. ..
Prob (y >= 6)
= 0.0250
if mean = 2.2
count (y)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17...
...
y prob(y|13.06)
0 0.0000
1 0.0000
2 0.0002
3 0.0008
4 0.0026
5 0.0067
6 0.0147
7 0.0274
.. ..
Prob (y <= 6)
= 0.0250
if mean = 13.06
UPPER
μ = 13.06
6 observed count

Confidence interval for μ

- For a given confidence level, there is one CI for each value of y.
- Each one can be worked out by trial and error, or – as has been done for the last 80 years – directly from the (exact) link between the tail areas of the Poisson and **Gamma** distributions.
- These CI's – for y up to at least 30 – were found in special books of statistical tables or in textbooks.
- As you can check, z-based intervals are more than adequate beyond this y. **Today**, if you have access to R (or `Stata` or SAS) you can obtain the first principles CIs directly **for *any* value of** y.

95% CI for mean count μ if we observe 6 events in a certain amount of experience

y	95%		90%		80%	
0	0.00	3.69	0.00	3.00	0.00	2.30
1	0.03	5.57	0.05	4.74	0.11	3.89
2	0.24	7.22	0.36	6.30	0.53	5.32
3	0.62	8.77	0.82	7.75	1.10	6.68
4	1.09	10.24	1.37	9.15	1.74	7.99
5	1.62	11.67	1.97	10.51	2.43	9.27
6	2.20	13.06	2.61	11.84	3.15	10.53
7	2.81	14.42	3.29	13.15	3.89	11.77
8	3.45	15.76	3.98	14.43	4.66	12.99
9	4.12	17.08	4.70	15.71	5.43	14.21
10	4.80	18.39	5.43	16.96	6.22	15.41
11	5.49	19.68	6.17	18.21	7.02	16.60
12	6.20	20.96	6.92	19.44	7.83	17.78
13	6.92	22.23	7.69	20.67	8.65	18.96
14	7.65	23.49	8.46	21.89	9.47	20.13
15	8.40	24.74	9.25	23.10	10.30	21.29
16	9.15	25.98	10.04	24.30	11.14	22.45
17	9.90	27.22	10.83	25.50	11.98	23.61
18	10.67	28.45	11.63	26.69	12.82	24.76
19	11.44	29.67	12.44	27.88	13.67	25.90
20	12.22	30.89	13.25	29.06	14.53	27.05
21	13.00	32.10	14.07	30.24	15.38	28.18
22	13.79	33.31	14.89	31.41	16.24	29.32
23	14.58	34.51	15.72	32.59	17.11	30.45
24	15.38	35.71	16.55	33.75	17.97	31.58

95% CI for mean count μ in R

- To obtain these in R we use the natural link between the Poisson and the *gamma* distributions.[2]

- In R, e.g., the 95% limits for μ based on $y = 6$ are obtained as
$\{\mu_L, \mu_U\} =$ `qgamma( c(0.025,0.975), c(6, 7) )`,
or, generically, for *any* y, as $\{\mu_L, \mu_U\} =$
`qgamma( c(0.025,0.975), c(y, y+1) )`.

- These limits can also be found using `stats::poisson.test` or (the less verbose) `survival::cipoisson` [both R functions use the *gamma* quantiles].

[2] details found here