Math 170E

Lecture Notes Section 2.7 [*†]

Poisson distribution

Instructor: Swee Hong Chan

NOTE: Materials that appear in the textbook but do not appear in the lecture notes might still be tested. Please send me an email if you find typos.

[*]Version date: Friday 29th January, 2021, 23:35.

[†]This notes is based on March Boedihardjo's and Jamie Haddock's notes from the past quarters, and I would like to thank them for their generosity. "*Nanos gigantum humeris insidentes* (I am but a dwarf standing on the shoulders of giants)".

1 Poisson distribution: Example

In a large city, telephone calls to Covid-19 hotline come on the average of 7 calls for every one hour.

Compute the probability that there is exactly one call to the hotline this hour.

2 Poisson distribution: Answer 1

The question only tells us that **mean** is 7. We need to make more assumptions for this question.

- Split this one-hour interval into **12** sub-intervals of equal length;
- Assume, for each sub-interval, there can be at most one call;
- Then number of phone calls in each sub-interval is a **Bernoulli trial** with $p = \frac{7}{12}$. (This means there is a call with probability p, and no phone call with probability $1 - p$.)

Let X be the number of phone calls in one hour. Then X is the **binomial RV** with $n = 12$ and $p = \frac{7}{12}$.

Recall that the pmf of a binomial distribution is

$$f(x) \quad = \quad \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x} \qquad \text{for } x \in \{0, 1, \ldots\}.$$

Hence, for $X = 1$, we have

$$P(X=1) = \frac{12!}{1!(12-1)!}\left(\frac{7}{12}\right)^1\left(\frac{5}{12}\right)^{11} \approx 0.00046.$$

3 Poisson distribution: Answer 2

Let's make a weaker assumption:

- Split this one-hour interval into **120** sub-intervals of equal length;
- We assume that, for each sub-interval, there can be at most one call;
- Then number of phone calls in each sub-interval is a **Bernoulli trial** with $\mathbf{p = \frac{7}{120}}$.

Under these assumptions, X is the binomial RV with $\mathbf{n = 120}$ and $\mathbf{p = \frac{7}{120}}$, so

$$P(X = 1) = \frac{120!}{1!(120-1)!} \left(\frac{7}{120}\right)^1 \left(\frac{113}{120}\right)^{119} \approx 0.0054.$$

4 Poisson distribution: Answer 3

Why stop with 120 sub-intervals?

- Suppose that there are now $\mathbf{n}$ sub-intervals of equal length;
- The number of phone calls in each sub-interval is a Bernoulli trial with $\mathbf{p} = \frac{\mathbf{7}}{\mathbf{n}}$.

Under these assumptions, X is the binomial RV with $\mathbf{n}$ and $\mathbf{p} = \frac{\mathbf{7}}{\mathbf{n}}$, so

$$\begin{aligned} P(X = 1) &= \frac{n!}{1!(n-1)!} \left(\frac{7}{n}\right)^1 \left(1 - \frac{7}{n}\right)^{n-1} \\ &= 7\left(1 - \frac{7}{n}\right)^{n-1}. \end{aligned}$$

Here are some values computed for different choices of n:

- For $n = 1,000$, we get $P(X = 1) \approx 0.006271873$;
- For $n = 10,000$, we get $P(X = 1) \approx 0.006372007$;
- For $n = 100,000$, we get $P(X = 1) \approx 0.006382057$;

5 Poisson distribution: Answer 4

We now let n increases to infinity.

Then the probability that $X = 1$ is

$$P(X = 1) \quad = \quad \lim_{n\to\infty} 7 \left(1 - \frac{7}{n}\right)^{n-1}.$$

Recall from calculus that

$$\lim_{n\to\infty} \left(1 - \frac{x}{n}\right)^{n} \quad = \quad e^{-x}.$$

Hence we have

$$\begin{aligned} P(X = 1) \quad &= \quad \lim_{n\to\infty} 7 \left(1 - \frac{7}{n}\right)^{n} \left(1 - \frac{7}{n}\right)^{-1} \\ &= \quad 7e^{-7} \quad \approx \quad 0.006383174. \end{aligned}$$

6 Poisson distribution: Answer 5

We are now interested with the probability that the number of calls in one hour is equal to nonnegative integer x.

Under the assumptions as before, X is the binomial RV with parameter n and $p = \frac{7}{n}$, so

$$P(X = x) \quad = \quad \frac{n!}{x!(n-x)!} \left(\frac{7}{n}\right)^x \left(1 - \frac{7}{n}\right)^{n-x}.$$

We now let n increases to infinity. Then

$$
\begin{aligned}
P(X=x) \quad &= \quad \lim_{n\to\infty} \frac{n!}{x!(n-x)!}\left(\frac{7}{n}\right)^x \left(1-\frac{7}{n}\right)^{n-x} \\
&= \quad \lim_{n\to\infty} \frac{n(n-1)\dots(n-x+1)}{x!}\frac{7^x}{n^x}\left(1-\frac{7}{n}\right)^{n-x} \\
&= \quad \lim_{n\to\infty} \frac{7^x}{x!}\ \frac{n}{n}\ \frac{n-1}{n}\ \dots\ \frac{n-x+1}{n}\ \left(1-\frac{7}{n}\right)^{n}\left(1-\frac{7}{n}\right)^{-x} \\
&= \quad \frac{7^x}{x!}\ 1\ 1\ \dots\ 1\ e^{-7}\ 1 \\
&= \quad \frac{7^x e^{-7}}{x!}.
\end{aligned}
$$

7 Poisson RV: Definition

The **Poisson random variable** X with **parameter** λ has the support and pmf

$$
\begin{aligned}
S &= \{0, 1, 2, 3, \ldots\}; \\
f(x) &= \frac{\lambda^x e^{-\lambda}}{x!} \qquad \text{for } x \in \{0, 1, 2, 3, \ldots\}.
\end{aligned}
$$

One should think of the Poisson RV as the **limit** of the binomial random variable as $n \to \infty$.

8 Poisson RV: Intuitive definition

- In a continuous interval of length 1, a finite number of points are randomly chosen;
- On average, λ points are chosen;
- For any two disjoint subintervals A and B, the number of points chosen from A is independent from the number of points chosen from in B (**most important assumption**);

The **Poisson random variable** X with **parameter** λ is the number of chosen points.

9 Poisson RV: Example Level 1

Flaws on a computer tape occured on the average of one flaw per 1200 feet.

Suppose that the number of flaws follow a Poisson distribution.

Compute the probability that there are no flaws in the 4800-foot roll.

Answer. Let X be the number of flaws in a 4800-feet roll.

Here the number of flaws is 4 per 4800 feet on average, so the rate λ is equal to 4. Hence

$$P(X = 0) \quad = \quad \frac{4^0 e^{-4}}{0!} \quad = \quad e^{-4} \quad \approx \quad 0.018.$$

$\square$

10 Poisson random variable: mgf, mean, and variance

Theorem 1. *Let X be the Poisson random variable with rate λ. The moment generating function of X is*

$$M(t) = e^{\lambda(e^t - 1)}.$$

The mean and variance of X are

$$\mu = \lambda; \qquad \sigma^2 = \lambda.$$

Please check the textbook for proofs.

11 Poisson RV: Self-replicating property

Let X be the number of points from a random process, chosen from an **interval of length 1**.

Suppose that X is a Poisson RV with parameter λ.

Theorem 2. *Now, let Y be the number of random points from the same process, but chosen from an* ***interval of length*** *t. Then Y is a Poisson RV with parameter $\lambda\mathbf{t}$.*

12 Poisson RV: Example Level 2

Let X be the number of falling stars in the sky above Los Angeles in one second.

Suppose that X is a Poisson random variable with the mean rate of 60 stars per second.

Let Y be the number of falling stars in the sky above Los Angeles in 5 seconds.

What is the probability that Y is less than or equal to 2?

13 Poisson RV: Example Level 2

Since there are 60 falling stars per second in the sky on average, this means there are 300 falling stars per 5 seconds on average.

By the self-replicating property, Y is the Poisson random variable with parameter $\lambda = 300$.

Thus the answer to the question is

$$\begin{aligned} P(Y \leq 2) &= P(Y=0) + P(Y=1) + P(Y=2) \\ &= \frac{(300)^0 e^{-300}}{0!} + \frac{(300)^1 e^{-300}}{1!} + \frac{(300)^2 e^{-300}}{2!} \\ &= 2.33 \times 10^{-126}. \end{aligned}$$

14 Approximating binomial distribution

Theorem 3. *Let X be the binomial distribution with parameter n and p.*

Let $\lambda = np$, and let Y be the Poisson distribution with mean rate λ.

For very large n and very small p, the probability $P(X = x)$, which is ***equal to***

$$\binom{n}{x} p^x (1-p)^{n-x},$$

can be ***approximated by***

$$P(Y = x) \quad = \quad \frac{\lambda^x e^{-\lambda}}{x!}.$$

15 Approximating binomial distribution: Example

A lightbulb company knows that 0.01% of its lightbulbs are flawed.

Let X be the number of flawed bulbs in a box of 10,000.

Approximate the probability that $X = 3$.

16 Approximating binomial distribution: Answer

Note that X has a binomial distribution with $n = 10,000$ and $p = 0.0001$.

It can be unrealistic to compute the probabilities

$$P(X = 3) \quad = \quad \binom{10000}{3}(0.0001)^3(0.9999)^{9997}.$$

Instead, let $\lambda = np = 1$, and we approximate X by the Poisson random variable with rate $\lambda = 1$, for which the probability is

$$P(Y = 3) \quad = \quad (1)^3\frac{e^{-1}}{3!} \quad \approx \quad 0.061.$$